

Individualized performance prediction of sleep-deprived individuals with the two-process model

Srinivasan Rajaraman,¹ Andrei V. Gribok,¹ Nancy J. Wesensten,² Thomas J. Balkin,² and Jaques Reifman¹

¹Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick; and ²Department of Behavioral Biology, Division of Psychiatry and Neuroscience, Walter Reed Army Institute of Research, Silver Spring, Maryland

Submitted 15 August 2007; accepted in final form 7 December 2007

Rajaraman S, Gribok AV, Wesensten NJ, Balkin TJ, Reifman J. Individualized performance prediction of sleep-deprived individuals with the two-process model. *J Appl Physiol* 104: 459–468, 2008. First published December 13, 2007; doi:10.1152/jappphysiol.00877.2007.—We present a new method for developing individualized biomathematical models that predict performance impairment for individuals restricted to total sleep loss. The underlying formulation is based on the two-process model of sleep regulation, which has been extensively used to develop group-average models. However, in the proposed method, the parameters of the two-process model are systematically adjusted to account for an individual's uncertain initial state and unknown trait characteristics, resulting in individual-specific performance prediction models. The method establishes the initial estimates of the model parameters using a set of past performance observations, after which the parameters are adjusted as each new observation becomes available. Moreover, by transforming the nonlinear optimization problem of finding the best estimates of the two-process model parameters into a set of linear optimization problems, the proposed method yields unique parameter estimates. Two distinct data sets are used to evaluate the proposed method. Results of simulated data (with superimposed noise) show that the model parameters asymptotically converge to their true values and the model prediction accuracy improves as the number of performance observations increases and the amount of noise in the data decreases. Results of a laboratory study (82 h of total sleep loss), for three sleep-loss phenotypes, suggest that individualized models are consistently more accurate than group-average models, yielding as much as a threefold reduction in prediction errors. In addition, we show that the two-process model of sleep regulation is capable of representing performance data only when the proposed individualized model is used.

total sleep deprivation; individualized modeling; performance prediction; parameter estimation; regularization

A PROMISING STRATEGY to better manage the detrimental effects of limited sleep on cognitive performance of civilian and military personnel involves biomathematical modeling. Biomathematical models can be used to estimate the performance of sleep-deprived individuals, to plan improved sleep/wake schedules, and to optimize the timing and dosing of fatigue countermeasures, e.g., naps and caffeine, so that performance peaks and is maintained during the desired times of the day.¹

To date, most mathematical models are limited to predicting group-average performance (18, 21, 26). Group-average mod-

eling, however, is incongruous with our knowledge that performance impairment due to total or partial sleep deprivation is strongly dependent on innate differences among individuals (25, 35, 39, 41). For example, on a psychomotor vigilance test (PVT) scale (19), interindividual differences accounted for nearly 70% of the total variance in the performance among individuals (35). What is needed, in particular for operational settings, is the ability to predict performance of specific individuals (26). However, prominent among the obstacles associated with individualizing model predictions are the interrelated problems of 1) how to deal with uncertainties in the initial states, i.e., initial homeostat and circadian values, of the individuals being modeled; and 2) how to determine the model parameters that would accurately reflect the substantial interindividual differences in performance impairment due to sleep loss (32).

Another obstacle in the development of individualized models relates to our inability to identify biomarkers capable of quantitatively predicting performance decrement due to sleep loss (3, 37). Given the absence of such independent predictor variables of performance, the next-best option for developing individual-specific models is to feed back past measures of observed performance $P(k)$, $k = k-1, k-2, \dots$, of an individual as inputs to the model to predict that individual's future performance $P(k)$, at time index k . For example, one could use past PVT measures to develop models capable of estimating future PVT values. As measures of performance from a specific individual become available, such a mechanism would allow the model parameters, which are assumed to be time invariant and fixed for a given individual,² to be systematically adapted to that individual, based on previous temporal patterns of performance, to more accurately predict future performance levels for that individual.

Presently, only two methods have been proposed for developing individualized performance prediction models (29, 36). Both methods used the concept of feeding current and previous measures of performance of a specific individual back into the model to predict that individual's future performance. The original work proposed by Olofsen and colleagues (29) assumes that a simple exponential model can be used to represent

² Of all the parameters, the trait parameters are innate and fixed for a given individual. The state parameters of a given individual, however, depend on the prior sleep/wake history and are, therefore, variable for different scenarios.

¹ The online version of this article contains supplemental material.
Address for reprint requests and other correspondence: J. Reifman, Senior Research Scientist, Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, MCMR-ZB-T, 363 Miller Dr., Fort Detrick, MD 21702 (e-mail: jaques.reifman@us.army.mil).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

daily-averaged performance measures of a group of individuals restricted to partial sleep deprivation over a 20-day period. Casting the exponential model in a mixed-effects regression framework allows the model to fit the group data so that the fixed effects (i.e., constant or time-varying parameters) are decoupled from the random effects (i.e., the inter- and intraindividual variability), where the group-average model parameters and their corresponding variances are estimated through the maximum likelihood principle. As new performance measures for an individual, not necessarily studied before, become available, a Bayesian approach makes use of the learned interindividual variability of the group and the individual's performance data to adapt the group-average model parameters and the model to that individual. More recently, Van Dongen and colleagues (36) extended this work by substituting the exponential model with Borbély's two-process model of sleep regulation (8, 14) and developed models for predicting (up to 24 h ahead) performance impairment levels of individuals in an 88-h total deprivation study.

Notwithstanding the significant contribution of these two methods that, for the first time, allowed for the development of individual-specific models of performance, they have some shortcomings. First, unique estimates of the two-process model parameters for a specific individual cannot be guaranteed because the Bayesian posterior distribution could be multimodal. The possible multimodal behavior of the posterior distribution could be due to the multimodal nature of the prior probability distribution of the group-average parameters or the multimodal likelihood function of an individual's trait parameters [both of which may occur because of nonlinear relationships of the 2-process model parameters with respect to the model's output $P(k)$]. Second, there is a high computational cost to maximize the Bayesian posterior distribution through a brute-force, grid-search approach, and the cost increases exponentially as the desired accuracy of the resulting parameter estimates increases. The multimodality of the posterior distribution, in particular, has serious implications. It may negate a key convergence property of adaptive (or learning) algorithms, in that the model parameters must converge asymptotically to their true values as the number of observations increases and the amount of noise in the data decreases (12).

In this study, we present an alternative method to develop individualized predictive models of performance for subjects restricted to total sleep deprivation. Similar to the method proposed by Van Dongen and colleagues (36), the five-parameter, two-process model of sleep regulation is used as the mathematical basis to predict an individual's temporal pattern of performance, and past measures of performance $P(k)$, $k = k-1, k-2, \dots$, are fed back as inputs to the model to estimate future performance $P(k)$. Unlike their approach, however, for predicting an individual's performance, the proposed method does not require the availability of performance data from a group of (like) individuals and the estimation of group-average model parameters. Instead, an individual's temporal variability is learned directly and solely from the individual's own data. A set of past performance data points are used to predict future levels of performance for that individual. Any new datum point that becomes available is combined with the individual's previous data (in a "batch" approach) and is used to adapt the model parameters and generate estimates of future performance, without requiring knowledge of an individual's initial

state, i.e., initial homeostat and circadian phase. In case a new datum point is either missing or unavailable at the "expected" sampling period, no adaptation of the parameters takes place, and the most recent parameter values are used to estimate future performance levels.

More importantly, for a given set of performance data, this approach guarantees a unique estimate of the five parameters of the two-process model and ensures that the model parameters asymptotically converge to their true values as the number of measurements increases and the amount of noise in the observed measurements reduces to zero. These are accomplished by transforming the underlying nonlinear optimization problem of finding the best estimate of $P(k)$, the solution of which may not be unique, into a set of linear optimization problems that are guaranteed to yield a unique solution.

METHODS

The two-process model of sleep regulation. In the two-process model of sleep regulation, two separate processes underlie sleep regulation (8). The first, Process S (Sleep homeostasis), which is dependent on prior sleep/wake history, increases exponentially with wake time and decreases exponentially with sleep/recovery time to a basal value (14). The rate of sleep pressure buildup and withdrawal is individual specific, constant, and has unknown values. The second, Process C (Circadian process), is a 24-h periodic, self-sustaining oscillator, independent of sleep/wake history (15, 17) that initiates in the suprachiasmatic nucleus of the hypothalamus. The equations comprising the two-process model at discrete sampling time index k can be expressed as (9, 14)

$$S(k) = 1 - \exp(-T_s/\tau_r)[1 - S(k-1)] \quad \text{during wakefulness,} \quad (1)$$

$$S(k) = \exp(-T_s/\tau_d)[S(k-1)] \quad \text{during sleep,} \quad (2)$$

$$C(k) = \sum_{i=1}^5 a_i \sin\left\{\frac{i2\pi}{\tau} [(k-1)T_s + \phi]\right\}, \quad (3)$$

where $S(k)$ denotes the value of the sleep homeostat, usually scaled between 0 and 1 (9); $C(k)$ denotes a five-harmonic sinusoidal equation, which can adequately approximate the circadian pacemaker under entrained conditions of performance and alertness (24); T_s represents the sampling period; τ_r and τ_d represent the time constants of Process S during wakefulness and sleep, respectively; τ denotes the time period of the circadian oscillator (~ 24 h); a_i , $i = 1, \dots, 5$, represents the amplitude of the five harmonics of the circadian process ($a_1 = 0.97$, $a_2 = 0.22$, $a_3 = 0.07$, $a_4 = 0.03$, and $a_5 = 0.001$); and ϕ denotes the initial circadian phase (1).

Performance modeling using the two-process model. The temporal pattern of cognitive performance is represented by the additive interaction of the homeostatic pressure for sleep and the circadian rhythm input (2). Therefore, performance $P(k)$, at time k , can be described by

$$P(k) = \alpha S(k) + \beta C(k), \quad (4)$$

where α and β are parameters that control the relative effect of the two processes on performance. In this study, we do not model the nonlinear interaction between Processes S and C, which has been shown to exist (13, 16, 23). Also, we choose to keep the shape of Process C identical among the individuals, with the coefficients of the sine functions held constant over time. However, our method can account for interindividual differences in the circadian process not only by modifying these coefficients proportionally but also independently. For total sleep deprivation, Eq. 4 can be rewritten as

$$P(k) = \alpha - \alpha S(0)\gamma^{k-1} + \sum_{i=1}^5 \bar{a}_i \sin\{i\omega[(k-1)T_s + \phi]\}, \quad (5)$$

where $\gamma = \exp(-\rho T_s)$, $\rho = 1/\tau_r$, $\omega = 2\pi/\tau$, and $\bar{a}_i = \beta a_i$. Equation 5 is a function of five unknown parameters [α , γ , β , $S(0)$, and ϕ] of

which α , γ , and β have been termed trait parameters, since they are innate individual characteristics, and $S(0)$ and ϕ have been termed state parameters, since they depend on prior sleep/wake history and environmental conditions, such as light exposure levels and circadian phase shift due to transmeridian traveling (36).

Performance $P(k)$ in Eq. 5 generates output values in a range that can be mapped (usually through linear mapping), by adjusting α and β , to predict subjective sleepiness and objective performance measures, such as the Karolinska Sleepiness Scale (4) and PVT lapses, respectively. Here, we map $P(k)$ to PVT lapses because this is the objective performance measure recorded in the laboratory study whose data are used in this work. Although PVT lapse is recognized as the most sensitive measure to sleep loss and, arguably, the most practical one for use in operational environments (5), it should be noted that PVT may be incapable of reflecting the true performance of individuals subjected to complex tasks experienced in real-world applications.

Subject-specific model development. The two-process model of sleep regulation is adopted as the underlying modeling framework for developing an individualized performance model. To date, no physiological predictor of sleep-related performance impairment has been identified. Therefore, information available from current and previous measurements of performance $P(k)$ from a specific individual is fed back into the model to predict that individual's future performance. Through this mechanism, with each new performance observation the five unknown model parameters [α , γ , β , $S(0)$, and ϕ] are simultaneously adjusted to account for the most recent information on the current state of the individual.

Regardless of the method used to adjust the performance model parameters in Eq. 5 for a specific individual, its solution involves the minimization of the errors between the performance model fit and a set of performance measurements from that specific individual. Because of the nonlinearity of the performance model with respect to its parameters, this minimization requires the solution of a nonlinear optimization problem. Such a problem, however, may have multiple solutions because of the possible nonconvexity of the objective function (33). Therefore, if the initial guess of the parameter estimates is not chosen in the region of attraction of the global minimum, traditional gradient-based methods, such as Levenberg-Marquardt or Gauss-Newton, may fail to converge to the optimum solution.

We propose a new approach for developing individualized performance models that overcomes this limitation. The central idea is to recast the otherwise nonlinear optimization problem into a set of linear optimization problems, whose solutions, if they exist, are guaranteed to be unique (7). This approach is motivated by the fact that performance $P(k)$ in Eq. 5 is composed of a linear combination of three basic components—a constant signal α ; an exponential signal

$\alpha S(0)\gamma^{k-1}$; and five sinusoidal signals $\bar{a}_i \sin\{i\omega[(k-1)T_s + \phi]\}$, $i = 1, \dots, 5$ —that can be decoupled from each other to form a set of three linear equations on $P(k)$. Additionally, each set of equations can be solved to uniquely obtain one or more of the five parameters of $P(k)$. The three basic components are decoupled from each other by separately applying a set of three first- and higher-order linear operators on $P(k)$, as follows.

Let Z denote a shift operator, such that $Z^n\{x(k)\} = x(k+n)$, where $x(k)$ denotes the value of the time series x at time k , with k and n being positive integers. Equivalent to differential operators in a continuous-time formulation, this notation has been widely used to compactly represent many difference operations in discrete-time formulation of time-series signals. For example, a first-order forward difference operation on $x(k)$, i.e., $x(k+1) - x(k)$, can be represented as $Z^1\{x(k)\} - Z^0\{x(k)\}$, which is equivalent to $(Z-1)\{x(k)\}$. To take this example further, the discrete approximation of the second-order derivative of $x(k)$, i.e., $x(k+2) - 2x(k+1) + x(k)$, can be represented as $Z^2\{x(k)\} - 2Z^1\{x(k)\} + Z^0\{x(k)\}$, or equivalently as $(Z-1)^2\{x(k)\}$. Therefore, a difference operation on $x(k)$ can be cast as an algebraic expression involving Z . Besides, the difference operators are commutative such that the order of operations does not affect the results. For example, the result of $(Z-2)(Z-1)\{x(k)\}$ is the same as $(Z-1)(Z-2)\{x(k)\}$.

Let us now define the following three operators

- I. $Z - 1$
- II. $Z^2 - 2 \cos(i\omega T_s)Z + 1$, $i = 1, 2, \dots, 5$
- III. $Z - \gamma$

so that, when individually applied to performance $P(k)$ in Eq. 5, each operator eliminates one of the three basic components. For example, by applying operator I to $P(k)$, we eliminate the constant signal α , since $(Z-1)\{\alpha\} = Z\{\alpha\} - \alpha = 0$. Similarly, operator II eliminates the sinusoidal signal $\sum_{i=1}^5 \bar{a}_i \sin\{i\omega[(k-1)T_s + \phi]\}$, and operator III eliminates the exponential signal $\alpha S(0)\gamma^{k-1}$ (see Supplement, available with the online version of this article). Hence, by successively applying a pair of different operators to $P(k)$ in Eq. 5, we eliminate two of the three basic components and end up with a linear equation on the parameters of the remaining signal that can be uniquely estimated.

Identification of parameters γ and $\alpha S(0)$ of the exponential signal. To uniquely identify parameters γ and $\alpha S(0)$ of the exponential signal (i.e., the homeostatic Process S), we eliminate the constant and sinusoidal components by successively applying operators I and II, respectively, to $P(k)$ in Eq. 5. The resulting exponential signal (Fig. 1) is characterized solely by $\alpha S(0)$ and γ , which can be estimated by solving an associated linear optimization problem, the solution of

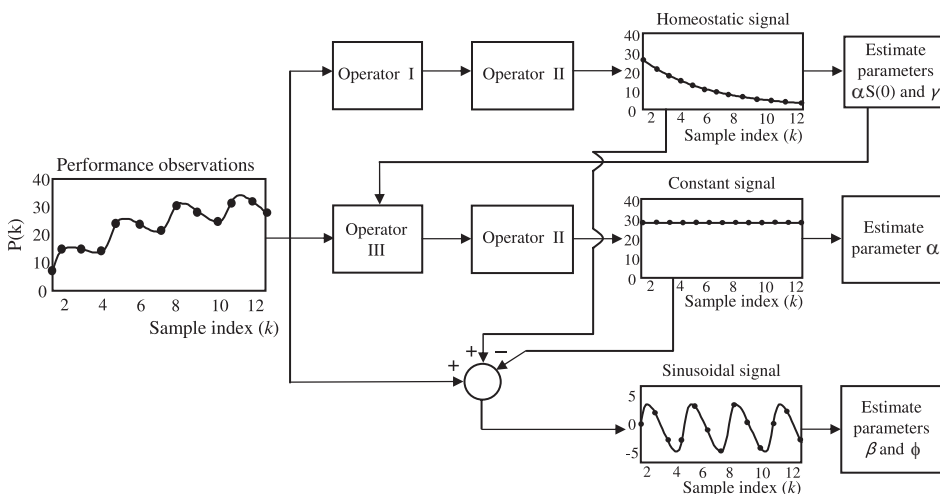


Fig. 1. Schematic diagram of the proposed method to uniquely estimate the 5 parameters of the 2-process model for a specific individual based on previous performance observations from that individual.

which is unique (see Supplement). Because *operators I* and *II* together represent an 11th-order polynomial in Z and two estimates of the exponential signal are required to uniquely obtain $\alpha S(0)$ and γ , this formulation requires at least 13 observations of $P(k)$ (see Supplement).

Identification of parameter α of the constant signal. To uniquely identify the parameter α of the constant signal, we eliminate the circadian and exponential components by successively applying *operators II* and *III*, respectively, to $P(k)$, where γ in *operator III* is obtained as described above. The resulting constant signal (Fig. 1) is characterized solely by α , which can be estimated by solving an associated linear optimization problem (see Supplement). Similar to *operators I* and *II*, *operators II* and *III* together represent an 11th-order polynomial in Z , but here one estimate of the constant signal requires only 12 observations of $P(k)$ to uniquely estimate α .

Identification of parameters β and ϕ of the sinusoidal signal. To uniquely estimate parameters β and ϕ of the sinusoidal signal, we could repeat the abovementioned procedure and apply *operators I* and *III* to $P(k)$. Another approach, resulting in a simpler solution, is to exploit the fact that performance $P(k)$ in Eq. 5 is composed of a linear combination of three basic components. Hence, once parameters of the exponential and constant components are computed, as discussed above, these components can be subtracted from $P(k)$ in Eq. 5 to yield an expression for the five harmonics of the sinusoidal component (see Supplement). The resulting nonlinear expression, which is a function of the two unknowns β and ϕ , is reformulated into a set of linear equations that can be uniquely solved based on at least 10 observations of $P(k)$, allowing for the unique estimation of β and ϕ (see Supplement).

Through this method, the five parameters of the two-process model in Eq. 5 can be uniquely estimated from at least 13 observations of $P(k)$ for the individual being modeled. In practice, however, due to the noisy nature of performance data, a larger number of observations is used to compute the first estimates of the parameters. Thereafter, as each new performance datum becomes available, it is added to the previous observations and together used to adapt the model parameters and generate future predictions.

The application of *operators I*, *II*, and *III* on observations $P(k)$ requires the computation of difference operations (i.e., first- and higher-order derivatives) on these observations. The calculation of these derivatives of $P(k)$ can be formulated as the solution of a least squares (LS) problem. For example, the first-order difference operation of an $N \times 1$ vector of measurements P with elements $P(k)$ can be solved by minimizing the functional $\|P - Uw\|_2^2$, where w denotes an $N \times 1$ vector of first-order differences and U denotes an $N \times N$ matrix representing the discrete analogue of integration (see Supplement) (22). It is well known that difference operations on noisy signals, such as PVT measurements, amplify the noise component and suppress the true underlying signal (20, 22, 34). The amplification of noise in the resulting difference operations leads to unreliable estimates of the underlying parameters of the constant, exponential, and sinusoidal signals, with the estimates having high variance.

To address this issue, we apply the regularization method to the LS problem so that a priori constraints, such as smoothness, are imposed on the elements of the solution w to limit noise amplifications (28, 30, 31, 34, 38). The regularized LS approach replaces the original LS functional, which we seek to minimize, with an augmented functional

$$\arg \min_w \|P - Uw\|_2^2 + \lambda^2 \|Lw\|_2^2, \quad (6)$$

where the regularization parameter λ is a real number that controls the trade-off between fitting the data P and satisfying an a priori constraint L , which is itself represented by a well-conditioned matrix, imposed on the nature of the solution (see Supplement). This aspect of regularization is similar to Bayesian inference, in which a priori knowledge is incorporated in the problem formulation to lend stability and precision to the solution (6, 10). The net effect of regularization

is a significant reduction on the variance of the solution, at the cost of introducing a small bias (see Supplement) (27).

RESULTS

We employ two data sets to test the proposed approach. The first consists of simulated performance data generated by running the two-process model in Eq. 5 with known parameters and superimposing random noise to emulate the variability observed in measured performance data. Simulated data permit demonstration of a key property of the proposed approach, i.e., the estimated model parameters asymptotically converge to their true (known) values 1) as the amount of noise in the data decreases and 2) as the number of performance data used to adapt the model parameters to a specific individual increases. The second data set consists of performance measures from a controlled laboratory study of nine healthy individuals restricted to 82 h of total sleep deprivation (40). These data allow testing of the approach within the context of between-subject and within-subject variability encountered in measured performance data and comparison of the predictive capabilities of individualized models with those otherwise provided by group-average models.

Simulated data set. The simulated data set is generated by running the two-process model in Eq. 5 with fixed (known) parameters and superimposing random noise to emulate the variability observed in performance data. The three trait parameters in Eq. 5, where $\alpha = 30.30$, $\beta = 6.35$, and $\rho = 0.03 \text{ h}^{-1}$ ($\gamma = 0.94$), are obtained by sampling the probability distribution of parameters estimated by Van Dongen and colleagues (36), while the state parameters, where $S(0) = 0.82$ and $\phi = 6.00 \text{ h}$, are randomly chosen. In each simulation, performance data are generated for 82 h of total sleep deprivation and sampled once every 2 h to duplicate the sampling rate of typical sleep studies. Data from each simulation are superimposed with random noise sampled from a normal distribution with zero mean and one of three (16, 4, or 1) levels of variance. These variances correspond to signal-to-noise ratios of 3.70 (59.10/16), 14.77 (59.10/4), and 59.10 (59.10/1), respectively, where 59.10 is the variance of the performance data signal simulated by running Eq. 5 with the abovementioned parameters for 82 h.

In each simulation, the first 42 h (or equivalently, the first 22 “training” data points) are used to obtain initial estimates of the five model parameters and to make the first prediction at some future time, in accordance with the desired prediction horizon. For example, for a 6-h prediction horizon, the first prediction, based on performance data up to 42 h, is made for performance at 48 h. Subsequently, each datum corresponding to a new performance observation is added to the existing data set one at a time, and the model parameters are adapted to take into account all information, including the latest performance datum, to make the next prediction. Following the example above, the performance datum “observed” at 44 h is added to the previous 22 data points to adapt the five model parameters and make a prediction at 50 h. This procedure emulates the temporal processes of data collection and prediction in potential field applications of individualized performance models and is used throughout this paper unless noted otherwise.

Figure 2 shows the simulated performance profile for a representative single individual with a signal-to-noise ratio of 14.77. The performance units are normalized to emulate results

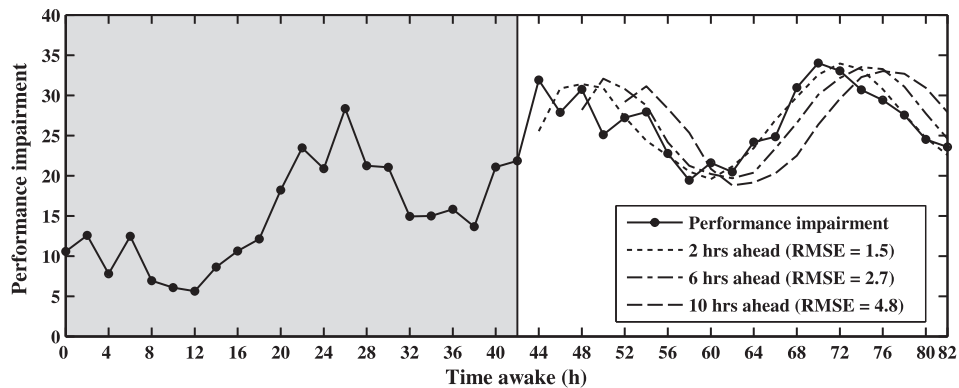


Fig. 2. Simulated performance data, normalized to a psychomotor vigilance test (PVT) scale, for 82 h of total sleep deprivation with a signal-to-noise ratio of 14.77. The solid circles show the simulated performance lapses every 2 h, whereas the dotted, dot-dashed, and dashed lines represent 2-, 6-, and 10-h-ahead predictions, respectively, obtained with individualized models. The shaded portion represents the amount of training data used to identify the initial model parameters after which predictions are made. Thereafter, the model parameters are adjusted every 2 h. The root mean square error (RMSE) between the predicted and simulated performance, calculated from 52 through 82 h, provides a metric of the predictions.

of a PVT scale. The shaded portion in Fig. 2 represents the amount of training data used to obtain the initial estimate of the parameters, after which predictions for the three selected prediction horizons of 2, 6, and 10 h began. The root mean square error (RMSE) is used as a metric to quantify the predictions. The smaller the RMSE between the predicted and the observed performance, the better is the resulting prediction. For consistency, the RMSEs are computed across overlapping prediction time intervals, for instance, between 52 and 82 h in this case. As expected, the RMSEs increase monotonically with increasing prediction horizons. This is because the model parameters are not updated over the prediction horizon. Unlike shorter prediction horizons, which employ intermediate observations to help the algorithm better learn the model parameters and yield improved predictions, longer horizons would not benefit from the use of intermediate observations. For the same reason, we also observe that the predictions are phase delayed with respect to the actual data and that such a delay increases with increasing prediction horizons. However, in the absence of noise in the observed data, if the algorithm is provided with sufficient data to fully learn the model parameters, the predictions would be approximately the same regardless of the prediction horizon.

Figure 3 shows the corresponding estimates for the five model parameters as a function of training points, starting at 42 h. The true value of the parameters is illustrated by the horizontal dashed lines. As expected, each one of the parameters asymptotically converges to its true value as the number of training data points increases, achieving 95% accuracy at the end of 82 h. Subsequent deviations of the parameter estimates from their true values can occur as a result of noise added to the simulated performance data.

To further illustrate the ability of the algorithm to consistently improve the estimation of the model parameters as a function of performance observations and reduced data noise, we perform three additional sets of simulations. Each set consists of 100 simulations, and each simulation in a set is superimposed with randomly sampled noise from a normal distribution with variances corresponding to signal-to-noise ratios of 59.10, 14.77, and 3.70. Each of the three sets of 100 simulations is superimposed with only one of these noise levels

so that the impact of each noise level on parameter estimation could be observed.

Table 1 shows the mean, the variance, and the mean squared error (MSE), i.e., the square of the bias plus the variance, for each of the five parameter estimates over the 100 simulations. For simulated data in which the bias in the estimate is known, the MSE, which accounts for both estimation accuracy (bias) and estimation precision (variance), is the most statistically meaningful metric. The parameter estimates are shown at two points in time, 62 and 82 h of wakefulness, for each of the three signal-to-noise ratios. For each of the five parameters, the results consistently indicate that, for a fixed time, the MSE increases as the signal-to-noise ratio decreases, and that, for a fixed signal-to-noise ratio, the MSE decreases with time.

Laboratory data set. The second data set used to test the proposed approach is based on data collected from a controlled laboratory study in which 48 healthy adults were kept awake for 85 consecutive hours to determine the effects of countermeasures on performance and alertness during prolonged sleep deprivation (40). During the entire wake period, PVT cognitive performance tests were measured every 2 h, starting after the first hour (0800) and ending 2 h before the end of the study, for a total of 42 PVT observations over 82 h. Eleven of these 48 individuals were administered placebo, while the others took stimulants, and data from 2 of those 11 contained clear outliers. Hence, the data for this set were from the remaining nine sleep-deprived individuals who had not been administered active fatigue countermeasures.

The laboratory study used in this work was approved by the Walter Reed Army Institute of Research Human Use Committee and the United States Army Medical Research and Materiel Command Human Subjects Review Board of the U.S. Army Surgeon General, and was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Written informed consent was obtained from all subjects prior to participation.

Figure 4 shows the mean performance profile and the SD of the nine individuals over the 82 h of PVT test data collection. Overall, the trend suggests that both PVT lapsing and variance (70% of which could be attributable to interindividual variability) (35) increase over time. Figure 4 also shows (dashed line)

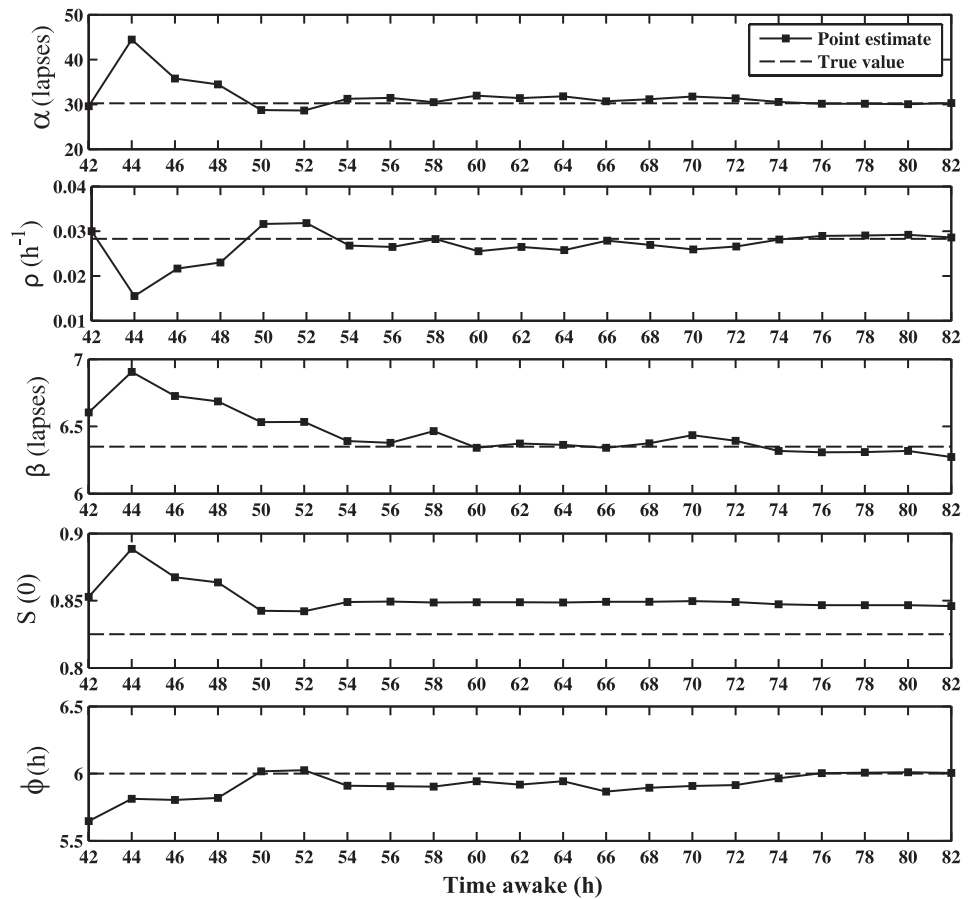


Fig. 3. Estimates and true value (horizontal dashed lines) of the 5 parameters [α , ρ , β , $S(0)$, and ϕ] of the 2-process model. These correspond to the simulated performance data illustrated in Fig. 2 and show that the estimates of model parameters asymptotically approach their true values as they are adjusted every 2 h (from the 42nd to 82nd hours of continuous wakefulness).

predictions obtained with the group-average two-process model using the model parameter values discussed above and depicted in Table 1. Visual inspection clearly indicates that the group-average model is not adequate for predicting the mean performance of the nine individuals. Another interesting observation is that the group-average model predictions are out of phase with the mean PVT observations. This indicates that the

group-average model with fixed, preset parameter values is not capable of accurately predicting mean performance of a group of individuals with uncertain initial conditions, i.e., unknown values for the initial homeostat and circadian phase settings. This is important as a practical matter because it is likely that such information will often be difficult to obtain in real-world settings.

Table 1. Parameter estimates for three Monte Carlo simulations of 100 trials each with three different signal-to-noise ratios

Signal-to-Noise Ratio	Time Awake, h	α , Lapses	ρ , h^{-1}	β , Lapses	$S(0)$	ϕ , h
<i>True value</i>						
		30.30	0.03	6.35	0.82	6.00
<i>Estimated values</i>						
59.10	62	30.66 ± 2.10 (MSE = 4.54)	0.03 ± 0.00 (MSE = 0.00)	6.35 ± 0.26 (MSE = 0.07)	0.83 ± 0.01 (MSE = 0.00)	6.01 ± 0.16 (MSE = 0.02)
	82	30.52 ± 0.90 (MSE = 0.87)	0.03 ± 0.00 (MSE = 0.00)	6.34 ± 0.24 (MSE = 0.06)	0.83 ± 0.01 (MSE = 0.00)	6.01 ± 0.13 (MSE = 0.02)
14.77	62	32.42 ± 9.72 (MSE = 98.90)	0.03 ± 0.01 (MSE = 0.00)	6.37 ± 0.52 (MSE = 0.27)	0.83 ± 0.03 (MSE = 0.01)	6.03 ± 0.31 (MSE = 0.10)
	82	30.92 ± 1.95 (MSE = 4.21)	0.03 ± 0.01 (MSE = 0.00)	6.35 ± 0.48 (MSE = 0.23)	0.83 ± 0.03 (MSE = 0.01)	6.02 ± 0.27 (MSE = 0.07)
3.70	62	34.16 ± 12.70 (MSE = 176.20)	0.03 ± 0.03 (MSE = 0.00)	6.07 ± 0.98 (MSE = 1.04)	0.84 ± 0.65 (MSE = 0.42)	0.70 ± 0.65 (MSE = 0.43)
	82	31.93 ± 6.53 (MSE = 45.30)	0.03 ± 0.01 (MSE = 0.00)	6.03 ± 0.96 (MSE = 1.02)	0.83 ± 0.61 (MSE = 0.38)	0.64 ± 0.56 (MSE = 0.31)

The estimates' mean ± SD and the mean square error (MSE), i.e., the square of bias plus variance, are shown at 2 time points (62 and 82 h of wakefulness) at three different signal-to-noise ratios (59.10, 14.77, and 3.70).

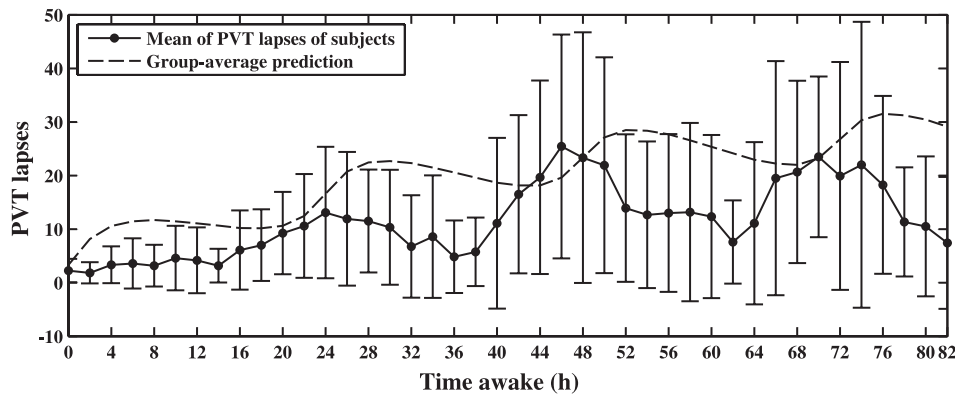


Fig. 4. Mean and SD of PVT performance measurements collected every 2 h for 9 individuals during 82 h of total sleep deprivation. The dashed line shows the performance predictions obtained with a group-average model.

Ideally, to compare the capability of the proposed (or any other) algorithm to develop individual-specific models of performance based on the two-process model, one would compare that algorithm's estimates against the true value of the model parameters. Unfortunately, except for simulated experiments, where the true parameter values are known, such a comparison is not possible. In the absence of such knowledge, a standard method to indirectly assess a predictive algorithm's capabilities is to compare the algorithm's predictions against observed (performance) data. However, because the predictive contribution of the model structure cannot be separated from those of the model parameter estimates, such a comparison inevitably considers the capability of the underlying model itself, in this case the two-process model, to represent the observed data. Hence, to test our individual-specific models (within the context of the 2-process model capability to represent performance data), we select three subjects representative of three distinct performance phenotypes: relatively vulnerable to sleep loss, relatively average sensitivity to sleep loss, and relatively resilient to sleep loss. Figure 5, *top*, *middle*, and *bottom*, illustrates the observed PVT lapses for the vulnerable, average, and resilient subjects, respectively, along with the group-average model predictions. Figure 5 also shows the individual-specific model predictions for 2-, 6-, and 10-hour prediction horizons in which the first 42 h of data (shaded) are employed to generate the initial estimates of the model parameters and make the initial prediction at the corresponding future time. For consistency, the RMSEs are computed between 52 and 82 h.

As expected, the RMSEs increase monotonically with increasing prediction horizons. In comparing the group-average with the individual-specific models, predictions from the individual-specific models consistently result in smaller RMSEs, even for the worst predictions attained with the 10-h prediction horizon. The most striking difference is observed for the resilient individual (Fig. 5, *bottom*), where the group-average model substantially overestimates performance impairment, yielding more than a threefold increase in the RMSE.

Although the differences in RMSE are not equally striking for the vulnerable and the average individuals, qualitatively there are significant differences. Regardless of the prediction horizons, the individual-specific models clearly capture the trends in the performance profiles, quasi-synchronously predicting the up and down variations of the observed data, while the group-average model fails to do so. The RMSE's failure to

describe this attribute of the predictions is pronounced in cases where data are significantly corrupted with noise—as in the data for the vulnerable and average individuals in Fig. 5—favoring “flat” predictions that fall within the variations of the data.

A model can be said to adequately describe a given set of measurements if the residual error, i.e., the difference between the model fit and the measurements, is an uncorrelated (white noise) signal (11). The whiteness of a residual-error signal can be inferred by computing its autocorrelation; for a purely white-noise signal, its autocorrelation coefficient, normalized between -1 and 1 , attains a value equal to one for a zero shift (i.e., zero delay) and a value equal to zero for all other shifts in the signal. For practical applications, where the residual error is not purely a white noise and the correlation coefficient for nonzero shifts is not zero, a degree of certainty of the whiteness of the residual-error signal can be inferred by computing approximate confidence error bounds of the autocorrelation coefficient.

This metric for model-fit characterization necessarily accounts for both the form of the model, i.e., the mathematical equations used to represent the underlying model, and the estimates of the model parameters. Hence, it affords us not only the ability to compare the performance of a group-average model with that of individual-specific models, but, for the first time, it provides the ability to quantify the power of the two-process model of sleep regulation in representing performance data for subjects restricted to prolonged, total sleep deprivation.

The top three panels in Fig. 6 show the autocorrelation function of the residual errors for both models for each of the three subjects (vulnerable, average, and resilient). The shaded areas illustrate the approximate 95% confidence bounds of the autocorrelation coefficients about zero, determined by the Portmanteau test (11). For the individual-specific models, we employ the model parameters attained at the end of the 82 h of wakefulness (since here we are testing for model fit instead of model prediction), and for the group-average model we employ the same parameter values used throughout the simulations (and illustrated in Table 1). The results clearly show that the individualized models describe the PVT lapse data better than the group-average model. In particular, for the vulnerable and average individuals, where the RMSEs in Fig. 5 between the individualized models and the group-average models are not markedly different, the autocorrelation coefficients reveal that the individualized models are indeed capable of capturing the

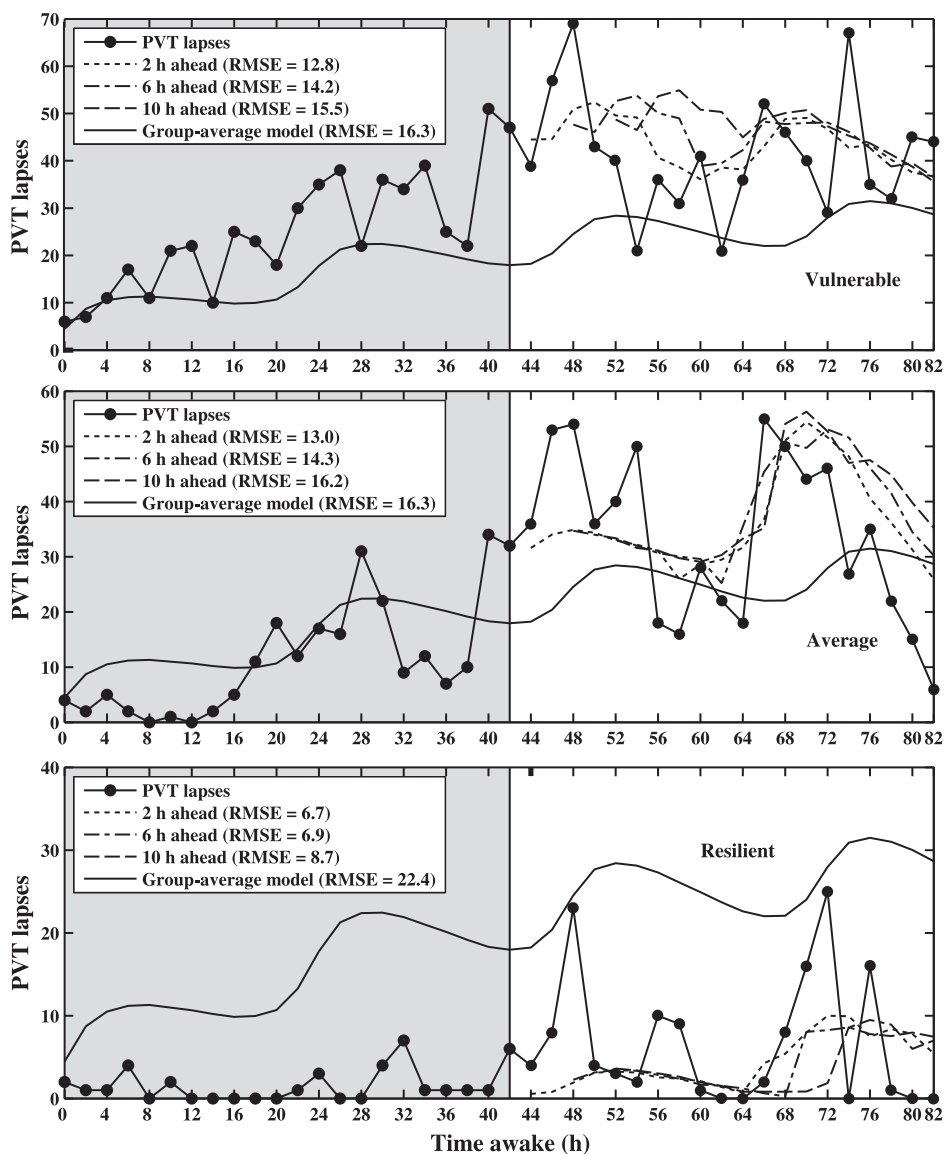


Fig. 5. Individualized model predictions for 3 subjects with different behavioral phenotypes [vulnerable to sleep loss (*top*), average sensitivity to sleep loss (*middle*), and resilient to sleep loss (*bottom*)]. The solid circles in each of the panels represent the observed PVT lapses measured every 2 h. The dotted, dot-dashed, and dashed lines represent the 2-, 6-, 10-h-ahead predictions, respectively, while the solid line represents the prediction using the group-average model.

correlations in the PVT lapse data, since their autocorrelation coefficients fall within the 95% bounds. This is not the case for the group-average model. Moreover, the top three panels in Fig. 6 strongly suggest that the two-process model of sleep regulation is capable of representing performance data for subjects restricted to prolonged, total sleep deprivation as long as the model parameters are customized for specific individuals using the proposed approach.

Because the group-average model is not expected to adequately predict performance for any one individual, we compute the residual error between the mean PVT lapses over the nine individuals and the group-average two-process model predictions illustrated in Fig. 4. Figure 6, *bottom*, clearly shows that the group-average model, assumed to represent an “average” individual, is not capable of accurately capturing the correlations in the mean PVT lapses.

DISCUSSION

In this study, we present a method based on the two-process model of sleep regulation to develop individual-specific mod-

els that predict decrements of cognitive performance over time during total sleep deprivation. Like other recently proposed individual-specific models (29, 36), the present method makes use of previous performance observations to predict future performance of individuals with uncertain initial state and unknown trait characteristics. However, unlike these approaches, the proposed method does not require additional performance data from a group of individuals. Instead, each individual’s temporal variability is captured solely from that individual’s own past performance data, and as each new observation becomes available these data are combined to adapt and customize the model parameters for that individual.

Another advantage of the proposed method is that, for a given set of performance data, it guarantees a unique estimate for the five parameters of the two-process model. This is achieved by transforming the nonlinear optimization problem of finding the optimal parameters of the two-process model, which may not have a unique solution, into a set of linear optimization problems whose solution, if it exists, is guaranteed to be unique. Through numerical simulations, we show that the model parameters

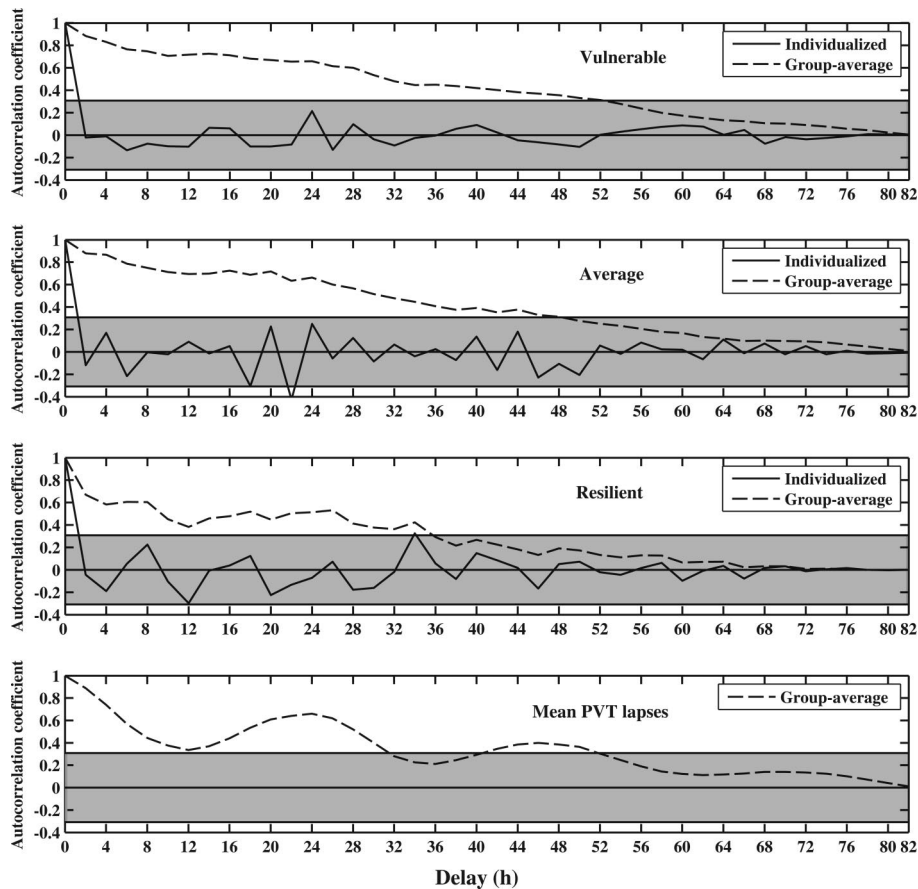


Fig. 6. Autocorrelation coefficients of the residual error as a function of time delay. Top 3 panels indicate the autocorrelation for the individualized and group-average model fit for each of the 3 individuals shown in Fig. 5. Bottom panel indicates the autocorrelation of the group-average model fit for the mean performance over the 9 individuals illustrated in Fig. 4. The shaded areas represent approximate 95% confidence intervals of the autocorrelation coefficients around zero.

asymptotically converge to their true values as the amount of training data increases and as the amount of noise in the performance data decreases, thereby confirming that our method satisfies key convergence properties of adaptive algorithms (12).

This work also addresses the as-yet-unanswered question of whether Borbély's two-process model (8, 14), which was developed to estimate slow-wave sleep propensity, can also serve as the basis for predicting the performance of sleep-deprived individuals. We analyze the amount of information content remaining in the residual error, i.e., the difference between the two-process model fit and the performance measurements, to provide a quantitative assessment of the model's ability to represent performance data. The whiteness of the residual errors suggests that, for individuals with vulnerable, average, and resilient phenotypes (Fig. 6, top three panels), as long as the parameters of the two-process model are systematically adapted to the particular individual as described here, the two-process model is capable of representing performance data for subjects restricted to prolonged, total sleep deprivation. The same is not true for the traditional fixed-parameter, group-average model (Fig. 6, bottom), which cannot accurately capture the correlations in the mean PVT lapses of a group of individuals with uncertain initial states, i.e., initial homeostat and circadian values.

Despite the advances of the proposed method to predict performance of sleep-deprived subjects at the individual level, it does have some limitations. The method requires the accumulation of a minimum number of past performance observations of an individual before accurate model-parameter estimates and performance predictions can be made for that

individual. In theory, a minimum of 13 observations are required. However, in practice, due to low signal-to-noise ratios in performance data, a larger number of observations is generally needed. This "batch" approach for learning the model parameters also precludes the direct use of the method to the modeling of chronic sleep deprivation, where large consecutive batches of data may not be available due to intermittent sleep/wake periods. Another limitation relates to the underlying assumption that measures of performance, such as PVT lapses, are available on a frequent basis. However, this assumption may not hold in certain operational settings because it may not be practical to discontinue work-related tasks for administering performance tests. Furthermore, we note that PVT lapses, selected as our predicted variable, may be incapable of reflecting the true performance of individuals subjected to complex cognitive tasks.

Considerable modeling efforts are still required to address the knowledge and capability gaps identified at the Department of Defense-sponsored Fatigue and Performance Modeling Workshop (18, 21, 26). The method described here can serve as a building block to address some of the gaps. For example, the "batch" approach for learning the model parameters can be reformulated into a "recursive" approach, where the model parameters are incrementally adjusted based solely on the most recent performance measurement. This model extension would allow for the prediction of performance for individuals restricted to chronic, partial sleep deprivation. Moreover, we could take advantage of the proposed linear representation of the two-process model and extend our method to analytically

estimate the reliability of the model predictions by providing statistical error bounds in the form of confidence and prediction intervals. This would allow for quantitatively determining the bounds within which model predictions may be trusted for a predefined coverage probability.

Along with the recent efforts by Van Dongen and colleagues (36), this work advances the capability of existing biomathematical performance models, shifting the focus away from traditional, group-average performance predictions into more accurate, operationally useful individual-specific predictions.

GRANTS

This work was funded, in part, by the Military Operational Medicine and the Combat Casualty Care Research Area Directorates of the U.S. Army Medical Research and Materiel Command, Ft. Detrick, Maryland.

DISCLAIMER

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 or of the U.S. Department of Defense. This paper has been approved for public release with unlimited distribution.

REFERENCES

- Achermann P, Borbely AA. Combining different models of sleep regulation. *J Sleep Res* 1: 144–147, 1992.
- Achermann P, Borbely AA. Simulation of daytime vigilance by the additive interaction of a homeostatic and a circadian process. *Biol Cybern* 71: 115–121, 1994.
- Afari N, Buchwald D. Chronic fatigue syndrome: a review. *Am J Psychiatry* 160: 221–236, 2003.
- Akerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. *Int J Neurosci* 52: 29–37, 1980.
- Balkin TJ, Bliese PD, Belenky G, Sing H, Thorne DR, Thomas M, Redmond DP, Russo M, Wesensten NJ. Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. *J Sleep Res* 13: 219–227, 2004.
- Berger JO. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag, 1985.
- Björck Å. *Numerical Methods for Least Squares Problems*. Philadelphia, PA: SIAM, 1996.
- Borbely AA. A two process model of sleep regulation. *Hum Neurobiol* 1: 195–204, 1982.
- Borbely AA, Achermann P. Sleep homeostasis and models of sleep regulation. *J Biol Rhythms* 14: 557–568, 1999.
- Casella G, Berger RL. *Statistical Inference*. Pacific Grove, CA: Duxbury/Thomson Learning, 2002.
- Chatfield C. *The Analysis of Time Series: An Introduction*. Boca Raton, FL: Chapman and Hall/CRC, 2004.
- Cherkassky VS, Mulier F. *Learning From Data: Concepts, Theory, and Methods*. New York: Wiley, 1998.
- Czeisler CA, Dijk DJ, Duffy JF. Entrained phase of the circadian pacemaker serves to stabilize alertness and performance throughout the habitual working day. In: *Sleep Onset: Normal and Abnormal Processes*, edited by Ogilvie RD and Harsh JR. Washington DC: Am. Psychol. Assoc., 1994, p. 89–110.
- Daan S, Beersma DG, Borbely AA. Timing of human sleep: recovery process gated by a circadian pacemaker. *Am J Physiol Regul Integr Comp Physiol* 246: R161–R183, 1984.
- Dijk DJ, Beersma DG, Daan S, Lewy AJ. Bright morning light advances the human circadian system without affecting NREM sleep homeostasis. *Am J Physiol Regul Integr Comp Physiol* 256: R106–R111, 1989.
- Dijk DJ, Duffy JF, Czeisler CA. Circadian and sleep/wake dependent aspects of subjective alertness and cognitive performance. *J Sleep Res* 1: 112–117, 1992.
- Dijk DJ, Visscher CA, Bloem GM, Beersma DG, Daan S. Reduction of human sleep duration after bright light exposure in the morning. *Neurosci Lett* 73: 181–186, 1987.
- Dinges DF. Critical research issues in development of biomathematical models of fatigue and performance. *Aviat Space Environ Med* 75: A181–A191, 2004.
- Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behav Res Meth Instr Comp* 17: 652–655, 1985.
- Engl HW, Hanke M, Neubauer A. *Regularization of Inverse Problems*. Boston, MA: Kluwer Academic, 2000.
- Friedl KE, Mallis MM, Ahlers ST, Popkin SM, Larkin W. Research requirements for operational decision-making using models of fatigue and performance. *Aviat Space Environ Med* 75: A192–A199, 2004.
- Hansen PC. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, PA: SIAM, 1998.
- Jewett ME, Dijk DJ, Kronauer RE, Boivin DB, Czeisler CA. Homeostatic and circadian components of subjective alertness interact in a non-additive manner during 48 hours of sleep deprivation. *J Sleep Res* 14: 493–499, 1996.
- Klerman EB, Hilaire MS. On mathematical modeling of circadian rhythms, performance, and alertness. *J Biol Rhythms* 22: 91–102, 2007.
- Leprout R, Colecchia EF, Berardi AM, Stickgold R, Kosslyn SM, Van Cauter E. Individual differences in subjective and objective alertness during sleep deprivation are stable and unrelated. *Am J Physiol Regul Integr Comp Physiol* 284: R280–R290, 2003.
- Mallis MM, Mejdal S, Nguyen TT, Dinges DF. Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviat Space Environ Med* 75: A4–A14, 2004.
- Marquardt DW. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics* 12: 591–612, 1970.
- Morozov VA, Nashed Z. *Methods for Solving Incorrectly Posed Problems*. New York: Springer-Verlag, 1984.
- Olofsen E, Dinges DF, Van Dongen HP. Nonlinear mixed-effects modeling: individualization and prediction. *Aviat Space Environ Med* 75: A134–A140, 2004.
- Ramsay JO. Principal differential analysis: data reduction by differential operators. *J R Stat Soc Series B* 58: 495–508, 1996.
- Ramsay JO, Silverman BW. *Functional Data Analysis*. New York: Springer, 2005.
- Reifman J, Gander P. Commentary on the three-process model of alertness and broader modeling issues. *Aviat Space Environ Med* 75: A84–A88; discussion A89, 2004.
- Schittkowski K. *Numerical Data Fitting in Dynamical Systems: A Practical Introduction with Applications and Software*. Boston, MA: Kluwer Academic, 2002.
- Tikhonov AN, Arsenin VY. *Solutions of Ill-Posed Problems*. New York: Halsted, 1977.
- Van Dongen HP, Baynard MD, Maislin G, Dinges DF. Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep* 27: 423–433, 2004.
- Van Dongen HP, Mott CG, Huang JK, Mollicone DJ, McKenzie FD, Dinges DF. Optimization of biomathematical model predictions for cognitive performance impairment in individuals: accounting for unknown traits and uncertain states in homeostatic and circadian processes. *Sleep* 30: 1129–1143, 2007.
- Viola AU, Archer SN, James LM, Groeger JA, Lo JC, Skene DJ, von Schantz M, Dijk DJ. PER3 polymorphism predicts sleep structure and waking performance. *Curr Biol* 17: 613–618, 2007.
- Wahba G. *Spline Models for Observational Data*. Philadelphia, PA: Soc. for Industrial and Applied Mathematics, 1990.
- Webb WB, Levy CM. Effects of spaced and repeated total sleep deprivation. *Ergonomics* 27: 45–58, 1984.
- Wesensten NJ, Killgore WD, Balkin TJ. Performance and alertness effects of caffeine, dextroamphetamine, and modafinil during sleep deprivation. *J Sleep Res* 14: 255–266, 2005.
- Wilkinson RT. Interaction of lack of sleep with knowledge of results, repeated testing, and individual differences. *J Exp Psychol* 62: 263–271, 1961.