Can a mathematical model predict an individual's trait-like response to both total and partial sleep loss?

SRIDHAR RAMAKRISHNAN[1], WEI LU[1], SRINIVAS LAXMINARAYAN[1], NANCY J. WESENSTEN[2], TRACY L. RUPP[2], THOMAS J. BALKIN[2] and JAQUES REIFMAN[1]

[1]Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, Fort Detrick, MD, USA; [2]Department of Behavioral Biology, Walter Reed Army Institute of Research, Silver Spring, MD, USA

Keywords
biomathematical model, psychomotor vigilance task, sleep-loss phenotype, trait preservation, two-process model

Correspondence
Jaques Reifman, PhD, Senior Research Scientist, DoD Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, ATTN: MCMR-TT, 504 Scott Street, Fort Detrick, MD 21702, USA.
Tel.: 301-619-7915;
fax: 301-619-1983;
e-mail: jaques.reifman.civ@mail.mil

Accepted in revised form 21 November 2014; received 18 July 2014

DOI: 10.1111/jsr.12272

SUMMARY

Humans display a trait-like response to sleep loss. However, it is not known whether this trait-like response can be captured by a mathematical model from only one sleep-loss condition to facilitate neurobehavioural performance prediction of the same individual during a different sleep-loss condition. In this paper, we investigated the extent to which the recently developed unified mathematical model of performance (UMP) captured such trait-like features for different sleep-loss conditions. We used the UMP to develop two sets of individual-specific models for 15 healthy adults who underwent two different sleep-loss challenges (order counterbalanced; separated by 2–4 weeks): (i) 64 h of total sleep deprivation (TSD) and (ii) chronic sleep restriction (CSR) of 7 days of 3 h nightly time in bed. We then quantified the extent to which models developed using psychomotor vigilance task data under TSD predicted performance data under CSR, and vice versa. The results showed that the models customized to an individual under one sleep-loss condition accurately predicted performance of the same individual under the other condition, yielding, on average, up to 50% improvement over non-individualized, group-average model predictions. This finding supports the notion that the UMP captures an individual's trait-like response to different sleep-loss conditions.

INTRODUCTION

Insufficient sleep impairs alertness and neurobehavioural performance. Results from Van Dongen *et al.* (2004) and Rupp *et al.* (2012) showed that substantial interindividual variability exists with regard to response to sleep loss. More importantly, both studies also showed that this response was trait-like. Van Dongen *et al.* (2004) evaluated the same 21 individuals under three separate total sleep deprivation (TSD) challenges of 36 h each and observed that the neurobehavioural deficits resulting from sleep loss were stable within individuals across the three challenges. Rupp *et al.* (2012) showed that an individual's neurobehavioural response to 64 h of TSD was correlated positively with that individual's response to chronic sleep restriction (CSR) [seven consecutive days of 3 h nightly time in bed (TIB)]. In both studies, they quantified individual response to sleep loss by averaging neurobehavioural performance [i.e. psychomotor vigilance task (PVT) data] over the last 12–24 h of the sleep-loss condition, and used the intraclass correlation coefficient (ICC) to assess the extent to which an individual's vulnerability rank among a group of individuals was preserved across sleep-loss conditions. Large ICC values observed in the two studies suggest a high degree of trait preservation. In other words, performance on the last day of a particular sleep-loss challenge can predict accurately the relative rank of an individual for a subsequent sleep-loss challenge. However, such analyses provide little or no information regarding the temporal dynamics of an individual's performance, which is critical in operational settings.

In the past, many biomathematical models have been developed to predict the impact of sleep/wake and circadian influences on human performance (Mallis *et al.*, 2004). Although these models predict the dynamics of performance

accurately under certain sleep-loss conditions, they fail to generalize and predict performance across different conditions (Van Dongen, 2004). Specifically, models developed from performance measures obtained under TSD conditions do not predict performance accurately under CSR, and vice versa, because such models (i) do not account for the sleep debt from the individual's prior sleep/wake history, (ii) have mis-specified the relationship between the lower and upper asymptotes of the sleep homeostatic process or (iii) have used a large number of parameters, making it difficult to estimate model parameters from limited data (Rajdev *et al.*, 2013).

To overcome these limitations, we recently developed a unified mathematical model of performance (UMP) to predict PVT performance more accurately across different sleep-loss conditions, ranging from short periods of acute TSD to longer periods of CSR (Rajdev *et al.*, 2013). Unified model parameters obtained by fitting the group-averaged TSD data from the Rupp *et al.* (2012) cross-over study could predict the group-averaged performance accurately under CSR, and vice versa. The purpose of this present work is to investigate the extent to which the UMP accurately predicts individual responses to sleep loss. In particular, we address four questions:

1. To what extent are the individual-specific model parameters preserved (similar) across the TSD and CSR conditions?
2. To what extent does the UMP developed for an individual from data derived under one sleep-loss condition predict the same individual's response under a different sleep-loss condition?
3. To what extent do individual-specific models increase prediction accuracy over group-average models?
4. Which of the two conditions (TSD or CSR) facilitates development of more generalizable models?

METHODS

Study data

We used PVT data from Rupp *et al.* (2012) in which 19 healthy adults underwent two sleep-loss challenges separated by 2–4 weeks: (i) 64 h of TSD and (ii) CSR consisting of seven consecutive nights of 3 h nightly TIB. Both challenges were preceded by a sleep-satiation stage of 7 in-laboratory nights with 10 h TIB and followed by 3 nights with 8 h TIB (recovery). For both TSD and CSR challenges, wakeup times were fixed at 07:00 hours.

During the entire wake period of TSD and CSR, 10-min PVTs were administered every 2 h. Using the response time (RT) data from each of these PVT sessions, we computed the following five commonly used performance statistics: (i) mean RT, (ii) median RT, (iii) slowest 10% RT, (iv) speed (=mean 1/RTs) and (v) lapses (number of RTs >500 ms).

We excluded four subjects from our analyses: one subject was excluded due to missing data, and three other subjects were excluded due to significant differences in RT distributions of their baseline sessions (first day of TSD/CSR) between the TSD and CSR conditions. We used the Wasserstein distance metric (Zhou and Shi, 2011) to quantify differences in RT distributions. For these three subjects, the Wasserstein distances were >2.5 interquartile distance from the median value computed across all subjects.

Unified model (UMP)

The UMP (Rajdev *et al.*, 2013) was developed as an extension of the classical two-process model wherein it accounts explicitly for sleep debt resulting from a known

Table 1 Biomathematical framework of the unified mathematical model of performance (UMP)

Governing equations

Performance impairment (P):

$$P(t) = S(t) + \kappa C(t), \tag{1}$$

where S and C denote the homeostatic and circadian processes of the two-process model at time t, respectively, and κ represents the circadian amplitude

Circadian process (C):

$$C(t) = \sum_{i=1}^{5} a_i \sin\left[i\frac{2\pi}{\tau}(t+\phi)\right], \tag{2}$$

where a_i, $i = 1, \ldots, 5$, represent the amplitude of the five harmonics ($a_1 = 0.97$, $a_2 = 0.22$, $a_3 = 0.07$, $a_4 = 0.03$, and $a_5 = 0.001$), τ denotes the fundamental period of the circadian clock (~24 h) and ϕ denotes the circadian phase

Homeostatic process (S):

$$\dot{S}(t) = \begin{cases} 1/\tau_w[U - S(t)] & \text{during wakefulness} \\ -1/\tau_s[S(t) - L(t)] & \text{during sleep,} \end{cases} \tag{3}$$

where U and L denote the upper and lower homeostatic asymptotes, respectively, τ_w and τ_s denote the wake- and sleep-time constants of the increasing and decreasing sleep pressure, respectively. $S(0) = S_0$ and $L(0) = L_0$ correspond to the initial state values for S and L, respectively

Lower asymptote (L) of process S:

$$L(t) = U \times D(t), \tag{4}$$

where D denotes the sleep debt.

Sleep debt (D):

$$\dot{D}(t) = -1/\tau_{LA}[D(t) - Loss(t)], \tag{5a}$$

$$\text{Loss}(t) = \begin{cases} 1 & \text{during wakefulness} \\ -2 & \text{during sleep,} \end{cases} \tag{5b}$$

where τ_{LA} denotes the time constant of the recursive filter that incorporates the exponential decay of the sleep-loss history.

sleep/wake history. To this end, it modulates the recovery capacity during sleep to vary inversely with extant sleep debt, with sleep losses incurred in the remote past having less effect on sleep debt. Table 1 summarizes the biomathematical equations (equations 1–5) governing the UMP. Five of the eight parameters of the model, U, τ_w, τ_s, τ_{LA} and κ, are termed trait parameters (parameters that reflect innate individual characteristics), while the remaining three parameters, S_0, L_0 and ϕ, are termed state parameters (parameters that depend on prior sleep/wake history) (Rajaraman *et al.*, 2009).

Individual-specific models

To obtain individual-specific models of performance, we fitted the UMP to each subject's PVT performance data obtained from each of the two sleep-loss conditions (TSD and CSR), resulting in two sets of model parameters for each subject, for each PVT statistic. Specifically, we minimized the following objective function to obtain the individual-specific model parameters $\Theta_i = (U, \tau_w, \tau_s, \tau_{LA}, \kappa, S_0, L_0, \phi)_i$ for the ith subject:

$$J(\Theta_i) = \frac{1}{T}\sum_{t=1}^{T}[P_{mi}(t) - P_i(t, \Theta_i)]^2, \quad (6)$$

where P_{mi} and P_i denote the measured performance data and the corresponding model fit of the ith subject, respectively, and T represents the total number of PVT measurements. Because the UMP output $P_i(t, \Theta_i)$ was insensitive to three trait parameters, τ_w, τ_s and τ_{LA} (i.e. the model output did not change appreciably with changes in these parameters), we fixed them to physiologically meaningful values, e.g. τ_w = 10 h, τ_s = 2 h and τ_{LA} = 7 days, for mean RT (Rajdev *et al.*, 2013; Rusterholz *et al.*, 2010). We reparameterized these parameters for each PVT statistic in order to constrain them to the following ranges: $0 < \tau_w < 40$ h, $0 < \tau_s < 4$ h and $0 < \tau_{LA} < 10$ days. We thus estimated only five individual-specific model parameters $\Theta_i = (U, \kappa, S_0, L_0, \phi)_i$ for each subject i.

Using the UMP parameters developed on the TSD data, we computed the individual-specific model fits (P_i) under TSD and the corresponding cross-condition predictions (P_{Xi}) under CSR, and vice versa.

Measures of parameter preservation

We computed the ICC for each of the five parameters to serve as a measure of the extent of parameter preservation. The ICC was computed as the ratio of between-subject variance to the sum of between-subject and within-subject variances, where we estimated the variances by performing a linear mixed-effects model fit on the parameters (Nakagawa and Schielzeth, 2010; Zhang *et al.*, 1998). Higher ICC values indicated greater agreement in the parameters across TSD and CSR conditions. We interpreted ICC agreement based on the following established ranges (Landis and Koch, 1977): slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80) and almost perfect (0.81–1.00) agreement.

Group-average models

To develop two sets of 15 group-average models for each of the two conditions, we fitted the UMP separately to group-averaged TSD data and group-averaged CSR data. To obtain a group-average model for predicting subject i, we excluded performance data from the ith subject when averaging the data used for model fitting, leading to 15 group-average models per condition. This ensured that the group-average model was developed independently of that subject's data and could be used as an unbiased predictor of that subject. We then used these models to compute within-condition predictions ($\bar{P}_i$) and cross-condition predictions ($\bar{P}_{Xi}$), generating four predictions for each subject i based on the two sets of group-average models.

Comparing model fits and predictions

To compare accuracies of the individual-specific model fits (P_i), group-average model predictions ($\bar{P}_i$ *and* $\bar{P}_{Xi}$) and individual-specific model predictions (P_{Xi}), we calculated the root mean squared error (RMSE) between each subject's performance data and the corresponding model fits or model predictions. We further compared the RMSEs generated by the individual-specific models and the group-average models for each subject, using the Wilcoxon paired, two-sided, signed-rank test (Zar, 1999).

RESULTS

Comparison of individual-specific models

We compared the individual-specific model parameters obtained under CSR and TSD conditions to determine whether the parameters were preserved across the two conditions. Table 2 lists the five UMP parameters obtained for each subject under each condition (entries within parentheses correspond to the TSD condition) for models developed using the PVT mean RT statistic. Also listed are the mean, standard deviation (SD) and ICC for each parameter.

We observed that for every subject, each of the parameters was similar across the two conditions, with the average difference being <15% for all parameters, except for ϕ, which exhibited an average absolute difference of <0.5 h. (Because ϕ exhibits a 24-h periodicity, it is improper to compute a percentage difference for it.) Furthermore, agreement for the two most sensitive parameters, U and κ, was almost perfect (ICC > 0.80). Two of the state parameters, S_0 and ϕ, showed moderate similarity across the two conditions (ICC > 0.40), but L_0 showed only slight agreement.

Table 2 Individual-specific parameters of the UMP for each of the 15 subjects based on PVT mean response time statistic under CSR and TSD (entries within parentheses) conditions. The intraclass correlation coefficient (ICC) indicates the degree of agreement in the parameters between the two conditions; a larger ICC indicates a greater degree of parameter preservation

	Trait parameters		*State parameters*		
Subject	*U (ms)*	κ *(ms)*	S_0 *(ms)*	L_0 *(ms)*	ϕ *(h)*
1	270 (296)	31 (37)	207 (200)	207 (180)	1.2 (1.0)
2	410 (385)	33 (30)	200 (200)	200 (180)	3.0 (2.4)
3	289 (281)	34 (32)	200 (200)	180 (180)	2.4 (4.5)
4	302 (283)	50 (48)	200 (200)	180 (180)	1.0 (1.5)
5	248 (238)	17 (15)	200 (209)	190 (181)	1.0 (4.1)
6	333 (289)	31 (44)	200 (200)	200 (180)	2.8 (5.7)
7	253 (254)	15 (18)	200 (200)	180 (180)	2.8 (2.1)
8	250 (236)	15 (15)	236 (214)	180 (214)	1.0 (1.0)
9	292 (271)	16 (30)	224 (200)	221 (200)	1.0 (1.5)
10	228 (230)	15 (15)	200 (204)	180 (180)	1.0 (1.0)
11	245 (237)	18 (20)	200 (200)	180 (180)	1.0 (1.5)
12	369 (395)	37 (50)	260 (222)	225 (188)	1.6 (1.0)
13	311 (366)	44 (50)	200 (200)	200 (180)	1.4 (1.0)
14	259 (253)	15 (15)	200 (200)	180 (184)	1.0 (1.0)
15	348 (385)	49 (50)	200 (200)	180 (188)	1.0 (1.0)
Mean	294 (293)	28 (31)	208 (203)	192 (185)	1.5 (2.0)
SD	52 (60)	13 (14)	18 (7)	16 (10)	0.8 (1.5)
ICC	0.90	0.89	0.52	0.10	0.46

CSR: chronic sleep restriction; L_0: lower homeostatic asymptote at time zero; PVT: psychomotor vigilance task; S_0: homeostatic state at time zero; SD: standard deviation; TSD: total sleep deprivation; *U*: upper asymptote of the homeostatic process; UMP: unified mathematical model of performance; ϕ: circadian phase; κ: circadian amplitude.

Individual-specific model fits versus predictions

Using the individual-specific UMP parameters obtained from CSR data, we computed the corresponding fits to CSR performance data and predictions of performance under the TSD condition, and vice versa. Fig. 1 shows the model fits (P_i) and cross-condition predictions (P_{Xi}) for three different subjects (subjects 3, 5 and 12), who showed different patterns of response to sleep loss. Individual-specific fits (blue solid lines) captured accurately the within- and across-day performance variations under both CSR and TSD. To a slightly lesser extent, the cross-condition predictions (red dashed lines) were also accurate, with RMSEs no greater than 5 ms compared to those of the fits. The fits and cross-condition predictions were also accurate during the recovery days, except for the first recovery day following TSD for subject 3 and for the second recovery day following CSR for subject 12; in these instances, the subjects appeared to recover faster than predicted by the models.

Individual-specific model predictions versus group-average model predictions

Fig. 2 shows the RMSEs of the individual-specific fits (P_i), cross-condition predictions based on individual-specific models (P_{Xi}) and cross-condition predictions based on group-average models ($\bar{P}_{Xi}$) for both the CSR and TSD conditions. As expected, for all subjects, RMSEs of P_i were smaller than those of P_{Xi}. However, the differences between them were not significant (for both conditions, mean difference over the 15 subjects = 4 ms, SD = 4 ms), implying that the individual-specific cross-condition model predictions were as good as the fits. In fact, on average, P_{Xi} yielded only 14 and 9% higher RMSEs than P_i under CSR and TSD conditions, respectively. In contrast, RMSEs of both P_i and P_{Xi} were significantly lower than those of $\bar{P}_{Xi}$ [$P < 0.05$, Wilcoxon's paired, two-sided, signed-rank test (Zar, 1999)], with P_{Xi} yielding, on average, 29 and 35% lower RMSEs than $\bar{P}_{Xi}$ under CSR and TSD conditions, respectively. Statistical comparisons of RMSEs of P_i and P_{Xi} with RMSEs of $\bar{P}_i$ (not shown) were no different from comparisons with RMSEs of $\bar{P}_{Xi}$.

Applicability of the UMP to other PVT statistics

To determine whether the results were affected by the choice of PVT statistic, we repeated the same procedure of developing individual-specific models and group-average models and comparing their corresponding RMSEs using four additional, frequently used PVT statistics: (i) median RT, (ii) slowest 10% RT, (iii) speed and (iv) lapses. Table 3 shows that, for each statistic, *U* and κ, the two most sensitive parameters, showed substantial to almost-perfect agreement across the two sleep-loss conditions (ICC > 0.60). In fact, the average ICC over *U* and κ was >0.75 for each statistic. The state parameters (S_0, L_0 and ϕ) varied from slight to moderate agreement across the two conditions for all statistics.

Table 4 lists the mean ($n = 15$) RMSEs of the individual-specific fits (P_i), cross-condition predictions based on individual-specific models (P_{Xi}) and cross-condition predictions based on group-average models ($\bar{P}_{Xi}$) obtained for the CSR and TSD conditions for each of the five PVT statistics. Also listed (within parentheses) are the percentage differences in the mean RMSEs of P_{Xi} compared to P_i and $\bar{P}_{Xi}$. Across all PVT statistics under the CSR condition, P_{Xi} yielded up to 17% higher RMSEs than P_i; under the TSD condition, P_{Xi} yielded up to 32% higher RMSEs than P_i. Under both CSR and TSD conditions, P_{Xi} yielded up to ~50% lower RMSEs than $\bar{P}_{Xi}$.

DISCUSSION

Individuals display a trait-like response to sleep loss. However, whether their unique response to sleep loss can be quantified via a mathematical model was not known. In this work, we used the previously developed UMP (Rajdev *et al.*, 2013) to characterize an individual's response to sleep loss and showed that the model parameters were preserved across two different sleep-loss conditions: (i) 64 h of TSD and (ii) 7 days of CSR of 3 h nightly TIB. That is, the model developed on the temporal dynamics of an individual's

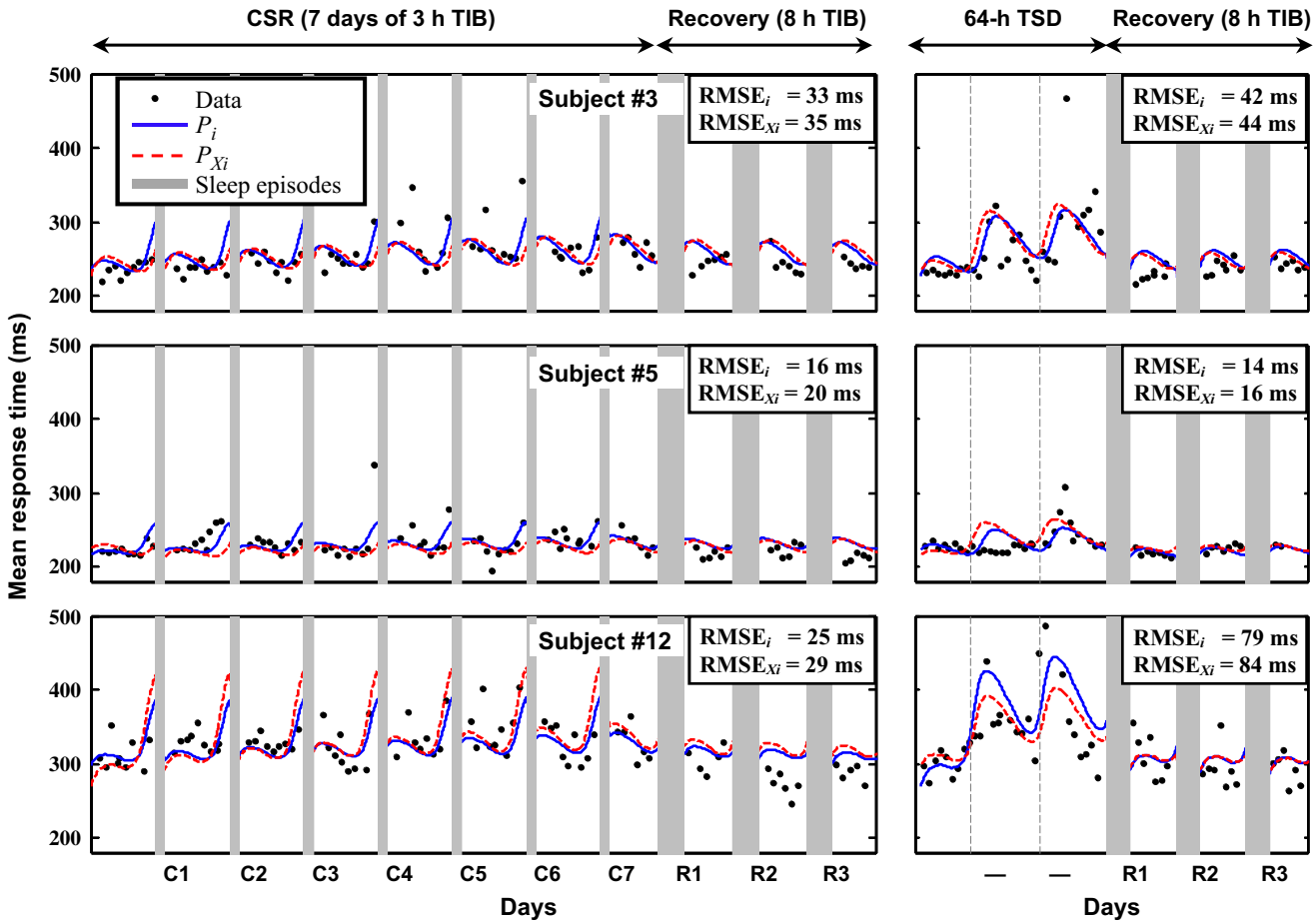

Figure 1. Individual-specific fits (P_i) and cross-condition predictions (P_{Xi}) of psychomotor vigilance task performance [mean response time (RT)] using the unified mathematical model of performance (UMP) for three different subjects challenged to two sleep-loss conditions: (i) chronic sleep restriction (CSR) for 7 days of 3 h nightly time in bed (TIB) and (ii) total sleep deprivation (TSD) for 64 h. Each condition was followed by 3 recovery days of 8 h nightly TIB. Mean RT data measured during wakefulness are represented by dots. The solid blue lines in each panel represent the individual-specific fits for that sleep-loss condition. The dashed red lines for the CSR condition represent the individual-specific cross-condition predictions based on models obtained by fitting the TSD data, and vice versa. The shaded regions represent the sleep episodes. RMSE: root mean squared error (ms).

performance under one sleep-loss condition predicted accurately the same individual's performance under the other condition. We also showed that these results were, to a large extent, independent of the choice of PVT statistic used to quantify performance impairment.

We intended to investigate four research questions. The first assessed the extent to which the individual-specific model parameters are preserved across the TSD and CSR conditions. We used the ICCs of the parameters to quantify the extent of parameter preservation. The two most sensitive trait parameters, U and κ (which denote the upper asymptote of the homeostatic process and the circadian amplitude, respectively), showed substantial to almost-perfect agreement (ICC > 0.60) for all PVT statistics (Table 3). Because U and κ are the two key parameters that represent an individual's vulnerability to sleep loss, this finding indicates that, for each individual, the UMP trait-like parameters are preserved across the two conditions studied here. Two state parameters, S_0 and ϕ, also showed fair to moderate agreement, probably because of the 7 nights of sleep satiation (10 h TIB) that preceded each of the sleep-loss challenges, which would have brought subjects to the same initial homeostatic sleep pressure states while synchronizing their circadian phases.

The second and third research questions investigated the extent to which the individual-specific models predicted PVT performance across the two conditions and assessed the improvement of these predictions over group-average model predictions. To answer these questions, we calculated the RMSEs between each individual's performance data and the corresponding model fits and predictions (Fig. 2 and Table 4). For each condition and across all PVT performance statistics, the individual-specific cross-condition model predictions (P_{Xi}) yielded 9–32% higher RMSEs than the corresponding fits (P_i). However, they yielded up to 50% lower RMSEs than the corresponding cross-condition group-average model predictions ($\bar{P}_{Xi}$). These findings suggest that the extent of parameter

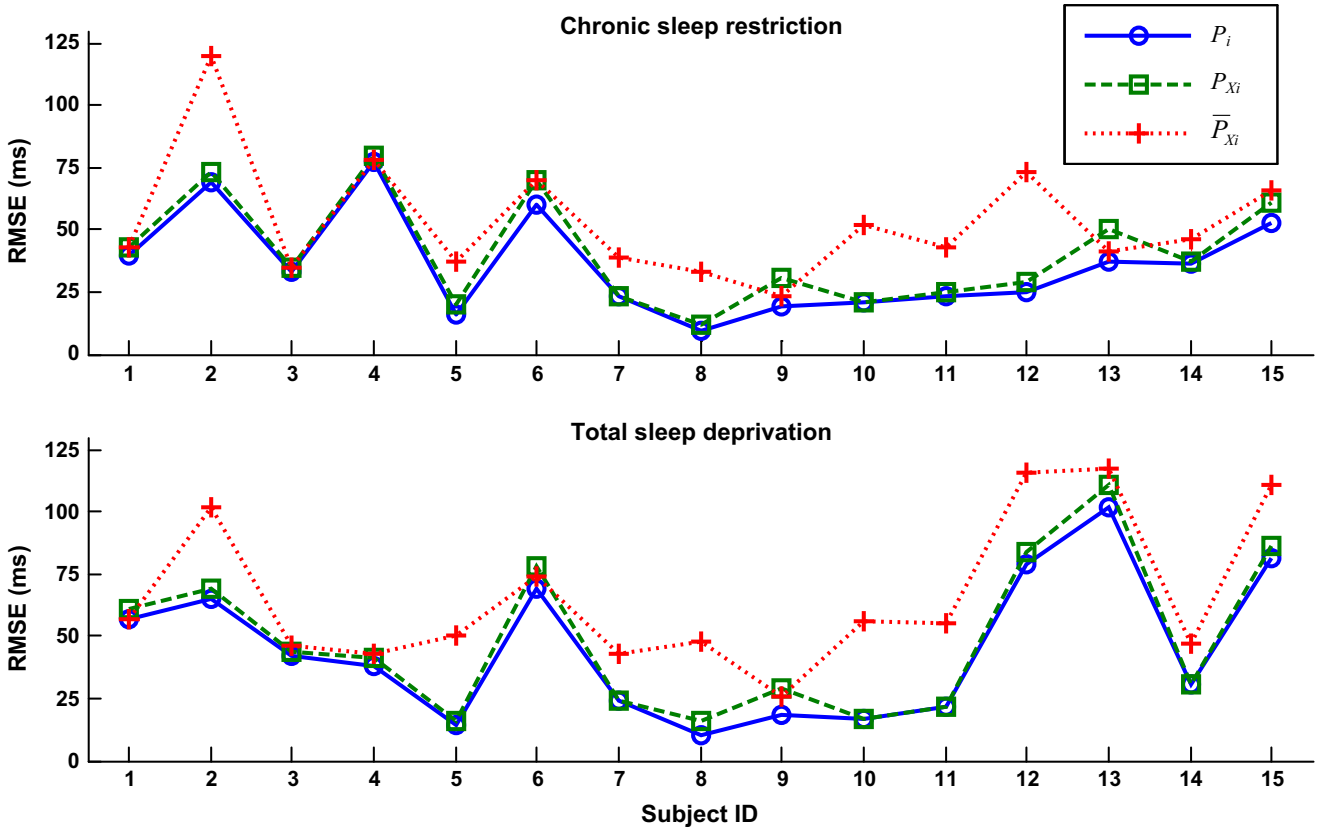

Figure 2. Root mean squared errors (RMSEs) of model fits (P_i), cross-condition individual-specific model predictions (P_{Xi}) and cross-condition group-average model predictions ($\bar{P}_{Xi}$) of psychomotor vigilance task performance (based on mean response time statistic) during chronic sleep restriction and total sleep deprivation for each of the 15 subjects. The lines joining the RMSEs across the subjects are provided as only a visual aid to differentiate between the RMSEs of P_i, P_{Xi} and $\bar{P}_{Xi}$.

Table 3 Intraclass correlation coefficients of the UMP parameters obtained in the chronic sleep restriction and total sleep deprivation conditions for five different PVT statistics

	Trait parameters		*State parameters*		
PVT statistic	U	κ	S_0	L_0	ϕ
Mean RT	0.90	0.89	0.52	0.10	0.46
Median RT	0.88	0.71	0.23	0.10	0.10
Slowest 10% RT	0.86	0.67	0.40	0.52	0.59
Speed	0.90	0.78	0.50	0.20	0.27
Lapses	0.84	0.86	0.40	0.37	0.31

L_0: lower homeostatic asymptote at time zero; PVT: psychomotor vigilance task; RT: response time; S_0: homeostatic state at time zero; U: upper asymptote of the homeostatic process; UMP: unified mathematical model of performance; ϕ: circadian phase; κ: circadian amplitude.

preservation across the two conditions is sufficient to advocate the use of individual-specific models over group-average models for predicting individuals' performance under either of the two sleep-loss conditions.

A key implication of this finding is that the UMP captures an individual's unique trait-like response to sleep loss. In other words, a model developed for an individual on one sleep-loss condition can be used to predict performance of the same individual under another sleep-loss condition. However, whether developing a model using performance data from one particular condition is better than another is not clear, and this formed our fourth research question. Results from our analyses suggested that the answer may depend on the PVT statistic used to quantify performance. For example, for median RT, speed and lapses, the individual-specific cross-condition predictions P_{Xi} were better (i.e. yielded smaller percentage differences in RMSEs relative to the model fits P_i) when developed on TSD data and tested under CSR conditions (Table 4). However, for the mean RT and slowest 10% RT, they were better when developed on CSR data. From a practical standpoint, developing models under TSD conditions is less time-consuming and burdensome; however, CSR conditions are more prevalent and realistic in operational environments.

We found that individual-specific models developed using 64 h of TSD data predicted performance accurately under CSR of 7 days of 3 h nightly TIB, and vice versa. However, whether the models predict individual performance accu-

Table 4 Mean (n = 15) RMSEs of model fits (P_i), cross-condition individual-specific model predictions (P_{Xi}) and cross-condition group-average model predictions ($\bar{P}_{Xi}$) of performance quantified using five different PVT statistics during chronic sleep restriction and total sleep deprivation; also indicated within parentheses are percentage differences in the mean RMSEs of P_{Xi} compared to P_i and $\bar{P}_{Xi}$

	Mean (n = 15) RMSE					
	Chronic Sleep Restriction			*Total Sleep Deprivation*		
PVT Statistic	P_i	P_{Xi}	$\bar{P}_{Xi}$	P_i	P_{Xi}	$\bar{P}_{Xi}$
Mean RT (ms)	36	41	53	45	49	66
	(14%)	(29%)		(9%)	(35%)	
Median RT (ms)	23	27	36	19	25	36
	(17%)	(33%)		(32%)	(44%)	
Slowest 10% RT (ms)	154	177	210	215	235	293
	(15%)	(19%)		(9%)	(25%)	
Speed (s^{-1})	0.34	0.39	0.58	0.32	0.40	0.60
	(15%)	(49%)		(25%)	(50%)	
Lapses (lapses)	3.41	3.77	4.34	2.76	3.33	4.06
	(11%)	(15%)		(21%)	(22%)	

PVT, psychomotor vigilance task; RMSE, root mean squared error; RT, response time

rately under less severe CSR conditions, such as 6 h nightly TIB, is not known. Further cross-over design studies, spanning from mild-to-severe CSR to acute TSD, are thus required to validate the model's suggested ability to capture an individual's trait-like response to sleep loss. Also, because the UMP was developed based on PVT performance, the extent to which its predictions generalize to other aspects of neurobehavioural performance is not known. For example, both Rupp *et al.* (2012) and Van Dongen *et al.* (2004) found that an individual's relative rank on the PVT was not the same as that individual's relative rank on other neurocognitive tasks (e.g. mathematical processing, running memory, and visual analogue scale of fatigue). However, the PVT is used more widely because it has been shown to be more sensitive to sleep loss than other neurobehavioural metrics (Balkin *et al.*, 2004). PVT-based model predictions could therefore serve only as indicators of the likelihood of near-future deficits in other aspects of neurobehavioural performance. Lastly, many prior studies have suggested that differences in the response to sleep loss are linked to genes that regulate the circadian and homeostatic processes (Landolt, 2008; Lo *et al.*, 2012; Rupp *et al.*, 2013). However, whether the UMP trait-like parameters U and κ are associated with an individual's genetic factors is not known.

In this work, we applied the UMP model on PVT performance data for individuals who underwent total and partial sleep loss under laboratory conditions. However, the model may also be applicable to individuals in operational environments, using available computational platforms. This would require information about the individual's sleep-wake history and performance data, which together would serve as inputs to the UMP model running on a computational platform [tablet, personal computer (PC) or smartphone] to customize the model and make predictions. For example, an individual's sleep-wake history could be inferred continuously via wrist-worn actigraphy and streamed to a computer. Similarly, brief, periodic PVT performance tests (Basner *et al.*, 2011) could be performed on a computer running the UMP model (Khitrov *et al.*, 2014).

In summary, this work has two key findings: (i) the UMP can explain the temporal dynamics of PVT performance under both TSD and CSR conditions at an individual level significantly better than a group-average model and (ii) the UMP parameters are preserved across TSD and CSR conditions, i.e. the model captures an individual's trait-like response to sleep loss, thereby facilitating accurate individual-specific performance predictions under other sleep-loss conditions. These two findings provide further evidence to support the mathematical formulation of our unified model. While challenges remain, these findings bring us a step closer to our long-term goal of incorporating models of fatigue due to sleep loss and models of fatigue countermeasures [e.g. models of the restorative power of caffeine (Ramakrishnan *et al.*, 2013, 2014)] into a computational tool to optimize neurobehavioural performance at an individual level.

ACKNOWLEDGEMENTS

This work was sponsored by the Military Operational Medicine Research Area Directorate of the US Army Medical Research and Materiel Command, Fort Detrick, MD, and by the US Department of Defense Medical Research and Development Program (grant no. DMRDP_13200). 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 US Army or of the US Department of Defense. This paper has been approved for public release with unlimited distribution.

CONFLICT OF INTERESTS

This was not an industry-supported study. All authors have reported no financial conflicts of interest.

AUTHOR CONTRIBUTIONS

SR, WL, SL and JR conceived research; SR and WL implemented the model and performed the computations; NJW, TLR and TJB provided data for modelling; SR and NJW wrote the paper, which was edited by JR.

REFERENCES

Balkin, T. J., Bliese, P. D., Belenky, G. *et al.* Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. *J. Sleep Res.*, 2004, 13: 219–227.

Basner, M., Mollicone, D. and Dinges, D. F. Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation. *Acta Astronaut.*, 2011, 69: 949–959.

Khitrov, M. Y., Laxminarayan, S., Thorsley, D. *et al.* PC-PVT: a platform for psychomotor vigilance task testing, analysis, and prediction. *Behav. Res. Methods*, 2014, 46: 140–147.

Landis, J. R. and Koch, G. G. The measurement of observer agreement for categorical data. *Biometrics*, 1977, 33: 159–174.

Landolt, H. P. Genotype-dependent differences in sleep, vigilance, and response to stimulants. *Curr. Pharm. Des.*, 2008, 14: 3396–3407.

Lo, J. C., Groeger, J. A., Santhi, N. *et al.* Effects of partial and acute total sleep deprivation on performance across cognitive domains, individuals and circadian phase. *PLoS One*, 2012, 7: e45987.

Mallis, M. M., Mejdal, S., Nguyen, T. T. and Dinges, D. F. Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviat. Space Environ. Med.*, 2004, 75: A4–A14.

Nakagawa, S. and Schielzeth, H. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol. Rev. Camb. Philos. Soc.*, 2010, 85: 935–956.

Rajaraman, S., Gribok, A. V., Wesensten, N. J., Balkin, T. J. and Reifman, J. An improved methodology for individualized performance prediction of sleep-deprived individuals with the two-process model. *Sleep*, 2009, 32: 1377–1392.

Rajdev, P., Thorsley, D., Rajaraman, S. *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.

Ramakrishnan, S., Rajaraman, S., Laxminarayan, S. *et al.* A biomathematical model of the restoring effects of caffeine on cognitive performance during sleep deprivation. *J. Theor. Biol.*, 2013, 319: 23–33.

Ramakrishnan, S., Laxminarayan, S., Wesensten, N. J., Kamimori, G. H., Balkin, T. J. and Reifman, J. Dose-dependent model of caffeine effects on human vigilance during total sleep deprivation. *J. Theor. Biol.*, 2014, 358C: 11–24.

Rupp, T. L., Wesensten, N. J. and Balkin, T. J. Trait-like vulnerability to total and partial sleep loss. *Sleep*, 2012, 35: 1163–1172.

Rupp, T. L., Wesensten, N. J., Newman, R. and Balkin, T. J. PER3 and ADORA2A polymorphisms impact neurobehavioral performance during sleep restriction. *J. Sleep Res.*, 2013, 22: 160–165.

Rusterholz, T., Durr, R. and Achermann, P. Inter-individual differences in the dynamics of sleep homeostasis. *Sleep*, 2010, 33: 491–498.

Van Dongen, H. P. Comparison of mathematical model predictions to experimental data of fatigue and performance. *Aviat. Space Environ. Med.*, 2004, 75: A15–A36.

Van Dongen, H. P., Baynard, M. D., Maislin, G. and Dinges, D. F. Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep*, 2004, 27: 423–433.

Zar, J. H. *Biostatistical Analysis*. Prentice Hall, Upper Saddle River, NJ, 1999.

Zhang, D., Lin, X., Raz, J. and Sowers, M. Semiparametric stochastic mixed models for longitudinal data. *J. Am. Stat. Assoc.*, 1998, 93: 710–719.

Zhou, D. and Shi, T. Statistical inference based on distances between empirical distributions with applications to AIRS level 3 data. In: *Proceedings of the 2011 NASA Conference on Intelligent Data Understanding*. Mountain View, CA, 2011: 129–143.