



Published in final edited form as:

J Theor Biol. 2009 January 21; 256(2): 227–239. doi:10.1016/j.jtbi.2008.09.012.

A new mathematical model for the homeostatic effects of sleep loss on neurobehavioral performance

Peter McCauley^a, Leonid V. Kalachev^b, Amber D. Smith^a, Gregory Belenky^a, David F. Dinges^c, and Hans P.A. Van Dongen^{a,*}

^a Sleep and Performance Research Center, Washington State University, PO Box 1495, Spokane, WA 99210, USA (pmccauley@wsu.edu; ambersmith@wsu.edu; belenky@wsu.edu; hvd@wsu.edu)

^b Department of Mathematical Sciences, University of Montana, Mathematics Building, Missoula, MT 59812, USA (kalachev@mso.umt.edu)

^c Division of Sleep and Chronobiology, Department of Psychiatry, University of Pennsylvania School of Medicine, 1013 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, USA (dinges@mail.med.upenn.edu)

Abstract

The two-process model of sleep regulation makes accurate predictions of sleep timing and duration for a variety of experimental sleep deprivation and nap sleep scenarios. Upon extending its application to waking neurobehavioral performance, however, the model fails to predict the effects of chronic sleep restriction. Here we show that the two-process model belongs to a broader class of models formulated in terms of coupled non-homogeneous first-order ordinary differential equations, which have a dynamic repertoire capturing waking neurobehavioral functions across a wide range of wake/sleep schedules. We examine a specific case of this new model class, and demonstrate the existence of a bifurcation: for daily amounts of wakefulness less than a critical threshold, neurobehavioral performance is predicted to converge to an asymptotically stable state of equilibrium; whereas for daily wakefulness extended beyond the critical threshold, neurobehavioral performance is predicted to diverge from an unstable state of equilibrium. Comparison of model simulations to laboratory observations of lapses of attention on a psychomotor vigilance test (PVT), in experiments on the effects of chronic sleep restriction and acute total sleep deprivation, suggests that this bifurcation is an essential feature of performance impairment due to sleep loss. We present three new predictions that may be experimentally verified to validate the model. These predictions, if confirmed, challenge conventional notions about the effects of sleep and sleep loss on neurobehavioral performance. The new model class implicates a biological system analogous to two connected compartments containing interacting compounds with time-varying concentrations as being a key mechanism for the regulation of psychomotor vigilance as a function of sleep loss. We suggest that the adenosinergic neuromodulator/receptor system may provide the underlying neurobiology.

*Corresponding author: Hans P.A. Van Dongen, PhD, Associate Research Professor, Sleep and Performance Research Center, Washington State University, Spokane, PO Box 1495, Spokane, WA 99210-1495, USA, Tel: +1 509 358 7755, Fax: +1 509 358 7810, E-mail: hvd@wsu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Fatigue and performance models; Two-process model of sleep regulation; Coupled non-homogeneous first-order ordinary differential equations; Cognitive performance; Adenosine receptors

1. Introduction

Sleep deprivation causes a wide range of neurobehavioral performance deficits (Dinges and Kribbs, 1991; Banks and Dinges, 2007). Various mathematical “fatigue and performance models” have been developed to predict such performance impairment (Mallis et al., 2004). However, scientific progress in this area has been limited by difficulties predicting performance under chronic conditions of partial sleep loss (Van Dongen et al., 2003; Van Dongen, 2004).

Most of the available fatigue and performance models are based on the seminal two-process model of sleep regulation (Borbély, 1982; Daan et al., 1984). This model posits that sleep and wakefulness are governed by two primary biological mechanisms: a homeostatic process that builds pressure for sleep during wakefulness and dissipates this pressure during sleep (Borbély and Achermann, 1999), and a circadian process that modulates sleep pressure as a function of time of day (Edgar et al., 1993). The two-process model has been successful in predicting various aspects of sleep and of waking neurobehavioral functions across a range of sleep and sleep deprivation paradigms (Borbély and Achermann, 1999; Achermann, 2004). For instance, it was shown that waking neurobehavioral functions could—in many instances—be predicted by the arithmetic difference between the homeostatic pressure for sleep and the circadian pressure for wakefulness (Achermann and Borbély, 1994).

Extending the two-process model from its original focus on sleep (Borbély, 1982) to include predictions of waking functions has been a goal for some time (Borbély and Achermann, 1999; Dinges and Achermann, 1999), but efforts to achieve this goal have not been universally successful. Several studies have shown that chronic sleep restriction leads to cumulative increases, progressing over days for a week or more, in sleep propensity and neurobehavioral impairment (Carskadon and Dement, 1981; Dinges et al., 1997; Belenky et al., 2003; Van Dongen et al., 2003)—see Fig. 1a. The two-process model does not accurately capture these increasing deficits, predicting instead a stabilization of waking neurobehavioral functions across days after just a few days of chronic sleep loss (Van Dongen et al., 2003)—see Fig. 1b. Other fatigue and performance models have similarly failed to predict the cumulative effects of chronic sleep restriction (Van Dongen, 2004).

Van Dongen et al. (2003) proposed a different model, shifting the emphasis from sleep loss to cumulative wake extension or “excess wakefulness”. This subtle conceptual difference provided a parsimonious explanation for the effects on waking functions of both acute total sleep deprivation and chronic partial sleep deprivation (Van Dongen et al., 2003; Van Dongen and Dinges, 2003b). Nevertheless, the excess wakefulness model is not useful for computational predictions of neurobehavioral impairment, because it does not explicitly state how recovery from the effects of prior sleep loss would be achieved.

Alternative solutions were introduced by Hursh et al. (2004) and by Johnson et al. (2004), who each included an additional regulatory process modulating their versions of the homeostatic process, in order for their models to account for the cumulative effects of chronic sleep restriction. Based on the approach proposed by Johnson et al. (2004), Avinash et al. (2005) then extended the original two-process model. Their objective was to capture the effects of chronic sleep restriction on waking neurobehavioral performance (Van Dongen et al., 2003)

while retaining the successes of the original two-process model in predicting other aspects of waking functions and sleep (Achermann, 2004). As the present paper builds upon this work, the mathematical basis is briefly reiterated here.

The homeostatic process of the original two-process model is typically represented as a pair of difference equations (Borbély and Achermann, 1999):

$$S_t = 1 - e^{-\frac{\Delta t}{\tau_r}} (1 - S_{t-\Delta t}) \quad \text{during wake,} \quad (1a)$$

$$S_t = e^{-\frac{\Delta t}{\tau_d}} S_{t-\Delta t} \quad \text{during sleep.} \quad (1b)$$

Here S is the homeostatic sleep pressure as a function of time t ; Δt is the time step; and $\tau_r > 0$ and $\tau_d > 0$ are time constants for the rise and decay of the homeostatic process during wakefulness and sleep, respectively. The reason the two-process model predicts excessively rapid stabilization of performance across days of sleep restriction is related to the asymptotic properties of Eqs. (1). Specifically, the wake equation tends to a steady state represented by an upper asymptote $U = 1$, while the sleep equation tends to a steady state represented by a lower asymptote $V = 0$. This asymptotic behavior can be demonstrated by rewriting Eqs. (1):

$$S_t - U_t = (S_{t-\Delta t} - U_{t-\Delta t}) e^{-\frac{\Delta t}{\tau_r}} \quad \text{during wake,} \quad (2a)$$

$$S_t - V_t = (S_{t-\Delta t} - V_{t-\Delta t}) e^{-\frac{\Delta t}{\tau_d}} \quad \text{during sleep.} \quad (2b)$$

The extension of the two-process model by Avinash et al. (2005) involved modulating the homeostatic process through manipulation of the asymptotes U and V in Eqs. (2), as follows:

$$U_t = U_{t-\Delta t} + \mu_r \Delta t \quad \text{during wake,} \quad (3a)$$

$$U_t = U_{t-\Delta t} + (1 - U_{t-\Delta t}) (1 - e^{-\frac{\Delta t}{\mu_d}}) \quad \text{during sleep,} \quad (3b)$$

$$V_t = U_t - 1. \quad (3c)$$

Here $\mu_r > 0$ represents the slope of a linear rise of the asymptotes during wakefulness, and $\mu_d > 0$ represents the time constant of an exponential decay of the asymptotes during sleep.

The model proposed by Avinash et al. (2005) performed better at capturing the cumulative deficits in neurobehavioral performance across days as induced by chronic sleep restriction, but at the cost of reduced accuracy in describing the magnitude of the effects across days of acute total sleep deprivation—see Fig. 1c. However, it can be shown that the model of Eqs. (3) belongs to a broader class of homeostatic models based on the same principles, which may offer further improvements in predicting performance impairment across days of sleep loss.

In this paper, we introduce this new model class—which to our knowledge has not been previously proposed in the published literature—and investigate its general dynamic properties. We do this first by examining, across wake/sleep cycles, the predicted levels of performance at the onset of wakefulness and at the onset of sleep. These predictions are notional—the predictions at wake onset do not account for transient performance impairment due to sleep inertia (e.g., Dinges et al., 1981; Dinges, 1990), and the predictions at sleep onset have no real meaning because the person is asleep. However, they completely describe the model behavior across wake/sleep cycles, and as such serve as useful anchor points to examine the pattern of neurobehavioral performance changes over consecutive days.

Next, for a specific case of the new model class, we compare model predictions to actual laboratory observations of lapses of attention on a psychomotor vigilance test (PVT; Dinges and Powell, 1985; Dorrian et al., 2005; Lim and Dinges, 2008), across consecutive days of acute total sleep deprivation or chronic partial sleep restriction. Finally, based on the modeling results for PVT lapses of attention, we infer a possible neurobiological mechanism underlying the dynamic effects of sleep and sleep loss on neurobehavioral performance.

2. A New Class of Models Formulated in Terms of Coupled Non-Homogeneous First-Order Ordinary Differential Equations

2.1. Defining the new model class

Beginning with the original two-process model (Achermann and Borbély, 1994), we can write model equations for neurobehavioral performance as:

$$p_n(t) = w_n(t) - c(t) \quad \text{for } t \in [t_n, t_n + W_n] \quad (\text{i.e., during wake}), \quad (4a)$$

$$q_n(t) = s_n(t) - c(t) \quad \text{for } t \in [t_n + W_n, t_n + T_n] \quad (\text{i.e., during sleep}). \quad (4b)$$

The variables w_n and s_n denote the homeostatic pressure during wakefulness and sleep, respectively, in the n^{th} wake/sleep cycle (i.e., day; $n = 0, 1, \dots$). The function $c(t)$ is the original circadian process (see Borbély and Achermann, 1999). Further, t_n denotes the time of the beginning of the n^{th} wake/sleep cycle, T_n is the total duration of the n^{th} cycle (such that $t_{n+1} = t_n + T_n$), and W_n is the duration of wakefulness in the n^{th} cycle. We require that $0 < W_n \leq T_n$, where $W_n = T_n$ corresponds to total sleep deprivation. Finally, p_n and q_n are the predictions for performance during wakefulness and sleep, respectively, in the n^{th} wake/sleep cycle. The predictions during sleep are notional; they are included strictly for continuity between consecutive wake/sleep cycles. Here p_n and q_n are coupled as follows:

$$p_n(t_n + W_n) = q_n(t_n + W_n), \quad (5a)$$

$$q_n(t_n + T_n) = p_{n+1}(t_{n+1}). \quad (5b)$$

The homeostatic process of Eqs. (1) may be written in the form of a system of first-order ordinary differential equations (ODEs):

$$\frac{dw_n}{dt} = \frac{-1}{\tau_r} (w_n - 1) \quad \text{for } t \in [t_n, t_n + W_n], \quad (6a)$$

$$\frac{ds_n}{dt} = \frac{-1}{\tau_d} s_n \quad \text{for } t \in [t_n + W_n, t_n + T_n]. \quad (6b)$$

Note that w_n and s_n are still functions of time t , but to reduce clutter in later differential equations this is no longer indicated explicitly. From Eqs. (6) it follows that Eqs. (4) can also be written as a system of first-order ODEs:

$$\frac{dp_n}{dt} = \frac{-1}{\tau_r} p_n + \beta(t) \quad \text{for } t \in [t_n, t_n + W_n], \quad (7a)$$

$$\frac{dq_n}{dt} = \frac{-1}{\tau_d} q_n + \gamma(t) \quad \text{for } t \in [t_n + W_n, t_n + T_n], \quad (7b)$$

where p_n and q_n are again coupled as per Eqs. (5). The non-homogeneities $\beta(t)$ and $\gamma(t)$ represent the circadian process, and may be generalized to include other non-homeostatic influences on performance.

The system of Eqs. (7) is an exact representation of the original two-process model (Borbély and Achermann, 1999). In the same manner, the extended two-process model of Avinash et al. (2005) can be written as a system of coupled non-homogeneous first-order ODEs:

$$\begin{bmatrix} \dot{p}_n \\ \dot{u}_n \end{bmatrix} = \begin{bmatrix} -1/\tau_r & 1/\tau_r \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_n \\ u_n \end{bmatrix} + \begin{bmatrix} \beta_1(t) \\ \beta_2(t) \end{bmatrix} \quad \text{for } t \in [t_n, t_n + W_n], \quad (8a)$$

$$\begin{bmatrix} \dot{q}_n \\ \dot{v}_n \end{bmatrix} = \begin{bmatrix} -1/\tau_d & (1/\tau_d - 1/\mu_d) \\ 0 & -1/\mu_d \end{bmatrix} \begin{bmatrix} q_n \\ v_n \end{bmatrix} + \begin{bmatrix} \gamma_1(t) \\ \gamma_2(t) \end{bmatrix} \quad \text{for } t \in [t_n + W_n, t_n + T_n]. \quad (8b)$$

Here u_n and v_n are the levels of the upper and lower asymptotes, respectively, in the n^{th} wake/sleep cycle. The non-homogeneities $\beta_1(t)$ and $\gamma_1(t)$ are bounded, oscillatory functions representing the circadian process and other non-homeostatic influences on performance. Likewise, $\beta_2(t)$ and $\gamma_2(t)$ are bounded, oscillatory functions representing any circadian or other non-homeostatic effects there might be on the levels of the upper and lower asymptotes. Note that in this notation, $\beta_1(t)$ and $\beta_2(t)$ have absorbed the parameter μ_r (i.e., the slope of the linear rise of the upper asymptote during wakefulness). Analogous to Eqs. (5), Eqs. (8) are coupled as follows:

$$\begin{bmatrix} p_n(t_n + W_n) \\ u_n(t_n + W_n) \end{bmatrix} = \begin{bmatrix} q_n(t_n + W_n) \\ v_n(t_n + W_n) + \delta \end{bmatrix}, \quad (9a)$$

$$\begin{bmatrix} q_n(t_n + T_n) \\ v_n(t_n + T_n) \end{bmatrix} = \begin{bmatrix} p_{n+1}(t_{n+1}) \\ u_{n+1}(t_{n+1}) - \delta \end{bmatrix}, \quad (9b)$$

where $\delta > 0$ is the distance between the two asymptotes. For the extended two-process model, $\delta = 1$ (Avinash et al., 2005).

When we write Eqs. (8) in generalized form, it becomes clear that there is an asymmetry in the extended two-process model of Avinash et al. (2005):

$$\begin{bmatrix} \dot{p}_n \\ \dot{u}_n \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_n \\ u_n \end{bmatrix} + \begin{bmatrix} \beta_1(t) \\ \beta_2(t) \end{bmatrix} \quad \text{for } t \in [t_n, t_n + W_n], \quad (10a)$$

$$\begin{bmatrix} \dot{q}_n \\ \dot{v}_n \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ 0 & \sigma_{22} \end{bmatrix} \begin{bmatrix} q_n \\ v_n \end{bmatrix} + \begin{bmatrix} \gamma_1(t) \\ \gamma_2(t) \end{bmatrix} \quad \text{for } t \in [t_n + W_n, t_n + T_n]. \quad (10b)$$

Namely, Eq. (10b) for sleep has one more parameter than Eq. (10a) for wakefulness. Adding the corresponding coefficient α_{22} in Eq. (10a) generates a useful new model with a bifurcation, which we will examine in detail in the next section.

Eqs. (10a) and (10b) also each have room for another parameter in their 2 by 2 coefficient matrices (i.e., α_{21} and σ_{21} , respectively). As such, we may define our new class of models, formulated in terms of coupled non-homogeneous first-order ODEs, by the following generalized equations (which incorporate the original and extended two-process models):

$$\begin{bmatrix} \dot{p}_n \\ \dot{u}_n \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} p_n \\ u_n \end{bmatrix} + \begin{bmatrix} \beta_1(t) \\ \beta_2(t) \end{bmatrix} \quad \text{for } t \in [t_n, t_n + W_n], \quad (11a)$$

$$\begin{bmatrix} \dot{q}_n \\ \dot{v}_n \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \begin{bmatrix} q_n \\ v_n \end{bmatrix} + \begin{bmatrix} \gamma_1(t) \\ \gamma_2(t) \end{bmatrix} \quad \text{for } t \in [t_n + W_n, t_n + T_n]. \quad (11b)$$

The coupling of these equations is given by Eqs. (9). Of the non-homogeneities $\beta_1(t)$, $\beta_2(t)$, $\gamma_1(t)$ and $\gamma_2(t)$ we only require that they are bounded, oscillatory functions. They co-determine the profiles of performance changes *within* wake/sleep cycles, in part through the circadian process, but this is beyond the focus of the present paper. Of primary interest are the α and σ coefficient matrices, as they determine the dynamic behavior of the system *across* wake/sleep cycles.

2.2. Dynamic properties of the new model class

For constant values of the α and σ coefficients, the general solution of the ODE system of Eqs. (11) is of the form (Derrick and Grossman, 1997):

$$\begin{bmatrix} p_n(t) \\ u_n(t) \end{bmatrix} = \psi_n(t) \psi_n^{-1}(t_n) \begin{bmatrix} p_n(t_n) \\ u_n(t_n) \end{bmatrix} + \int_{t_n}^t \psi_n(t) \psi_n^{-1}(s) \begin{bmatrix} \beta_1(s) \\ \beta_2(s) \end{bmatrix} ds, \quad (12a)$$

$$\begin{bmatrix} q_n(t) \\ v_n(t) \end{bmatrix} = \phi_n(t) \phi_n^{-1}(t_n + W_n) \begin{bmatrix} q_n(t_n + W_n) \\ v_n(t_n + W_n) \end{bmatrix} + \int_{t_n + W_n}^t \phi_n(t) \phi_n^{-1}(s) \begin{bmatrix} \gamma_1(s) \\ \gamma_2(s) \end{bmatrix} ds, \quad (12b)$$

where $\psi_n(t)$ and $\phi_n(t)$ are the respective fundamental solutions of the homogeneous parts of Eqs. (11). These fundamental solutions depend on the eigenvalues λ_i and the eigenvectors $[k_{i1} \ k_{i2}]$ of the α and σ coefficient matrices. The eigenvalues λ_1 and λ_2 and the corresponding eigenvectors of the coefficient matrix are found by solving:

$$\det \left(\begin{bmatrix} \alpha_{11} - \lambda_i & \alpha_{12} \\ \alpha_{21} & \alpha_{22} - \lambda_i \end{bmatrix} \right) = 0; \tag{13a}$$

$$\begin{bmatrix} \alpha_{11} - \lambda_i & \alpha_{12} \\ \alpha_{21} & \alpha_{22} - \lambda_i \end{bmatrix} \begin{bmatrix} k_{i1} \\ k_{i2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{13b}$$

The process is analogous for the eigenvalues λ_3 and λ_4 and the corresponding eigenvectors of the σ coefficient matrix.

The fundamental solution $\psi_n(t)$ depends on the real and distinct eigenvalues λ_1 and λ_2 found by solving Eqs. (13); and the fundamental solution $\phi_n(t)$ depends on the likewise derived real and distinct eigenvalues λ_3 and λ_4 , as follows (Derrick and Grossman, 1997):

$$\psi_n(t) = \begin{bmatrix} k_{11}e^{\lambda_1 t} & k_{21}e^{\lambda_2 t} \\ k_{12}e^{\lambda_1 t} & k_{22}e^{\lambda_2 t} \end{bmatrix}, \tag{14a}$$

$$\phi_n(t) = \begin{bmatrix} k_{31}e^{\lambda_3 t} & k_{41}e^{\lambda_4 t} \\ k_{32}e^{\lambda_3 t} & k_{42}e^{\lambda_4 t} \end{bmatrix}. \tag{14b}$$

Note that while Eqs. (14) are sensitive to shifting of the origin of the time variable t , the functions $\psi_n(t)$ and $\phi_n(t)$ end up being used only in products with their respective inverses, and these so-called principal matrix solutions are invariant to time translation.

Having found the general solution of the ODE system of Eqs. (11), difference equations can be derived for the predicted level of performance at the onset of each wake period and at the onset of each sleep period. Although these predictions for wake onset do not account for transient effects of sleep inertia (Dinges, 1990), and the predictions for sleep onset are notional (since the person is asleep), they completely describe the model behavior across wake/sleep cycles. They therefore serve as useful anchor points to examine the pattern of neurobehavioral performance changes across days.

Using Eqs. (9) and (12), the difference equations for performance at wake onset $p_n(t_n)$, and for performance at sleep onset $q_n(t_n + W_n)$, can be shown to be given by:

$$\begin{bmatrix} p_{n+1}(t_{n+1}) \\ u_{n+1}(t_{n+1}) \end{bmatrix} = \phi_n(t_n + T_n) \phi_n^{-1}(t_n + W_n) \psi_n(t_n + W_n) \psi_n^{-1}(t_n) \begin{bmatrix} p_n(t_n) \\ u_n(t_n) \end{bmatrix} + F_n, \tag{15a}$$

$$\begin{bmatrix} q_{n+1}(t_{n+1} + W_{n+1}) \\ v_{n+1}(t_{n+1} + W_{n+1}) \end{bmatrix} = \psi_{n+1}(t_{n+1} + W_{n+1}) \psi_{n+1}^{-1}(t_{n+1}) \phi_n(t_n + T_n) \phi_n^{-1}(t_n + W_n) \begin{bmatrix} q_n(t_n + W_n) \\ v_n(t_n + W_n) \end{bmatrix} + G_n, \tag{15b}$$

$$F_n = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \phi_n(t_n + T_n) \phi_n^{-1}(t_n + W_n) \right) \begin{bmatrix} 0 \\ \delta \end{bmatrix} + \int_{t_n + W_n}^{t_n + T_n} \phi_n(t_n + T_n) \phi_n^{-1}(s) \begin{bmatrix} \gamma_1(s) \\ \gamma_2(s) \end{bmatrix} ds + \phi_n(t_n + T_n) \phi_n^{-1}(t_n + W_n) \int_{t_n}^{t_n + W_n} \psi_n(t_n + W_n) \psi_n^{-1}(s) \begin{bmatrix} \beta_1(s) \\ \beta_2(s) \end{bmatrix} ds. \quad (15c)$$

$$G_n = - \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \psi_{n+1}(t_{n+1} + W_{n+1}) \psi_{n+1}^{-1}(t_{n+1}) \right) \begin{bmatrix} 0 \\ \delta \end{bmatrix} + \int_{t_{n+1}}^{t_{n+1} + W_{n+1}} \psi_{n+1}(t_{n+1} + W_{n+1}) \psi_{n+1}^{-1}(s) \begin{bmatrix} \beta_1(s) \\ \beta_2(s) \end{bmatrix} ds + \psi_{n+1}(t_{n+1} + W_{n+1}) \psi_{n+1}^{-1}(t_{n+1}) \int_{t_n + W_n}^{t_n + T_n} \phi_n(t_n + T_n) \phi_n^{-1}(s) \begin{bmatrix} \gamma_1(s) \\ \gamma_2(s) \end{bmatrix} ds. \quad (15d)$$

Changes in neurobehavioral performance *across* wake/sleep cycles depend entirely on this system of difference equations for performance at the onsets of wakefulness and sleep.

Of particular interest is whether the pattern of changes in neurobehavioral performance across wake/sleep cycles can display a steady state or equilibrium—that is, whether the performance profile within days can be found to repeat itself across days or across clusters of days when a particular wake/sleep schedule is maintained. This condition of fixed wake duration W and fixed wake/sleep cycle duration T is described by:

$$p_{n+m}(t_{n+m}) = p_n(t_n), \quad (16a)$$

$$q_{n+m}(t_{n+m} + W) = q_n(t_n + W), \quad (16b)$$

where $m = 1, 2, \dots$ is the number of wake/sleep cycles after which the performance pattern repeats itself. If the oscillation period τ of the non-homogeneities $\beta_1(t)$, $\beta_2(t)$, $\gamma_1(t)$ and $\gamma_2(t)$ equals the wake/sleep cycle duration T , as is the case under conditions of circadian entrainment, then the steady state performance pattern would be expected to repeat itself every day (i.e., $m = 1$). If $\tau \neq T$, then a beat phenomenon could occur in which the performance pattern repeats itself every m days. Forced desynchrony protocols (e.g., Dijk and Czeisler, 1994) are based on this latter idea.

Indeed, for fixed wake duration W and fixed wake/sleep cycle duration T , and assuming that the non-homogeneities oscillate with period $\tau = T$, Eqs. (14), (15c) and (15d) may become repetitive across wake/sleep cycles n (where $m = 1$). The equilibrium state for this case, which we denote as $[p(t_n) \ u(t_n)]$ and $[q(t_n + W) \ v(t_n + W)]$ for wakefulness and for sleep, respectively, can be derived by solving Eqs. (15), which results in:

$$\begin{bmatrix} p(t_n) \\ u(t_n) \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \Phi \Psi \right)^{-1} F, \quad (17a)$$

$$\begin{bmatrix} q(t_n + W) \\ v(t_n + W) \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \Psi \Phi \right)^{-1} G, \quad (17b)$$

$$\Psi = \psi(t_n + W)\psi^{-1}(t_n), \tag{17c}$$

$$\Phi = \phi(t_n + T)\phi^{-1}(t_n + W). \tag{17d}$$

Because of the matrix inversions embedded in Eqs. (17a) and (17b), a state of equilibrium can only occur if:

$$\det \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \Phi\Psi \right) \neq 0, \text{ and} \tag{18a}$$

$$\det \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \Psi\Phi \right) \neq 0. \tag{18b}$$

We will examine this condition for a specific case of the models defined by Eqs. (11), later in this paper.

Provided a state of equilibrium is shown to exist, the question arises whether it is stable, that is, whether the model predictions would converge to this state for a repetitive wake/sleep schedule. We can say that the model is asymptotically stable (for $m = 1$) or asymptotically periodic (for $m > 1$), if (Kelly and Peterson, 2001):

$$\lim_{n \rightarrow \infty} \begin{bmatrix} p_{n+m}(t_{n+m}) \\ u_{n+m}(t_{n+m}) \end{bmatrix} = \begin{bmatrix} p(t_n) \\ u(t_n) \end{bmatrix}, \tag{19a}$$

$$\lim_{n \rightarrow \infty} \begin{bmatrix} q_{n+m}(t_{n+m} + W) \\ v_{n+m}(t_{n+m} + W) \end{bmatrix} = \begin{bmatrix} q(t_n + W) \\ v(t_n + W) \end{bmatrix}, \tag{19b}$$

even if the starting values $[p_0(t_0) \ u_0(t_0)]$ and $[q_0(t_0) \ v_0(t_0)]$ are not already at equilibrium.

Because Eqs. (15) are linear in $[p_n(t_n) \ u_n(t_n)]$ and $[q_n(t_n + W) \ v_n(t_n + W)]$, it can be shown (Kelly and Peterson, 2001) that states of equilibrium are asymptotically stable or periodic if all eigenvalues Λ_i of the system of Eqs. (15), whether real or complex, are inside the unit circle (i.e., $|\Lambda_i| < 1$). For fixed wake duration W and fixed wake/sleep cycle duration T , and assuming again that the non-homogeneities oscillate with period $\tau = T$ (so that $m = 1$), these eigenvalues are found by solving the characteristic equations:

$$\det \left(\Phi\Psi - \Lambda_i \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = 0, \tag{20a}$$

$$\det \left(\Psi \Phi - \Lambda_i \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = 0. \tag{20b}$$

The eigenvalues derived from Eq. (20a) are identical to those derived from Eq. (20b); and the equilibrium states $[p(t_n) \ u(t_n)]$ and $[q(t_n + W) \ v(t_n + W)]$ are either both asymptotically stable, or both unstable. When both are asymptotically stable, it means that the predicted levels of performance at wake onset, at sleep onset, and by extension at all points in between, converge across days to a steady state repeated from day to day. When both are unstable, it implies that the predicted levels of performance at wake onset, at sleep onset, and in between, diverge across days toward infinity. Note that over days it grows progressively more difficult to maintain a wake/sleep schedule that induces diverging increases in performance deficits. In practice, sleep tends to break through into scheduled wakefulness (e.g., Doran et al., 2001) before performance becomes catastrophically impaired.

3. A Model with a Bifurcation

We now consider a particular case of the model of Eqs. (11):

$$\begin{bmatrix} \dot{p}_n \\ \dot{u}_n \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ 0 & \alpha_{22} \end{bmatrix} \begin{bmatrix} p_n \\ u_n \end{bmatrix} + \begin{bmatrix} \beta_1(t) \\ \beta_2(t) \end{bmatrix} \quad \text{for } t \in [t_n, t_n + W_n], \tag{21a}$$

$$\begin{bmatrix} \dot{q}_n \\ \dot{v}_n \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ 0 & \sigma_{22} \end{bmatrix} \begin{bmatrix} q_n \\ v_n \end{bmatrix} + \begin{bmatrix} \gamma_1(t) \\ \gamma_2(t) \end{bmatrix} \quad \text{for } t \in [t_n + W_n, t_n + T_n], \tag{21b}$$

where $\alpha_{11} < 0$ and $\sigma_{11} < 0$, and where $\alpha_{11} \neq \alpha_{22}$ and $\sigma_{11} \neq \sigma_{22}$. The coupling of these equations is given by Eqs. (9). As before, we require that the non-homogeneities $\beta_1(t)$, $\beta_2(t)$, $\gamma_1(t)$ and $\gamma_2(t)$ are bounded, oscillatory functions. Expediently, we also assume that the non-homogeneities oscillate with period $\tau = T$.

Per Eqs. (13), the (real and distinct) eigenvalues of the α and σ coefficient matrices are: $\lambda_1 = \alpha_{11} < 0$; $\lambda_2 = \alpha_{22}$; $\lambda_3 = \sigma_{11} < 0$; and $\lambda_4 = \sigma_{22}$. Through Eqs. (14), these eigenvalues determine the existence of states of equilibrium as assessed using Eqs. (18). Under conditions of fixed wake duration W and fixed wake/sleep cycle duration T , Eqs. (18) reduce to the following sole inequality: $(1 - e^{\alpha_{11}W} e^{\sigma_{11}(T-W)})(1 - e^{\alpha_{22}W} e^{\sigma_{22}(T-W)}) \neq 0$. If both α parameters and both σ parameters are negative, this inequality is satisfied and thus states of equilibrium exist for all $0 < W \leq T$ (both for performance at wake onset and for performance at sleep onset). If either $\alpha_{22} \geq 0$ or $\sigma_{22} \geq 0$, however, there may be a critical amount of daily wakefulness W_c , with $0 < W_c \leq T$, for which no equilibrium exists:

$$W_c = \frac{\sigma_{22}}{\sigma_{22} - \alpha_{22}} T. \tag{22}$$

To assess the stability of the states of equilibrium when they do exist, we solve Eqs. (20), which results in the following eigenvalues:

$$\Lambda_1 = e^{\alpha_{11}W} e^{\sigma_{11}(T-W)}, \tag{23a}$$

$$\Lambda_2 = e^{\alpha_{22}W} e^{\sigma_{22}(T-W)}. \tag{23b}$$

If all α and σ parameters in Eqs. (23) are negative, then $0 < \Lambda_i < 1$ for both eigenvalues, meaning that the equilibrium states (which then exist for all $0 < W \leq T$) are always asymptotically stable. Since the model system considered here is linear, this stability is global (i.e., the predictions converge to equilibrium regardless of initial conditions) (Verhulst, 2000).

If α_{22} (the key parameter distinguishing the model given by Eqs. (21) from the extended two-process model) is positive, there are three possibilities for Λ_2 . In order of increasing amount of wake extension (i.e., greater sleep restriction), these possibilities are:

- For $W < W_c$, we find that $0 < \Lambda_2 < 1$, which implies globally asymptotically stable states of equilibrium;
- For $W = W_c$, no equilibrium exists (see above);
- For $W > W_c$, we find that $\Lambda_2 > 1$, which implies that the states of equilibrium are unstable.

Thus, for $\alpha_{22} > 0$, the model behavior is such that if the amount of wakefulness W in each wake/sleep cycle exceeds a critical threshold W_c , the model flips from a state in which performance predictions converge toward an asymptotically stable equilibrium, to a state in which performance predictions diverge away from an unstable equilibrium. This qualitative change in dynamic behavior implies a bifurcation, as illustrated in Fig. 2.

It is instructive to study the model behavior when daily wakefulness is kept constant at the bifurcation value: $W = W_c$. Here, the generalized iterative system of Eqs. (15) assumes the following specific form:

$$\begin{bmatrix} p_{n+1}(t_{n+1}) \\ u_{n+1}(t_{n+1}) \end{bmatrix} = \begin{bmatrix} e^{\frac{\alpha_{22}\sigma_{11}-\alpha_{11}\sigma_{22}}{\sigma_{22}-\sigma_{22}}T} & f \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_n(t_n) \\ u_n(t_n) \end{bmatrix} + F, \tag{24a}$$

$$\begin{bmatrix} q_{n+1}(t_{n+1}+W_c) \\ v_{n+1}(t_{n+1}+W_c) \end{bmatrix} = \begin{bmatrix} e^{\frac{\alpha_{22}\sigma_{11}-\alpha_{11}\sigma_{22}}{\sigma_{22}-\sigma_{22}}T} & g \\ 0 & 1 \end{bmatrix} \begin{bmatrix} q_n(t_n+W_c) \\ v_n(t_n+W_c) \end{bmatrix} + G, \tag{24b}$$

$$f = \frac{\sigma_{12}(e^{\sigma_{11}(T-W_c)} - e^{\sigma_{22}(T-W_c)})e^{\alpha_{22}W_c}}{\sigma_{11} - \sigma_{22}} + \frac{\alpha_{12}(e^{\alpha_{11}W_c} - e^{\alpha_{22}W_c})e^{\sigma_{11}(T-W_c)}}{\alpha_{11} - \alpha_{22}}, \tag{24c}$$

$$g = \frac{\sigma_{12}(e^{\sigma_{11}(T-W_c)} - e^{\sigma_{22}(T-W_c)})e^{\alpha_{11}W_c}}{\sigma_{11} - \sigma_{22}} + \frac{\alpha_{12}(e^{\alpha_{11}W_c} - e^{\alpha_{22}W_c})e^{\sigma_{22}(T-W_c)}}{\alpha_{11} - \alpha_{22}}, \tag{24d}$$

The solution of this system tends to a straight line as $n \rightarrow \infty$. The change across days for performance at wake onset and sleep onset is defined, respectively, by slopes M_p and M_q :

$$M_p = \frac{f \cdot F_2}{1 - e^{-\frac{\sigma_{22}\sigma_{11} - \sigma_{11}\sigma_{22}}{\sigma_{22} - \sigma_{22}} T}}, \quad (25a)$$

$$M_q = \frac{g \cdot G_2}{1 - e^{-\frac{\sigma_{22}\sigma_{11} - \sigma_{11}\sigma_{22}}{\sigma_{22} - \sigma_{22}} T}}, \quad (25b)$$

where F_2 and G_2 are the second element of vectors F and G in Eqs. (15). From Eqs. (25) it follows that the slopes of change across days are not necessarily the same for performance at wake onset and performance at sleep onset.

4. Model Simulations

4.1. Comparison to data from sleep restriction experiments

To compare the model given by Eqs. (21) to actual performance observations under conditions of sleep loss, we fit it to group-average data of performance lapses on a psychomotor vigilance test (PVT; Dinges and Powell, 1985; Dorrian et al., 2005; Lim and Dinges, 2008) from a study of healthy young adults exposed to chronic sleep restriction or total sleep deprivation—with $W = 16\text{h}, 18\text{h}, 20\text{h},$ or 24h (Van Dongen et al., 2003). These data are shown in Fig. 1a.

For the non-homogeneities, we make use of the circadian process $c(t)$ defined by Borbély and Achermann (1999), applied to the performance predictions p_n and q_n (but not the asymptotes u_n and v_n), as follows:

$$\begin{bmatrix} \beta_1(t) \\ \beta_2(t) \end{bmatrix} = \begin{bmatrix} \gamma c(t - \theta) + \mu \\ 0 \end{bmatrix} \quad \text{for } t \in [t_n, t_n + W_n], \quad (26a)$$

$$\begin{bmatrix} \gamma_1(t) \\ \gamma_2(t) \end{bmatrix} = \begin{bmatrix} \gamma c(t - \theta) + \mu \\ 0 \end{bmatrix} \quad \text{for } t \in [t_n + W_n, t_n + T_n]. \quad (26b)$$

Here γ and μ are parameters scaling the circadian process, and θ is a phase parameter shifting it in time. For the initial conditions $[p_0(t_0) u_0(t_0)]$ we estimate the values corresponding to the equilibrium state at $W = 16\text{h}$, which characterizes the baseline condition in the study. Further, $t_0 = 7.5\text{h}$ (i.e., 07:30), and T and τ are fixed at 24h .

Using least-squares regression on all 404 data points shown in Fig. 1a, we find the following parameter estimates:

$$\left\{ \begin{array}{l} \left[\begin{array}{cc} \alpha_{11} & \alpha_{12} \\ 0 & \alpha_{22} \end{array} \right] = \left[\begin{array}{cc} -0.0135 & 0.000929 \\ 0 & 0.00743 \end{array} \right] \\ \left[\begin{array}{cc} \sigma_{11} & \sigma_{12} \\ 0 & \sigma_{22} \end{array} \right] = \left[\begin{array}{cc} -2.17 & 0.872 \\ 0 & -0.0397 \end{array} \right] \\ \delta=19.8 \\ \gamma=5.86 \\ \mu=0.472 \\ \theta=12.7 \\ p_0(t_0)=4.49 \\ u_0(t_0)=29.9 \end{array} \right. \quad (27)$$

The resulting PVT performance predictions are shown in Fig. 1d, and the predictions for the total sleep deprivation condition ($W = 24\text{h}$) are explored in more detail in Fig. 3. With the parameter estimates of Eqs. (27), the model explains 72.4% of the variance in the group-average data of Fig. 1a. It fits substantially better to the data than the original two-process model (Fig. 1b, explained variance 22.6%) and the extended two-process model (Fig. 1c, explained variance 38.4%).

Evaluation of Eq. (22) given the parameter estimates in Eqs. (27) indicates that there must be a bifurcation at $W_c = 20.2\text{h}$. That is, the model should flip from a state of convergence to a state of divergence when daily wakefulness is increased to more than 20.2h (i.e., when daily sleep is reduced to less than 3.8h).

This property can be verified by comparing model predictions to the group-average observations of PVT performance in another study of chronic sleep restriction, with $W = 15\text{h}$, 17h, 19h or 21h (Belenky et al., 2003). These data are shown in Fig. 4a. We use the model of Eqs. (21) again, apply the non-homogeneities defined in Eqs. (26), set $t_0 = 7.0\text{h}$ (i.e., 07:00) in accordance with the study design (Belenky et al., 2003), and fix all model parameters at their previously estimated values given in Eqs. (27). Applying linear scaling to account for any irrelevant differences in absolute performance outcomes (e.g., due to variations in population characteristics or performance testing conditions), we find the scaling factor to be 1.17—suitably close to 1. The corresponding performance predictions are shown in Fig. 4b. They explain 72.2% of the variance in the data, and fit well to the observed performance changes across days.

Note that the $W = 21\text{h}$ condition shows a divergent profile in both observations and predictions (Figs. 4a and 4b), which is not seen in the $W \leq 19\text{h}$ conditions in this study. This qualitative difference indicates the presence of a bifurcation. Indeed, on the basis of fitting the model to the data in Fig. 1a, we had predicted that a bifurcation should occur at $W_c = 20.2\text{h}$ (see above). The goodness-of-fit of our model to the data in Fig. 4a is consistent with this prediction, and provides a first validation of the model introduced in this paper.

4.2. New predictions for chronic sleep restriction and recovery

The value of a new model is determined, in part, by any falsifiable new predictions it makes. Here we present three specific predictions that can be tested in laboratory experiments, and that will have considerable theoretical and/or practical impact if confirmed.

The first new prediction focuses on the effectiveness of nap sleep as a means to sustain performance across days. It has previously been shown that a single 2h nap can mitigate performance impairment across extended periods of wakefulness (Dinges et al., 1988), but it is not known whether a 2h nap taken every day (i.e., $W = 22\text{h}$) can maintain performance at

reasonable levels across days. Our new model defined by Eqs. (21) and (26) with the parameter values given by Eqs. (27) predicts that since daily wake duration exceeds the bifurcation threshold ($W_c = 20.2\text{h}$), performance deficits should escalate, and thus a daily 2h nap would not suffice to maintain reasonable levels of performance.

As shown in Fig. 5, which compares the predicted effects of the 2h nap schedule to those of total sleep deprivation, the daily nap would mitigate performance impairment substantially in the first few days (cf. Van Dongen and Dinges, 2003b), but in later days this beneficial effect would increasingly diminish. After 8 days with a 2h nap each day, the predicted level of performance impairment approaches that of 3 days with total sleep deprivation; and if it were possible to continue the nap schedule much longer, the effectiveness of the 2h daily nap would essentially be lost. Thus, our prediction is that a daily 2h nap is not effective as a means to sustain performance across days.

The second new prediction pits the new model defined by Eqs. (21) against the only other quantitative, sleep/wake physiology-based model of the effects of chronic sleep restriction on neurobehavioral performance: the excess wakefulness model (Van Dongen et al., 2003). In that model, performance impairment across days is posited to be proportional to the cumulative amount of wakefulness exceeding a maximum period of stable wakefulness ζ (where $\zeta \approx 16\text{h}$ if prior sleep duration exceeds $\sim 4\text{h}$). This is similar to the concept of “cumulative sleep debt” (e.g., Dement, 2006), but conceptually distinct from the modeling framework introduced in the present paper.

Our prediction involves the important question of how much sleep is needed to recover from performance impairment induced by prior chronic sleep restriction (e.g., Lamond et al., 2007). The excess wakefulness model would predict that as long as wake duration exceeds ζ , performance will continue to deteriorate. On the contrary, the new model defined by Eqs. (21) and (26) with the parameter values given by Eqs. (27) would predict that when wake duration is less than the bifurcation threshold W_c , performance levels should converge to a state of equilibrium, and thus some recovery could occur if wake duration is shorter than what was maintained in the prior days of chronic sleep loss.

As a specific example, consider a scenario involving 5 days of wake extension to 20h per day (i.e., 4h sleep daily), followed by a day with wake extension to just 18h (i.e., 6h sleep). The opposing model predictions are illustrated in Fig. 6. The excess wakefulness model predicts that performance deteriorates progressively across the 5 days with 20h wakefulness, and continues to deteriorate—albeit at a slower rate—following the day with only 18h wakefulness. Our new model also predicts progressive performance degradation across the 5 days with 20h wakefulness. However, the state of equilibrium for 18h awake occurs at a lower level than the performance degradation reached after 5 days with 20h wakefulness. Therefore, the new model forecasts some degree of recovery after the 18h wakefulness day. This prediction may seem counterintuitive considering that staying awake for 18h following multiple baseline days with 16h wakefulness (8h sleep) actually leads to performance degradation (Van Dongen et al., 2003). Yet, preliminary evidence from an ongoing laboratory study (Banks et al., 2005) suggests that some recovery does occur with 18h wakefulness (6h sleep) in this chronic sleep restriction scenario, supporting the new model over the excess wakefulness model.

The third new prediction concerns the “recycle” issue, which derives from the question of whether or not there is any carry-over of performance impairment from past sleep restriction when beginning a new period of sleep restriction following limited time for recovery. Let’s consider a laboratory study currently underway (Banks et al., 2007b), which involves a period of 5 days with wake extension to 20h per day (i.e., 4h sleep daily), followed by a day with only 14h of wakefulness (i.e., 10h time for recovery sleep), followed by another period of 5 days

with wake extension to 20h per day. Initial experimental evidence would suggest that the intervening 14h wake/10h sleep day should provide (near-)complete recovery to baseline performance (Banks et al., 2007a), effectively undoing the impairment incurred by the prior sleep loss. Thus, the performance profile seen during the second 5-day period of wake extension might be expected to be similar to that seen during the first 5-day period of wake extension.

The dynamics of the new model, however, imply that the single 14h wake/10h sleep day should be seen as an intermittent perturbation in an extended series of days with wake extension to 20h per day. Thus, the model predicts that the 10h recovery sleep confers only a short-lasting performance improvement, after which performance further deteriorates as it continues to converge to the asymptotically stable equilibrium associated with $W = 20$ h. This prediction is illustrated in Fig. 7. Preliminary evidence from the laboratory study examining the scenario at hand suggests that indeed there is substantial carry-over of performance impairment from the first 5-day period with daily wake extension to the second (Banks et al., 2007b), providing tentative support for the new model.

5. Discussion

5.1. Model implications

The regulation of sleep, wakefulness and performance involves an array of possible neurobiological mechanisms (e.g., Porkka-Heiskanen et al., 1997; Krueger and Obál, 2003; Fuller et al., 2006), and is not fully understood. Nonetheless, at the behavioral level, the circadian component has been captured by models with relatively few degrees of freedom (see Indic et al., 2006). We believe the same may be possible for the sleep homeostatic component. Using evidence from laboratory studies with multiple days of sleep loss (Figs. 1a, 4a), we showed that the homeostatic regulation of neurobehavioral performance can be described by means of a system of coupled non-homogeneous first-order ODEs with only a few additional degrees of freedom relative to the homeostatic process postulated in the original two-process model (Borbély and Achermann, 1994, 1999).

Our new model does include an additional component, modulating the homeostatic process across days and weeks, as prompted by findings from chronic sleep restriction experiments demonstrated to be incongruent with the original two-process model (Van Dongen et al., 2003; Van Dongen, 2004). Yet, the model structure introduced in this paper is essentially still composed of a homeostatic process and a circadian process. Conceptually, therefore, the new model remains compatible with the principles of sleep regulation instantiated in the original two-process model (Borbély, 1982). The dynamics of the new model *across* days are principally governed by the α and σ coefficient matrices in the homogeneous part of the differential equations (the homeostatic process), while the changes *within* days are primarily governed by the non-homogeneities (the circadian process). These model components also interact, in agreement with laboratory observations of a nonlinear interaction between the homeostatic and circadian processes (Dijk et al., 1992; Van Dongen and Dinges, 2003a).

Two seminal laboratory studies first highlighted the need for fundamentally new model development beyond the two-process model in order to account for the waking neurobehavioral consequences of chronic sleep loss (Belenky et al., 2003; Van Dongen et al., 2003). However, these two studies previously drew markedly different conclusions about the dynamics of neurobehavioral impairment across days of sleep restriction. In their study with 7 days of systematic sleep restriction, Belenky et al. (2003) reported a plateau of cognitive impairment when sleep was restricted to 7h or 5h per day, as well as incomplete recuperation at the end of the study after 3 days with 8h time in bed for recovery sleep. They hypothesized that chronic sleep loss induces long-lasting adaptive changes in the brain's response to sleep loss, leading to stabilized reduced performance under conditions of sleep restriction at the cost of diminished

maximal performance capacity following recovery sleep. In contrast, in their study with 14 days of sleep restriction, Van Dongen et al. (2003) noted that performance continued to degrade when sleep was restricted to 6h or 4h per day, with no evidence of adaptation across the study period.

In the present paper, the two data sets (Figs. 1a and 4a) are examined in a single analytical framework. Using PVT performance lapses as a well-validated outcome measure (Dorrian et al., 2005) for both studies, no convincing evidence of an impairment plateau is found in either data set. Yet, our modeling results indicate that stabilization of performance impairment would occur eventually, beyond the duration of the two experiments. Furthermore, the modeling outcomes suggest that several days with recovery sleep would be needed to restore performance to baseline levels. Experiments currently underway (Wesensten et al., 2005; Banks et al., 2007a, 2007b) will shed further light on the time course of post-deprivation recovery.

Mathematical examination of the dynamics of the new model defined by Eqs. (21) revealed an unanticipated emergent model property: a bifurcation involving a critical amount of wakefulness which, if exceeded, changes the model behavior from a state of convergence toward an asymptotically stable equilibrium, to a state of divergence away from an unstable equilibrium (as illustrated in Fig. 2). This feature, previously alluded to (Belenky et al., 2003; Van Dongen and Dinges, 2003b) but as yet not considered explicitly, turned out to capture an essential aspect of the nature of performance impairment due to sleep loss. Using data from the chronic sleep restriction and total sleep deprivation experiments documented by Van Dongen et al. (2003) (Fig. 1a), we estimated the critical wakefulness threshold to occur at 20.2h (i.e., at 3.8h daily sleep). This estimate was supported by data from the chronic sleep restriction study of Belenky et al. (2003), who observed escalating performance impairment when wakefulness was extended to 21h per day (3h daily sleep condition in Fig. 4a).

The significance of the bifurcation in the new model implies that other two-process-based models of performance impairment due to chronic sleep loss (Hursh et al., 2004; Johnson et al., 2004; Avinash et al., 2005), which do not possess the bifurcation property, must have a more limited range of applicability than the new model. The excess wakefulness model (Van Dongen et al., 2003), which is based on the fundamentally different conjecture that performance impairment across days is proportional to the cumulative amount of wakefulness in excess of a ration determined by the preceding sleep period, does not *a priori* have this same limitation of scope (Van Dongen and Dinges, 2003b). However, the excess wakefulness model and the model introduced in the present paper make contradictory predictions for performance impairment after a period of chronic sleep restriction followed by a limited amount of recovery sleep (Fig. 6). This juxtaposition entails one of three testable new predictions by which our present model can be validated.

Another new prediction we made is that a daily 2h nap by itself, despite being effective to mitigate performance impairment from sleep deprivation on the short term (Dinges et al., 1988), cannot maintain reasonable levels of performance across multiple days (Fig. 5). The third new prediction, which also has real-life relevance, concerns the longevity of the performance improvement conferred by a single prolonged recovery night (“sleeping in”) preceded and followed by periods of chronic sleep restriction (Fig. 7). Our model predicts that a single recovery night interrupting a series of consecutive wake extensions to, say, 20h per day constitutes a mere temporary perturbation, after which performance levels continue to decline and converge to an asymptotically stable equilibrium. Confirmation of this prediction by experimental evidence currently being obtained (Banks et al., 2007b) will have significant implications—both theoretically, for our understanding of sleep and performance regulation, and practically, with regard to sleep/wake/work scheduling in operational settings.

The relationship between sleep and wakefulness on the one hand and neurobehavioral performance on the other hand is often conceptualized as involving depletion of performance capability by wakefulness and replenishment of performance capability by sleep (e.g., Hursh et al., 2004). This would imply that performance levels can be estimated by keeping a running tally of prior sleep and wakefulness (weighted for recency)—and that perhaps this also reflects the underlying brain mechanisms. Our examination of model system dynamics has revealed that this latter idea may need to be refined.

In our new model class (which encompasses the original two-process model and other models based upon it), performance is actually a function of the prevailing wake/sleep ratio. Provided the duration of wakefulness does not exceed the critical threshold W_c , performance predictions converge across days from the present performance level to the applicable steady state (i.e., asymptotically stable equilibrium). Sleep/wake history is fully represented by the current performance level (p or q) and the current level of the asymptote (u or v), and past amounts of sleep and wake have no further impact beyond the present. In this view, the brain does not need to maintain a running tally of sleep and wakefulness—it does not need to keep track of a “sleep debt” (cf. Dement, 2006).

Rather, based on the dynamic behavior of our new model, it seems that the effects of sleep loss and the effects of recovery sleep on waking neurobehavioral performance should be interpreted in the context of underlying physiologic balance shifts. The time constants for convergence to homeostatic balance appear to be in the order of weeks (Figs. 1d and 4b), and a state of equilibrium may seldom be achieved in practice. Still, conceptualizing the effects of sleep and sleep loss on waking performance in terms of physiologic balance shifts renders irrelevant the ostensibly irresolvable question of which components of sleep are most important for recuperation (e.g., Lubin et al., 1974).

5.2. Possible underlying neurobiological mechanisms

The dynamics of the model defined by Eqs. (21) may provide insight into the nature of the underlying neurobiological mechanisms. Conceptually, the model resembles a system of two connected compartments containing interacting substances with time-varying concentrations—one with longer time constants than the other. In this regard, our model could be a mathematical representation of the interaction between a neurotransmitter or neuromodulator and its receptor, with the density of both changing dynamically across time awake and time asleep. However, the model’s dynamic behavior and the parameter estimates we obtained (notably the finding that $\alpha_{12} > 0$ and $\sigma_{12} > 0$) point to positive feedback regulation in the system, which is not typical in neurotransmitter/neuromodulator mechanisms. Yet, such a regulatory process may be taking place in the adenosinergic system.

Adenosine is a (by)product of brain energy metabolism (Porkka-Heiskanen et al., 2002), and has been reported to induce sleepiness and impair waking functions, particularly through the cholinergic system in the basal forebrain (Basheer et al., 2000). Hence, the adenosinergic system might be a final pathway in the homeostatic regulation of sleep and waking neurobehavioral functions (Benington and Heller, 1995), and could be the temporal bottleneck that determines the time constants across days in our model. In accordance with the dynamic structure of the model, it has been observed that both extracellular adenosine level and adenosine A_1 receptor density change dynamically in response to sleep loss (Yanik and Radulovacki, 1987; Basheer et al., 2004, 2007; Elmenhorst et al., 2007; Porkka-Heiskanen et al., 2000). Moreover, sleep deprivation-induced increases in extracellular adenosine lead to concomitant increases in A_1 receptor expression, implicating positive feedback regulation (Basheer et al., 2007)—in agreement with the present model.

Based on these considerations, we propose an explanation for the effects of sleep loss on PVT performance lapses in particular, and on neurobehavioral performance in general, in terms of adenosine binding to receptors that are up- and downregulated dynamically across wake/sleep cycles. We postulate that wakefulness and sleep induce adenosine receptor upregulation and downregulation, respectively, as represented in the model by increases and decreases of the asymptotes (u and v). Thus, increased adenosine production during extended wakefulness would cause both increased sleep homeostatic pressure, inducing waking neurobehavioral impairment, and receptor upregulation. This would effectively enhance sensitivity to sleep loss on subsequent days (Basheer et al., 2007), which would serve a protective function by restraining further sleep restriction. Should additional sleep loss occur anyway, physiologic balance would shift as the rates of adenosine receptor upregulation during wakefulness and downregulation during sleep establish a new equilibrium.

However, if wakefulness is extended to more than the critical amount W_c , which we estimated here to be 20.2h, then a physiologic balance may no longer be achievable. This bifurcation, observed in both the model predictions and the experimental observations, may suggest a role for slow wave activity (SWA; $\sim 0.5\text{--}4.5$ Hz) observed in the EEG of non-REM sleep. SWA during sleep is substantially preserved when wake duration is no greater than ~ 20 h per day (Brunner et al., 1993; Van Dongen et al., 2003). However, when daily wakefulness is extended beyond ~ 20 h, then insufficient time for sleep remains to fully express SWA (see Van Dongen and Dinges, 2003b). The reduction in SWA could be related to the qualitative change in the effects of sleep loss on neurobehavioral performance when wakefulness is extended beyond the critical wake duration W_c .

Also, a connection between SWA and adenosinergic mechanisms has been noted (see Landolt, 2008). For instance, stimulation of adenosine A_1 receptors affects SWA expression in the same manner as does acute total sleep deprivation (e.g., Benington et al., 1995). Here, we hypothesize more specifically that SWA is a physiological correlate of adenosine receptor downregulation during sleep. This could explain why homeostatic balance can be achieved when wake duration is no more than approximately 20h per 24h day, as it allows enough time for sleep to preserve SWA. However, if daily wakefulness is extended beyond the bifurcation threshold, then despite SWA intensification, the overall expression of SWA is curtailed. The hypothesized adenosine receptor downregulation may thus no longer be sufficient to counter the upregulation during prior wakefulness, and a physiologic balance may no longer be achievable. As a result, adenosinergic sensitivity to sleep loss would escalate, which in turn would cause the accelerating neurobehavioral impairment that has been observed under such extreme sleep restriction conditions (Belenky et al., 2003; Van Dongen et al., 2003; Van Dongen and Dinges, 2003b).

Our proposed account of the waking neurobehavioral effects of sleep deprivation across days, postulated to be governed by dynamic changes in both adenosine production and adenosine receptor expression, may have noteworthy implications for the role of caffeine as a countermeasure for neurobehavioral impairment due to sleep loss. Caffeine's main mode of action is as an adenosine receptor antagonist (e.g., Biaggioni et al., 1991). As such, in addition to mitigating the neurobehavioral consequences of sleep loss (e.g., Penetar et al., 1993), it might also block the sleep deprivation-mediated adenosine receptor upregulation. It may thus be hypothesized that regular consumption of moderate amounts of caffeine could help to prevent increasing sensitivity to sleep loss across days of sleep restriction, which would offer a strategy for managing chronic sleep loss. Although this may already be practiced by millions of individuals around the world, how this could be effective had not really been understood mechanistically (and still needs to be confirmed through direct experimental evidence).

At higher doses, caffeine may interfere with the expression of SWA (e.g., LaJambe et al., 2005), which by extension of our hypothesis could hamper the sleep-related downregulation of adenosine receptors. Thus, depending on dose (and timing), caffeine may also be counterproductive in mitigating the waking neurobehavioral consequences of sleep deprivation. In that sense, effective use of caffeine as a countermeasure for sleep loss may not be straightforward. In safety-critical scenarios, therefore, it may be advisable to target caffeine administration with the help of a biology-based model of its physiological effects. Such a model could be developed using the mathematical framework introduced in the present paper.

5.3. Further work

We have put forth a new model formulated in terms of coupled non-homogeneous first-order ODEs, with a dynamic repertoire capturing sleep homeostatic changes in waking neurobehavioral functions across a wide range of wake/sleep schedules. Further work is needed to integrate our model with a state-of-the-art mathematical model of the circadian component (e.g., Jewett et al., 1999; St. Hilaire et al., 2007), and to deal with sleep inertia (e.g., Åkerstedt and Folkard, 1997; Jewett and Kronauer, 1999). In addition, trait-like individual differences in vulnerability to sleep loss (Van Dongen et al., 2004a) have yet to be accounted for in the new model. This will be resolved in a follow-up project using modern statistical modeling tools (e.g., Van Dongen et al., 2004b), which can also yield improved model parameter estimates (and their standard errors) as well as confidence intervals for model predictions (see Van Dongen et al., 2007; Smith et al., in press).

Finally, it should be recognized that the effects of sleep loss on waking neurobehavioral performance depend in part on which aspects of cognitive functioning are considered (Durmer and Dinges, 2005). Our present focus on PVT performance lapses entails a well-validated (Dorrian et al., 2005; Lim and Dinges, 2008) but incomplete account of neurobehavioral responses to sleep loss (e.g., see Van Dongen et al., 2004a). Ongoing efforts to connect fatigue and performance models with computational models of cognition (Gunzelmann et al., 2007) represent a promising strategy to address this issue.

Acknowledgements

We thank the Department of Computer Science at the University of Montana for use of computer resources. This work was supported by AFOSR grants FA9550-05-1-0086 and FA9550-06-1-0055; and in part by AFOSR grants F49620-95-1-0388 and FA9550-05-1-0293, DURIP grant FA9550-06-1-0281, NIH grants NR04281 and RR00040, DOT grant DTFH61-94-Y-00090, the National Space Biomedical Research Institute through NASA NCC 9-58, and USAMRMC award W81XWH-05-1-0099.

Role of funding sources

This work was entirely funded by federal agencies: the Air Force Office of Scientific Research, the National Institutes of Health, the U.S. Department of Transportation, the National Space Biomedical Research Institute, and the U.S. Army Medical Research and Materiel Command (see acknowledgments above for grant numbers). The funding agencies had no involvement in study design, in data collection and analysis, in data interpretation, in writing the manuscript, or in submission of the paper. The views and conclusions expressed in this manuscript do not necessarily represent those of the federal agencies that supported the research.

References

- Achermann P. The two-process model of sleep regulation revisited. *Aviat Space Environ Med* 2004;75:A37–A43. [PubMed: 15018264]
- 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. [PubMed: 8068773]
- Åkerstedt T, Folkard S. The three-process model of alertness and its extension to performance, sleep latency, and sleep length. *Chronobiol Int* 1997;14:115–123. [PubMed: 9095372]

- Avinash, D.; Crudele, CP.; Amin, DD.; Robinson, BM.; Dinges, DF.; Van Dongen, HPA. Parameter estimation for a biomathematical model of psychomotor vigilance performance under laboratory conditions of chronic sleep restriction. In: Ruigt, GSF.; Van Bommel, AL.; DeBoer, T.; Hofman, WF.; Van Luitelaar, G., editors. *Sleep-Wake Research in the Netherlands*. Vol. 16. Dutch Society for Sleep-Wake Research; The Netherlands: 2005. p. 39-42.
- Banks S, Dinges DF. Behavioral and physiological consequences of sleep restriction. *J Clin Sleep Med* 2007;3:519–528. [PubMed: 17803017]
- Banks S, Van Dongen H, Dinges DF. How much sleep is needed to recover from sleep debt? The impact of sleep dose on recovery. *Sleep* 2005;28:A138.
- Banks S, Van Dongen H, Dinges D. Can a sleep debt be recovered with one night of sleep? *Sleep Biol Rhythms* 2007a;5:A194.
- Banks S, Van Dongen H, Dinges D. Neurobehavioral response to sleep restriction is influenced by pre-existing sleep debt. *Sleep Biol Rhythms* 2007b;5:A103.
- Basheer R, Bauer A, Elmenhorst D, Ramesh V, McCarley RW. Sleep deprivation upregulates A₁ adenosine receptors in the rat basal forebrain. *NeuroReport* 2007;18:1895–1899. [PubMed: 18007182]
- Basheer R, Porkka-Heiskanen T, Strecker RE, Thakkar MM, McCarley RW. Adenosine as a biological signal mediating sleepiness following prolonged wakefulness. *Biol Signals Recept* 2000;9:319–327. [PubMed: 11025338]
- Basheer R, Strecker RE, Thakkar MM, McCarley RW. Adenosine and sleep-wake regulation. *Prog Neurobiol* 2004;73:379–396. [PubMed: 15313333]
- Belenky G, Wesensten NJ, Thorne DR, Thomas ML, Sing HC, Redmond DP, Russo MB, Balkin TJ. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *J Sleep Res* 2003;12:1–12. [PubMed: 12603781]
- Benington JH, Heller HC. Restoration of brain energy metabolism as the function of sleep. *Prog Neurobiol* 1995;45:347–360. [PubMed: 7624482]
- Benington JH, Kodali SK, Heller HG. Stimulation of A₁ adenosine receptors mimics the electroencephalographic effects of sleep deprivation. *Brain Res* 1995;692:79–85. [PubMed: 8548323]
- Biaggioni I, Paul S, Puckett A, Arzubiaga C. Caffeine and theophylline as adenosine receptor antagonists in humans. *J Pharmacol Exp Ther* 1991;258:588–593. [PubMed: 1865359]
- Borbély AA. A two process model of sleep regulation. *Human Neurobiol* 1982;1:195–204.
- Borbély AA, Achermann P. Sleep homeostasis and models of sleep regulation. *J Biol Rhythms* 1999;14:557–568. [PubMed: 10643753]
- Brunner DP, Dijk DJ, Borbély AA. Repeated partial sleep deprivation progressively changes the EEG during sleep and wakefulness. *Sleep* 1993;16:100–113. [PubMed: 8446828]
- Carskadon MA, Dement WC. Cumulative effects of sleep restriction on daytime sleepiness. *Psychophysiol* 1981;18:107–113.
- Daan S, Beersma DGM, Borbély AA. Timing of human sleep: recovery process gated by a circadian pacemaker. *Am J Physiol* 1984;246:R161–R178. [PubMed: 6696142]
- Dement WC. Talking about our sleep debt. *Sleep Rev* 2006;7:10–13.
- Derrick, W.; Grossman, S. *Elementary Differential Equation with Boundary Value Problems*. Vol. 4. Addison-Wesley: Reading; 1997.
- Dijk DJ, Czeisler CA. Paradoxical timing of the circadian rhythm of sleep propensity serves to consolidate sleep and wakefulness in humans. *Neurosci Lett* 1994;166:63–68. [PubMed: 8190360]
- Dijk DJ, Duffy JF, Czeisler CA. Circadian and sleep/wake dependent aspects of subjective alertness and cognitive performance. *J Sleep Res* 1992;1:112–117. [PubMed: 10607036]
- Dinges, DF. Are you awake? Cognitive performance and reverie during the hypnopompic state. In: Bootzin, RR.; Kihlstrom, JF.; Schacter, DL., editors. *Sleep and Cognition*. American Psychological Association; Washington, DC: 1990. p. 159-175.
- Dinges DF, Achermann P. Future considerations for models of human neurobehavioral function. *J Biol Rhythms* 1999;14:598–601. [PubMed: 10643757]

- Dinges, DF.; Kribbs, NB. Performing while sleepy: Effects of experimentally-induced sleepiness. In: Monk, TH., editor. *Sleep, Sleepiness and Performance*. John Wiley & Sons; Chichester: 1991. p. 97-128.
- Dinges, DF.; Orne, EC.; Evans, FJ.; Orne, MT. Performance after naps in sleep-conducive and alerting environments. In: Johnson, LC.; Tepas, DI.; Colquhoun, WP.; Colligan, MJ., editors. *Biological Rhythms, Sleep and Shift Work*. Spectrum; New York: 1981. p. 539-552.
- Dinges DF, Pack F, Williams K, Gillen KA, Powell JW, Ott GE, Aptowicz C, Pack AI. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. *Sleep* 1997;20:267–277. [PubMed: 9231952]
- 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.
- Dinges DF, Whitehouse WG, Orne EC, Orne MT. The benefits of a nap during prolonged work and wakefulness. *Work Stress* 1988;2:139–153.
- Doran SM, Van Dongen HPA, Dinges DF. Sustained attention performance during sleep deprivation: Evidence of state instability. *Arch Ital Biol* 2001;139:253–267. [PubMed: 11330205]
- Dorrian, J.; Rogers, NL.; Dinges, DF. Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss. In: Kushida, CA., editor. *Sleep Deprivation. Clinical Issues, Pharmacology, and Sleep Loss Effects*. Marcel Dekker; New York: 2005. p. 39-70.
- Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. *Semin Neurol* 2005;25:117–129. [PubMed: 15798944]
- 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. [PubMed: 8441003]
- Elmenhorst D, Meyer PT, Winz OH, Matusch A, Ermert J, Coenen HH, Basheer R, Haas HL, Zilles K, Bauer A. Sleep deprivation increases A1 adenosine receptor binding in the human brain: A positron emission tomography study. *J Neurosci* 2007;27:2410–2415. [PubMed: 17329439]
- Fuller PM, Gooley JJ, Saper CB. Neurobiology of the sleep-wake cycle: Sleep architecture, circadian regulation, and regulatory feedback. *J Biol Rhythms* 2006;21:482–493. [PubMed: 17107938]
- Gunzelmann, G.; Gluck, KA.; Price, S.; Van Dongen, HPA.; Dinges, DF. Decreased arousal as a result of sleep deprivation: The unraveling of cognitive control. In: Gray, WD., editor. *Integrated models of cognitive systems*. Oxford University Press; New York: 2007. p. 243-253.
- Hursh SR, Redmond DP, Johnson ML, Thorne DR, Belenky G, Balkin TJ, Storm WF, Miller JC, Eddy DR. Fatigue models for applied research in warfighting. *Aviat Space Environ Med* 2004;75:A44–A53. [PubMed: 15018265]
- Indic P, Gurdziel K, Kronauer RE, Klerman EB. Development of a two-dimension manifold to represent high dimension mathematical models of the intracellular mammalian circadian clock. *J Biol Rhythms* 2006;21:222–232. [PubMed: 16731662]
- Jewett ME, Forger DB, Kronauer RE. Revised limit cycle oscillator model of human circadian pacemaker. *J Biol Rhythms* 1999;14:493–499. [PubMed: 10643746]
- Jewett ME, Kronauer RE. Interactive mathematical models of subjective alertness and cognitive throughput in humans. *J Biol Rhythms* 1999;14:588–597. [PubMed: 10643756]
- Johnson ML, Belenky G, Redmond DP, Thorne DR, Williams JD, Hursh SR, Balkin TJ. Modulating the homeostatic process to predict performance during chronic sleep restriction. *Aviat Space Environ Med* 2004;75:A141–A146. [PubMed: 15018276]
- Kelly, W.; Peterson, A. *Difference Equations: An Introduction with Applications*. Vol. 2. Academic Press; San Diego: 2001.
- Krueger JM, Obál F. Sleep function. *Front Biosci* 2003;8:d511–d519. [PubMed: 12700033]
- LaJambe CM, Kamimori GH, Belenky G, Balkin TJ. Caffeine effects on recovery sleep following 27 h total sleep deprivation. *Aviat Space Environ Med* 2005;76:108–113. [PubMed: 15742825]
- Lamond N, Jay SM, Dorrian J, Ferguson SA, Jones C, Dawson D. The dynamics of neurobehavioural recovery following sleep loss. *J Sleep Res* 2007;16:33–41. [PubMed: 17309761]
- Landolt HP. Sleep homeostasis: A role for adenosine in humans? *Biochem Pharmacol* 2008;75:2070–2079. [PubMed: 18384754]

- Lim J, Dinges DF. Sleep deprivation and vigilant attention. *Ann N Y Acad Sci* 2008;1129:305–322. [PubMed: 18591490]
- Lubin A, Moses JM, Johnson LC, Naitoh P. The recuperative effects of REM sleep and Stage 4 sleep on human performance after complete sleep loss: Experiment 1. *Psychophysiol* 1974;11:133–146.
- Mallis MM, Mejdal S, Nguyen TT, Dinges DF. Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviat Space Environ Med* 2004;75:A4–A14. [PubMed: 15018262]
- Penetar D, McCann U, Thorne D, Kamimori G, Galinski C, Sing H, Thomas M, Belenky G. Caffeine reversal of sleep deprivation effects on alertness and mood. *Psychopharmacol* 1993;112:359–365.
- Porkka-Heiskanen T, Alanko L, Kalinchuk A, Stenberg D. Adenosine and sleep. *Sleep Med Rev* 2002;6:321–332. [PubMed: 12531135]
- Porkka-Heiskanen T, Strecker RE, McCarley RW. Brain site-specificity of extracellular adenosine concentration changes during sleep deprivation and spontaneous sleep: An in vivo microdialysis study. *Neuroscience* 2000;99:507–517. [PubMed: 11029542]
- Porkka-Heiskanen T, Strecker RE, Thakkar M, Bjørkum AA, Greene RW, McCarley RW. Adenosine: A mediator of the sleep-inducing effects of prolonged wakefulness. *Science* 1997;276:1265–1268. [PubMed: 9157887]
- Smith AD, Genz A, Freiburger DM, Belenky G, Van Dongen HPA. Efficient computation of confidence intervals for Bayesian model predictions based on multidimensional parameter space. *Meth Enzymol*. in press
- St Hilaire MA, Klerman EB, Khalsa SB, Wright KP, Czeisler CA, Kronauer RE. Addition of a non-photic component to a light-based mathematical model of the human circadian pacemaker. *J Theor Biol* 2007;247:583–599. [PubMed: 17531270]
- Van Dongen HPA. Comparison of mathematical model predictions to experimental data of fatigue and performance. *Aviat Space Environ Med* 2004;75:A15–A36. [PubMed: 15018263]
- 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* 2004a;27:423–433. [PubMed: 15164894]
- Van Dongen HPA, Dinges DF. Investigating the interaction between the homeostatic and circadian processes of sleep-wake regulation for the prediction of waking neurobehavioural performance. *J Sleep Res* 2003a;12:181–187. [PubMed: 12941057]
- Van Dongen HPA, Dinges DF. Sleep debt and cumulative excess wakefulness. *Sleep* 2003b;26:249.
- 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;26:117–128. [PubMed: 12683469]
- 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. [PubMed: 17910385]
- Van Dongen HPA, Olofsen E, Dinges DF, Maislin G. Mixed-model regression analysis and dealing with interindividual differences. *Meth Enzymol* 2004;384:139–171. [PubMed: 15081686]
- Verhulst, F. *Nonlinear Differential Equations and Dynamical Systems*. Vol. 2. Springer; New York: 2000.
- Wesensten N, Reichardt R, Balkin T. Chronic sleep restriction and resatiation. I Recovery of psychomotor vigilance performance. *Sleep* 2005;28:A135.
- Yanik G, Radulovacki M. REM sleep deprivation up-regulates adenosine A1 receptors. *Brain Res* 1987;402:362–364. [PubMed: 3030496]

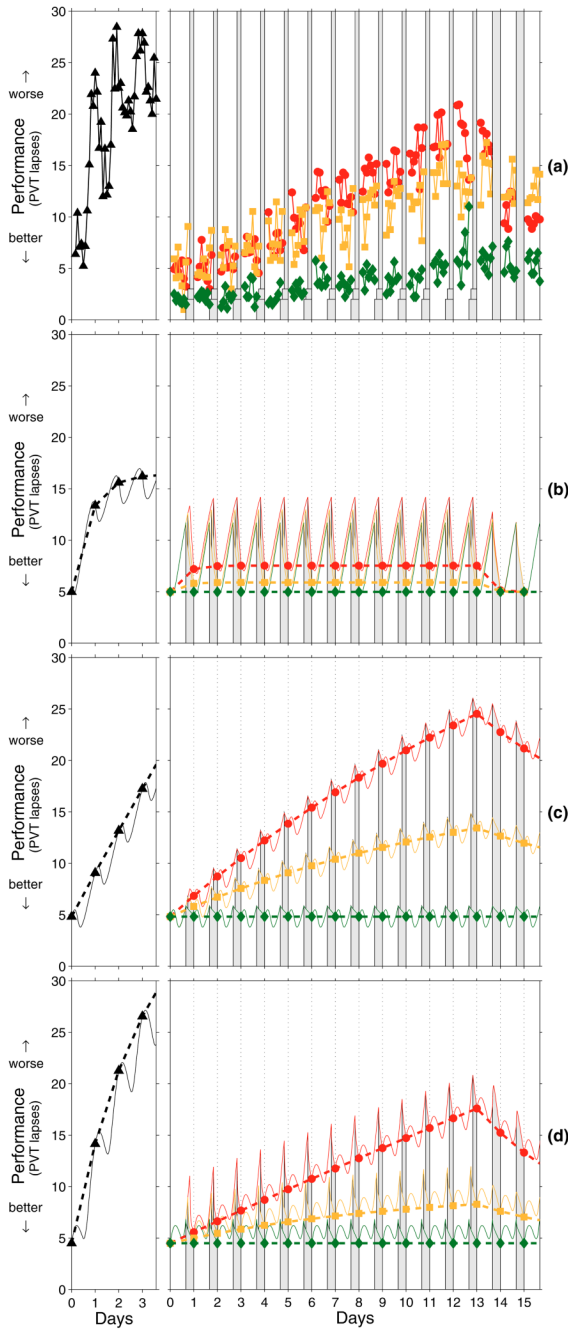


Fig. 1. Neurobehavioral performance observations and predictions by different models. A total of 48 healthy young adults were subjected to one of four laboratory sleep deprivation protocols (Van Dongen et al., 2003). Each protocol began with several baseline days involving 16h scheduled wake time (SWT)/8h time in bed (TIB); the last of these baseline days is labeled here as day 0. Subsequently, 13 subjects were kept awake (24h SWT/0h TIB) for three additional days, for a total of 88h awake (left panels), after which they received varied amounts of recovery sleep (not shown). The other subjects underwent various doses of sleep restriction for 14 consecutive days, followed by two recovery days with 16h SWT/8h TIB (right panels). The sleep restriction schedule involved 20h SWT/4h TIB per day for 13 subjects (circles; red); 18h SWT/6h TIB

per day for another 13 subjects (boxes; yellow); and 16h SWT/8h TIB per day for the remaining 9 subjects (diamonds; green). Awakening was scheduled at 07:30 each day. Neurobehavioral performance was tested every 2h during scheduled wakefulness using the PVT, for which the number of lapses (reaction times greater than 500ms) was recorded.

(a) Observed neurobehavioral performance (PVT lapses) for each test bout (dots represent group averages). The first two test bouts of each waking period are omitted in order to avoid confounds from sleep inertia. Gray bars indicate scheduled sleep periods.

(b) Corresponding performance predictions according to the original two-process model (Borbély and Achermann, 1999), linearly scaled to the data. Data points represent performance predictions at wake onset. Thin curves represent predictions within days, but the focus here is on changes across days (dashed lines). Note the rapid stabilization across days predicted to occur in the chronic sleep restriction conditions (right panel), which does not match the observations shown in (a).

(c) Corresponding predictions according to the extended two-process model (Avinash et al., 2005), linearly scaled to the data. Note the under-prediction of performance impairment in the total sleep deprivation condition (left panel) and the over-prediction of the impairment build-up across days in the 20h SWT/4h TIB condition (right panel), relative to the actual data shown in (a).

(d) Corresponding predictions according to the new model introduced in this paper as defined by Eqs. (21) and (26). Note the improved fit to the experimental observations across days for total sleep deprivation (left panel), which is explored in more detail in Fig. 3, as well as for the 20h SWT/4h TIB condition (right panel). Performance impairment in the 18h SWT/6h TIB and 16h SWT/8h TIB conditions (right panel) is under-predicted. However, the group-average impairment levels observed for these conditions are inflated due to a few outliers (Van Dongen et al., 2003).

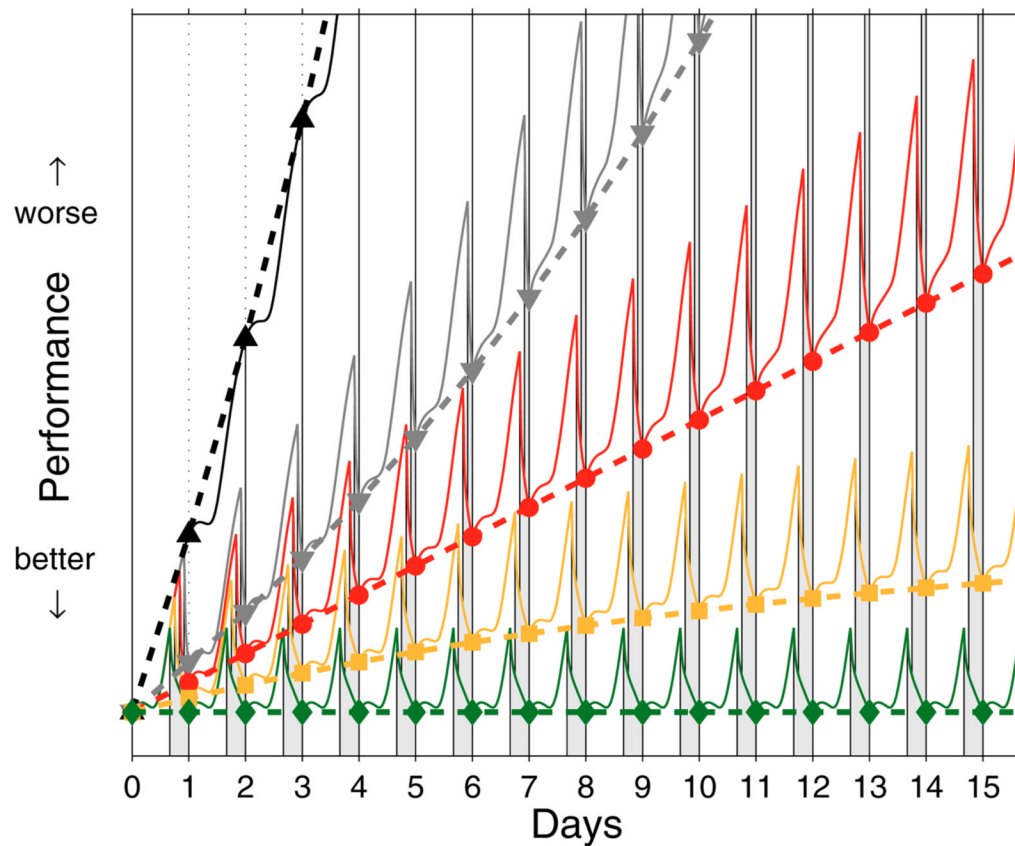


Fig. 2. Illustration of the model given by Eqs. (21) using parameter values selected to illustrate its bifurcating dynamic behavior. The figure shows model predictions at wake onset (data points) for 16 days ($n = 0, 1, \dots, 15$) of fixed duration $T = 24\text{h}$, assuming a constant period $\tau = 24\text{h}$ for the non-homogeneities. The thin curves represent the predictions within days using the non-homogeneities given by Eqs. (26)—but the profile of changes *across* days (dashed lines) as determined by the α and σ coefficient matrices in Eqs. (21) is of primary interest here. Each prediction curve corresponds to a different amount of daily wakefulness: $W = 16\text{h}$ (diamonds; green), $W = 18\text{h}$ (boxes; yellow), $W = 20\text{h}$ (circles; red), $W = 22\text{h}$ (downward triangles; gray), and $W = 24\text{h}$ (i.e., total sleep deprivation) (upward triangles; black). Light gray areas indicate nocturnal sleep periods. In this illustration, the model parameter values are intentionally selected such that the bifurcation threshold occurs at $W_c = 20\text{h}$ (i.e., 4h daily sleep). For daily wake durations below this bifurcation threshold (green and yellow), the model converges to an asymptotically stable equilibrium, meaning that performance impairment ultimately levels off. For daily wake durations beyond the bifurcation threshold (gray and black), the model diverges from an unstable equilibrium, meaning that performance impairment tends to escalate. At exactly the bifurcation value $W = W_c$ (red), there is no equilibrium state, resulting in an asymptotically linear build-up of performance impairment across days.

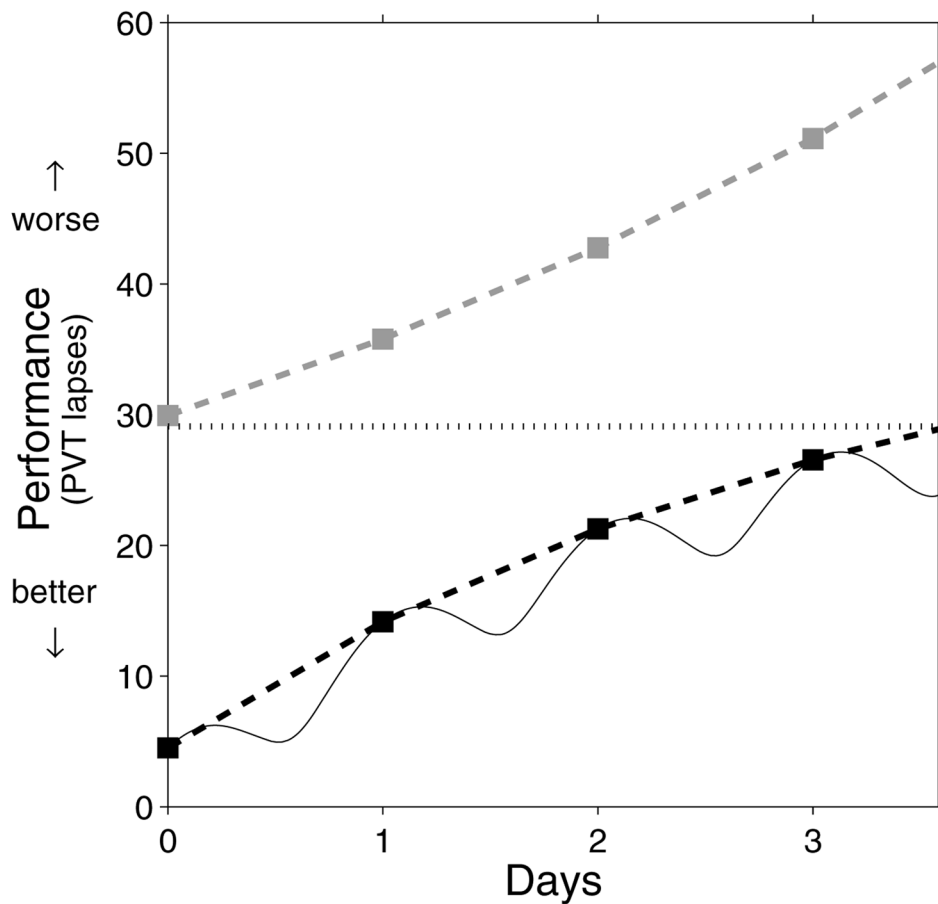


Fig. 3.

Detailed examination of the performance predictions under conditions of total sleep deprivation. The new model defined by Eqs. (21) and (26) with the parameter estimates given by Eqs. (27) has a bifurcation at $W_c = 20.2h$, implying that predictions for performance in the total sleep deprivation condition (i.e., $W = 24h > W_c$; see Fig. 1a, top left panel) should exhibit diverging (i.e., escalating) performance impairment across days. However, the actual predictions (see Fig. 1d, bottom left panel) would seem to suggest a converging pattern. This can be explained by simultaneously considering the performance predictions p_n (black dashed curve), the level of the unstable equilibrium state p (dotted horizontal line), and the upper asymptote u_n (gray dashed curve). Since $\alpha_{22} > 0$, the upper asymptote u_n increases exponentially across days. Thus, within waking episodes, performance p_n is increasingly drawn upwards. On the other hand, the equilibrium level p is located above the initial performance value $p_0(t_0)$. Thus, divergence from this unstable equilibrium would entail a drive downwards. Here, the net result is that performance impairment is predicted to increase across days, but in a decelerating manner (cf. Van Dongen et al., 2003). If wakefulness were maintained for additional days, though, the performance predictions would cross the unstable equilibrium state and then diverge from it upwards, exposing the typical escalating behavior for $W > W_c$ in this model (see the illustration in Fig. 2, black upward triangles).

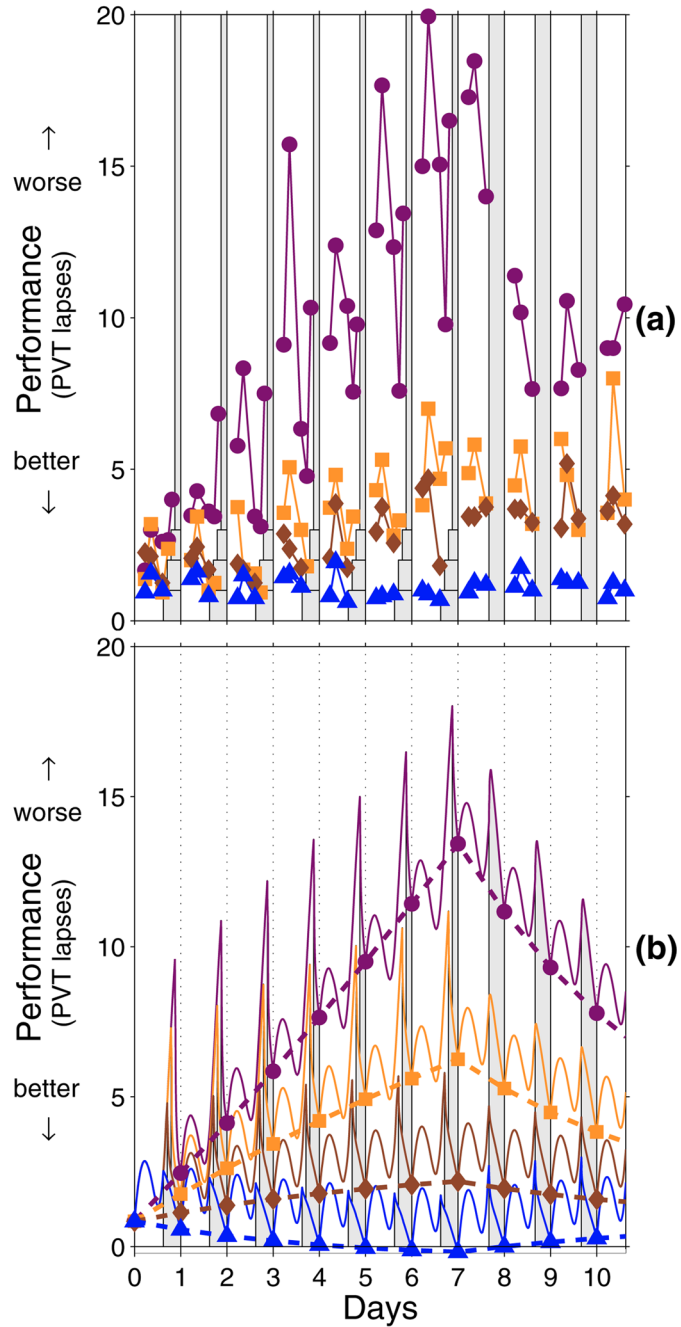


Fig. 4. Experimental observations and predictions by our new model for neurobehavioral performance impairment. A total of 66 healthy young adults were subjected to one of four laboratory sleep deprivation protocols (Belenky et al., 2003). Each protocol began with several baseline days involving 16h scheduled wake time (SWT)/8h time in bed (TIB); the last of these baseline days is labeled here as day 0. The subjects subsequently underwent various doses of sleep restriction for seven consecutive days, followed by three recovery days with 16h SWT/8h TIB. The sleep restriction schedule involved 21h SWT/3h TIB per day for 13 subjects (circles; purple); 19h SWT/5h TIB per day for 13 subjects (boxes; orange); 17h SWT/7h TIB per day for 14 subjects (diamonds; brown); and 15h SWT/9h TIB per day for 16 subjects (triangles; blue). Awakening

was scheduled at 07:00 each day. Neurobehavioral performance was tested daily at 09:00, 12:00, 15:00 and 21:00 using the PVT. In the 19h SWT/5h TIB condition an additional test bout occurred at midnight, and in the 21h SWT/3h TIB condition yet another one took place two hours after midnight.

(a) Observed neurobehavioral performance (PVT lapses) for each test bout (dots represent group averages). The first test bout of each waking period is omitted in order to avoid confounds from sleep inertia. Gray bars indicate scheduled sleep periods.

(b) Corresponding performance predictions according to the new model defined by Eqs. (21) and (26). Parameter estimates are fixed at the values of Eqs. (27), as previously estimated using the data in Fig. 1a. Data points represent predictions at wake onset; thin curves represent predictions within days. The focus here is on changes across days (dashed lines). Note that the model predictions across the seven days of sleep restriction accurately capture the qualitative change from convergence (i.e., leveling off of performance impairment) in the 15h, 17h and 19h SWT conditions, to divergence (i.e., disproportionately rapid escalation of performance impairment) in the 21h SWT/3h TIB condition.

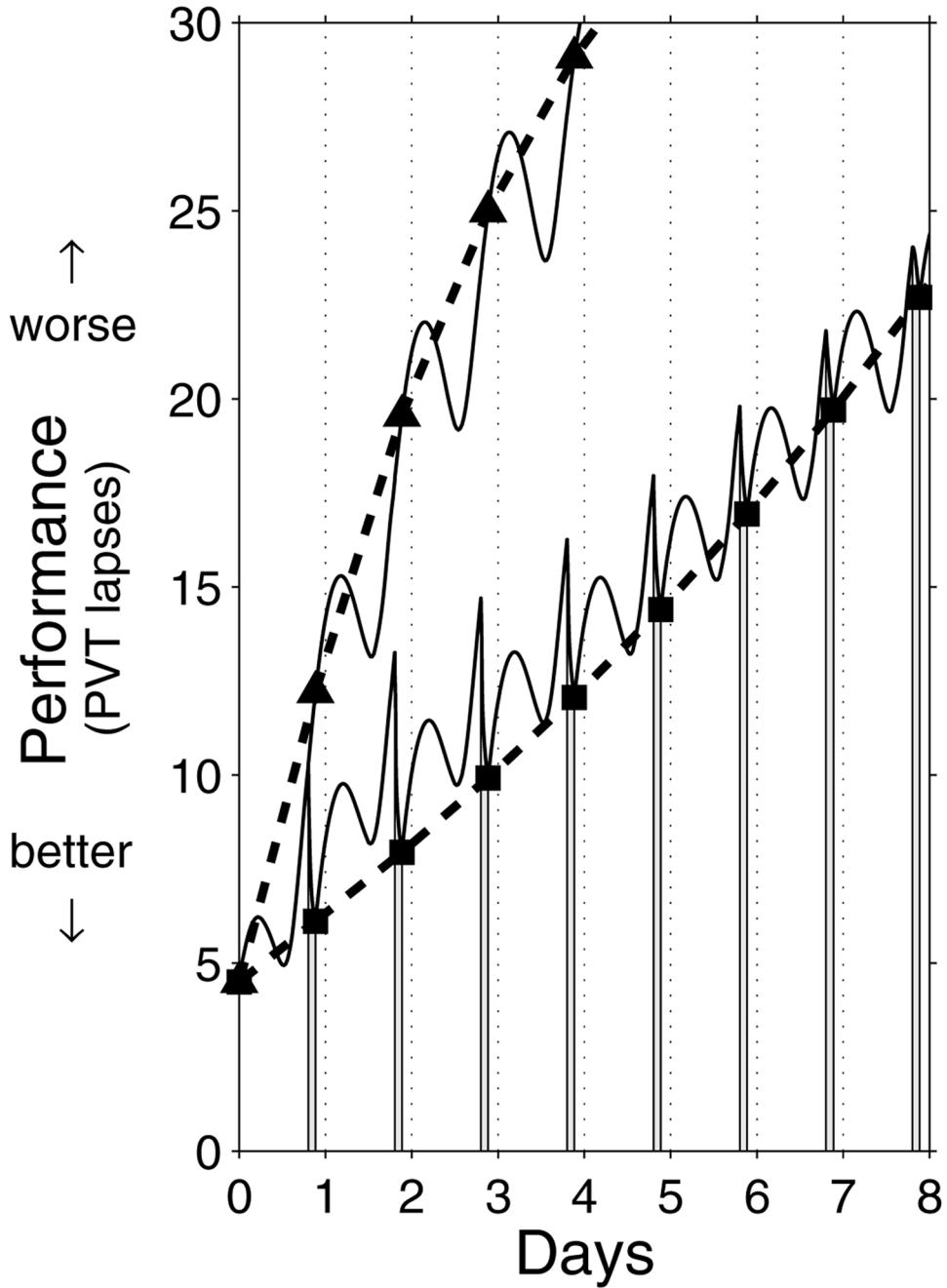


Fig. 5. Prediction of the effectiveness of a single 2h nocturnal period of nap sleep each day for maintaining performance across days. The figure shows predicted performance at wake onset (boxes) and within each of the days (thin curve) across 8 days with a nap scheduled from 02:45 until 04:45 daily (gray bars). For comparison, predicted performance across 4 days of total sleep deprivation is shown as well (triangles represent performance at 04:45, which is the same time as scheduled wake onset in the nap condition). Both conditions are initiated after awakening from 8h baseline sleep at 07:30 (dashed vertical lines indicate 07:30 at 24h increments). The performance predictions, derived from the new model given by Eqs. (21), (26) and (27) and expressed in terms of the number of lapses on the PVT, indicate that a daily

2h nap is not effective at maintaining reasonable levels of performance across multiple days. (For a discussion of the predictions for the total sleep deprivation condition, see Fig. 3, but note that the timing of wake onset is not the same.)

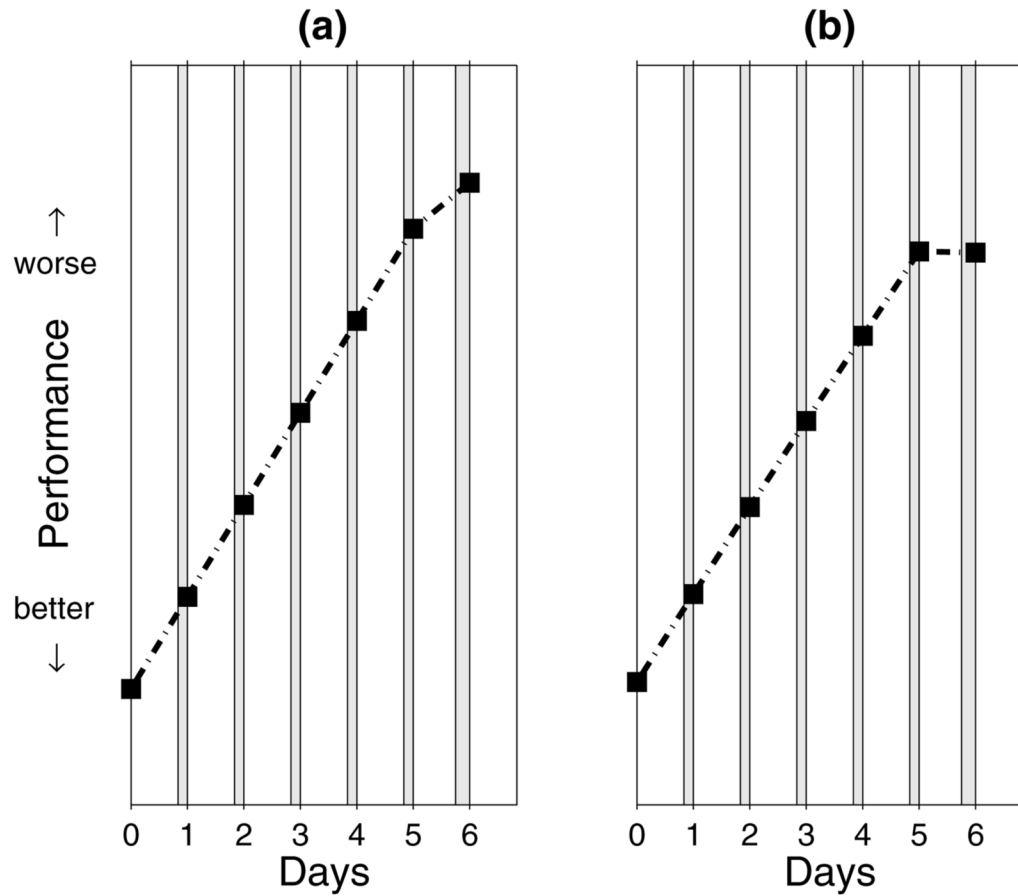


Fig. 6. Opposing predictions from two models regarding recovery following chronic sleep restriction. The figure shows performance predictions at wake onset (boxes) for five days with 20h wakefulness and 4h sleep per day, followed by one day with 18h wakefulness and 6h recovery sleep. Gray areas indicate nocturnal sleep periods.
 (a) Predictions for performance changes across days according to the excess wakefulness model (Van Dongen et al., 2003). This model predicts that performance deteriorates progressively across the five days with 20h wake/4h sleep, and continues to deteriorate at a slower rate following the day with 18h wake/6h sleep.
 (b) Predictions for performance changes across days according to the model defined by Eqs. (21), (26) and (27). This new model also predicts that performance deteriorates progressively across the five days with 20h wake/4h sleep, but forecasts a modest relative performance improvement following the day with 18h wake/6h sleep. (Note that sleep inertia is not accounted for in these predictions.)

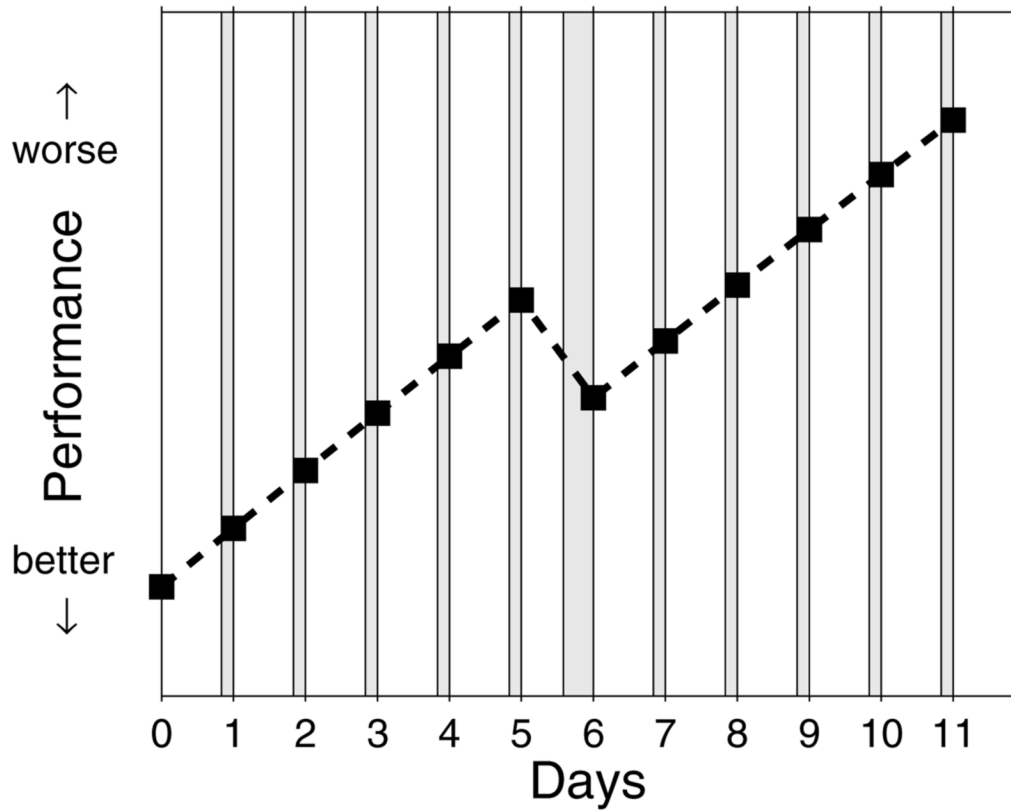


Fig. 7. New prediction for rapid recycling after a period of chronic sleep restriction. The figure shows predicted performance at wake onset (boxes) over days, during a period of five days with wake extension to 20h per day (i.e., 4h sleep daily), followed by one day with 14h scheduled wakefulness (i.e., 10h for recovery sleep), followed by recycling into a second period of five days with wake extension to 20h per day (4h sleep daily). Gray areas indicate nocturnal sleep periods. The performance predictions, derived from the new model given by Eqs. (21), (26) and (27), indicate that the intermittent recovery sleep should confer only a short-lasting benefit—in the second period of sleep restriction, performance is predicted to further deteriorate (while converging towards the asymptotically stable equilibrium for $W = 20h$ with a time constant extending far beyond the period displayed in the graph).