

# Current Approaches and Challenges to Development of an Individualized Sleep and Performance Prediction Model

Erik Olofsen<sup>1</sup>, Hans P.A. Van Dongen<sup>2</sup>, Christopher G. Mott<sup>3</sup>, Thomas J. Balkin<sup>4</sup> and David Terman<sup>\*,5</sup>

<sup>1</sup>Department of Anaesthesiology, Leiden University Medical Center, Leiden, The Netherlands

<sup>2</sup>Sleep and Performance Research Center, Washington State University, Spokane, WA, USA

<sup>3</sup>Pulsar Informatics Inc., Norfolk, VA, USA

<sup>4</sup>Department of Neurobiology and Behavior, Walter Reed Army Institute of Research, Silver Spring, MD, USA

<sup>5</sup>Department of Mathematics, The Ohio State University, Columbus, OH, USA

**Abstract:** Workplace safety and productivity will be enhanced considerably with the development and application of individualized sleep and performance prediction models. These are models that predict an individual's operational performance based on his/her unique sleep schedule, individual sleep requirements, and individual pattern of responses to sleep loss across a variety of cognitive performance domains. Progress in the individualization of such models will occur as the result of integrated efforts based on (a) an expanding understanding of the relevant physiological processes underlying the sleep/circadian/performance interactions as well as (b) novel empirical and statistical approaches. In the present paper, an overview is presented of the state of the art of the individualization of sleep and performance models with sections on current efforts in model integration, the application of Bayesian forecasting techniques to the problem of model individualization, construction of Bayesian confidence intervals for predicted performance, and the problem of generalizability of individualized model predictions – i.e., the problem of using models constructed with performance data from one cognitive domain to predict performance in another cognitive domain. Success in model individualization will ultimately be facilitated by concerted, coordinated efforts involving multiple scientific entities and stakeholders.

**Keywords:** Fatigue and performance modeling, flip-flop sleep switch, individualized prediction, generalizability to other metrics, Bayesian confidence intervals, noise sources.

## INTRODUCTION

The present paper, which springs forth from the Air Force Office of Scientific Research (AFOSR) sponsored meeting “New Approaches to Modeling Sleep/Wake Dynamics and Cognitive Performance” (October 26–27, 2006; Columbus, OH, USA), presents an overview of mathematical modeling approaches to predicting the performance capacity of individuals during sleep loss.

For over a century of sleep deprivation research [1], a primary focus has been on cataloguing the behavioral and physiological effects of sleep loss (e.g., [2]). In the classic sleep deprivation paradigm, a group of subjects provides baseline measurements on a battery of tests (at specific times of day, to control for circadian rhythm effects on alertness and performance). This is followed by one or more administrations of a test battery (again at the specified times of day) across one or more days of total sleep deprivation. Results are then reported as mean effects of sleep loss on the dependent variables of interest.

Although it has long been recognized that individuals vary significantly in their ability to maintain alertness and performance during periods of sleep loss (e.g., [3,4]), such individual differences are seen merely as a source of error variance in the typical sleep deprivation study/analysis. However, with recent efforts to develop mathematical models that can be used in operational environments (see [5]) to predict performance deficits, optimize work/rest schedules, and maximize the efficacy of fatigue countermeasures (e.g., through optimal timing and dosing [6]), it is clear that systematic individual differences in resilience to the effects of sleep loss need to be moved from the realm of error variance to the forefront of efforts to understand, quantify, and predict the consequences of sleep loss. This involves identifying factors that determine inter- and intra-individual variability, and incorporating them in mathematical prediction models. The origin of these factors may be genetic, environmental, psychological, motivational, drug use related, etc., and may or may not be different between the sexes. The practical utility of sleep/performance prediction models will ultimately depend upon the extent to which they accurately predict the performance capacity of each individual in the operational environment, e.g., in small-team operations where each individual's performance contribution is important for overall mission success. It is likely that the ability to predict the effects of sleep loss on the individual will be a leading yard-

\*Address correspondence to this author at the Department of Mathematics, 231 W. 18th Avenue, The Ohio State University, Columbus, OH 43210-1174, USA; Tel: +1-614-292-5285; Fax: +1-614-292-1479; E-mail: terman@math.ohio-state.edu

stick against which progress in sleep/performance model development is measured over the next several years.

There are two general approaches to facilitate development of mathematical models to predict the performance capacity of individuals during sleep loss. The first involves basic research with the aim of discovering the physiological underpinnings of individual differences in resilience during sleep loss. From this knowledge, specification of an individual's position within the "sleep deprivation resilience-sensitivity distribution" would be possible, which would in turn allow appropriate adjustments to be made to model parameters to optimize predictions on an individual-by-individual basis. Furthermore, such research could have the potential added benefit of contributing to general knowledge about the physiological underpinnings of the sleep homeostat and/or circadian rhythm functioning and modulation.

The second approach involves development and validation of techniques in which the aim is to enhance prediction at the level of the individual, but without (or with incomplete) understanding of the physiological basis of individual differences in susceptibility to sleep loss. This approach has the advantage of being more efficient – i.e., with a greater likelihood of producing valid, useful sleep/performance models within a shorter time frame, and at less expense. And since it has been shown that resilience/sensitivity to sleep loss is a trait-like characteristic that remains stable over time [7], a process in which individualized model parameters are determined on the basis of an empirical, initial test of susceptibility to sleep loss could be viable.

The effects of caffeine, like those of sleep deprivation, on cognitive performance show large inter-individual differences that are trait-like and genetically determined [8]. Also, like with vulnerability to sleep deprivation, there are changes in sensitivity to caffeine as a function of age [9]. But, at least in part, the same adenosinergic mechanisms underlie the effects of sleep deprivation and those of caffeine [10]. This instills some confidence that modeling approaches successfully predicting cognitive responses to sleep deprivation may be readily adaptable to include prediction of caffeine effects.

Although preliminary evidence suggests that individual differences in performance impairment during sleep deprivation are predictable by baseline brain metabolism as measured by functional magnetic resonance imaging (fMRI) [11-13] or polymorphisms in genes related to sleep and circadian rhythms [14,15], to date no practicable predictors of susceptibility to sleep loss have yet been identified [16].

In the absence of baseline measures, it is still possible to predict performance of an individual without knowledge of that individual's unique physiology, by means of statistical approaches that make use of prior information regarding the distribution of responses to sleep deprivation across the population. Such approaches, which invariably involve some aspect of Bayesian statistical modeling, utilize a small amount of performance information about the individual at hand, and relate that to the variability in performance profiles previously obtained from a sample of the population (e.g., in a laboratory or field experiment) – see [17]. Given that performance responses to sleep loss involve a trait [7], the performance of a specific individual observed *during* a sustained operation (i.e., when little or no sleep is obtained),

combined with the prior information estimated from the population, allows for reliable prediction of future performance.

Both approaches have advantages and disadvantages. A general disadvantage of the statistical approach is that because it does not necessarily result in increased insight regarding the physiology underlying susceptibility to sleep loss (i.e., a physiological process that, if discovered and understood, could itself be modeled), the potential for breakthroughs that might result in significant leaps in model development is relatively limited. Breakthroughs may come from physiological modeling (e.g., [18]); however, many fundamental issues associated with the nature of sleep are poorly understood. Despite considerable progress in neurobiological research [19-21], the sleep homeostat and how it interacts with the circadian process remains mysterious. Moreover, it is not known how individual differences are reflected in these processes. For this reason, it is difficult at this time to envision construction of a model that is detailed enough to account for susceptibility to sleep loss, but simple enough to be useful.

Thus, in the long run, there is a need to invest in both approaches: physiology-based modeling to inform understanding of the principles underlying individual variability, and statistics-based modeling to deal with whatever relevant physiological questions remain unresolved. Of course, these two approaches are not mutually exclusive. Even a partial understanding of some component of the physiology may suggest new ways to formulate statistical models. Conversely, insights from the statistical models may also help constrain the physiological models and suggest new hypotheses that can be tested experimentally.

Presented in the next section is a neurobiological model of the human sleep/wake cycle that integrates the two-process model of sleep regulation [22] with the flip-flop sleep switch model [23]. This results in a platform that could conceivably be used to identify the physiological parameters that underlie individual differences in resilience during sleep deprivation. In the following section, a statistical (Bayesian) approach to addressing the challenges associated with constructing an individualized model (when knowledge of the underlying, relevant physiology is unavailable or insufficient) is described. In the next section, the focus is on the factors that impact the width and the ultimate utility of (Bayesian) confidence intervals as applied to individualized performance prediction models. This is followed by a section in which the generalizability of predictions across different performance metrics is examined. Finally, in the discussion, the various current challenges to the prediction of performance in individual subjects are summarized, and suggestions for addressing these challenges are provided.

## A NEUROBIOLOGICAL MODEL OF THE HUMAN SLEEP/WAKE CYCLE (TERMAN)

Current models for sleep timing and for interactions between the circadian and homeostatic processes are generally based on the two-process model of sleep regulation [22]. In this model, two processes, *S* and *C*, regulate sleep. The *S* variable represents a homeostatic process that accrues during waking hours and decays exponentially during sleep. It is correlated with slow-wave EEG activity (delta power) during

sleep. The other variable,  $C$ , represents the circadian process that is typically presumed to be independent of the homeostatic process. The sum of  $S$  and  $C$  represents the total sleep propensity, which in many models is also considered a direct reflection of the level of waking performance impairment. Sleep episodes begin and end spontaneously when the sleep propensity falls above or below one of two fixed thresholds.

The two-process model has been successful in accounting for features of the sleep/wake cycle; however, it has been less successful in predicting waking performance, especially under conditions of chronic partial sleep restriction [24,25]. Also, the success of the two-process model in predicting the timing of wake and sleep comes despite a lack of a detailed understanding of the biological mechanisms that underlie the observed sleep dynamics. It remains a mystery why the timing of waking up and falling asleep should occur when some combination of the (poorly understood) homeostat and circadian drive reaches two somewhat arbitrarily constructed thresholds.

In the present paper, we propose a more detailed, biological model for the sleep/wake cycle and interactions between the circadian and homeostatic processes – one that reproduces many of the features of the two-process model. Because this model is based on findings from recent neurophysiological experiments, it might prove useful as a platform for formulation and testing of hypotheses regarding the physiological underpinnings of individual differences in performance during sleep loss.

This mathematical model is based on the flip-flop sleep switch model of sleep/wake regulation [26], in which it is proposed that sleep-promoting neurons within the ventral lateral preoptic nucleus (VLPO) and wake-promoting neurons in the monoaminergic cell groups (AMIN) inhibit each other, resulting in both stable wakefulness and sleep. Here we describe the mathematical model, and demonstrate that features of the two-process model are successfully reproduced. In fact, using singular perturbation methods, it can be shown that, in a rigorous mathematical sense, the present model can be reduced to the two-process model. Details of this analysis are given in [26].

A schemata of our model for the sleep-wake switch is shown in Fig. (1A). The model includes the sleep-promoting neurons (VLPO), the wake-promoting monoaminergic cell groups (AMIN), orexin neurons (ORX), a circadian pacemaker corresponding to activity within the suprachiasmatic nucleus (SCN), and input from cortical areas that depend on sleep/wake homeostasis. Note that the sleep-promoting and wake-promoting cells inhibit each other. We assume that output of SCN inhibits VLPO. While experiments have demonstrated that direct connections from SCN onto VLPO are sparse, SCN does project onto other cells groups including the dorsomedial hypothalamic nucleus, and this has an inhibitory effect on VLPO [27]. In our model, AMIN receives excitatory input from ORX, which, in turn, is excited by SCN; moreover, ORX is inhibited by VLPO [28].

The VLPO and AMIN are each modeled as a system of two ordinary differential equations. The equations for AMIN can be written as

$$\delta_A x_A' = f(x_A, y_A) - I_{VLPO} + I_{ORX} + I_{CTX}, \quad (1)$$

$$y_A' = g(x_A, y_A), \quad (2)$$

while the equations for VLPO can be written as

$$\delta_V x_V' = f(x_V, y_V) - I_{AMIN} - I_{SCN} + I_{CTXV}, \quad (3)$$

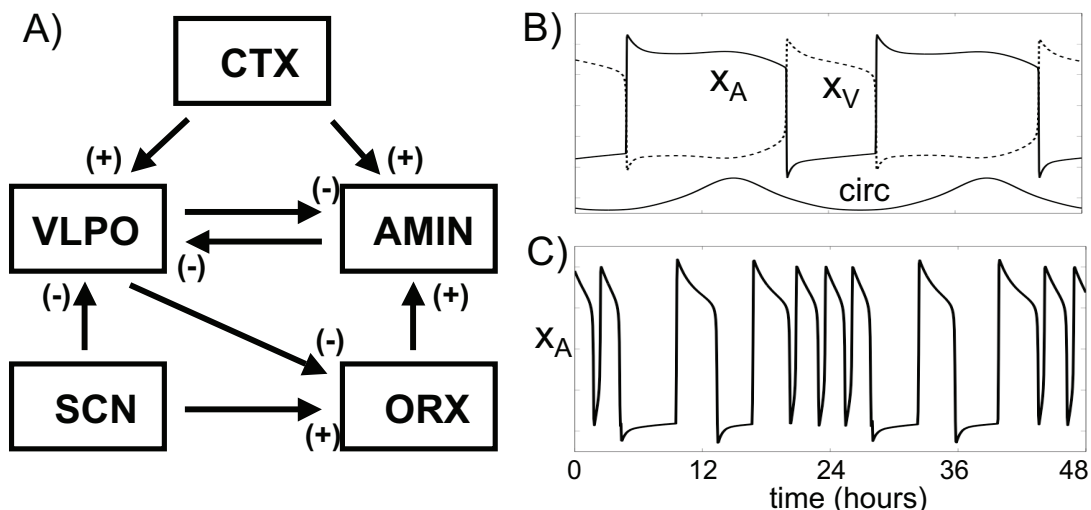
$$y_V' = g(x_V, y_V). \quad (4)$$

Here,  $x_A$  and  $x_V$  represent the overall population activity of the wake-promoting and sleep-promoting cells, respectively,  $y_A$  and  $y_V$  are recovery variables, and  $\delta_A$  and  $\delta_V$  are constants. The nonlinear functions  $f$  and  $g$  are of the form

$$f(x, y) = 3x - x^3 + 2 - y, \quad (5)$$

$$g(x, y) = \varepsilon (\gamma H_\infty(x) - y) / \tau(x), \quad (6)$$

where  $\varepsilon$  and  $\gamma$  are constants.  $H_\infty(x) = 0.5 \tanh(x / 0.01)$  is a smooth approximation of the Heaviside step-function; that is,  $H_\infty(x) \approx 0$  if  $x < 0$  and  $H_\infty(x) \approx 1$  if  $x > 0$ . Furthermore,  $\tau(x)$  is a step function that is of the form  $\tau(x) = \tau_1 + (\tau_2 - \tau_1) H_\infty(x)$ . The precise parameter values can be found in Rempe *et al.* [26].



**Fig. (1).** A. The network model. B. A solution under normal sleep/wake conditions. C. Without ORX input, the network exhibits unstable switching between the wake and sleep states.

$x_A$ , overall population activity of the wake-promoting cells;  $x_V$ , overall population activity of the sleep-promoting cells; circ, SCN output.

The terms  $I_{VLPO}$ ,  $I_{AMIN}$ ,  $I_{SCN}$ , and  $I_{ORX}$  correspond to the inputs from VLPO, AMIN, SCN and ORX, respectively. We assume that

$$I_{VLPO} = g_{VLPO} H_{\infty}(x_V), \quad (7)$$

$$I_{AMIN} = g_{AMIN} H_{\infty}(x_A), \quad (8)$$

$$I_{SCN} = g_{SCN} C(t), \quad (9)$$

where  $g_{VLPO}$ ,  $g_{AMIN}$  and  $g_{SCN}$  are constants.  $C(t)$  represents the circadian pacemaker, which is modeled as a sinusoid with harmonics (see [29]). The orexin neurons receive excitatory input from SCN and inhibitory input from VLPO. Experiments have demonstrated that these cells are typically inactive while asleep [28]. To simplify the model, we do not include a separate differential equation for ORX. Instead, we assume that

$$I_{ORX} = I_{SCN} (1 - H_{\infty}(x_V)). \quad (10)$$

This implies that ORX is silent while VLPO is active; on the other hand, if VLPO is silent, then ORX follows SCN.

Finally,  $I_{CTXA}$  and  $I_{CTXV}$  correspond to input from the cortex (CTX) to AMIN and VLPO, respectively. We assume that

$$I_{CTXA} = I_A - g_{\text{hom}} h, \quad (11)$$

$$I_{CTXV} = I_V + g_{\text{hom}} h, \quad (12)$$

where  $I_A$  and  $I_V$  represent some background cortical drive,  $g_{\text{hom}}$  is constant, and  $h$  represents sleep/wake homeostasis. In the model,  $h$  decays at some exponential rate while the system is ‘asleep’ and increases similarly while ‘awake’. The variables can be scaled so that the system is defined to be asleep if  $x_A < 0$  and awake if  $x_A > 0$ . Then  $h$  satisfies equations of the form  $h' = a_h (h_{\text{max}} - h)$  while awake and  $h' = -b_h h$  while asleep. Here,  $a_h$ ,  $b_h$  and  $h_{\text{max}}$  are constants. Note that while awake, cortical drive to AMIN decreases while cortical drive to VLPO increases; that is, there is increasing pressure to fall asleep. This model can be considered a biology-based extension of the two-process model of sleep regulation [22] that includes a mathematical biology for the sleep/wake and wake/sleep transitions.

A solution of the model is shown in Fig. (1B). In this solution, the system ‘sleeps’ for approximately eight hours during the trough of the circadian cycle. For the solution shown in Fig. (1C), input from ORX to AMIN has been removed. This produces rapid, unstable switching between the sleep and wake states – similar to that observed in SCN-lesioned rodents [30] – a model feature that is consistent with the hypothesis that input from orexin/hypocretin neurons acts to stabilize the sleep-wake switch [23].

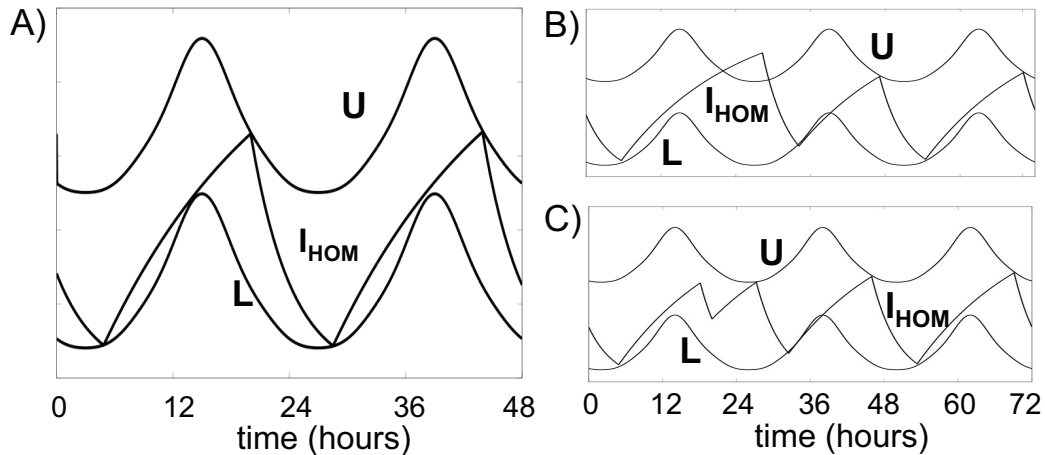
Fig. (2A) shows the sleep/wake homeostatic input  $I_{\text{HOM}} = g_{\text{hom}} h$  along with two other curves that are given by  $L = I_{\text{SCN}} + K_A$  and  $U = I_{\text{SCN}} + K_B$ , where  $K_A$  and  $K_B$  are constants. We will explain later how these constants are chosen. Note that the system falls asleep when  $I_{\text{HOM}}$  crosses the lower threshold  $L$  and wakes up when  $I_{\text{HOM}}$  crosses the upper threshold  $U$ . This behavior is just like the two-process model of sleep regulation [29]. For the solution shown in Fig. (2B), 24 hours of continuous wakefulness was simulated *via* an increasing cortical drive. As before, the network is designed to initiate sleep when  $I_{\text{HOM}}$  crosses  $L$  and to awaken when  $I_{\text{HOM}}$  crosses  $U$ , but in this example extended wakefulness is maintained by increasing AMIN activation. In Fig. (2B), the network is put to sleep for a two hour nap and then awakened by changing the cortical drive to VLPO and AMIN. The network then falls asleep when  $I_{\text{HOM}}$  crosses  $U$  and then wakes up when  $I_{\text{HOM}}$  crosses  $L$ .

A mathematical analysis of the model helps to explain the consistency of the model presented here with the two-process model. We analyze the model using phase plane analysis. To illustrate this method, we begin by assuming that the total inputs to each of the cell groups, AMIN and VLPO, are constant. That is, let

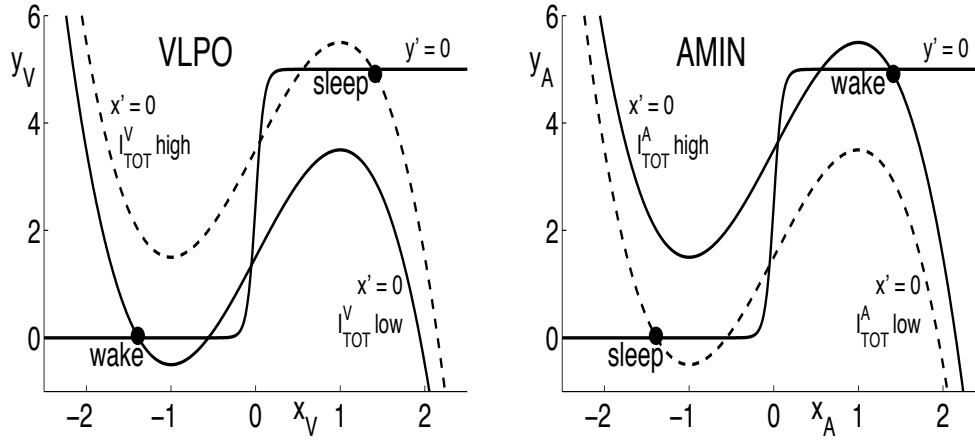
$$I_{\text{TOT}}^A = -I_{VLPO} + I_{ORX} + I_{CTXA}, \quad (13)$$

$$I_{\text{TOT}}^V = -I_{AMIN} - I_{SCN} + I_{CTXV}, \quad (14)$$

and assume, for the moment, that both of these are constant in time – i.e., we consider the system in a ‘frozen’ state. The models for AMIN and VLPO are then two-dimensional systems and we can consider the corresponding phase planes.



**Fig. (2).** A. The sleep homeostatic input  $I_{\text{HOM}}$  oscillates between an upper curve  $U$ , when the system falls asleep, and a lower curve  $L$ , when the system wakes up. The thresholds are modulated by circadian input. B. The network is forced to stay awake for 24 hours by increasing the cortical drive to AMIN for 24 hours. C. The network is put to sleep at time 18 (i.e., 6 PM) by decreasing the cortical drive to AMIN. After two hours, the network is awakened by both increasing the cortical drive to AMIN and decreasing the drive to VLPO. The parameters are then reset to their default values.



**Fig. (3).** Phase planes for the separate VLPO and AMIN systems when there are constant total inputs. If the total inputs to VLPO and AMIN are high and low, respectively, then there are stable fixed points along the right and left branches of the respective cubic-shaped  $x_V$ - and  $x_A$ -nullclines (dashed curves). In this case the system is said to be ‘asleep’. If the total inputs to VLPO and AMIN are low and high, respectively, then there are stable fixed points along the left and right branches of the respective  $x_V$ - and  $x_A$ -nullclines (solid curves). In this case the system is said to be ‘awake’.

Fig. (3) illustrates two possibilities. If  $I_{TOT}^V$  is high and  $I_{TOT}^A$  is low, then there is a stable fixed point along the right branch of the cubic-shaped  $x_V$ -nullcline and a stable fixed point along the left branch of the  $x_A$ -nullcline. This corresponds to the system being ‘asleep’. If  $I_{TOT}^V$  is low and  $I_{TOT}^A$  is high, then there is a stable fixed point along the left branch of the  $x_V$ -nullcline and a stable fixed point along the right branch of the  $x_A$ -nullcline. This corresponds to the system being ‘awake’.

We now consider the full system and step through a complete sleep/wake cycle. The total inputs to AMIN and VLPO change as the circadian and homeostatic inputs change. From Eqs. (1) and (3), we see that changes in the total inputs will either raise or lower the  $x_A$ - and  $x_V$ -nullclines. As these nullclines change, stable fixed points of the frozen system may appear or disappear through bifurcations. As we shall see, these bifurcations are responsible for the system either ‘waking up’ or ‘falling asleep’.

We begin with the system ‘asleep’ so that VLPO is at a stable fixed point along the right branch of the  $x_V$ -nullcline and AMIN is at a stable fixed point along the left branch of the  $x_A$ -nullcline. While asleep, the cortical drive to VLPO decreases. This has the effect of lowering the  $x_V$ -nullcline. Eventually, the right knee of the  $x_V$ -nullcline just touches the  $y_V$ -nullcline (see Fig. 4); this corresponds to a saddle-node bifurcation of the frozen VLPO system. At this time, the system ‘wakes up’, by which we mean the following: Immediately after the saddle-node bifurcation, there is no longer a stable fixed point along the right branch of the  $x_V$ -nullcline, and  $(x_V, y_V)$  jumps down to the silent phase, as shown in Fig. (4). This then ‘releases’ AMIN from inhibition; that is,  $I_{VLPO} = g_{VLPO} H_\infty(x_V)$  switches from  $g_{VLPO}$  to 0. When  $I_{VLPO} = 0$ , the  $x_A$ - and  $y_A$ -nullclines no longer intersect along the left branch of the  $x_A$ -nullcline; however, they do intersect along the right branch of the  $x_A$ -nullcline. Hence,  $(x_A, y_A)$  must jump up to the active phase and approach this stable fixed point. The system is now ‘awake’.

We continue to step through the sleep/wake cycle, now assuming that the system is awake. While awake, the cortical

drive to VLPO increases. This has the effect of raising the  $x_V$ -nullcline. Eventually, there is a saddle-node bifurcation when the left knee of the  $x_V$ -nullcline just touches the  $y_V$ -nullcline. Immediately after this,  $(x_V, y_V)$  jumps up to the active phase, resulting in inhibition of AMIN, and  $(x_A, y_A)$  jumps down to the silent phase. That is, the system ‘falls asleep’.

Note that the system wakes up or falls asleep when there are saddle-node bifurcations of the frozen VLPO system; the bifurcation parameter is the total input to VLPO. Since the total input involves both a circadian and a homeostatic component, the positions of the saddle-node bifurcations define relationships between the circadian and homeostatic inputs. It is these relationships that define the upper and lower threshold in the two-process model. Suppose that the saddle-node bifurcations at the left and right knees of the  $x_V$ -nullcline occur when  $I_{TOT}^V = I_L$  and  $I_{TOT}^V = I_R$ , respectively. We note that it is easy to determine these constants explicitly from the precise forms of the nonlinear functions  $f$  and  $g$ . Now,

$$I_{TOT}^V = -I_{AMIN} - I_{SCN} + I_{CTXV} = -g_{AMIN} H_\infty(x_A) - I_{SCN} + I_V + I_{HOM}. \quad (15)$$

When the system is awake,  $H_\infty(x_A) = 1$ , and when the system is asleep,  $H_\infty(x_A) = 0$ . It follows that the system falls asleep when

$$I_{TOT}^V = -g_{AMIN} - I_{SCN} + I_V + I_{HOM} = I_L, \quad (16)$$

or

$$I_{HOM} = g_{AMIN} + I_V + I_L + I_{SCN} = K_B + I_{SCN}, \quad (17)$$

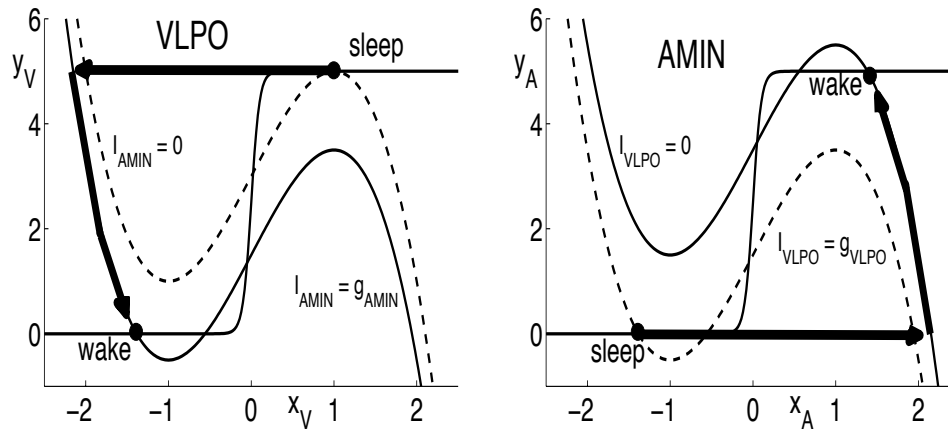
where  $K_B = g_{AMIN} + I_V + I_L$ . This defines the upper threshold curve  $U$  in the two-process model. The system wakes up when

$$I_{TOT}^V = -I_{SCN} + I_V + I_{HOM} = I_R, \quad (18)$$

or

$$I_{HOM} = I_V + I_R + I_{SCN} = K_A + I_{SCN}, \quad (19)$$

where  $K_A = I_V + I_R$ . This defines the lower threshold curve  $L$  in the two-process model.



**Fig. (4).** The mechanism for waking up. **(Left)** Waking up is initiated when there is a saddle-node bifurcation in the active phase of the VLPO system. At this time, the  $x_V$ -nullcline (dashed curve) just touches the  $y_V$ -nullcline. **(Right)** When VLPO jumps down, this releases AMIN from inhibition and it jumps up.

We note that there are other mathematical models of sleep/wake cycles motivated, in part, by the hypothalamic sleep switch [18,31,32]. In particular, Behn *et al.* [32] model the activity of wake- and sleep-promoting cell groups as relaxation oscillators, as is done here. Their model has the ability to produce polyphasic sleep rhythms with frequent brief awakenings, features more prominent in the sleep of mice than of humans. The model of Phillips and Robinson [18], also based on the hypothalamic switch, describes hysteresis in sleep and wake states. A recent extension of the model includes orexinergic cells, and has been used to model subjective fatigue during sleep deprivation [33].

It must be stressed that both the behavior of, and the connections between, the building blocks of our model are still simplifications of the underlying physiology. For example, inter-individual differences may be present in the VLPO activity due to polymorphisms in genes modulating sleep/wake homeostasis and circadian rhythm, such as PER3 [34]. Also, the model does not yet incorporate the pathogenesis of, e.g., primary insomnia [35]. However, by the incorporation of recent neurophysiological findings, this model and future extensions of it may prove useful as platforms for formulation and testing of hypotheses regarding the physiological underpinnings of inter-individual differences during sleep loss.

In conclusion, in this section we demonstrated that the physiological mechanism of a flip-flop sleep switch proposed by Saper *et al.* [23] explains generation and maintenance of a stable sleep state consistent with the two-process model of sleep regulation [22,36] – a demonstration of the compatibility of these models that serves to enhance and substantiate both.

#### BAYESIAN ANALYSIS OF PSYCHOMOTOR VIGILANCE LAPSES (OLOFSEN)

In this section, we make improvements to a Bayesian approach for individualization and forecasting, as applied to biomathematical models of performance impairment during conditions of sleep deprivation. This approach, as described previously [37], utilizes the distribution of model parameters in the population at large (i.e., a population model) without requiring prior knowledge about where specific individuals fall within this distribution. The approach was applied to

data from a study in which healthy volunteers spent 20 days in a sleep laboratory [24]. After three baseline days with 8 hours time in bed, subjects' nocturnal sleep was restricted to 4 hours time in bed for 14 consecutive days. Cognitive performance was tested every 2 hours during wakefulness using a psychomotor vigilance test (PVT; see [38]). This simple reaction time task requires subjects to respond as quickly as possible to a stimulus appearing at random intervals from 2 to 10 seconds. The test takes 10 minutes to complete, during which time approximately 80 to 90 stimuli are presented. The outcome measure used was the 'number of lapses per test bout', with a lapse defined as a reaction time greater than 500 ms. Daily averages of the number of PVT lapses were computed over the trials performed between 09:30 and 23:30.

Across days, the daily averages were described with a mixed-effects regression model of the form

$$y(t) = \alpha + \beta t^\theta + \varepsilon, \quad (20)$$

where  $y$  represents the daily average number of lapses,  $t$  denotes time in days (last baseline day 0, restriction days 1 through 14), and  $\varepsilon$  denotes normally distributed independent noise. For the random effects across subjects, baseline and slope parameters  $\alpha$  and  $\beta$  were assigned lognormal distributions across subjects, and nonlinearity parameter  $\theta$  was assigned a normal distribution across subjects. The distribution parameters of these three parameters were estimated by fitting the mixed-effects regression model to the data from 10 study subjects. This provided the population model underlying the Bayesian forecasting approach.

The data from another three subjects who participated in the study were set aside as "previously unstudied individuals". Bayesian forecasting was subsequently employed to predict the performance of these subjects across all days of the study. That is, data collected from the previously unstudied individuals was used to make likelihood-based adjustments [39] to the subjects' model parameter distributions "on the fly". In this manner, one-day-ahead predictions were made daily using the individualized model parameters. These predictions were compared to the subsequently observed actual data for the subject. For 38 of the 42 one-day-ahead predictions (14 predictions for each of the 3 subjects), individualized predictions were more accurate than predictions

**Table 1. Parameter Estimates of the Lapses Model (4 Hours Sleep per Day Condition)**

	Estimate	SE	$\omega^2$	SE
$\alpha$	0.0176	0.0102	2.26	1.45
$\beta$	0.00846	0.00394	1.39	0.906
$\theta$	1.28	0.0806	–	–
$\sigma^2$	0.187	0.0272		

SE, standard error;  $\omega^2$ , estimated variance for parameter;  $\sigma^2$ , residual error variance.

based on the corresponding group-average model (i.e., > 90%). Thus, Bayesian forecasting was shown to be a useful tool for tailoring biomathematical model predictions of performance impairment to individual subjects. However, at that time no estimates were available for the uncertainty of the predictions, nor had a measure of the speed of convergence for the model individualization process been derived. These issues are addressed here.

First, it is recognized that the number of lapses  $y$  in Eq. (20) goes to infinity when time  $t$  goes to infinity (for positive  $\beta$  and  $\theta$ ), but  $y$  should really be bounded by the number of stimulus trials actually presented. Therefore, we reformulate the model as follows:

$$P_{\text{lapse}}(t) = (P_0 + \beta t^\theta) / (1 + \beta t^\theta), \quad (21)$$

where  $P_{\text{lapse}}$  stands for the probability of experiencing a lapse for any given stimulus trial, and  $P_0$  is the baseline probability of a lapse, with

$$P_0 = \alpha / (1 + \alpha). \quad (22)$$

For a fixed number of stimulus trials and if  $\beta t^\theta \ll 1$ , the models in Eqs. (20) and (21) are equivalent, and the interpretation of  $\theta$  is unchanged.

The number of stimulus trials resulting in a lapse can now be described by a binomial distribution, with its one parameter  $P_{\text{lapse}}$  given by Eq. (21). In this manner, intra-individual variability is accounted for, but for a given number of stimulus trials  $N_{\text{trials}}$  it is fixed:

$$\text{VAR}(P_{\text{lapse}}) = N_{\text{trials}} P_{\text{lapse}} (1 - P_{\text{lapse}}). \quad (23)$$

The population model described by Eqs. (20) through (23) was implemented in the software BUGS.<sup>1</sup> The objective was to implement Bayesian forecasting as previously described [37], but to add the computation of 95% prediction intervals. A 95% prediction interval is defined as the Bayesian 95% confidence interval for a predicted observation, with the (common sense and desired) interpretation that there is a 95% chance of a future observation lying in that interval. This is in contrast with a classical (frequentist) 95% confidence interval, for which the interpretation is reversed: if the construction of the confidence interval is repeated, 95% of those intervals would contain the observation [39].

The BUGS code representing the model described in Eqs. (20) through (23) is given in Source Code Appendix 1. Non-informative priors were used for the unknown parameters in a standard (BUGS) way. BUGS provides samples of the dis-

tributions of the parameters of interest, such as the model parameters and the output. Parameter estimates from BUGS using the experimental data are listed in Table 1. Fig. (5) shows the data with median and 95% prediction intervals – the data seem to be well-described. Differences between the present parameter estimates and those reported earlier [37] are mainly caused by the specification of the model by Eq. (21). The inter-individual variability in  $\theta$  could now be ignored (its estimation error was larger than the value itself).

Simulated data were generated from the model with parameters equal to those given in Table 1, but with 100 individuals instead of 10 as in the experimental data. Furthermore, the number of trials was fixed at 750 (approximately the average total number of trials per day). Parameter estimates from BUGS for the simulated data are listed in Table 2. By comparing the estimates with the values used for simulation, and by inspecting the standard errors, it can be concluded that the parameters are well-identified. In Fig. (6) the data are plotted with median and 95% prediction intervals.

An analysis that included only the first ten simulated subjects revealed considerable uncertainty, and bias in the estimate of  $\beta$  (see Table 3) as well as in the estimates of inter-individual variability. This makes it likely that the estimates in Table 1 are biased as well, suggesting that more data are needed to estimate the model parameters and their distributions. (Note that Bayesian model individualization and forecasting may be applied regardless of the amount of data available; cf. [17]).

A visual posterior predictive check was performed to compare the distribution of the model output with the observed data [40]. It was performed using the BUGS program given in Source Code Appendix 2, which makes predictions using informative priors (all data were set to “not available”). The informative priors were constructed using the means and standard deviations from the BUGS analysis of the experimental data; see Table 4. The results of the visual posterior predictive check are presented in Fig. (7). About 5% of the data are outside the 95% prediction intervals, but these are all under the lower limit. This is probably because the postulated (lognormal) distributions of the model parameters are too skewed.

Results for one-day-ahead predictions of the experimental and simulated data are presented in Figs. (8) and (9), respectively. Available data were restricted to those acquired up to the beginning of the one-day prediction horizon; data were considered “not available” for subsequent days. For the experimental data there were 8 one-day-ahead predictions outside the 95% prediction intervals (mainly below the

<sup>1</sup>Bayesian inference Using Gibbs Sampling; see <http://www.mrc-bsu.cam.ac.uk/bugs>

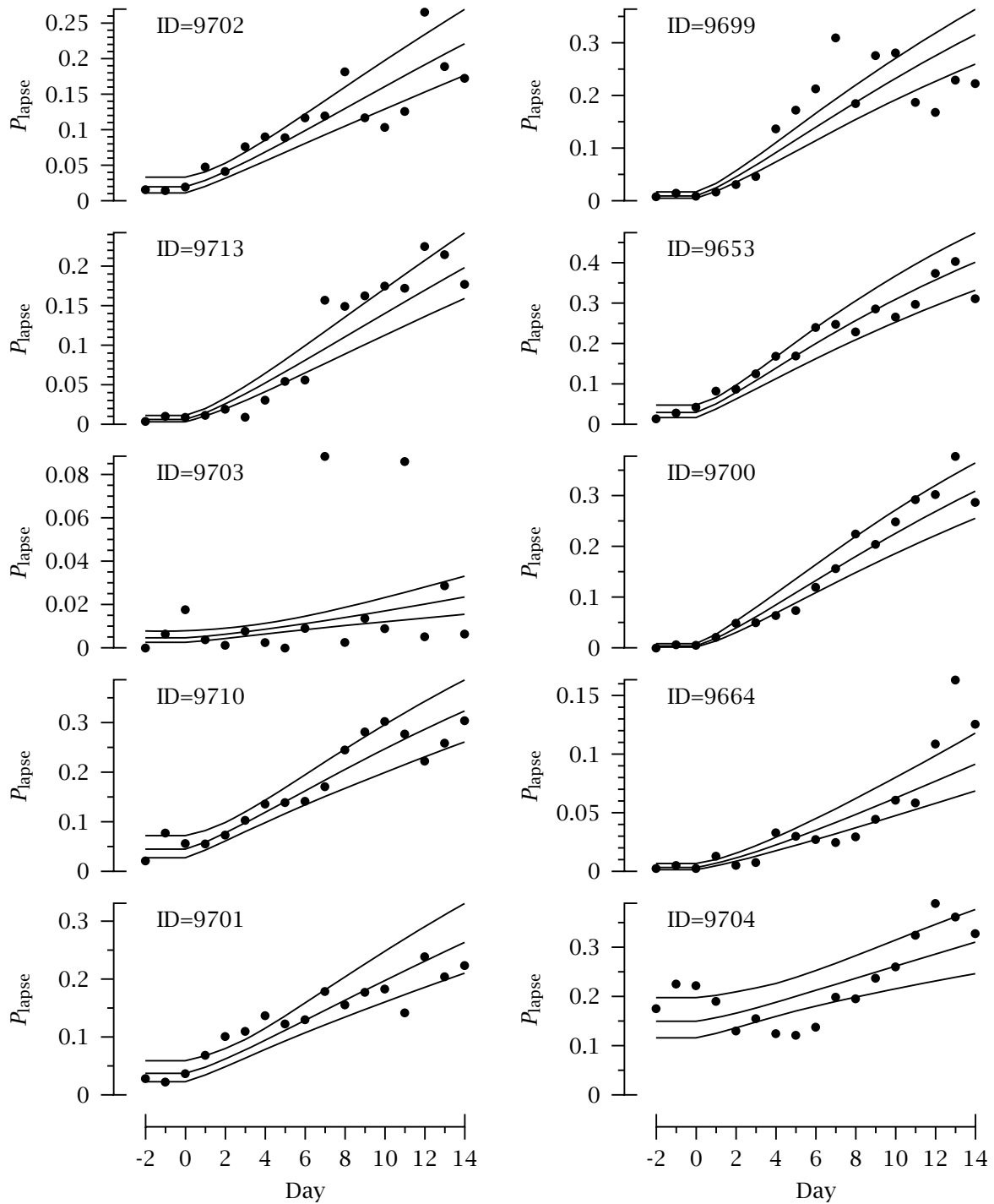


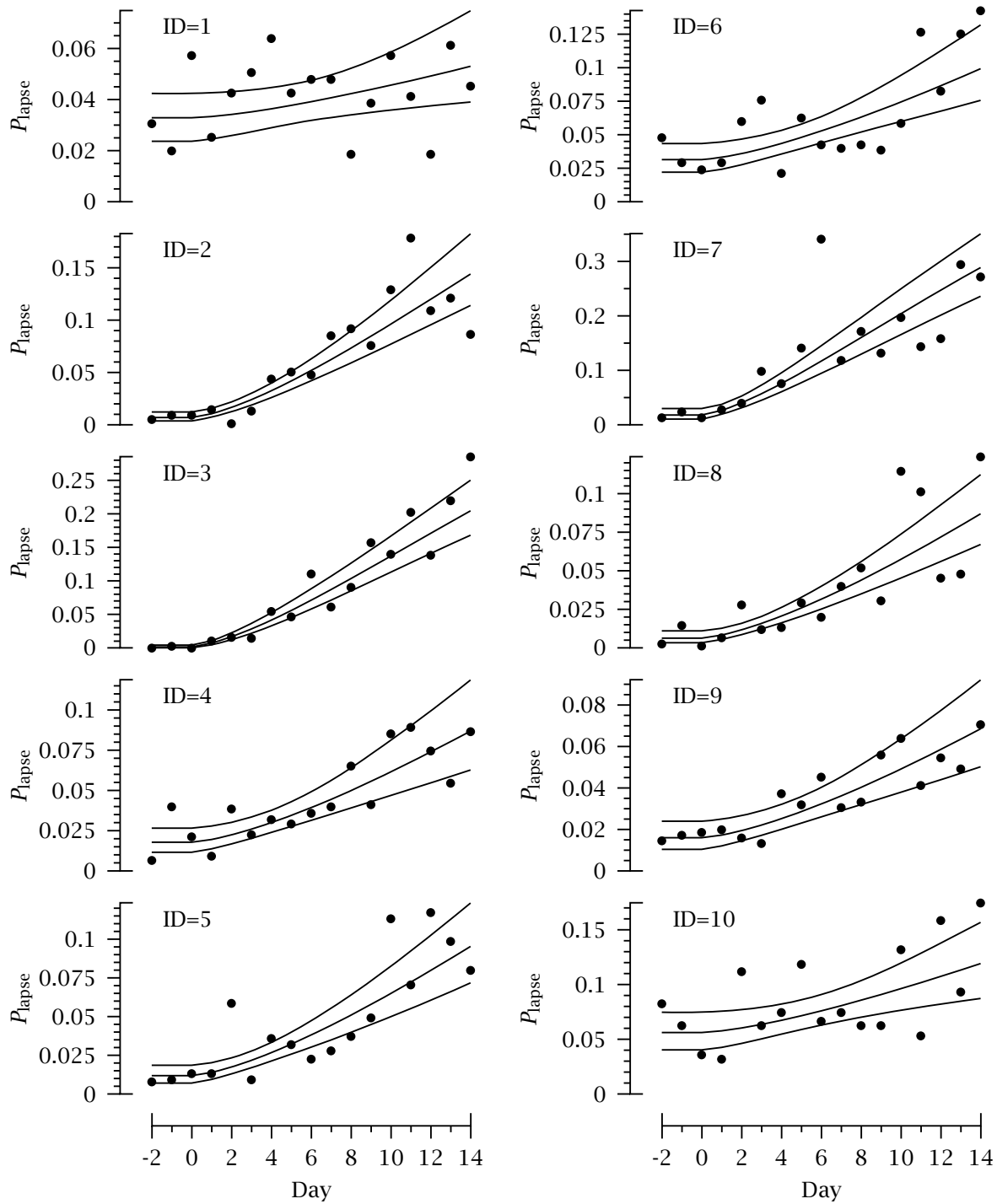
Fig. (5). Experimental data (dots: fraction of lapses over number of trials) and model fits (median and 95% prediction intervals).

Table 2. Parameter Estimates of the Lapses Model (100 Simulated Subjects)

	Estimate	SE	$\omega^2$	SE
$\alpha$	0.0185	0.00273	2.06	0.325
$\beta$	0.00759	0.000999	1.39	0.214
$\theta$	1.30	0.0245	–	–
$\sigma^2$	0.183	0.00790		

SE, standard error;  $\omega^2$ , estimated variance for parameter;  $\sigma^2$ , residual error variance.





**Fig. (6).** Simulated data (dots: fraction of lapses over number of trials) and model fits (median and 95% prediction intervals).

**Table 3. Parameter Estimates of the Lapses Model (First 10 Simulated Subjects Only)**

	Estimate	SE	$\omega^2$	SE
$\alpha$	0.0153	0.00626	1.44	1.02
$\beta$	0.00220	0.000918	1.01	0.831
$\theta$	1.45	0.100	–	–
$\sigma^2$	0.165	0.0248		

SE, standard error;  $\omega^2$ , estimated variance for parameter;  $\sigma^2$ , residual error variance.

Table 4. Prior Information

Parameter	Mean	SD	$t$	$r$	$\mu$
malpha	-4.163	0.4859	4.236		
mbeta	-4.862	0.4189	5.699		
ptheta	0.2447	0.06329	249.6		
talpha	0.5795	0.2883		4.040	6.972
tbeta	0.9533	0.4804		3.938	4.131
terr	5.461	0.7859		48.28	8.842

SD, standard deviation;  $t = 1 / SD^2$ ;  $\mu = \text{mean} / SD^2$ ;  $r = \text{Mean} \cdot \mu$ . See the BUGS manual<sup>1</sup> for a description of  $t$ ,  $r$  and  $\mu$ .

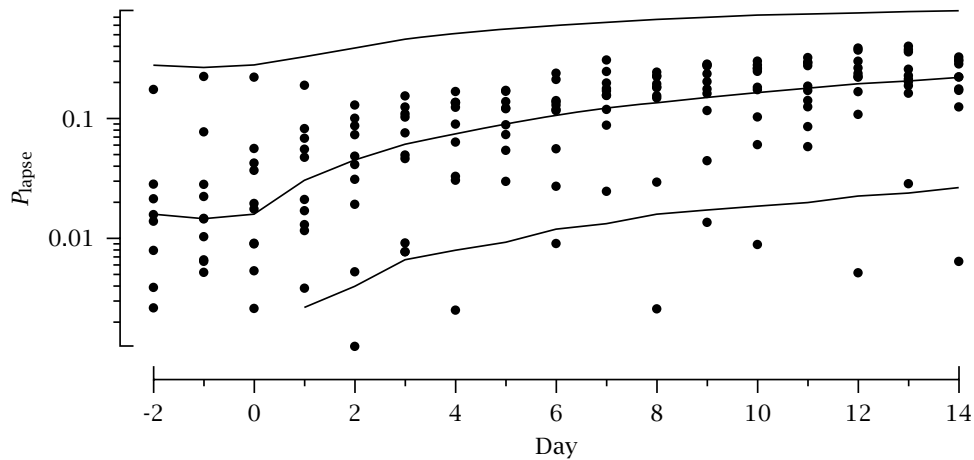


Fig. (7). Posterior predictive check.

lower limit for subject 9703), and for the simulated data there were 6 one-day-ahead predictions outside the 95% prediction intervals, whereas a number of  $0.05 \text{ (probability)} \times 10 \text{ (subjects)} \times 17 \text{ (days)} \approx 9$  was expected.

To assess the speed of convergence, another set of simulated data including 100 subjects and a total of 25 days of sleep restricted to 4 hours per day was considered. For each individual, Bayesian 95% confidence intervals of individualized  $\alpha$ ,  $\beta$  and  $P_{\text{lapse}}$  (for parameters  $\theta$  and  $\sigma^2$  there was no inter-individual variability in the model) were calculated and normalized by dividing by the median. The means of the normalized confidence intervals are presented in Fig. (10). Information about parameter  $\alpha$  is mainly present at the baseline measurements; in contrast, information about parameter  $\beta$  is only present after the baseline measurements (this follows by definition from the model specification). For the probability of a lapse, the uncertainty is largest for the first baseline measurement and again large for the first day of sleep deprivation, because no information on  $\beta$  has yet been obtained for this individual. After an initially fast convergence, subsequent convergence is approximately proportional to the square root of the number of days (accumulation of precision if information about the parameter is present in the data [37]).

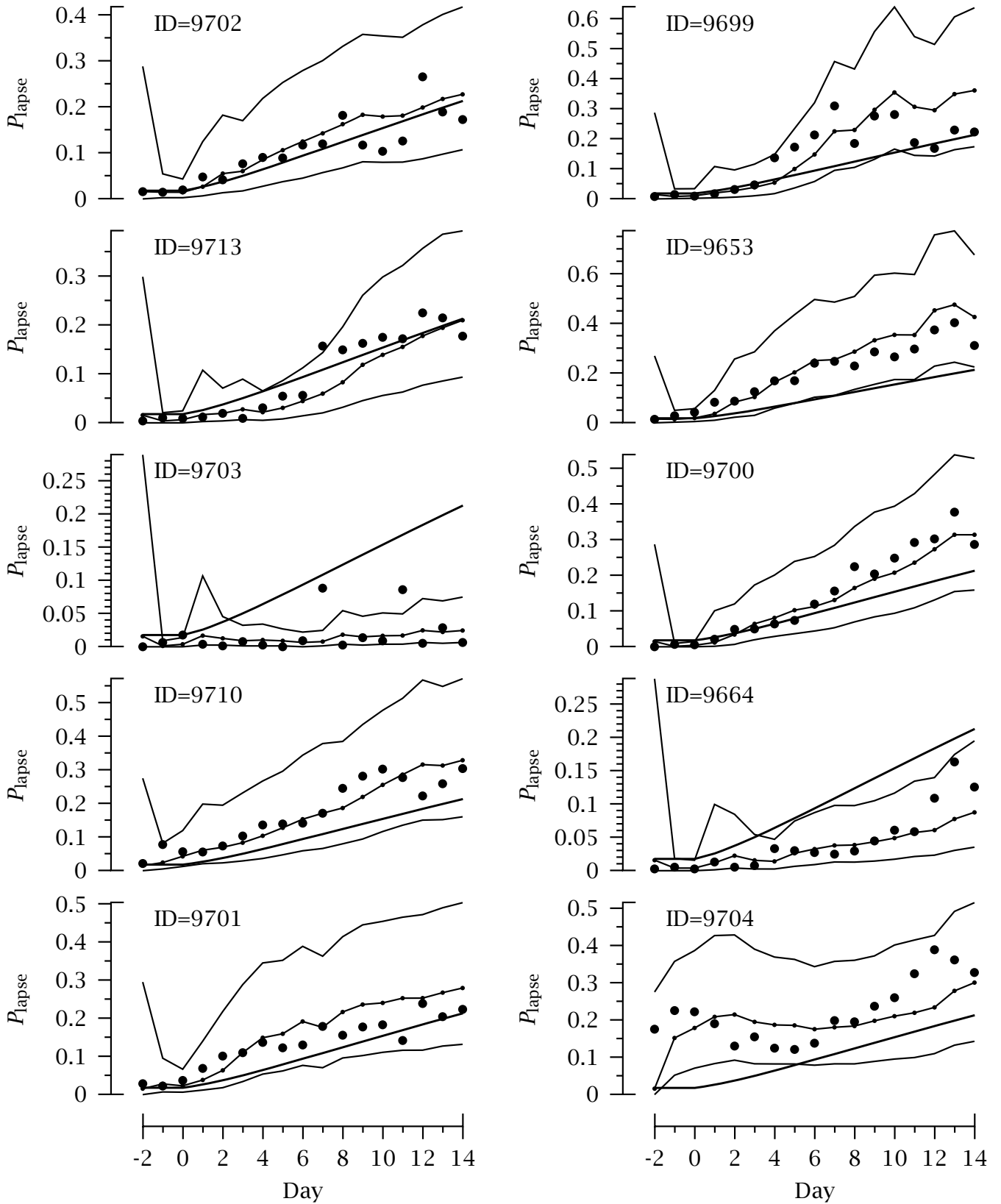
In conclusion, full Bayesian individualization and forecasting was implemented in a BUGS model. The BUGS software provides samples of not only the model parameters, but also the distribution of the forecasts, from which relevant

characteristics can be computed, such as 95% prediction intervals for performance measures.

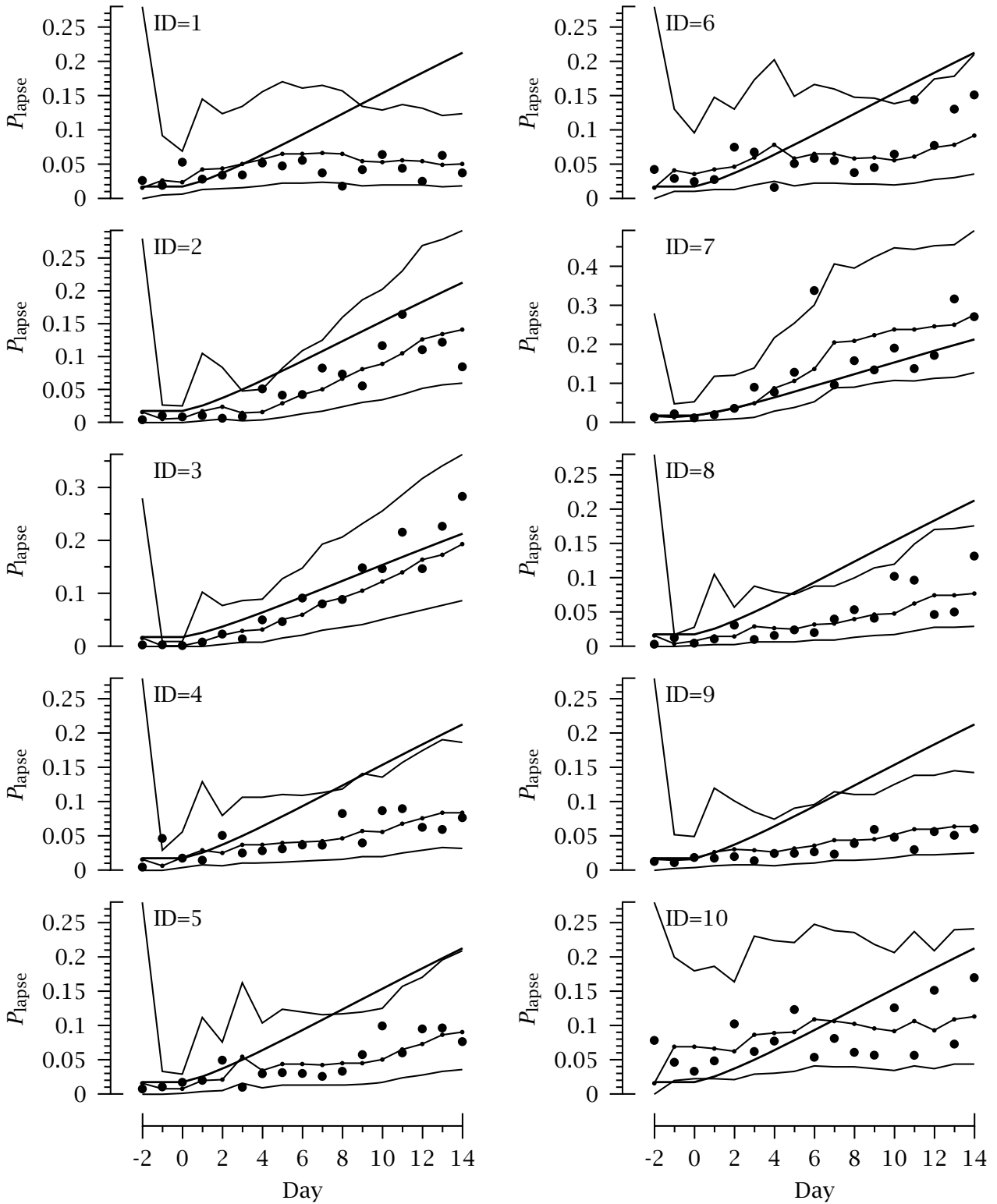
**SOURCES OF NOISE AND BAYESIAN CONFIDENCE INTERVALS IN INDIVIDUALIZED PREDICTIONS (MOTT)**

When predicting performance for specific individuals, it is axiomatic that individualized performance models should provide more accurate predictions than population average models. However, even in individualized models there may be considerable unexplained variance in operationally relevant cognitive capabilities such as ‘problem solving’ and ‘decision making’. It is generally thought that unexplained variance, like that typically observed on performance tests such as the PVT, is at least partly attributable to measurement noise (here defined as noise due to inconsistencies in the testing process, environment, and equipment) and partly due to biological noise (variability in the underlying neurobiological processes that mediate performance of the task at hand).

Laboratory experiments – which constitute a primary source of data from which current human performance prediction models have been constructed – are typically conducted in controlled environments so as to minimize known sources of measurement noise. Unexplained variability in laboratory studies is often attributed mainly to endogenous neurobiological noise. In field studies there is a greater potential for exogenous measurement noise (e.g., the presence



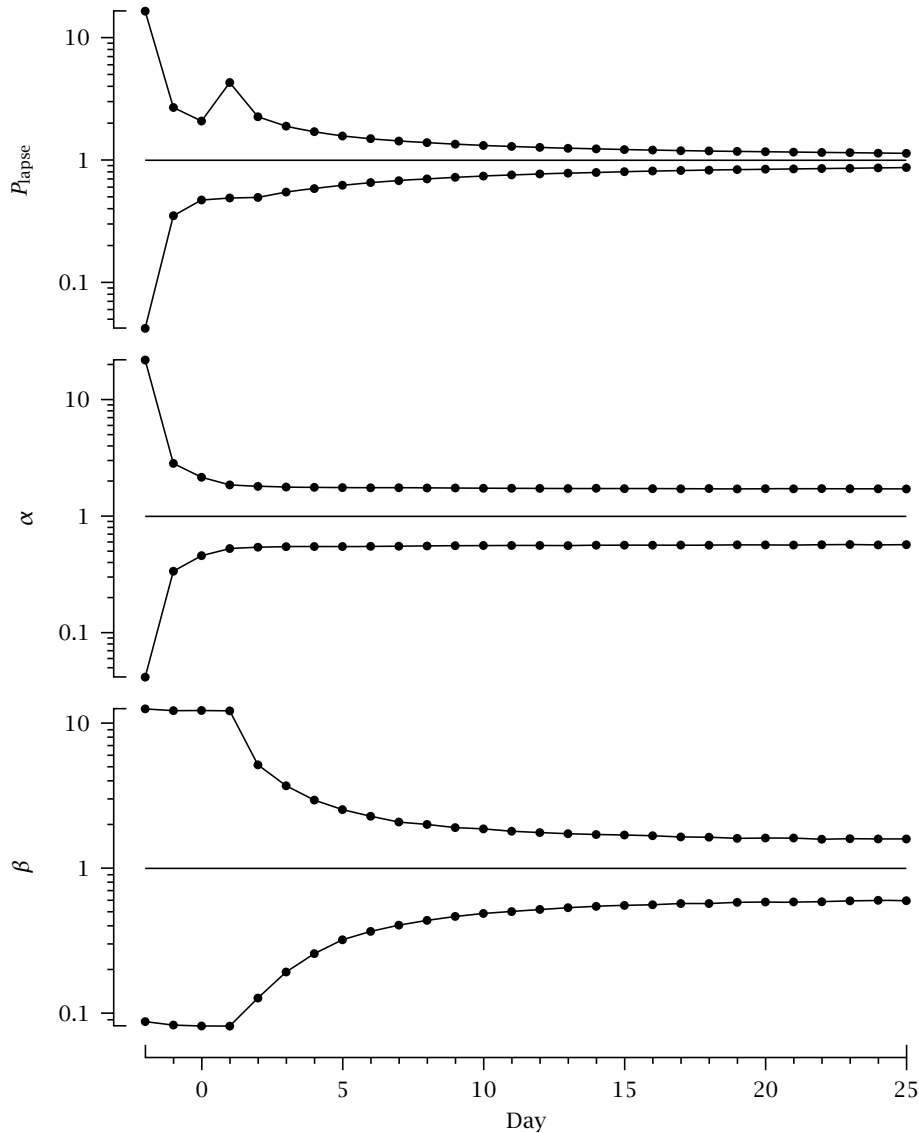
**Fig. (8).** One-day-ahead predictions (small dots) of the experimental data (dots), 95% prediction interval boundaries (lines), and the population predictions (thick lines).



**Fig. (9).** One-day-ahead predictions (small dots) of the simulated data (dots), 95% prediction interval boundaries (lines), and the population predictions (thick lines).

of distracters that can impact test performance). Whereas variance observed from endogenous neurobiological noise reflects actual variability in cognitive capabilities, variance from exogenous sources is not indicative of the actual state

of cognitive capabilities. When generating performance predictions, the prediction intervals (as defined in the previous section) should normally include endogenous noise and exclude exogenous noise. Because the field environment is the



**Fig. (10).** Average normalized Bayesian 95% confidence intervals of individualized  $P_{\text{lapse}}$ ,  $\alpha$  and  $\beta$ .

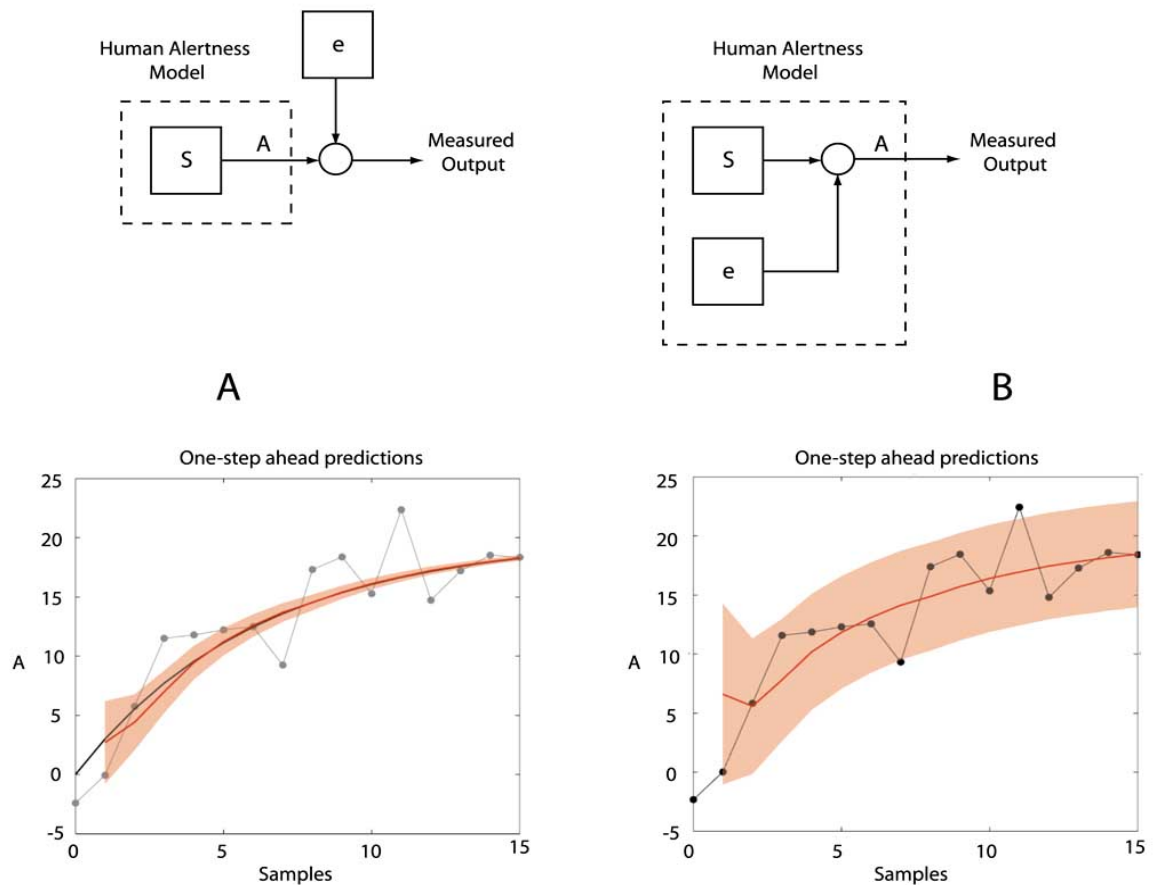
typical target venue for deployment of predictive models, it will be important to differentiate endogenous and exogenous noise sources and account for them in the modeling and prediction framework.

To illustrate the effects of the noise model on 95% prediction intervals, it is useful to consider the hypothetical case of an individual whose measured performance over a single period of extended, continuous wakefulness follows an exponential decay with normally distributed noise. If the noise is modeled as exogenous measurement noise, then predictions of true performance will have narrow intervals that forecast the location of the underlying exponential trajectory. If, on the other hand, the noise is modeled as endogenous system noise, then predictions of true performance will have wider intervals. These two situations are illustrated in Fig. (11). Note that the mean performance prediction is the same in both cases, and it is just the prediction interval that varies. Alternate noise models could include a combination of both system and measurement noise that would determine the

width of the prediction intervals as a continuum between the two cases presented here.

Additional experimental work may be required to develop noise models that provide the distinction between endogenous versus exogenous noise variance involves comparing data from a test performed in both a controlled laboratory environment (assuming purely endogenous noise) and a typical operation environment (assuming superposition of endogenous plus exogenous noise). The additional variance observed in the operational environment could be attributed to exogenous noise. While endogenous noise sources may be reasonably estimated from controlled laboratory studies, we would expect exogenous noise sources to be highly context dependent and therefore more difficult to generalize.

In previous modeling work, the size of confidence intervals may have been misestimated by not properly accounting for the different possible origins of noise. New noise models will be required, and the utility of individualized perform-



**Fig. (11).** One-step ahead predictions with 95% confidence intervals generated on the same data set, but with **A.** variability modeled as an exogenous noise source (measurement noise), and **B.** variability modeled as an endogenous random process (system noise). S, hypothetical exponential decay process; e, noise process; A, predicted performance; connected dots, measured values; thick black curve, actual values; red curve, predicted means; shaded orange areas, 95% prediction intervals.

ance prediction models in operational environments will be enhanced by prediction intervals that, to the extent possible, specify the proportion of noise that results from endogenous vs. exogenous sources.

Tolerance for human performance deficits on operational tasks might range from very low (such as piloting an airliner, in which case mistakes might lead to catastrophe and loss of life) to very high (such as writing a report, in which case there may be considerable tolerance for error since such errors would often be of little consequence). Therefore, the availability of meaningful individual predictions with confidence intervals that are applicable to real-world operations will also benefit from estimation/specification of the extent to which operational success depends upon accurate and timely performance of particular tasks – that is, specification of the threshold that delineates acceptable from unacceptable operational task performance.

#### GENERALIZABILITY OF INDIVIDUALIZED PREDICTIONS TO DIFFERENT NEUROBEHAVIORAL PERFORMANCE MEASURES (VAN DONGEN)

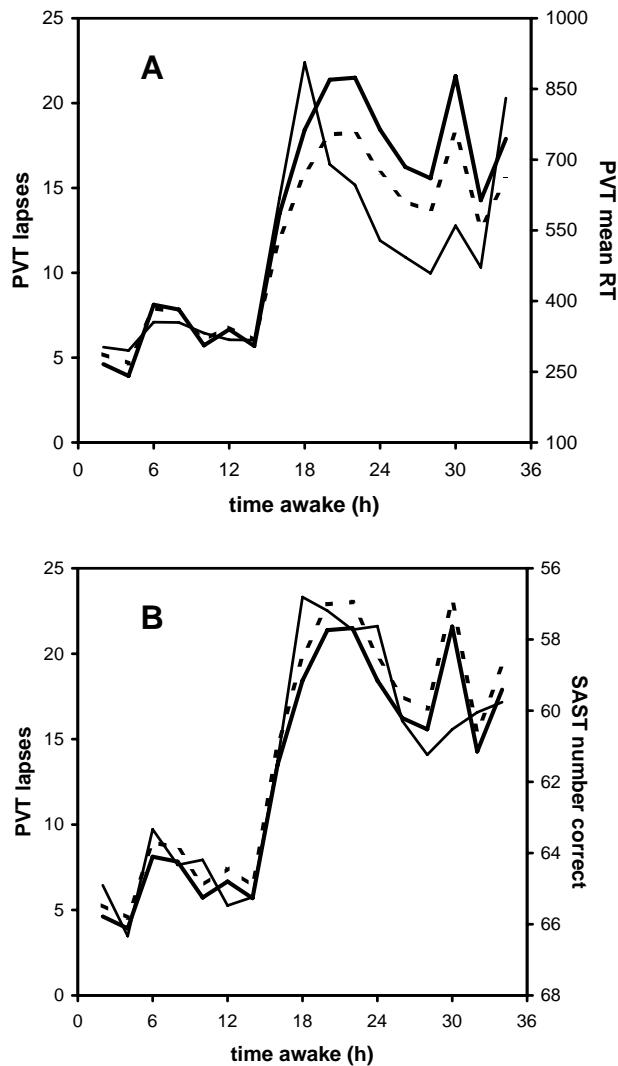
It is a well-recognized problem that performance predictions targeting one neurobehavioral outcome measure do not readily translate into predictions for another outcome measure [41,42]. This is especially relevant in operational envi-

ronments [43,44], where predictions may be needed for performance on a wide variety of tasks using models that were calibrated on a single standardized alertness measure [45,46]. One way this issue has been dealt with is to formulate transformation functions, which mathematically convert predictions for one outcome variable into predictions for another outcome variable. This approach has been applied successfully to group-average performance predictions (e.g., [47]).

Advances in biomathematical modeling of human performance (such as the techniques described in the present paper) have been focused on improvement of individualized predictions [17,37]. While individualization of model predictions represents a critical determinant of the ultimate usefulness of biomathematical modeling in small-crew, safety-critical settings (e.g., trucking, aviation, space flight), the critical issue of generalizability remains relatively unexplored. In fact, the applicability of metric transformations in the context of individualized performance predictions has never before been evaluated. A priori, mathematical concerns arise if such transformations are non-linear, since this may produce incompatible alterations of the shape of the distribution of the individual differences. Even if the transformations are linear, however, individual differences in one measure of performance may not be congruent with individual differences in other measures [7], raising doubts about

whether individualized predictions are amenable to transformation.

To examine this issue, data were used from a study involving 21 healthy adults (10 males and 11 females; ages 22–40). As part of a larger protocol [48,49], these subjects underwent 36 hours of total sleep deprivation in a controlled laboratory environment. Every 2 hours from 10:00 until 22:00 the next day, they were tested on a series of neurobehavioral tasks including the psychomotor vigilance test (PVT; [50]) and the serial addition/subtraction task (SAST; [51]). From every test session (except the first one, which was discarded because of potential sleep inertia), the number of lapses (reaction times > 500 ms) on the PVT, the mean reaction time (RT) on the PVT, and the number of correct responses on the SAST were determined. Temporal profiles of these variables, averaged over subjects, were assessed using mixed-model ANOVA in SAS 9.1.3 (SAS Institute Inc., Cary, NC) – see Fig. (12).



**Fig. (12).** Group-average temporal profiles of **A.** PVT mean RT and **B.** SAST number correct (thin curves) plotted in conjunction with PVT lapses (thick curve) as observed every 2 hours during 36 hours of total sleep deprivation. Also shown are estimations of group-average PVT mean RT (top panel, dotted curve) and SAST number correct (bottom panel, dotted curve) based on linear transformations of PVT lapses.

Polynomial functions were utilized to derive transformation functions from PVT lapses to PVT mean RT and from PVT lapses to SAST number correct, using the partial least squares (PLS) procedure in SAS on the group-average temporal profiles. Cross-validation revealed that a linear function constituted the optimal transformation for both PVT mean RT and SAST number correct. Higher order polynomials did not contribute significantly to goodness-of-fit ( $P > 0.1$ ). For PVT mean RT, the optimal transformation was

$$RT = 27.83x + 158.92, \quad (24)$$

and for SAST number correct (NC), the optimal transformation was

$$NC = -0.505x + 67.797, \quad (25)$$

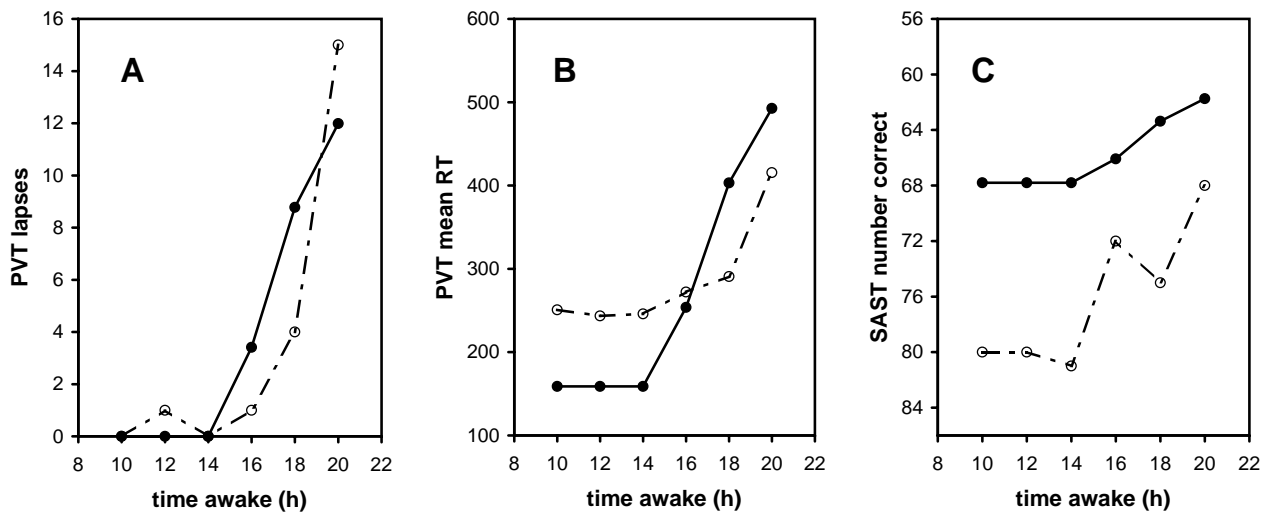
where  $x$  represents PVT lapses. See Fig. (12).

The group-average temporal profile of PVT lapses during total sleep deprivation is accurately predicted by the two-process model of sleep regulation [52] as the postulated result of an accumulating homeostatic drive for sleep combined with circadian rhythm [53]. The distributions across individuals of the parameter values for the two-process model, as used to predict performance, have recently been estimated in a sample of the healthy adult population [17]. That information served as a basis for implementing a Bayesian forecasting procedure [37] to make person-specific performance predictions with the two-process model [54]. This approach produces predictions of an individual's future performance on the basis of the individual's past performance, evaluated with regard to the population distributions of the two-process model's parameters. (For further details, see [17,37]).

Individualized performance predictions were made for a new subject (female, age 35 y) who also participated in the laboratory study with 36 hours of total sleep deprivation described above. At 8 hours awake (6 PM), PVT lapses were predicted 12 hours ahead (in steps of 2 hours). Fig. (13A) shows the predicted number of lapses, as well as the subsequently observed actual number of lapses, across the 12-hour prediction period. The predictions matched the observations fairly well – both predictions and observations showed a steep increase in PVT lapses after 16 hours of wakefulness, in agreement with the temporal profile of performance impairment from sleep deprivation documented for highly vulnerable individuals (Fig. 1 in [55]).

The individual's predictions for PVT lapses were transformed to predictions for PVT mean RT and SAST number correct by means of Eqs. (18) and (19). Fig. (13B) shows the generated predictions for PVT mean RT, as well as the actually observed mean RT values. The match between the two was less close than in the case of PVT lapses – daytime mean RT was underestimated, and the nighttime decrement was overestimated. However, in an absolute sense, the discrepancies were small. This is not surprising, since PVT lapses and mean RT have been found to be highly correlated within and between individuals [56,57].

Fig. (13C) shows the predictions for SAST number correct calculated by linear transformation of the predictions for PVT lapses, as well as the actual observations for SAST number correct. The predictions did not correspond well to the observations. This finding is in line with previous results



**Fig. (13).** Person-specific performance predictions (solid curves), and actual performance as subsequently observed (dashed curves), for one healthy adult exposed to total sleep deprivation. At 8 hours awake (6 PM), Bayesian forecasting was applied to predict PVT lapses over the next 12 hours, in steps of 2 hours. **A.** Predictions for PVT lapses and actual observations. **B.** Predictions for PVT mean RT calculated by linear transformation of the predictions for PVT lapses, and actual PVT mean RT observations. **C.** Predictions for SAST number correct calculated by linear transformation of the predictions for PVT lapses, and actually observed results for SAST number correct.

indicating that individual responses to sleep deprivation on the PVT are distinct from individual responses on the SAST (Table 4 in [7]). In this particular case, the individual's SAST performance was better than predicted throughout the 12-hour prediction period. This suggests the possibility of a greater-than-average aptitude for the SAST that compensated for the high vulnerability to sleep deprivation that was revealed by the PVT. If this interpretation is correct, then the implications for the strategy of utilizing transformations to accurately predict a particular individual's performance on one task based on that individual's performance on another task becomes exceedingly complicated, since such a transformation would require a-priori knowledge of the pattern of that individual's relative cognitive capabilities.

The results in Fig. (13) illustrate that individualized performance predictions do not, in the case of the PVT and the SAST, translate well from one performance task to the other. Nevertheless, it is unlikely that responses to sleep deprivation for every possible performance task are entirely idiosyncratic (i.e., that there is no shared variance). Thus, further research is needed to tease apart the independent dimensions of systematic individual differences in neurobehavioral responses to sleep deprivation [16]. The usefulness of metric transformations to generalize subject-specific performance predictions across different tasks will ultimately depend on such knowledge. In the meantime, it may be fruitful to investigate the extent to which relevant covariates (e.g., baseline performance capability) help to improve metric transformations for the individualized prediction of different types of neurobehavioral performance.

### SYNTHESIS (BALKIN)

With the successful individualization of human fatigue/performance prediction models, initiation of a new era of enhanced workplace safety and productivity will become possible. There are two general approaches to individualization of such models: (a) model development that follows, and is based upon, scientific progress in understanding the

physiological underpinnings of human performance and how it is interactively mediated by the sleep homeostat and the circadian rhythm of alertness, and (b) a statistical approach in which performance prediction models based on empirical data are derived without, or with only incomplete, understanding of underlying physiological processes. These approaches are not mutually exclusive, and an iterative process in which advances on both fronts are compared and integrated is recommended. Such an integrative approach is illustrated in the first part of the present paper, in which two existing models – the two-process model of sleep regulation and an extension of the flip-flop sleep switch model – are integrated in a manner that demonstrates the mechanisms by which the sleep homeostat and the mechanisms by which state switching and state stability is achieved can realistically interact to account for the real-world phenomenon of relative sleep/wake state stability.

Also detailed here is a Bayesian approach to the process of model individualization – an approach that is likely to constitute a continuing, significant component of this ongoing effort, as well as an important component of the final product – since even in the event that in-depth knowledge of the physiological processes that underlie sleep, circadian rhythms, and performance is eventually achieved, a Bayesian (or Bayesian-like) approach will most likely remain necessary to estimate each individual's position within the distributions of relevant physiological states and traits. Part and parcel of the statistical approach is the need to estimate the (Bayesian) confidence intervals for individualized performance predictions. The intricacies of this issue are discussed in the present paper as well.

Perhaps the biggest 'elephant in the kitchen' is the issue we addressed last: the transduction of performance model predictions into meaningful predictions of actual, real-world operational performance. Current models are largely based on performance (e.g., PVT) data collected in laboratory settings. However, the process of predicting performance in one cognitive domain from measures in another cognitive do-



main may not be straightforward. As suggested here, it is likely that future progress will depend upon the extent to which independent dimensions of systematic individual differences in neurobehavioral responses to sleep deprivation are delineated [16]. However, in parallel with that effort, it should also be possible to apply existing models to various operational settings and determine in a post-hoc manner the extent to which model predictions empirically relate to various aspects of operational performance (e.g., accident rates), even if the nature of the mathematical relationship between the predictor and the predicted variables is not yet fully understood [58,59]. Such a two-pronged approach would, in many ways, be similar to (and an extension of) the two general approaches to model individualization that are currently underway – with one based on our understanding of the relevant, underlying physiological processes and relationships, and the other based on whatever empirical data can be obtained.

Given the scope and complexity of the problems associated with individualization of human sleep/performance prediction models, it is likely that success will ultimately depend upon cooperative, concerted effort from multiple stakeholders and scientific entities. The purpose of this paper, and the purpose of the AFOSR-sponsored meeting that spawned this paper, has been to explicate the current approaches and describe the current state-of-the-art, in an attempt to garner momentum in the effort to individualize sleep and performance prediction models.

## APPENDICES

For these BUGS programs, “Y” is the number of lapses, “aux” is the number of stimulus trials, “err” stands for residual error variance (to account for deviations from Eq. (22)), “N” is the number of subjects, “j” represents the day in the experiment (the three baseline days are included), and “M” is the duration of the experiment in days. Parameters “malpha” and “w2alpha” (the population mean and variance of random effects), etc., are computed for compatibility with earlier results [37].

- Source Code Appendix 1: BUGS code for the model described by Eqs. (20) through (23).

```
model
{
  for( i in 1:N ) {
    for( j in 1:3 ) {
      Y[i,j] ~ dbin(pr0[i,j],aux[i,j])
      pr0[i,j] <- exp(pr1[i,j])/(1+exp(pr1[i,j]))
      pr1[i,j] <- pr2[i,j] + err[i,j]
      pr2[i,j] <- logit(pr[i,j])
      pr[i,j] <- p0[i,j]
      p0[i,j] <- alpha[i]/(1+alpha[i])
      mu[i,j] <- pr[i,j]*aux[i,j]
      err[i,j] ~ dnorm(0,terr)
    }
    for( j in 4:M ) {
      Y[i,j] ~ dbin(pr0[i,j],aux[i,j])
      pr0[i,j] <- exp(pr1[i,j])/(1+exp(pr1[i,j]))
      pr1[i,j] <- pr2[i,j] + err[i,j]
      pr2[i,j] <- logit(pr[i,j])
      pr[i,j] <- (p0[i,j] + fac[i,j]) / (1 + fac[i,j])
      fac[i,j] <- beta[i]*pow(x[j],ptheta)
      p0[i,j] <- alpha[i]/(1+alpha[i])
      mu[i,j] <- pr[i,j]*aux[i,j]
      err[i,j] ~ dnorm(0,terr)
    }
  }
  alpha[i] ~ dlnorm(malphi,talphi)
```

## DISCLAIMER

This material has been reviewed by the Walter Reed Army Institute of Research, and there is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the position of the Department of the Army or the Department of Defense.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the AFOSR, and especially Dr. Willard Larkin, for their vision and sponsorship of the workshop “New Approaches to Modeling Sleep/Wake Dynamics and Cognitive Performance” held at the Mathematical Biosciences Institute at The Ohio State University (Columbus, OH, USA), October 26–27, 2006. The workshop was funded by AFOSR award GRT00005382.

The experimental data for the Bayesian analysis of psychomotor vigilance lapses were kindly provided by David Dinges, Department of Psychiatry, University of Pennsylvania School of Medicine. Hans Van Dongen was supported by NIH grant HL70154, and in part by AFOSR grant FA9550-09-1-0136, federally directed funding W81XWH-05-1-0099 and DURIP grant FA9550-06-1-0281. David Terman was partially funded by the NSF under agreement 0112050 and by AFOSR grant FA9550-06-1-0033.

```

    beta[i] ~ dlnorm(mbeta, tbeta)
  }
  malpha ~ dnorm(0, 1.0E-6)
  talpha ~ dgamma(0.001, 0.001)
  mbeta ~ dnorm(0, 1.0E-6)
  tbeta ~ dgamma(0.001, 0.001)
  ptheta ~ dlnorm(0, 1.0E-6)
  terr ~ dgamma(0.001, 0.001)
  palpha <- exp(malpha)
  pbeta <- exp(mbeta)
  mtheta <- log(ptheta)
  w2alpha <- 1/talpha
  w2beta <- 1/tbeta
  s2err <- 1/terr
}

```

- Source Code Appendix 2: BUGS code for individualized performance prediction.

model

```

{
  for( j in 1:3 ) {
    Y[j] ~ dbin(pr0[j], aux[j])
    pr0[j] <- exp(pr1[j]) / (1 + exp(pr1[j]))
    pr1[j] <- pr2[j] + err[j]
    pr2[j] <- logit(pr[j])
    pr[j] <- p0[j]
    p0[j] <- alpha / (1 + alpha)
    mu[j] <- pr[j] * aux[j]
    err[j] ~ dnorm(0, terr)
  }
  for( j in 4:M ) {
    Y[j] ~ dbin(pr0[j], aux[j])
    pr0[j] <- exp(pr1[j]) / (1 + exp(pr1[j]))
    pr1[j] <- pr2[j] + err[j]
    pr2[j] <- logit(pr[j])
    pr[j] <- (p0[j] + fac[j]) / (1 + fac[j])
    fac[j] <- beta * pow(x[j], ptheta)
    p0[j] <- alpha / (1 + alpha)
    mu[j] <- pr[j] * aux[j]
    err[j] ~ dnorm(0, terr)
  }
  alpha ~ dlnorm(malpha, talpha)
  beta ~ dlnorm(mbeta, tbeta)
  malpha ~ dnorm(-4.163, 4.236)
  talpha ~ dgamma(4.040, 6.972)
  mbeta ~ dnorm(-4.862, 5.699)
  tbeta ~ dgamma(3.938, 4.131)
  ptheta ~ dlnorm(0.2447, 249.6)
  terr ~ dgamma(48.28, 8.842)
}

```

## REFERENCES

- [1] Patrick GTW, Gilbert JA. On the effects of loss of sleep. *Psychol Rev* 1896; III(5): 469-483.
- [2] Naitoh P. Sleep deprivation in human subjects: a reappraisal. *Wak Sleep* 1976; 1: 53-60.
- [3] Wilkinson RT. Interaction of lack of sleep with knowledge of results, repeated testing, and individual differences. *J Exp Psychol* 1961; 62: 263-271.
- [4] Webb WB, Levy CM. Effects of spaced and repeated total sleep deprivation. *Ergonomics* 1984; 27: 45-58.
- [5] Neri DF. Preface: Fatigue and Performance Modeling Workshop, June 13-14, 2002. *Aviat Space Environ Med* 2004; 75: A1-A3.
- [6] Balkin TJ, Kamimori GH, Redmond DP, *et al.* On the importance of countermeasures in sleep and performance models. *Aviat Space Environ Med* 2004; 75: A155-A157.
- [7] Van Dongen HPA, Baynard MD, Maislin G, Dinges DF. Systematic interindividual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability. *Sleep* 2004; 27: 423-433.
- [8] Landolt HP. Genotype-dependent differences in sleep, vigilance, and response to stimulants. *Curr Pharm Des* 2008; 14: 3396-3407.
- [9] Stephenson PE. Physiologic and psychotropic effects of caffeine on man. A review. *J Am Diet Assoc* 1977; 171: 240-247.
- [10] Rétey JV, Adam M, Gottselig JM, Khatami R, Dürr R, Achermann P & Landolt H-P. Adenosinergic mechanisms contribute to indi-

- vidual differences in sleep deprivation-induced changes in neurobehavioral function and brain rhythmic activity. *J Neurosci* 2006; 26: 10472-10479.
- [11] Caldwell JA, Mu Q, Smith JK, *et al.* Are individual differences in fatigue vulnerability related to baseline differences in cortical activation? *Behav Neurosci* 2005; 119: 694-707.
- [12] Mu Q, Mishory A, Johnson KA, *et al.* Decreased brain activation during a working memory task at rested baseline is associated with vulnerability to sleep deprivation. *Sleep* 2005; 28: 433-446.
- [13] Chuah YML, Venkatraman V, Dinges DF, Chee MWL. The neural basis of interindividual variability in inhibitory efficiency after sleep deprivation. *J Neurosci* 2006; 26: 7156-7162.
- [14] Rétey JV, Adam M, Gottselig JM, *et al.* Adenosinergic mechanisms contribute to individual differences in sleep deprivation-induced changes in neurobehavioral function and brain rhythmic activity. *J Neurosci* 2006; 26: 10472-10479.
- [15] Viola AU, Archer SN, James LM, *et al.* PER3 polymorphism predicts sleep structure and waking performance. *Curr Biol* 2007; 17: 613-618.
- [16] Van Dongen HPA, Vitellaro KM, Dinges DF. Individual differences in adult human sleep and wakefulness: Leitmotif for a research agenda. *Sleep* 2005; 28: 479-496.
- [17] Van Dongen HPA, Mott CG, Huang JK, Mollicone DJ, McKenzie FD, Dinges DF. Optimization of biomathematical model predictions for cognitive performance impairment in individuals: Accounting for unknown traits and uncertain states in homeostatic and circadian processes. *Sleep* 2007; 30: 1129-1143.
- [18] Phillips AJK, Robinson PA. A quantitative model of sleep-wake dynamics based on the physiology of the brainstem ascending arousal system. *J Biol Rhythms* 2007; 22: 167-179.
- [19] Saper CB, Scammell TE, Lu J. Hypothalamic regulation of sleep and circadian rhythms. *Nature* 2005; 437: 1257-1263.
- [20] Tononi G, Cirelli C. Sleep function and synaptic homeostasis. *Sleep Med Rev* 2006; 10: 49-62.
- [21] Krueger JM, Rector DM, Roy S, Van Dongen HPA, Belenky G, Panksepp J. Sleep as a fundamental property of neuronal assemblies. *Nat Rev Neurosci* 2008; 9: 910-919.
- [22] Borbély AA. A two process model of sleep regulation. *Hum Neurobiol* 1982; 1: 195-204.
- [23] Saper CB, Chou TC, Scammell TE. The sleep switch: hypothalamic control of sleep and wakefulness. *Trends Neurosci* 2001; 24: 726-731.
- [24] Van Dongen HPA, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep* 2003; 26: 117-126.
- [25] McCauley P, Kalachev LV, Smith AD, Belenky G, Dinges DF, Van Dongen HPA. A new mathematical model for the homeostatic effects of sleep loss on neurobehavioral performance. *J Theor Biol* 2009; 256: 227-239.
- [26] Rempe M, Best J, Terman D. A neurological model of the human sleep/wake cycle. 2008 [cited 2009 February 7] Technical Report 72. Columbus, Ohio: Mathematical Biosciences Institute. Available from: [http://mbi.osu.edu/publications/reports/techreport\\_72.pdf](http://mbi.osu.edu/publications/reports/techreport_72.pdf)
- [27] Chou TC, Bjorkum AA, Gaus SE, Lu J, Scammell TE, Saper CB. Afferents to the ventrolateral preoptic nucleus. *J Neurosci* 2002; 22: 977-990.
- [28] Sakurai T. The neural circuit of orexin (hypocretin): maintaining sleep and wakefulness. *Nat Rev Neurosci* 2007; 8: 171-181.
- [29] Borbély AA, Achermann P. Sleep homeostasis and models of sleep regulation. *J Biol Rhythms* 1999; 14: 557-568.
- [30] Edgar DM, Dement WC, Fuller CA. Effect of SCN lesions on sleep in squirrel monkeys: Evidence for opponent processes in sleep-wake regulation. *J Neurosci* 1993; 13: 1065-1079.
- [31] Takakawa Y, Karashima A, Koyama Y, Katayama N, Nakao M. A quartet neural system model orchestrating sleep and wakefulness mechanisms. *J Neurophysiol* 2006; 95: 2055-2069.
- [32] Behn CGD, Brown EN, Scammell TE, Kopell NJ. Mathematical model of network dynamics governing mouse sleep-wake behavior. *J Neurophysiol* 2007; 97: 3828-3840.
- [33] Fulcher BD, Phillips AJK, Robinson PA. Quantitative physiologically based modeling of subjective fatigue during sleep deprivation. *J Theor Biol* 2010; 264: 407-419.
- [34] Dijk DJ, Archer SN. PERIOD3, circadian phenotypes, and sleep homeostasis. *Sleep Med Rev* 2010; 14(3): 151-160.
- [35] Richardson GS. Human physiological models of insomnia. *Sleep Med* 2007; 8(Suppl 4): S9-S14.
- [36] Daan S, Beersma DGM, Borbély AA. Timing of human sleep: recovery process gated by a circadian pacemaker. *Am J Physiol* 1984; 246: R161-R178.
- [37] Olofsen E, Dinges DF, Van Dongen HPA. Nonlinear mixed-effects modeling: Individualization and prediction. *Aviat Space Environ Med* 2004; 75: A134-A140.
- [38] Dorrian J, Rogers NL, Dinges DF. Psychomotor vigilance performance: A neurocognitive assay sensitive to sleep loss. In: Kushida C, Ed. *Sleep deprivation: Clinical issues, pharmacology and sleep loss effects*. New York: Marcel Dekker 2005; 39-70.
- [39] Lee M. *Bayesian statistics - an introduction*. London: Arnold, 2004.
- [40] Holford N. The visual predictive check - superiority to standard diagnostic (Rorschach) plots. Proceedings of the 14th meeting of the Population Approach Group in Europe; 2005 June 16-17; Pamplona, Spain. 2005 [cited 2009 February 7] Abstract 738. Available from: <http://www.page-meeting.org/default.asp?abstract=738>.
- [41] Dinges DF. Critical research issues in development of biomathematical models of fatigue and performance. *Aviat Space Environ Med* 2004; 75: A181-A191.
- [42] Maislin G. Commentary on a model to predict work-related fatigue based on hours of work. *Aviat Space Environ Med* 2004; 75: A70-A71.
- [43] Friedl KE, Mallis MM, Ahlers ST, Popkin SM, Larkin W. Research requirements for operational decision-making using models of fatigue and performance. *Aviat Space Environ Med* 2004; 75: A192-A199.
- [44] Van Dongen HPA, Balkin TJ. Commentary on modeling performance and alertness: The QinetiQ approach. *Aviat Space Environ Med* 2004; 75: A104-A105.
- [45] Dijk DJ, Larkin W. Fatigue and performance models: General background and commentary on the Circadian Alertness Simulator for fatigue risk assessment in transportation. *Aviat Space Environ Med* 2004; 75: A119-A121.
- [46] Reifman J, Gander P. Commentary on the three-process model of alertness and broader modeling issues. *Aviat Space Environ Med* 2004; 75: A84-A88.
- [47] Belyavin AJ, Spencer MB. Modeling performance and alertness: the QinetiQ approach. *Aviat Space Environ Med* 2004; 75: A93-A103.
- [48] Vitellaro KM, Levin AA, Ballas C, Dinges DF, Van Dongen HPA. Neurobehavioral performance under varying workload conditions during repeated exposure to sleep deprivation. In: Van Bommel AL, Beersma DGM, Hofman WF, Ruigt GSF, Vos PJE, Eds. *Sleep-wake research in the Netherlands*, Dutch Society for Sleep-Wake Research 2003; vol. 14: 106-109.
- [49] Tucker AM, Dinges DF, Van Dongen HPA. Trait interindividual differences in the sleep physiology of healthy young adults. *J Sleep Res* 2007; 16: 170-180.
- [50] Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behav Res Meth Instr Comp* 1985; 17: 652-655.
- [51] Thorne DR, Genser SG, Sing GC, Hegge FW. The Walter Reed performance assessment battery. *Neurobehav Toxicol Teratol* 1985; 7: 415-418.
- [52] Van Dongen HPA, Dinges DF. Sleep, circadian rhythms, and psychomotor vigilance. *Clin Sports Med* 2005; 24: 237-249.
- [53] Achermann P, Borbély AA. Simulation of daytime vigilance by the additive interaction of a homeostatic and a circadian process. *Biol Cybern* 1994; 71: 115-121.
- [54] Kronauer RE, Gunzelmann G, Van Dongen HPA, Doyle FJ III, Klerman EB. Uncovering physiologic mechanisms of circadian rhythms and sleep/wake regulation through mathematical modeling. *J Biol Rhythms* 2007; 22: 233-245.
- [55] Van Dongen HPA, Maislin G, Dinges DF. Dealing with inter-individual differences in the temporal dynamics of fatigue and performance: Importance and techniques. *Aviat Space Environ Med* 2004; 75: A147-A154.
- [56] Doran SM, Van Dongen HPA, Dinges DF. Sustained attention performance during sleep deprivation: Evidence of state instability. *Arch Ital Biol* 2001; 139: 253-267.
- [57] Frey DJ, Badia P, Wright KP Jr. Inter- and intra-individual variability in performance near the circadian nadir during sleep deprivation. *J Sleep Res* 2004; 13: 305-315.

- [58] Åkerstedt T, Connor J, Gray A, Kecklund G. Predicting road crashes from a mathematical model of alertness regulation - the Sleep/Wake Predictor. *Accid Anal Prev* 2008; 40: 1480-1485.
- [59] Van Dongen HPA, Hursh SR. Fatigue, performance, errors and accidents. In: Kryger MH, Roth T, Dement WC, Eds. *Principles and practice of sleep medicine*, 5<sup>th</sup> ed. Philadelphia: Elsevier; 2010, in press.

---

Received: March 22, 2009

Revised: July 29, 2009

Accepted: May 11, 2010

© Olofsen *et al.*; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.