SLEEPJ, 2021, 1–13

doi: 10.1093/sleep/zsaa263
Advance Access Publication Date: 29 November 2020
Original Article

ORIGINAL ARTICLE

Models for predicting sleep latency and sleep duration

Francisco G. Vital-Lopez[1,2], Thomas J. Balkin[3,4] and Jaques Reifman[1,*]

[1]Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Development Command, Fort Detrick, MD, [2]The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, MD, [3]Behavioral Biology Branch, Walter Reed Army Institute of Research, Silver Spring, MD and [4]Oak Ridge Institute for Science and Education, Oak Ridge, TN

*Corresponding author. Jaques Reifman, Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Development Command, ATTN: FCMR-TT, 504 Scott Street, Fort Detrick, MD 21702. Email: jaques.reifman.civ@mail.mil.

Abstract

Study Objectives: Planning effective sleep–wake schedules for civilian and military settings depends on the ability to predict the extent to which restorative sleep is likely for a specified sleep period. Here, we developed and validated two mathematical models, one for predicting sleep latency and a second for predicting sleep duration, as decision aids to predict efficacious sleep periods.

Methods: We extended the Unified Model of Performance (UMP), a well-validated mathematical model of neurobehavioral performance, to predict sleep latency and sleep duration, which vary nonlinearly as a function of the homeostatic sleep pressure and the circadian rhythm. To this end, we used the UMP to predict the time course of neurobehavioral performance under different conditions. We developed and validated the models using experimental data from 317 unique subjects from 24 different studies, which included sleep conditions spanning the entire circadian cycle.

Results: The sleep-latency and sleep-duration models accounted for 42% and 84% of the variance in the data, respectively, and yielded acceptable average prediction errors for planning sleep schedules (4.0 min for sleep latency and 0.8 h for sleep duration). Importantly, we identified conditions under which small shifts in sleep onset timing result in disproportionately large differences in sleep duration—knowledge that may be applied to improve performance, safety, and sustainability in civilian and military operations.

Conclusions: These models extend the capabilities of existing predictive fatigue-management tools, allowing users to anticipate the most opportune times to schedule sleep periods.

Statement of Significance

Millions of workers across the world are unable to sleep at night on a regular basis, with negative consequences to their health and work performance. To help mitigate these effects, shift workers should sleep at those available times that are likely to result in the best sleep. Here, we describe two mathematical models that can accurately predict sleep-onset latency and sleep duration for a variety of sleep conditions spanning the entire circadian cycle—information that can be used to optimize the efficiency (and thus the restorative value) of scheduled sleep periods for workers who are not following a typical day–night, wake–sleep schedule.

Key words: alertness impairment; shift work; sleepiness; sleep deprivation; sleep loss

Submitted: 1 May, 2020; **Revised:** 6 October, 2020

Introduction

Certain workers, such as healthcare providers, first responders, pilots, and Service members engaged in military operations, are often on irregular work schedules that preclude adequate sleep. For example, night shifts that necessitate daytime sleep lead to fatigue and impaired cognitive performance, including deficits in alertness and sustained attention [1]. There are multiple factors that can potentially contribute to poor sleep in shift workers, including a non-sleep-conducive bedroom environment, childcare and social demands, and poorly timed ingestion of caffeinated beverages. However, the primary problem is not attaining the necessary sleep duration by placing the sleep period during the ascending phase of the circadian rhythm of alertness. For example, after being awake for 24 h, an individual falling asleep at around 07:00 is likely to remain asleep for only about 4.5 h [2]. Another challenge is that sleep periods that lead to long sleep durations may not necessarily lead to optimal neurobehavioral performance during work periods. Therefore, planning effective sleep schedules requires knowing whether an individual can actually fall asleep at the proposed time and stay asleep for the desired duration as well as determining the time course of neurobehavioral performance for a given sleep schedule.

To address these questions, a few computational models have been proposed. In a seminal study, Borbély [3] proposed the two-process model, which postulates that slow-wave sleep is regulated by (1) a homeostatic process *S* that depends on sleep history and (2) a sleep-independent circadian process *C*. Process *S* (the homeostatic need for sleep) monotonically increases as a function of time awake, and monotonically decreases as a function of time asleep. Process *C* reflects the influence of the circadian rhythms on sleep, which varies non-monotonically during the day. Together, the combined effects of processes *S* and *C* determine the sleep propensity that is manifest at any time. In subsequent work, Borbély et al. [4] extended the two-process model to account for the propensity for sleep initiation (i.e. sleep latency), which they defined as the difference between the homeostatic process *S* and an upper threshold *H* that accounts for circadian variation. They proposed a sinusoidal function to represent *H*, and chose its amplitude and skewness so that the simulated propensity for sleep initiation is relatively low during the first 16 h of wakefulness (after 8 h of nighttime sleep) and relatively high during sleep deprivation, based on observed sleep latencies [2]. However, this model provided an estimation of the propensity for sleep initiation using an arbitrary scale, instead of the actual time to fall asleep. Subsequently, Akerstedt and Folkard [5] proposed a model of alertness regulation (i.e. the three-process model of alertness [TPMA]) linking processes *S* and *C* to a subjective measure of alertness, and then extended this model to predict sleep latency [6]. They validated the model using data from three studies that included 23 distinct time points from 44 subjects.

Borbély [3] also used the two-process model to predict sleep duration. Specifically, he based the model on empirical observations [2, 7, 8] that showed that sleep duration was not a monotonically increasing function of previous time awake, but that instead sleep duration may decrease with increasing time awake, depending on the circadian phase at sleep onset. Accordingly, in the two-process model, sleep ceases when process *S* decreases below a notional sleep termination threshold that accounts for circadian variation. He fitted the model to account for the variation in sleep duration as a function of sleep-onset time in a study [2] in which bedtime was delayed in 4-h increments over a 24-h period. However, the model was never validated using an independent study. Inspired by Borbély's work, Akerstedt and Folkard extended their sleep-latency model to predict sleep duration, by assuming that sleep terminates spontaneously when alertness reaches an upper threshold that accounts for circadian variation. They validated the model using data from three studies that included five distinct time points from 36 subjects [9]. Other modeling efforts have attempted to address related questions using different modeling frameworks. For example, more recently, Phillips et al. proposed a mathematical model based on the biochemistry of the adenosine system to simulate the effects of sleep history on cognitive performance and the regulation of sleep [10]. The authors showed that their model was able to fit sleep duration data in a sleep-extension study, however, the model was not validated against independent datasets.

Building upon the theoretical groundwork described above, here, we extended the Unified Model of Performance (UMP) to predict sleep latency and sleep duration, and provide computational tools for planning efficient and feasible sleep schedules within a single modeling framework. While previous approaches are based on subjective measures of alertness [6, 9], the UMP is based on an objective measure, as assessed by the mean response time (RT) of the widely used psychomotor vigilance test (PVT). Subjective measures of alertness do not seem to reflect the full extent of neurobehavioral performance impairment under total sleep deprivation (TSD) and restricted sleep conditions [11, 12]. Another potential benefit of the UMP is that it considers extant sleep debt as a function of a known sleep history, with recent history exerting greater influence in the predictions. This capability, which is unique to the UMP, improves model predictions as it allows us to capture an individual's capacity to recover during sleep as a function of sleep debt and naturally bridges the continuum of sleep deprivation conditions [13]. The proposed models are intended to predict sleep latency and sleep duration for young adults that carry sleep debt similar to or larger than that associated with habitual sleep of 8 h or less, common in modern society [1, 14]. They are not intended to predict individuals under sleep-satiated conditions or to predict sleep onset time.

In this report, we also address another limitation of previous reports in which at most three studies are used to validate the model predictions. Here, we comprehensively validated our results by using a total of 23 studies to independently validate the model predictions, including 298 unique subjects not used in developing the models. Of these, we used 20 studies to validate the sleep-latency model and 11 studies to validate the sleep-duration model. Then, we investigated why, under certain conditions, sleep bouts starting at slightly different times lead to distinctively short or long sleep durations.

Methods

Sleep-latency datasets

For model development and validation, we searched the literature for studies that reported sleep schedules and measured sleep latency under different levels of sleep debt and across the circadian cycle. We collected sleep-latency data from 21 distinct

studies conducted in 18 laboratories, collectively involving 278 unique subjects. Specifically, for each study, we retrieved the average sleep latency of a group of subjects at each measurement time, amounting to 147 time points. Table 1 provides a brief description of the studies, including the number of subjects, sex, age, the method used to measure sleep latency, and the number of data points collected from each study. Figure 1 shows the sleep periods and the time of the sleep-latency measurements (upper triangles and diamonds indicate whether sleep latency alone or sleep latency and sleep duration were measured, respectively) for each study. We used data from Study D1 [4] (14 time points from 8 subjects) to estimate the parameters of the sleep-latency model, and data from Studies D2 [15] and V1 to V19 [2, 6, 16–32] to independently validate the model predictions (20 studies, including 133 time points from 270 subjects). In Study D1, sleep latency was measured using the multiple sleep latency test [33] (MSLT), in which each test ended after sleep onset or 20 min, if subjects did not fall asleep within this time period. In the five validation studies that used MSLT, the test was limited to 20 min in four studies (V2, V3, V17, and V18, Table 1) and 30 min in one study (V5). Note that in the other 15 validation studies (73 out of 133 time points) subjects were free to take as much time as they needed to fall asleep, without constraining the observed values to 20 min as in the MSLT.

Sleep-duration datasets

For model development and validation, we searched the literature for studies that reported sleep schedules and measured sleep duration under different levels of sleep debt and across the circadian cycle. We collected sleep-duration data from 12 previous studies conducted in nine laboratories, collectively involving 165 unique subjects. Specifically, for each study, we retrieved the average sleep duration of a group of subjects at each measurement time, amounting to 45 time points. Table 1 and Figure 1 provide information about the studies, where in Figure 1 gray and hatch bars indicate studies in which subjects stayed in bed for a fixed time or slept ad libitum, respectively. We used data from Study D2 (6 time points from 11 subjects) to estimate the parameters of the sleep-duration model, and data from Studies V12 to V22 [2, 26–32, 34–36] (11 studies, including 39 time points from 154 subjects) to validate the model predictions. In Study D2, after three baseline nights of habitual sleep, subjects remained awake for 40 h, and then completed 21–25 28-h cycles of forced desynchrony. Each cycle consisted of a 9 h 20 min sleep period and an 18 h 40 min wakefulness period. The sleep periods started at 23:00, 03:00, 07:00, 11:00, 15:00, or 19:00, with each period occurring at the same time of day every six cycles. During sleep periods, subjects remained in bed in darkness and slept in one or more bouts. We used the reported

Table 1. Experimental studies used to develop and validate the sleep-latency and sleep-duration models.

Study	Condition	Number of subjects (women)*	Age (years), mean ± SEM or range	Measurement method		Number of measurements	
				Sleep latency	Sleep duration	Sleep latency	Sleep duration
D1 [4]	Extended wakefulness	8	N/A	MSLT		14†	
D2 [15]	Forced desynchrony	11 (0)	21–30	PSG	PSG	6	6‡
V1 [16]	Extended wakefulness	12	18–31	PSG		2	
V2 [17]	Sleep restriction	24 (12)	22.5 ± 0.8	MSLT		20	
V3 [18]	Sleep restriction	32 (0)	20–35	MSLT		16	
V4 [6]	Irregular schedule	8	N/A	EEG		18	
V5 [19]	Habitual sleep	8 (4)	19–23	MSLT§		8	
V6 [20]	Habitual sleep	12	20–31	PSG		2	
V7 [21]	Habitual sleep	12 (6)	23.8 ± 0.7	PSG		1	
V8 [22]	Habitual sleep	9 (0)	22.4 ± 0.4	PSG		2	
V9 [23]	Nap	10	20–30	PSG		1	
V10 [24]	Nap	10 (5)	20–22	PSG		3	
V11 [25]	Nap	7 (3)	21–24	Actigraphy		3‖	
V12 [2]	Extended wakefulness	6	29–45	E&E	E&E	7	7
V13 [26]	Extended wakefulness	34 (20)	38.2 ± 2.5	PSG	PSG	2	1
V14 [27]	Extended wakefulness	12 (6)	24.2 ± 1.0	PSG	PSG	2	1
V15 [28]	Sleep restriction	8 (0)	20–47	E&E	E&E	4	4
V16 [29]	Sleep extension	17 (10)	21.8 ± 0.9	PSG	PSG	1	6
V17 [30]	Sleep extension	15 (4)	20.1 ± 0.3	MSLT	Actigraphy	10	3
V18 [31]	Night shift	15 (7)	19–30	MSLT/PSG	PSG	10¶	5
V19 [32]	Irregular schedule	8 (8)	18–34	E&E	E&E	15	7
V20 [34]	Extended wakefulness	16 (0)	21–26		PSG		2
V21 [35]	Extended wakefulness	9 (0)	22–26		EEG		1
V22 [36]	Sleep restriction	14 (7)	27.4 ± 1.1		PSG		2
Total		**317 (92)**				**147**	**45**

E&E, electroencephalograms and electrooculograms; EEG, electroencephalography; MSLT, multiple sleep latency test; PSG, polysomnography; SEM, standard error of the mean.

*Sex of subjects was not provided in studies without parenthesis.

†Data used to develop the sleep-latency model.

‡Data used to develop the sleep-duration model.

§Modified MSLT with a 30-min test duration, as opposed to the standard 20 min.

‖One of the three sleep-latency measurements was obtained during a 20-min sleep period.

¶Five of the 10 sleep-latency measurements were obtained using PSG.

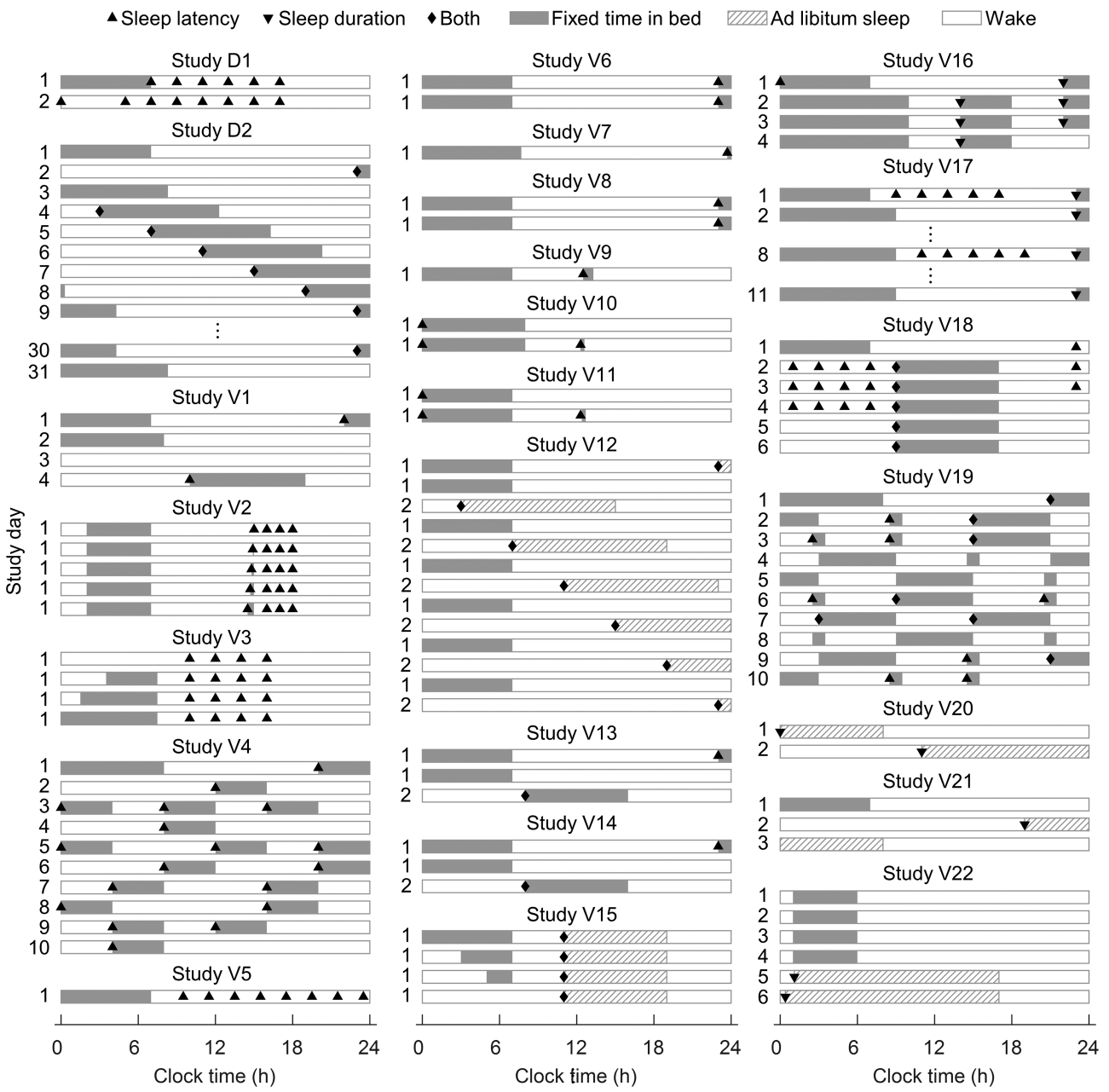

Figure 1. Sleep schedules and time of sleep-latency measurements for 24 studies used to develop and validate the sleep-latency and sleep-duration models. Study day *N* indicates the *N*th day of an arm of a study. For studies with more than one arm, the start of each arm is indicated by study day 1. For example, Study V13 had one arm of 1 day (first bar) and another arm of 2 days (second and third bars). In Study D2, 24 sleep-latency and sleep-duration measurements were obtained at six different times of day (four measurements at each time of day). We used the average sleep duration at each time of day to develop the sleep-duration model, and the average sleep latency at each time of day to validate the sleep-latency model. Study V17 reported average sleep duration for three periods (days 2–4, 5–8, and 9–11). Study V18 reported 10 sleep latency values, five corresponding to the sleep periods and five corresponding to the average of three values measured at the same time of day during wakefulness.

average sleep efficiency at each time of day and the duration of the sleep periods (9 h 20 min) to compute the average sleep duration.

The UMP

Based on the principles postulated by Borbély [3], we previously developed the UMP to quantitatively predict the temporal patterns of alertness for a given sleep–wake schedule [13, 37]. Figure 2, A shows process *S* (solid blue line) and process *C* (dashed orange line), which are additively combined to predict an objective measure of alertness impairment *P* (dotted purple line):

$$P(t) = S(t) + \kappa C(t), \quad (1)$$

where κ denotes the circadian amplitude, *t* represents time, and *P* is provided as an estimate of the mean RT statistics of the PVT [38]. We refer the reader to Tables S1 and S2 in Supplementary

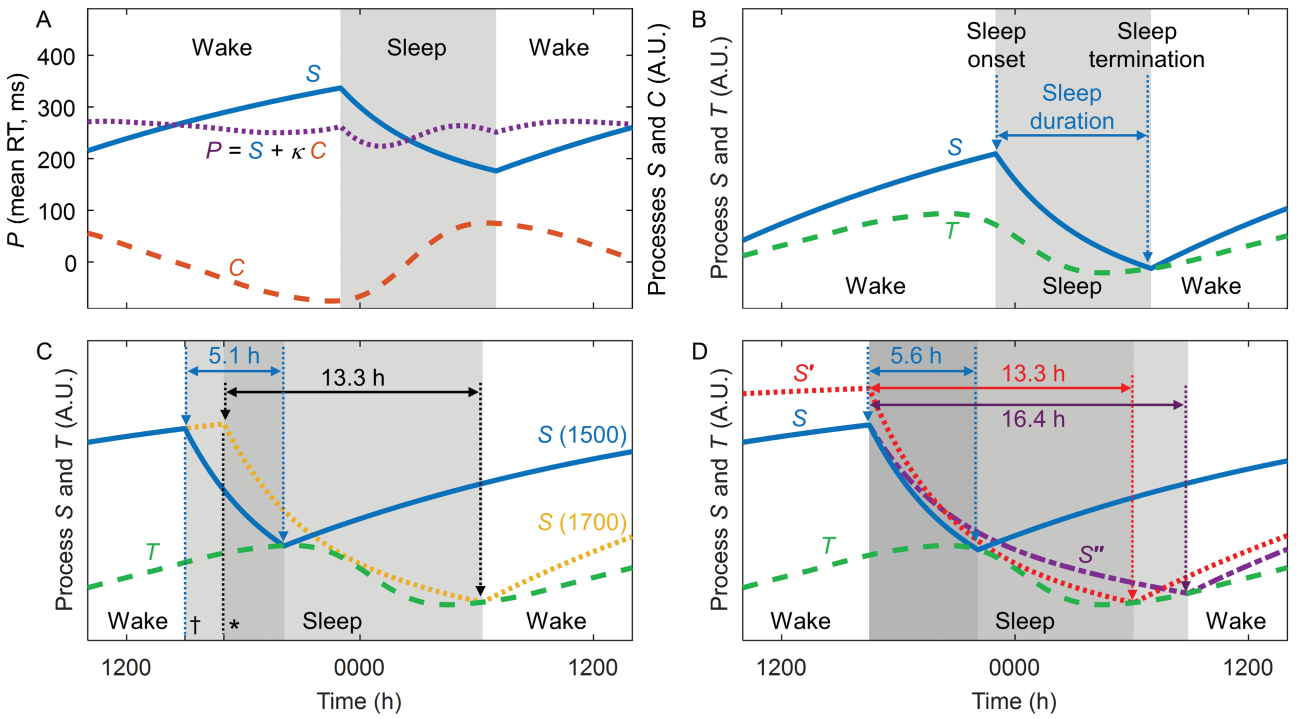

Figure 2. Unified Model of Performance (UMP) and its extension to predict sleep duration. (A) Simulation of the homeostatic process *S* (solid blue line) and the circadian process *C* (dashed orange line), which added together yield a quantitative measure of alertness impairment (*P*, dotted purple line), as determined by the mean response time (RT) on the psychomotor vigilance test. (B) Extension of the UMP to predict sleep duration. Process *S* increased during wakefulness and started to decrease after sleep onset until it reached the sleep-termination threshold *T* (dashed green line), which only depends on the time of day. (C) Discontinuity of predicted sleep duration. Process *S*, for a sleep period starting at 15:00 (†) in Study V12 (solid blue line), intersected with the threshold *T* (dashed green line) on the first upswing, resulting in a sleep duration of 5.1 h. In contrast, process *S* (dotted yellow line), for a sleep period starting at 17:00 (*), did not intersect *T* until the second upswing, resulting in a much longer sleep duration of 13.3 h. These simulations show that a relatively small difference in sleep-onset time (i.e. 2 h) may result in a large difference (8.2 h) in predicted sleep duration. (D) Effect of the initial level of process *S* and sleep debt on sleep duration. With the level of sleep debt and sleep-onset time held constant, the higher the level of process *S* (compare dotted red line [*S′*] with solid blue line [S]), the longer the sleep duration (13.3 h vs. 5.6 h). Similarly, with the initial level of process *S* and sleep-onset time held constant, the higher the sleep debt (compare dashed-dotted purple line [*S″*] with solid blue line [S]), the slower the rate at which *S* decreases and, hence, the longer the sleep duration (16.4 h vs. 5.6 h).

material for the equations and model parameters, respectively, for processes *S* and *C*.

In the UMP, process *S* is modulated as a function of sleep debt [13], limiting an individual's capacity to fully recover during 8 h of sleep in a 24-h period when the individual carries sleep debt [39]. This allows the UMP to seamlessly predict alertness levels under both TSD and chronic sleep restriction (CSR) conditions in one unified model [13].

The UMP was developed and validated using a comprehensive array of experimental datasets from 27 studies, including 59 different sleep and caffeine conditions and data from nearly 900 subjects (mostly healthy young adults). The sleep conditions included CSR (3–5 h of sleep per night for up to 7 days), TSD (28–88 h), combinations of CSR and TSD, diurnal sleep, and sleep banking, whereas the caffeine conditions included single and multiple caffeine doses ranging from 50 to 600 mg. In particular, Ramakrishnan et al. [37] showed that, in 87% of the time, group-average predictions of the UMP were indistinguishable from the experimental results.

Sleep-latency model

As proposed by Akerstedt and Folkard [6], we used the following equation to describe the association between sleep latency (*SL*) and alertness impairment *P*:

$$SL(t) = A_{SL} e^{-k_{SL} P(t)}, \quad (2)$$

where *t* denotes the time of day (in hours), A_{SL} represents a scaling factor (in minutes), and k_{SL} denotes the rate at which *SL* decreases with *P* (in ms^{-1}), which, for a given sleep–wake schedule, was computed using the UMP. For example, Figure 2, A shows the predicted alertness impairment *P* (dotted purple line) for a habitual, 8-h nocturnal sleep (from 23:00 to 07:00). To predict the associated sleep latency at bedtime, we took the value of *P* at the start of the sleep period (gray area) and used equation (2) to compute the sleep latency. We estimated the values of A_{SL} and k_{SL} by fitting the sleep-latency model (equation 2) to the Study D1 data, and then validated the resulting model by comparing its predictions to experimental values from Studies D2 and V1 to V19.

Sleep-duration model

Following Borbély's procedure [3], the sleep-duration model has two components, one that takes into account sleep history (the homeostatic process *S*) and one that considers the effect of the circadian (process *C*). Thus, to predict sleep duration, we assumed that sleep spontaneously ends when the homeostatic

process S decreases to a circadian-regulated, sleep-termination threshold T representing the propensity of an individual to wake up. Accordingly, we determined sleep duration as the period between sleep onset (Figure 2, B, dotted blue arrow, left) and the time (dotted blue arrow, right) at which process S (solid blue line) intersected with threshold T (dashed green line), defined as:

$$T(t) = A_{SD} - \kappa_{SD} C(t + \varphi_{SD}), \quad (3)$$

where κ_{SD} represents the amplitude of the threshold T (in ms), φ_{SD} indicates a phase shift (in hours) of the threshold T with respect to process C, and A_{SD} denotes a constant (in ms) whose value is set so that process S reaches T at 07:00, after sleep onset at 23:00 under rested conditions. We estimated κ_{SD} and φ_{SD} by fitting the sleep-duration model (equation 3) to the Study D2 data, and then validated the sleep-duration model by comparing its predictions to experimental values from Studies V12 to V22.

In summary, here, we used the UMP with the same model parameter values as the ones estimated in Ramakrishnan et al. [37] to infer the values of processes S and C, and the resulting alertness impairment P, as a function of sleep history and time of day. Moreover, based on previous efforts [3, 6], we formulated equations to predict sleep latency and threshold T, and used two studies that broadly span the entire circadian cycle to estimate the model parameters, each to estimate parameters of one model.

Sleep-latency and sleep-duration inputs and predictions

To predict sleep latency and sleep duration for a given study, the extended UMP takes as inputs the sleep-schedule history during the baseline nights, the nominal sleep schedule we wish to predict (i.e. the time-in-bed period [TIB] or only the TIB start times for ad libitum sleep), and the start times of the sleep-latency tests, when tests were performed. It then predicts sleep latency and sleep duration for the first sleep period in the schedule, and updates the nominal duration of this sleep period with the predicted value. For studies with fixed sleep durations, when the predicted duration was longer than the scheduled TIB, we used the TIB as the updated sleep duration to conform to the study schedule. Next, the model predicts sleep latency and sleep duration for the second sleep period in the nominal schedule, and updates the duration of the sleep period with the predicted value. This process continues until we predict all sleep-latency and sleep-duration periods in the schedule. In addition, the model outputs the predicted sleep latency values for the corresponding tests in the study.

Goodness of fit

To assess the goodness of fit of the sleep-latency model predictions, we calculated the root mean square error (RMSE) between the model predictions and each group-average sleep-latency data point from each validation study. To provide the overall RMSE across all studies, we obtained the RMSE for each study and then averaged these values over the number of validation studies (20, D2 and V1–V19, in this case). Alternatively, we also calculated the pooled RMSE by computing the RMSE across all group-average sleep-latency data points from all validation studies and averaging over the total number of data points (133, in this case). While the former method provides equal weight to each predicted study, the latter provides equal weight to each predicted data point and, consequently, more weight to studies (and group of subjects) with more measurements. In addition, we calculated the coefficient of determination (R^2) to estimate the proportion of variance in the sleep-latency data captured by the model, and the concordance correlation coefficient to quantify the agreement between the model predictions and the experimental data [40]. We estimated both coefficients by pooling the group-average sleep-latency data points from each of the studies used to develop and validate the model (147 data points from Studies D1, D2, and V1 to V19). We carried out similar calculations to assess the goodness of fit of the sleep-duration model predictions, using data from 11 studies (39 data points from Studies V12 to V22) to compute the overall and pooled RMSEs, and from all 12 studies (45 data points from Studies D2 and V12 to V22) to compute R^2 and the concordance correlation coefficient.

Results

Sleep-latency model

By fitting the sleep-latency model (equation 2) to the Study D1 data, we estimated the scaling factor A_{SL} as 272.4 (standard error = 58.5) min and the rate constant k_{SL} as 0.012 (8.1×10^{-4}) ms^{-1}. Subsequently, we validated the sleep-latency model by computing the RMSEs between the sleep-latency predictions and the corresponding values for each study (D2 and V1–V19, Table 2). Note that we used data from Study D2 for two purposes, one to validate the sleep-latency model and one to develop the sleep duration model. The average RMSE was 4.0 min, with all but two studies yielding RMSEs of less than 5.2 min and eight studies yielding errors of less than 3.0 min. For the former, the exceptions were Studies V6 (RMSE = 6.5 min, habitual sleep) and V19 (RMSE = 12.3 min, irregular schedule). For comparison, the half width of the 95% confidence interval computed with the average standard error of mean sleep latency was 3.0 min (based on Studies D2, V1, V3, V7–V11, V13, V14, V17, V18, and V20, for which data were available). Pooling all the predictions across all time points in each of the 20 validation studies yielded an overall RMSE of 5.4 min, whereas educated guesses of sleep latency of 5 or 10 min yielded RMSEs of 8.4 and 6.8 min, respectively.

Figure 3 shows the observed mean latency for Study D1 (black stars) and the observed values for the validation studies (Study D2 and V1–V19) plotted against the alertness impairment P predicted by the UMP, as well as the sleep-latency model predictions (solid black line). For 79 of the 133 time points in the validation set (60%), the error between the predicted and experimental values was less than 3.0 min, and was larger than 10.0 min for only seven time points (one in Studies V4 and V18, and five in Study V19). The predictions for the 22 time points outside the range of the data used to fit the model (above the horizontal dashed line in Figure 3) had an average RMSE of 7.9 min. Overall, the sleep-latency model accounted for 42% of the variance in the observed values (R^2 = 0.42) and had a concordance correlation coefficient of 0.65. Note that the model predicted reasonably well the validation studies, with exception of Study V19 (RMSE = 12.3 min, Table 2). In fact, for this study, the TPMA approach proposed by Akerstedt and Folkard [6] also yielded a large RMSE (10.8 min; see Supplementary material).

Table 2. Prediction errors of the sleep-latency and sleep-duration models.

	RMSE	
Study	**Sleep latency (min)**	**Sleep duration (h)**
D1	1.3	
D2	5.1	0.6
V1	4.5	
V2	3.4	
V3	1.3	
V4	5.0	
V5	2.3	
V6	6.5	
V7	3.9	
V8	4.4	
V9	3.4	
V10	4.8	
V11	2.7	
V12	1.5	1.7
V13	4.0	0.6
V14	2.5	0.3
V15	2.6	0.7
V16	2.9	0.8
V17	1.9	0.5
V18	5.2	0.5
V19	12.3	1.4
V20		1.1
V21		0.2
V22		0.6
Average*	4.0	0.8

RMSE, root mean square error.
*The average values included only the validation studies (Studies D2 and V1–V19 for sleep latency and Studies V12–V22 for sleep duration).

Without including this study, our model accounted for 61% of the variance in the data and the RMSE for the pooled datasets reduced to 3.7 min, instead of 5.4 min.

Sleep-duration model

By fitting the sleep-duration model to the Study D2 data, we estimated the circadian amplitude κ_{SD} of the sleep-termination threshold T (equation 3) as 41.2 (standard error = 3.8) ms and the phase shift φ_{SD} as 2.0 (0.2) h. The value of the constant A_{SD} (211.6 ms) was determined so that the homeostatic process S reached the threshold T at 07:00, after sleep onset at 23:00 under rested conditions. The relatively small phase shift (φ_{SD} = 2.0 h) between process C and the threshold T indicates that the crest of the propensity to wake up nearly coincides with the circadian boost in alertness during the evening, while the trough in the propensity to wake up nearly coincides with the circadian nadir in alertness during the night (dashed orange line in Figure 2, A and dashed green line in Figure 2, B).

We then validated the sleep-duration model by computing the RMSEs between the predicted and experimental sleep durations for each study (V12–V22, Table 2). The average RMSE for the 11 validation studies was 0.8 h, whereas eight of the studies had RMSEs of less than 1.0 h and Study V12 yielded the largest RMSE (1.7 h). Pooling the validation datasets yielded an overall RMSE of 1.1 h. Figure 4 shows the model predictions plotted against the observed mean sleep durations for Study D2 used to develop the model and for the validation studies. The majority of these plotted data are near the diagonal, with a high concordance correlation coefficient of 0.93, indicating that the model accurately predicted sleep duration. Overall, the model accounted for 84% of the variance in the experimental values (R^2 = 0.84), even though most of the validation-study data (rectangles depicted by the dashed lines) were outside the range of the data used to fit the model.

Abrupt changes in sleep duration

The predictions for Study V12 yielded the largest deviations from the observed values, where the model tended to over-predict sleep duration despite correctly reflecting the general trend. In contrast, the model also under-predicted sleep duration in 3 of 4 cases (three in Study V19 and one in Study V12) when sleep started around 15:00 (points labeled "a" in Figure 4). In particular, the data point from Study V12 (right-most "a") showed large experimental variance, with measured sleep durations ranging from 4 to 13 h [2, 41].

To shed light into this intriguing observation, using Study V12 as a benchmark, we performed simulations to investigate how time of day of sleep onset, level of process S, and sleep debt affect sleep duration. To this end, we started by performing simulations where we varied the time of day of sleep onset, and analyzed the time course of process S and the termination threshold T (Figure 2, C). When sleep started at ~15:00 (symbol †) as in Study V12, we found that process S (solid blue line) reached threshold T (dashed green line) when the propensity to wake up was at its crest, resulting in sleep cessation 5.1 h after sleep onset. In contrast, when sleep onset occurred at 17:00 (symbol * in Figure 2, C), process S (dotted yellow line) remained higher than T (dashed green line) even when the circadian propensity to wake up reached its crest. Moreover, because T decreased faster than S after cresting, sleep cessation was delayed until 13.3 h after sleep onset. These simulations suggest that, for sleep periods starting at ~15:00 in this specific study, relatively small variations in sleep-onset time, or individual differences in process S or threshold T, may result in either short (~5 h) or long bouts of sleep (~10 h), with the resulting experimental sleep duration in Figure 4 reflecting the average of such bimodal distribution.

This is consistent with the findings in Study V12. When sleep started at 15:00, sleep duration was long for four subjects (7–13 h) but short for the other two subjects (~4 h) [41]. Moreover, when the latter two subjects started a sleep period 4 h later at 19:00, they remained asleep for 11 and 14 h. This suggests that the sleep-onset time at which their sleep duration transitioned from short to long was later in the day than that of the other four subjects. Consistent with our analysis, these four subjects remained asleep for 9 to 12 h for the rest period starting at 19:00 [41]. Borbély [3] as well as Akerstedt and Folkard [9] also noted this duality in their simulations of Study V12, although they did not examine the issue in detail.

To further characterize this phenomena, we carried out simulations by systematically varying the variables that affect sleep duration: (1) time of day of sleep onset, (2) level of process S, and (3) sleep debt. The time of day determines the propensity to wake up during sleep (i.e. the value of T), which changes non-monotonically with time, with small differences in time of sleep onset resulting in large differences in sleep duration (Figure 2,

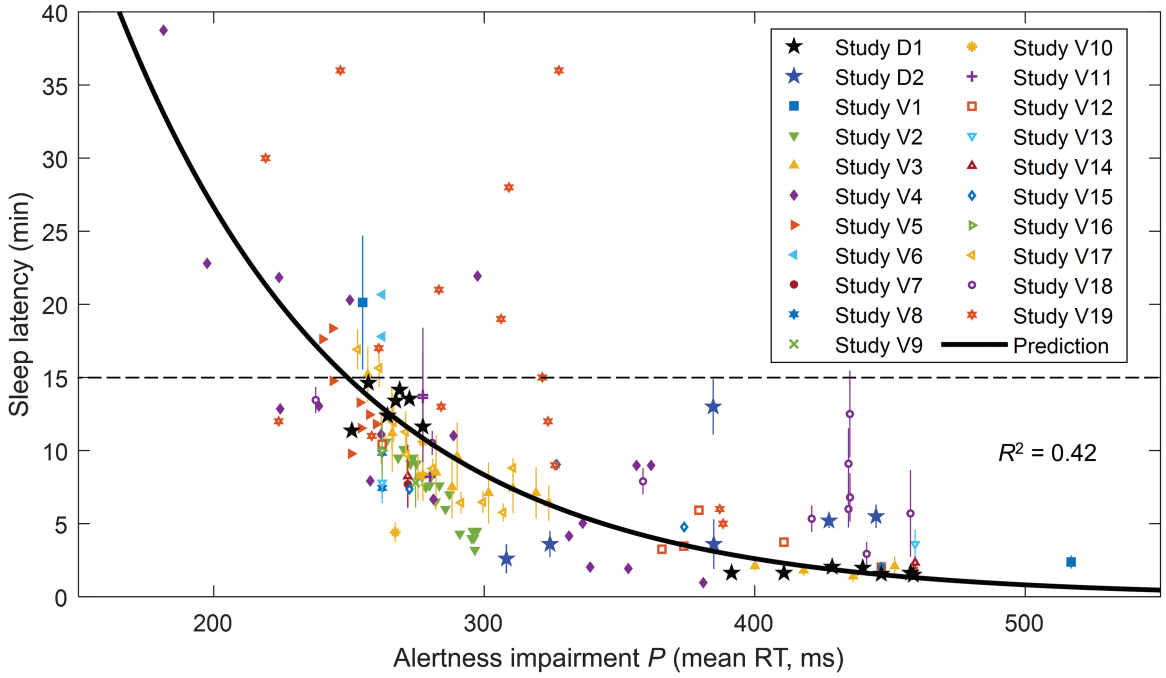

Figure 3. Sleep latency as function of predicted alertness impairment *P*. The black line represents the predicted sleep latency as a function of alertness impairment *P* (i.e. the predicted mean response time [RT]), with the sleep-latency model fitted using Study D1 data (black stars). Using the sleep schedule of each study as the input, we used the Unified Model of Performance to predict values for *p* at the time of the sleep-latency measurements. Twenty-two points outside the range of those of the study used to fit the sleep-latency model lie above the horizontal dashed line. The vertical bars represent the 95% confidence intervals of the observed mean sleep latencies (i.e. 1.96 × standard error of the mean [SEM]) for the studies that reported SEMs. See Figure 1 and Table 1 for details on sleep schedules, number of subjects, and sleep-latency measurements of each study. R^2: coefficient of determination.

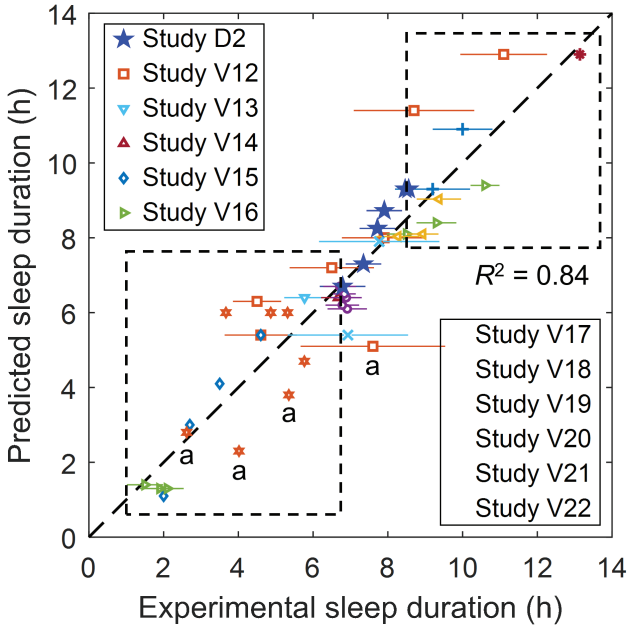

Figure 4. Observed and predicted sleep durations for different sleep studies. We used the sleep-duration measurements in Study D2 (blue stars) to estimate the parameters of the sleep-duration model, and measurements from Studies V12 to V22 to validate the model predictions. The horizontal bars represent the 95% confidence intervals of the experimental mean sleep durations (i.e. 1.96 × standard error of the mean [SEM]) for the studies that reported SEMs. The points labeled "a" correspond to sleep periods that started at 15:00. The two rectangles drawn with dashed lines contain data points outside the range of those of the study used to estimate the parameters of the sleep-duration model. See Figure 1 and Table 1 for details on sleep schedules, number of subjects, and sleep-duration measurements for each study. R^2: coefficient of determination.

C). The other two variables determine the time course of the monotonic decrease of process *S*, with higher initial levels of *S* prolonging sleep duration because of the larger reduction in *S* required to reach the sleep-termination threshold *T* (Figure 2, D, *S′* vs. *S*). Similarly, higher sleep debt prolongs sleep duration because *S* decreases at a lower rate (Figure 2, D, *S″* vs. *S*).

Figure 5 shows the effect of each of the three variables on sleep duration, while keeping the other two constant. For example, Figure 5, A shows how sleep duration varies as function of sleep-onset time, with *S* set to 420 and sleep debt set to 0.41 (values equivalent to those for a well-rested individual after 32 h of continuous wakefulness). In this case, the sleep duration was ~7 h for a sleep-onset time of 06:00, and gradually decreased as the sleep-onset time increased because *S* reached *T* at progressively higher values. This continued until the sleep-onset time reached 17:00, at which point *S* became higher than *T* at its zenith, resulting in a sleep duration of 13 h; beyond this point, the sleep duration gradually decreased again for the same reason. Increasing sleep debt (to 0.46) yielded a similar pattern, but with longer sleep durations and a shift toward an earlier sleep-onset time for the pivot point that determines whether sleep duration will be short or long (Figure 5, B). In contrast, reducing the initial value of *S* (to 350) resulted in shorter sleep durations and a shift toward a later sleep-onset time for the abrupt change in sleep duration (Figure 5, C). In other words, the model suggests that as sleep loss increases (i.e. as *S* and sleep debt increase), the pivot point that determines whether sleep duration will be short or long occurs earlier in the day.

Figure 5, D shows how sleep duration changes as a function of process *S*, for a sleep-onset time of 18:00 and a sleep debt

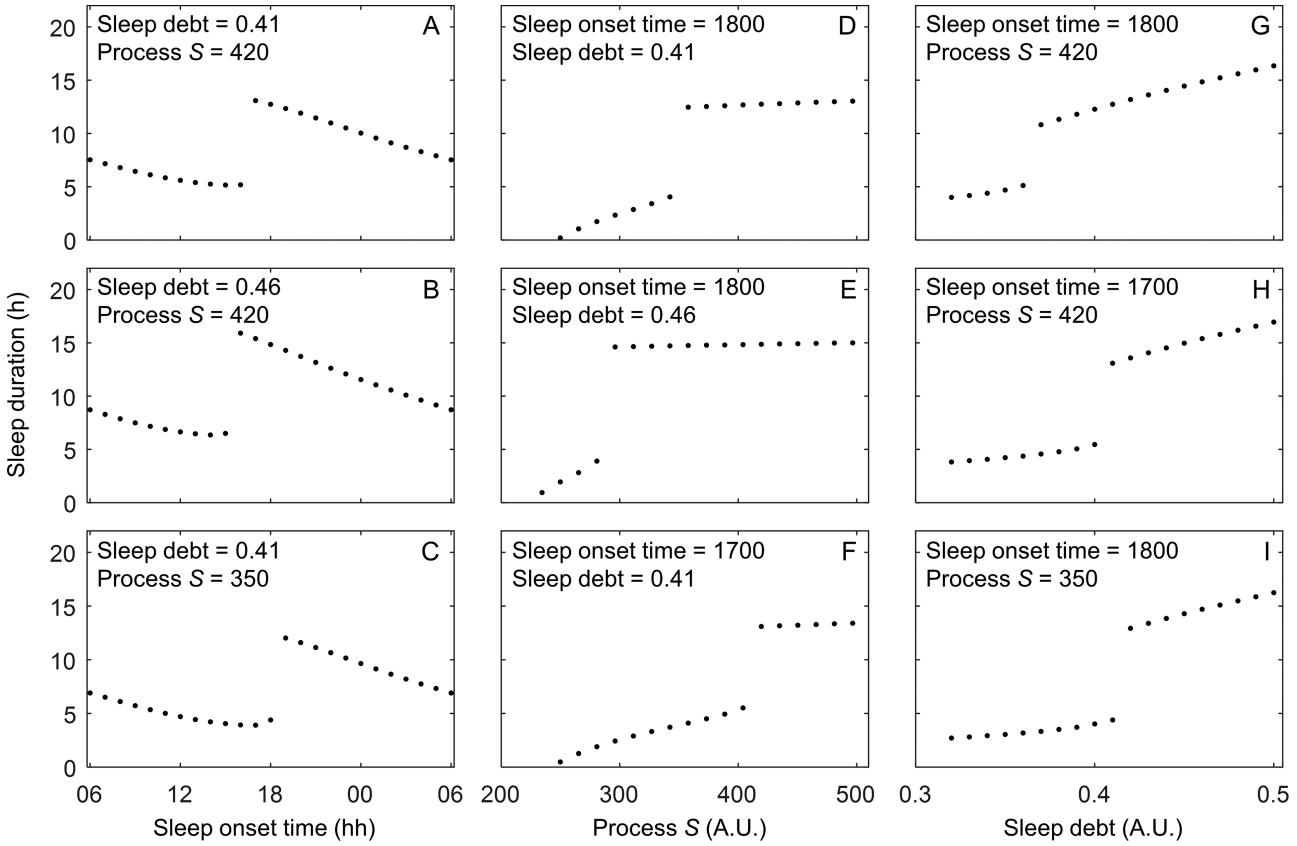

Figure 5. Sleep duration as a function of sleep-onset time, process S, or sleep debt. (A–C) Sleep duration as a function of sleep-onset time, with the level of process S and sleep debt fixed, as indicated in each panel. (D–F) Sleep duration as a function of process S, with sleep-onset time and sleep debt fixed, as indicated in each panel. (G–I) Sleep duration as a function of sleep debt, with sleep-onset time and level of process S fixed, as indicated in each panel. Values of sleep debt of 0.41 and 0.46 correspond to a well-rested individual after 32 and 48 h of continuous wakefulness, respectively. Values of process S of 350 and 420 correspond to a well-rested individual after 18 and 32 h of continuous wakefulness, respectively.

of 0.41. As expected, the sleep duration gradually increased as the initial values of process S increased, until S reached approximately 350 and became higher than T at its zenith, at which point the sleep duration sharply increased. As was the case for sleep-onset time, increasing the sleep debt (to 0.46) yielded a similar pattern, but with longer sleep durations and a reduction in the initial value of S at which the sleep duration sharply increased (Figure 5, E). Conversely, shifting the sleep-onset time to 17:00 resulted in shorter sleep durations and an increase in the initial value of S at which the sleep duration sharply increased (Figure 5, F). The effect of sleep debt on sleep duration was qualitatively similar to that of the initial value of S. When we set S to 420 and the sleep-onset time to 18:00, the sleep duration increased gradually as the sleep debt increased, with an abrupt change in sleep duration occurring at a sleep debt of 0.37 (Figure 5, G). Shifting the sleep-onset time earlier (Figure 5, H) or reducing the initial value of process S (Figure 5, I) had the same effect: both changes increased the value of sleep debt at which the sleep duration abruptly increased.

We also analyzed sleep duration by simultaneously varying process S and the sleep-onset time, with the sleep debt fixed at 0.41 attained after 32 h of wakefulness (Figure 6, A). This analysis resulted in a decision boundary (solid black line) that separated sleep durations into short and long. To the left of this boundary, sleep duration is short and gradually increases for earlier sleep-onset times, whereas to the right of the boundary, sleep duration is long and gradually decreases for later sleep-onset times. The figure also shows the predicted sleep duration for the sleep-onset time of 15:00 in Study V12 as an open square, whose proximity to the boundary could help explain the large variability in sleep duration across individuals. Figure 6, B shows the boundary for different levels of sleep debt. The minimum sleep debt for which there is an abrupt change in the sleep duration is 0.32 (equivalent to that accumulated during 10 h of wakefulness). As the sleep debt increases, the boundary shifts to earlier times and the range of possible values for process S decreases because sleep debt restricts the recovery of process S (i.e. it increases the lower bound of S).

Discussion

Planning efficient sleep schedules for shift workers requires the ability to predict the extent to which individuals will find it difficult to initiate and maintain sleep. Here, we attempted to accomplish this by extending the UMP to predict sleep latency and sleep duration, as a function of sleep history and time of day. We validated these models' predictions using an extensive set of experimental studies, probing sleep–wake conditions that spanned the entire circadian cycle and a wide range of sleep debt levels, with wakefulness ranging from 2 to 40 h, sleep restriction from 2 to 6 h of sleep per night, and naps of 5–60 min. In total, these 24 studies provided

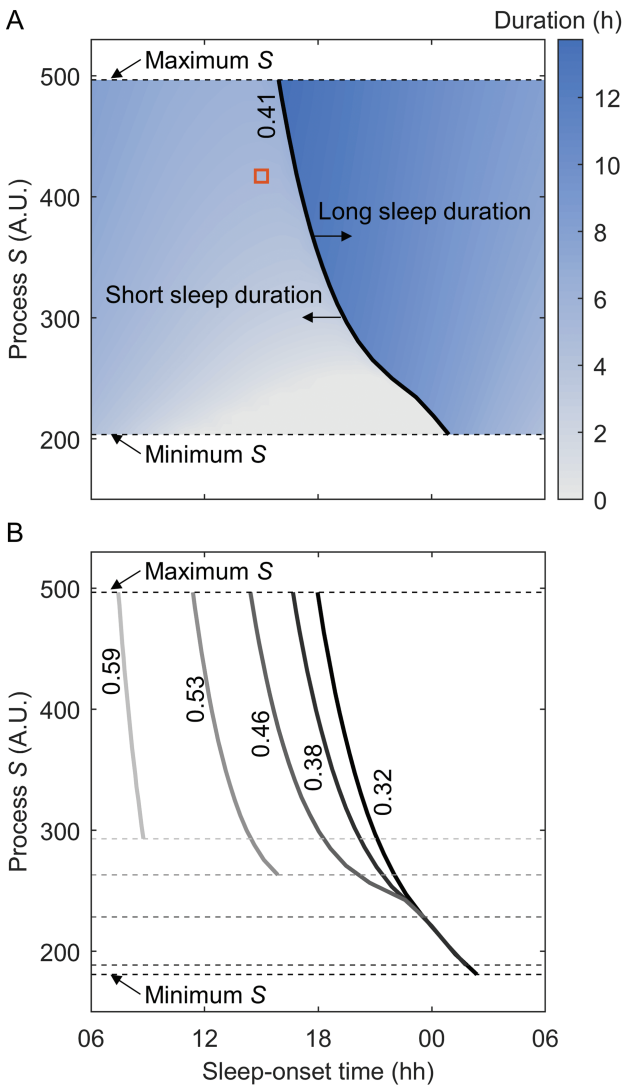

Figure 6. Sleep duration as a function of sleep-onset time, process S, and sleep debt. (A) Boundary between sleep periods of short and long durations (solid black line) as a function of sleep-onset time and process S, with the sleep debt fixed at 0.41. Sleep duration is indicated by the color bar. The open square indicates the predicted level of process S when sleep started at 15:00 in Study V12. Dashed lines at the top and bottom denote the maximum and minimum levels of process S, respectively. (B) Boundary between sleep periods of short and long durations (to the left and to the right of the solid lines, respectively) as a function of sleep-onset time and level of process S for different levels of sleep debt, as indicated by the number next to each line. Sleep debts of 0.32, 0.38, 0.46, 0.53, and 0.59 correspond to continuous wakefulness periods of 10, 24, 48, 72, and 96 h, respectively, starting from a well-rested (no sleep debt) condition. The minimum level of S depends on sleep debt (as described in Supplementary material), with lighter shades of gray indicating higher levels of sleep debt.

192 distinct time points from 317 subjects. As such, the resulting sleep-latency and sleep-duration models produced accurate predictions across a wide array of sleep histories.

Validation of the sleep-latency model based on 20 studies (D2 and V1–V19, Table 2) resulted in an average RMSE of 4.0 min between the predicted and observed values. This is an acceptable error level for planning sleep schedules, given its small magnitude relative to potential sleep durations. For the sleep-duration predictions over 11 validation studies (V12–V22, Table 2), the average RMSE was 0.8 h. This relatively high average error was largely driven by two studies (V12 and V19), which yielded an average error of 1.5 h. Nevertheless, comparisons between model predictions and experimental data indicated a large concordance correlation coefficient (0.93). Notably, for Study V18, which simulates night-shift work, the prediction errors for sleep latency and sleep duration were 5.2 min and 0.5 h, respectively. Although these results suggest that the models may be capable of predicting real-world scenarios, further validation of the sleep-latency and sleep-duration models against data from actual shift-work studies is needed.

We developed the sleep-duration model using data from a forced desynchrony study (D2) [15], whereas the 11 studies used for model validation collected sleep-duration data under conditions other than forced desynchrony. Yet the overall prediction error for the validation studies was only slightly larger than that of Study D2 (0.8 vs. 0.6 h). Moreover, the prediction error for sleep-latency in Study D2 (5.1 min) was only slightly larger than the overall error (4.0 min) across all validation studies. These suggest that the extended UMP can predict sleep latency and sleep duration under forced-desynchrony conditions as well as other conditions without the need for condition-specific model modifications.

In the development of the sleep-duration model, the estimated circadian phase of the termination threshold T had a small shift (2.0 h) with respect to that in process C. We previously estimated the circadian phase of process C in the UMP [37] using alertness data from a sleep dose–response study [39]. However, to develop the sleep-duration model we could not use the same study because, for modeling sleep duration, we needed a study in which subjects slept for prolonged periods, with sleep periods starting at different phases of the circadian cycle. To this end, we estimated the parameters of T in equation (3) using the study by Dijk et al. [15], which resulted in different values for the circadian phase in T and C. This procedure is similar to the one used by Akerstedt and Folkard in their TPMA method for predicting alertness regulation and sleep duration [9].

To assess the performance of the extended UMP with previous models, we compared its predictions against those obtained with the TPMA [6, 9]. In doing so, we excluded from the comparisons the studies we used to develop our models (Studies D1 and D2, Table 1) and the studies they used to develop their models (Studies V12 and V15). While the two approaches produced similar prediction errors for sleep latency (4.2 min), the UMP prediction error for sleep duration was 70% smaller (30 min, 0.7 vs. 1.2 h) than that of the TPMA (see Supplementary material). In terms of the coefficient of determination R^2, the extended UMP produced slight but consistently superior results. For example, for sleep latency, the UMP and the TPMA captured 35% and 28% of the variance in the data, respectively, whereas for sleep duration these values increased to 90% and 80%, respectively.

Using our model and results from additional studies, we found that the discontinuity in sleep duration that was previously observed at specific sleep onset times [2, 41] can actually occur throughout the circadian cycle and can be explained by the confluence of three factors: (1) time of day of sleep onset, (2) level of process S, and (3) sleep debt. Three independent studies support this hypothesis [8, 34, 42]. In a study where individuals were allowed to sleep at libitum, and sleep- and body-temperature cycles were decoupled in the absence of external time cues, Zulley et al. also observed a discontinuity in sleep

duration [8]. Specifically, this study revealed that sleep bouts starting approximately 15 h after the trough of the body temperature were either short (4–8 h) or long (12–18 h). In a similar study, Strogatz et al. observed the same phenomenon, although the discontinuity in sleep duration was observed when sleep started approximately 9 h after the trough in body temperature. Sleep bouts starting around this time resulted in either short (3–10 h) or long (11–22 h) sleep durations [42]. Dijk and Beersma observed a similar binomial distribution in sleep duration in Study V20 [34]. In the study, where eight subjects initiated sleep at 11:00 after 27 h of wakefulness, five of them slept for 4–6 h and the other three slept for 9–10 h. This finding suggests that when predicting sleep duration, it is necessary to determine whether a relative small change in the sleep-onset time can result in large changes in sleep duration, that is, determine the proximity of the sleep-onset time to the pivotal discontinuity boundary (Figure 6). To this end, we can use the sleep-duration model to perform simulations around the desired sleep-onset time to determine the presence of such discontinuity.

With the model extensions, we can use the UMP to address different sleep-optimization tasks. For example, we can use it to identify sleep periods that maximize sleep duration. More importantly, these new capabilities allow us to use the UMP to design sleep schedules that maximize alertness during work periods. Accordingly, we can use the UMP to systematically generate thousands of sleep–wake schedules similar to the approach used by Vital-Lopez et al. to optimize caffeine consumption [43], assess the feasibility of each schedule using the sleep-latency and sleep-duration models, and then rank order the feasible schedules by comparing the predicted time course of alertness during work periods.

Our work has limitations. First, we developed the sleep-latency and sleep-duration models using data from laboratory studies in which a homogenous population of healthy young adults served as subjects. The extent to which we can extrapolate the models' predictions to heterogeneous, older populations is unknown. Second, the models do not account for individual variations due to differences in chronotype, sex, or age. In particular, women tended to be under-represented in the studies used here (169 men vs. 92 women in the studies that provided sex information). However, it is not clear whether sex has a determinant effect on predicting sleep latency or sleep duration, because we were able to predict sleep duration for studies involving both men and women based on a model entirely developed using data from men (Study D2). Tentatively, we could address this limitation, in part, by developing tailored models to each individual, wherein we individualize the model by customizing the two-process model parameters based on the individual's PVT data [44, 45]. For example, Rusterholz et al. [46] found considerable variation in the time constants associated with the increase and decrease of process *S*, whereas Liu et al. [44] found that the upper asymptote of process S as well as the amplitude and phase of process *C* were the more important parameters to capture individual variations. Third, the models do not account for the effects of exogenous factors that affect sleep, such as stimulants (e.g. caffeine), soporifics, and light exposure. Although the UMP accounts for the effects of caffeine on alertness [47], it remains to be seen whether the developed models can predict sleep latency and sleep duration following caffeine consumption. In addition, the sleep-latency model may not be suitable for conditions in which individuals are forced to spend prolonged periods in bed, because all studies used to developed and validate the models involved subjects who carried some level of sleep debt. However, we do not believe that this constitutes a severe limitation because such conditions are rarely encountered in real-world situations.

In summary, we developed and validated mathematical models to predict sleep latency and sleep duration as a function of sleep history and time of day. By being able to determine whether an individual can actually fall asleep at the proposed time and stay asleep for the desired duration, these capabilities allow us to assess the feasibility of potential sleep–wake schedules, laying the foundation for developing algorithms that identify the best sleep times to optimize alertness during work periods. By combining these new capabilities with existing, publicly available resources, such as the *2B-Alert* Web [48], we will be able to deploy a comprehensive set of fatigue-management tools to predict and optimize sleep and alertness, as well as the effects of caffeine consumption on neurobehavioral performance [43].

Funding

This work was sponsored by the Military Operational Medicine Research Program of the U.S. Army Medical Research and Development Command (USAMRDC), Ft. Detrick, MD. TJB was supported by an appointment to the Research Participation Program at the Walter Reed Army Institute of Research administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and USAMRDC.

Disclosure Statement

Financial disclosure: This was not an industry-supported study. The authors are co-inventors of the *2B-Alert* technology, which has been licensed to the private sector. However, the model extensions reported in the current manuscript are not part of the Licensing Agreement.

Non-financial disclosure: The authors have indicated no conflicts of interest. The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army, the U.S. DoD, or The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70–25. This paper has been approved for public release with unlimited distribution.

References

1. Akerstedt T. Sleepiness as a consequence of shift work. *Sleep*. 1988;**11**(1):17–34.
2. Akerstedt T, *et al*. The circadian variation of experimentally displaced sleep. *Sleep*. 1981;**4**(2):159–169.
3. Borbély AA. A two process model of sleep regulation. *Hum Neurobiol*. 1982;**1**(3):195–204.
4. Borbély AA, *et al*. Sleep initiation and initial sleep intensity: interactions of homeostatic and circadian mechanisms. *J Biol Rhythms*. 1989;**4**(2):149–160.

5. Akerstedt T, *et al*. Validation of the S and C components of the three-process model of alertness regulation. *Sleep*. 1995;**18**(1):1–6.
6. Akerstedt T, *et al*. Predicting sleep latency from the three-process model of alertness regulation. *Psychophysiology*. 1996;**33**(4):385–389.
7. Czeisler CA, *et al*. Human sleep: its duration and organization depend on its circadian phase. *Science*. 1980;**210**(4475):1264–1267.
8. Zulley J, *et al*. The dependence of onset and duration of sleep on the circadian rhythm of rectal temperature. *Pflugers Arch*. 1981;**391**(4):314–318.
9. Akerstedt T, *et al*. Predicting duration of sleep from the three process model of regulation of alertness. *Occup Environ Med*. 1996;**53**(2):136–141.
10. Phillips AJK, *et al*. Modeling the adenosine system as a modulator of cognitive performance and sleep patterns during sleep restriction and recovery. *PLoS Comput Biol*. 2017;**13**(10):e1005759.
11. Leproult R, *et al*. Individual differences in subjective and objective alertness during sleep deprivation are stable and unrelated. *Am J Physiol Regul Integr Comp Physiol*. 2003;**284**(2):R280–R290.
12. Zhou X, *et al*. Mismatch between subjective alertness and objective performance under sleep restriction is greatest during the biological night. *J Sleep Res*. 2012;**21**(1):40–49.
13. Rajdev P, *et al*. A unified mathematical model to quantify performance impairment for both chronic sleep restriction and total sleep deprivation. *J Theor Biol*. 2013;**331**:66–77.
14. Bonnet MH, *et al*. We are chronically sleep deprived. *Sleep*. 1995;**18**(10):908–911.
15. Dijk DJ, *et al*. Ageing and the circadian and homeostatic regulation of human sleep during forced desynchrony of rest, melatonin and temperature rhythms. *J Physiol*. 1999;**516**(Pt 2):611–627.
16. Paech GM, *et al*. Caffeine has minimal effects on daytime recovery sleep following severe sleep deprivation. *Sleep Biol Rhythms*. 2016;**14**(2):149–156.
17. Brooks A, *et al*. A brief afternoon nap following nocturnal sleep restriction: which nap duration is most recuperative? *Sleep*. 2006;**29**(6):831–840.
18. Rosenthal L, *et al*. Level of sleepiness and total sleep time following various time in bed conditions. *Sleep*. 1993;**16**(3):226–232.
19. Richardson GS, *et al*. Circadian variation of sleep tendency in elderly and young adult subjects. *Sleep*. 1982;**5**(Suppl 2):S82–S94.
20. Nicholson AN, *et al*. Heterocyclic amphetamine derivatives and caffeine on sleep in man. *Br J Clin Pharmacol*. 1980;**9**(2):195–203.
21. Drapeau C, *et al*. Challenging sleep in aging: the effects of 200 mg of caffeine during the evening in young and middle-aged moderate caffeine consumers. *J Sleep Res*. 2006;**15**(2):133–141.
22. Landolt HP, *et al*. Caffeine intake (200 mg) in the morning affects human sleep and EEG power spectra at night. *Brain Res*. 1995;**675**(1–2):67–74.
23. Takahashi M, *et al*. Brief naps during post-lunch rest: effects on alertness, performance, and autonomic balance. *Eur J Appl Physiol Occup Physiol*. 1998;**78**(2):93–98.
24. Hayashi M, *et al*. The effects of a 20-min nap at noon on sleepiness, performance and EEG activity. *Int J Psychophysiol*. 1999;**32**(2):173–180.
25. Hayashi M, *et al*. The alerting effects of caffeine, bright light and face washing after a short daytime nap. *Clin Neurophysiol*. 2003;**114**(12):2268–2278.
26. Carrier J, *et al*. Effects of caffeine are more marked on daytime recovery sleep than on nocturnal sleep. *Neuropsychopharmacology*. 2007;**32**(4):964–972.
27. Carrier J, *et al*. Effects of caffeine on daytime recovery sleep: a double challenge to the sleep-wake cycle in aging. *Sleep Med*. 2009;**10**(9):1016–1024.
28. Akerstedt T, *et al*. Sleep duration and the power spectral density of the EEG. *Electroencephalogr Clin Neurophysiol*. 1986;**64**(2):119–122.
29. Klerman EB, *et al*. Interindividual variation in sleep duration and its association with sleep debt in young adults. *Sleep*. 2005;**28**(10):1253–1259.
30. Kamdar BB, *et al*. The impact of extended sleep on daytime alertness, vigilance, and mood. *Sleep Med*. 2004;**5**(5):441–448.
31. Muehlbach MJ, *et al*. The effects of caffeine on simulated night-shift work and subsequent daytime sleep. *Sleep*. 1995;**18**(1):22–29.
32. Akerstedt T, *et al*. Regulation of sleep and naps on an irregular schedule. *Sleep*. 1993;**16**(8):736–743.
33. Carskadon MA, *et al*. Effects of total sleep loss on sleep tendency. *Percept Mot Skills*. 1979;**48**(2):495–506.
34. Dijk DJ, *et al*. Effects of SWS deprivation on subsequent EEG power density and spontaneous sleep duration. *Electroencephalogr Clin Neurophysiol*. 1989;**72**(4):312–320.
35. Dijk DJ, *et al*. Time course of EEG power density during long sleep in humans. *Am J Physiol*. 1990;**258**(3 Pt 2):R650–R661.
36. Depner CM, *et al*. Ad libitum weekend recovery sleep fails to prevent metabolic dysregulation during a repeating pattern of insufficient sleep and weekend recovery sleep. *Curr Biol*. 2019;**29**(6):957–967.e4.
37. Ramakrishnan S, *et al*. A unified model of performance: validation of its predictions across different sleep/wake schedules. *Sleep*. 2016;**39**(1):249–262.
38. Basner M, *et al*. Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep*. 2011;**34**(5):581–591.
39. Belenky G, *et al*. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *J Sleep Res*. 2003;**12**(1):1–12.
40. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;**45**(1):255–268.
41. Daan S, *et al*. Timing of human sleep: recovery process gated by a circadian pacemaker. *Am J Physiol*. 1984;**246**(2 Pt 2):R161–R183.
42. Strogatz SH, *et al*. Circadian regulation dominates homeostatic control of sleep length and prior wake length in humans. *Sleep*. 1986;**9**(2):353–364.
43. Vital-Lopez FG, *et al*. Caffeine dosing strategies to optimize alertness during sleep loss. *J Sleep Res*. 2018;**27**(5):e12711.
44. Liu J, *et al*. Real-time individualization of the unified model of performance. *J Sleep Res*. 2017;**26**(6):820–831.
45. Reifman J, *et al*. *2B-Alert* App: a mobile application for real-time individualized prediction of alertness. *J Sleep Res*. 2019;**28**(2):e12725.
46. Rusterholz T, *et al*. Inter-individual differences in the dynamics of sleep homeostasis. *Sleep*. 2010;**33**(4):491–498.
47. Ramakrishnan S, *et al*. A unified model of performance for predicting the effects of sleep and caffeine. *Sleep*. 2016;**39**(10):1827–1841.
48. Reifman J, *et al*. *2B-Alert* Web: an open-access tool for predicting the effects of sleep/wake schedules and caffeine consumption on neurobehavioral performance. *Sleep*. 2016;**39**(12):2157–2159.