Method to determine whether sleep phenotypes are driven by endogenous circadian rhythms or environmental light by combining longitudinal data and personalised mathematical models

Anne C. Skeldon[1,2,*], Thalia Rodriguez Garcia[1,2,3], Sean F. Cleator[1,2,3], Ciro della Monica[2,3], Kiran K.G. Ravindran[2,3], Victoria L. Revell[2,3], and Derk-Jan Dijk[2,3]

[1]Department of Mathematics, University of Surrey, Guildford, UK
[2]UK Dementia Research Institute, Care Research and Technology Centre at Imperial College, London and the University of Surrey, Guildford, UK
[3]Surrey Sleep Research Centre, Department of Clinical and Experimental Medicine, University of Surrey, Guildford, UK
[*]Corresponding author

June 14, 2023

Abstract

Sleep timing varies between individuals and can be altered in mental and physical health conditions. Sleep and circadian sleep phenotypes, including circadian rhythm sleep-wake disorders, may be driven by endogenous physiological processes, exogeneous environmental light exposure along with social constraints and behavioural factors. Identifying the relative contributions of these driving factors to different phenotypes is essential for the design of personalised interventions.

The timing of the human sleep-wake cycle has been modelled as an interaction of a relaxation oscillator (the sleep homeostat), a stable limit cycle oscillator with a near 24-hour period (the circadian process), man-made light exposure and the natural light-dark cycle generated by the Earth's rotation. However, these models have rarely been used to quantitatively describe sleep at the individual level. Here, we present a new Homeostatic-Circadian-Light model (HCL) which is simpler, more transparent and more computationally efficient than other available models and is designed to run using longitudinal sleep and light exposure data from wearable sensors. We carry out a systematic sensitivity analysis for all model parameters and discuss parameter identifiability.

We demonstrate we can describe individual sleep phenotypes in each of 34 older participants (65-83y) by feeding individual participant light exposure patterns into the model and fitting two parameters that capture individual average sleep duration and timing. The fitted parameters describe endogenous drivers of sleep phenotypes.

We then quantify exogenous drivers using a novel metric which encodes the circadian phase dependence of the response to light. Combining endogenous and exogeneous drivers better explains individual mean mid-sleep (adjusted R-squared 0.62) than either driver on its own (adjusted R-squared 0.07 and 0.17 respectively).

Critically, our model and analysis highlights that different people exhibiting the *same* sleep phenotype may have *different* driving factors and opens the door to personalised interventions to regularize sleep-wake timing that are readily implementable with current digital health technology.

Summary

Disrupted sleep has long term health consequences and affects our day-to-day ability to function physically, mentally and emotionally. But what determines when and how long we sleep?

It is well-known that daily light exposure patterns determine the timing of the body clock. However, creating mathematical models that can take realistic light exposure patterns and predict plausible sleep timing has been challenging. Furthermore, nearly all previous studies have focused on developing models for average behaviour, yet sleep timing and duration are highly individual.

In this paper, we present a simple model that combines sleep regulatory and circadian processes. The model can take individual light exposure patterns and, by fitting physiologically plausible parameters, describe individual mean sleep timing and duration. We test our model on data collected from 34 older participants. Our modelling approach suggests that some of the participants slept late because of physiological factors, while for other individuals, late sleep was a consequence of their light environment.

This approach of combining a model with longitudinal data could be implemented in digital health technology such that your smart watch could tell you not only how you slept last night, but also how to change your light environment to sleep better tomorrow.

1 Introduction

Sleep is a recurring physiological state in many species. In humans, it is a major determinant of quality of life, and disorders and disturbances of sleep are predictive of adverse health outcomes [1]. Sleep timing and duration change across the human life span but also show considerable individual differences in every age group. When individual differences in sleep timing are extreme, they are considered disorders; circadian rhythm sleep-wake disorders are prototypical examples [2]. Sleep disorders, including those associated with timing, are associated with psychiatric disorders and evidence indicates the relationship may be bidirectional [3, 4]. Interventions to ameliorate sleep timing are therefore desirable. Understanding, at an individual level, the endogenous versus exogeneous factors determining different sleep phenotypes [5] will facilitate the design of more effective, personalised treatments.

Sleep fulfils an essential need that accrues during wakefulness. Sleep homeostasis refers to the process that keeps track of this build up and dissipation of sleep need and this process is a determinant of sleep timing. Sleep timing in humans is also governed by a second physiological process, namely circadian rhythmicity. In the context of sleep regulation, circadian rhythmicity refers to the approximately 24-hour variation in sleep-wake propensity which persists even when sleep need and environmental factors such as light are kept constant. Circadian sleep propensity is high during the circadian night and low during the circadian day [6].

The alignment of circadian rhythms with respect to the 24-hour day is regulated by environmental light exposure patterns. The circadian pacemaker, located in the suprachiasmatic nuclei (SCN) in the brain, generates a near 24-hour rhythm with a near 24-hour period (24.15 h, standard deviation 0.2 h) [7]. This non-24-hour rhythm is synchronised to the 24-hour day by light input via a well-defined neural pathway connecting retinal photoreceptors to the SCN. The effects of light on the pacemaker depend on the phase of the circadian cycle at which light information reaches the SCN. Light in the biological morning speeds up the clock and light in the biological evening slows down the clock [8, 9]. Entrainment of our non-24-hour clocks to the 24-hour day occurs when there is an appropriate balance between ‘speeding up’ and ‘slowing down’.

Environmental light exposure patterns were once primarily driven by the natural light-dark cycle. However, in industrialised societies, most people spend approximately 90%

of their time indoors [10] and, access to electric light means light is available at any time of day or night. Since indoor light is typically of lower intensity than outdoor light and we self-select when to turn lights on and off, the strength of the light-dark signal as a signal that can entrain the biological clock to the 24-hour day has been greatly degraded. The consequence of a less robust 24-hour light-dark signal is that for most people, i.e. those with an intrinsic period longer than 24-hours (77%), circadian rhythms and preferred sleep timing delay [11, 12, 13, 14]. Social factors, such as work, school schedules and socialising, also influence sleep timing [15]. In sum, sleep homeostasis, circadian rhythmicity, the light-dark cycle, and social constraints interact to determine the timing of wake and sleep across the 24-hour day.

Understanding of the genetic molecular determinants and correlates of the sleep homeostat, circadian rhythmicity and the effects of light has grown rapidly in the past few decades (e.g. [16, 17]). At the same time laboratory studies have indicated that long and short sleepers differ with respect to the average level of homeostatic sleep pressure [18] whereas early and late sleepers differ in their intrinsic circadian period [19, 20]. However, at the individual level, understanding how these different regulatory mechanisms result in different sleep phenotypes in daily life is more limited. The difficulty of ascertaining causal factors driving sleep phenotypes is exacerbated by the fact that the same phenotype may have different underlying causes. For example, late sleep timing may be a consequence of too little bright light in the morning or too much bright light in the evening or a long intrinsic period (slow clock), or a combination of all of these. Parameters of the clock and sleep homeostasis can be quantified in laboratory protocols but these are labour intensive and are not scalable.

Mathematical models are one way to unravel the relative contributions of different factors. Several system-level mathematical models for the human sleep-wake cycle have been developed. Kronauer and colleagues modelled the sleep-wake cycle as an interaction between a strong circadian oscillator (driving core body temperature) and a weak sleep-wake oscillator. Entrainment to the 24-hour day occurred via forcing the strong oscillator by the effects of a so-called zeitgeber ('time giver') which entered the system via the weak oscillator [21]. However, the zeitgeber was not based on available data for the effects of light on the human circadian timing system. Daan, Beersma and Borbély modelled the interaction of the homeostatic and circadian process (hence the so-called 'two-process' model) but did not model the effects of light on the circadian process and hence could not simulate entrainment to the 24-hour day. Parameters of

the homeostatic process of the two-process model were based on electroencephalography (EEG) slow wave activity (SWA, EEG power density in the 0.75-4.5 Hz range) and it was assumed that this variable played a key role in sleep timing. Parameters for the circadian process were motivated by average sleep duration following displacement of sleep with respect to the day by extending the wake period [22, 23].

The two-process model is a phenomenological model. Building on insights on the neuronal regulatory mechanisms and the conceptual flip-flop model [24], Phillips and Robinson developed a more physiologically-based model by considering mutual inhibitory sleep and wake promoting nuclei whose firing rates were modulated by circadian and homeostatic processes [25]. Later, they incorporated Kronauer's model for the phase-shifting effects of light in phase response experiments [26, 27], by including a self-sustained oscillator and light input [28]. Previously, we have shown that an adapted version of the Phillips-Robinson model can accurately model sleep duration and timing across the lifespan [29] and have systematically investigated the impact of light and imposed social constraints (such as getting up for work) [14]. We have shown that combining light exposure data, collected longitudinally, with the Phillips-Robinson model can replicate average sleep timing and sleep duration within people living with schizophrenia and healthy unemployed individuals [30]. Using a similar model, Postnova and co-workers have modelled the effect of imposed wake schedules due to shift-working [31] and in forced desynchrony experiments [32], on sleepiness, cognitive performance and 'circadian dynamics'. Hong et al. have introduced the notion of 'circadian sufficient sleep' to quantify the degree of alignment of sleep and circadian rhythms [33]. Concurrently neuronal models capturing wake, rapid eye movement (REM) sleep and non-REM sleep have been developed [34], which, consistent with observations, have modelled the differences between long and short sleepers by differences in mean levels of a sleep homeostat [35].

Neuronal models are specifically designed to incorporate more physiological insight into mathematical models of sleep-wake regulation than the phenomenological two-process model. However, consequently, they are more mathematically complex and have more parameters than are needed to explain currently available data on sleep timing and duration.

Here, we take a different approach and present a new, more minimalist, model that incorporates the effects of sleep homeostasis, circadian rhythmicity and light on sleep

timing, which has a lower compute time and where the role of each parameter is more transparent. This new model has the intuitive appeal of the two-process model but also includes the effect of light on circadian timing. We undertake a systematic sensitivity analysis to find which parameters primarily regulate sleep timing and duration and discuss which parameters are identifiable from light and sleep timing data. In older participants (N=34, age 65-83 y), we demonstrate that we can extract individual parameters describing endogenous physiological processes by combining the model with data from wearable light sensors and fitting for individual average sleep duration and timing. Associations between model parameters and multiple sleep measures, including standard biomarkers of sleep homeostasis collected in the laboratory, are considered. The exogenous driver of light is quantified using both standard metrics of light exposure and a novel light metric that quantifies both the intensity and time-of-day dependent effects of light on the biological clock. Finally, using the extracted parameters and the novel light metric, we propose a method to identify the endogenous physiological versus the exogenous environmental factors underlying different phenotypes and discuss the implications for the design of light interventions delivered by digital health technology applications.

2 Methods

2.1 A new mathematical model for realistic sleep timing

Our Homeostatic-Circadian-Light (HCL) model is motivated by elements of the original two-process model for sleep-wake regulation [22, 23], the neuronal model of Phillips-Robinson model [25, 28] and a model of the interaction of light with the human biological clock [27]. The model is illustrated in Fig. 1 and described in detail below.

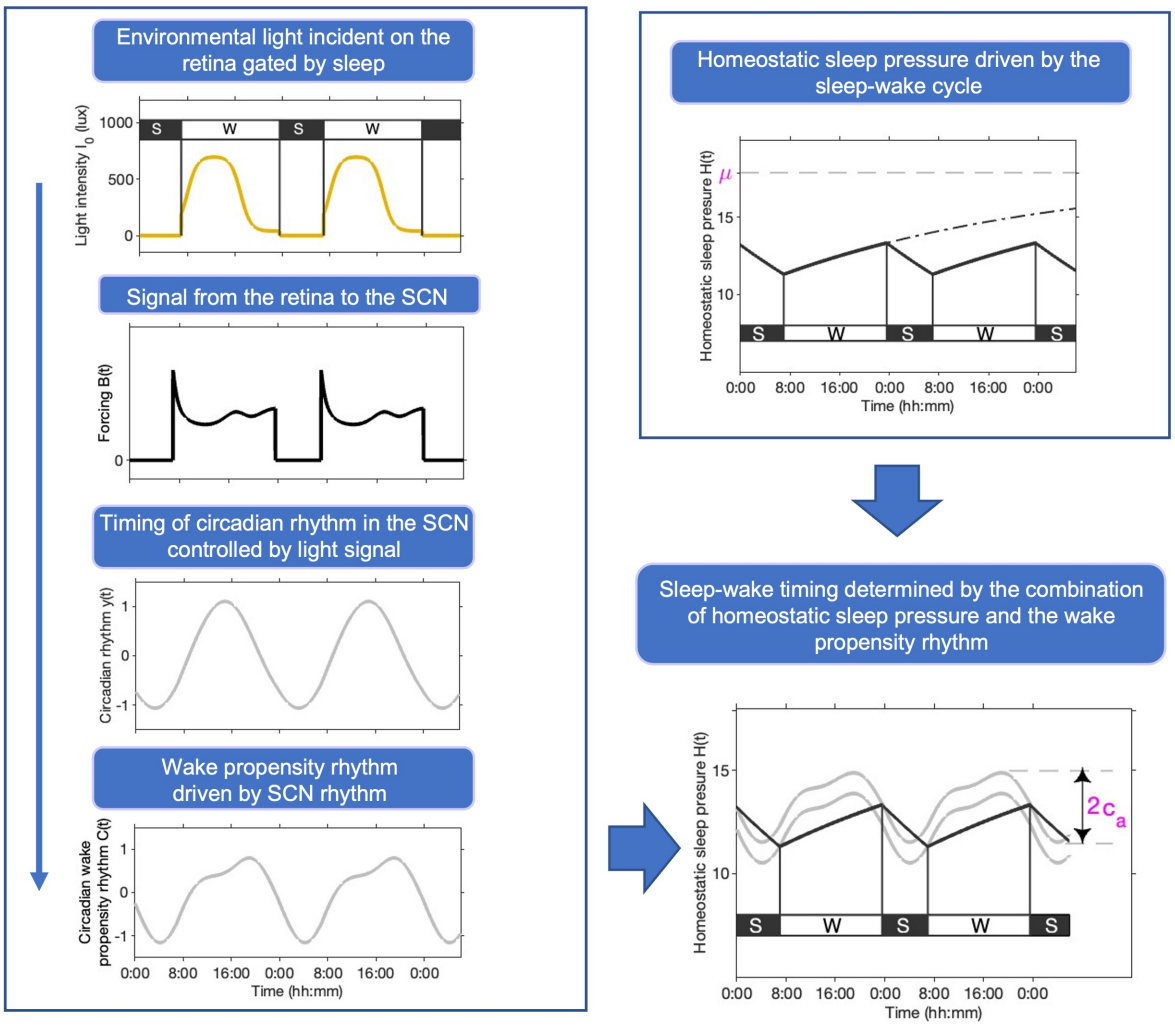

Figure 1: **Homeostat-Circadian-Light (HCL) model.** Patterns of available environmental light gated by the sleep-wake cycle (since we turn off the lights and close our eyes when we go to sleep), result in a rhythmic signal that passes from the retina to the suprachiasmatic nuclei (SCN). Provided the signal is of sufficient "strength", it controls the timing of rhythms in the SCN, which in turn drives the 24-hour rhythm in the circadian drive for wakefulness and sleep. Homeostatic sleep pressure H increases during wake and decreases during sleep. If kept awake, homeostatic sleep pressure asymptotes to an asymptote μ. The larger the value of μ, the faster the rise rate of homeostatic sleep pressure during wake. Switching between wake (W) and sleep (S) states occurs at thresholds modulated by the circadian rhythm of wake-sleep propensity which has peak-to-peak amplitude $2c_a$.

Sleep homeostasis

As in the two-process model, it is assumed that there is a sleep pressure signal $H(t)$ which increases during wake and decreases exponentially during sleep. Specifically,

$$H(t) = \begin{cases} \mu + (H(t_{\rm off}) - \mu) \exp\left(-\frac{t-t_{\rm off}}{\chi}\right), & \text{during wake,} \\ H(t_{\rm on}) \exp\left(-\frac{t-t_{\rm on}}{\chi}\right), & \text{during sleep,} \end{cases} \tag{1}$$

where t is time, $t_{\rm on}$ and $t_{\rm off}$ are the times of sleep onset and offset respectively, χ describes the rate of decay of sleep pressure during sleep whereas the rise rate during wake is governed by both μ and χ. The parameter μ is the 'upper asymptote' for the homeostatic sleep pressure signal since during wake $H(t) \to \mu$ as $t \to \infty$. In other words, if kept awake for a long time, homeostatic sleep pressure approaches an upper limit that cannot be surpassed, see Fig. 1. Similarly, homeostatic sleep pressure approaches a lower asymptote of zero during sleep.

In differential equation form,

$$\chi \frac{dH}{dt} = -H + (1 - S)\,\mu, \tag{2}$$

where S takes the value 1 during sleep and 0 during wake.

Circadian rhythmicity

As in the two-process model, we assume that spontaneous switching from wake to sleep and from sleep to wake occurs at thresholds which are periodically modulated to mimic the effect of circadian rhythmicity. Hence, switching from wake to sleep occurs at $H(t) = H^+(t)$ where

$$H^+(t) = \overline{H}_0 + \frac{1}{2}\Delta + c_{\rm a} C(t), \tag{3}$$

and switching from sleep to wake occurs at $H(t) = H^-(t)$ where

$$H^-(t) = \overline{H}_0 - \frac{1}{2}\Delta + c_{\rm a} C(t). \tag{4}$$

The constant $\overline{H}_0$ is the mean level of upper and lower thresholds and Δ is the separation between the thresholds. The wave form of the time dependent circadian modulation $C(t)$ is derived from experimental data [36] and is discussed in greater detail below. Paradoxically, but consistent with data and the understanding that sleep and circadian

processes act in opposition, the circadian wake propensity signal increases during the 'day', reaching a maximum shortly before bedtime and a minimum in the second half of the night [37].

Equations (2)-(4) may be considered as a lower dimensional representation of a neuronal model in which switching between wake and sleep occurs on a fast timescale. The neuronal interpretation of $\overline{H}_0$ is that it represents the drive to be awake, for example from orexin [24], whereas Δ represents the strength of the mutual inhibition between sleep and wake promoting neurons. Further discussion of the link between $\overline{H}_0$ and Δ and parameters in neuronal models such as [28] is given in the Supplementary Material, Section S1.

Light coupling to circadian rhythmicity

The function $C(t)$ describes the circadian wake propensity rhythms and has amplitude a peak-to-peak amplitude of approximately two. Here, the circadian rhythm $C(t)$ is constructed from a model of the effect of light on circadian phase and data describing the circadian drive for wakefulness [36]. Specifically, following Kronauer [26], the circadian timing system is modelled as the interaction of Process L (L for light), and Process P (P for pacemaker) [26]. In Process L, light is considered to activate photoreceptors in the eye and result in a driving 'force', $\hat{B}$, on the circadian pacemaker,

$$\tilde{I} = (1-S)\, I(t), \tag{5}$$

$$\frac{dn}{dt} = \alpha_0 \left(\frac{\tilde{I}}{I_0}\right)^p (1-n) - \beta n, \tag{6}$$

$$\hat{B} = G\alpha_0 \left(\frac{\tilde{I}}{I_0}\right)^p (1-n), \tag{7}$$

where $I(t)$ is the available light, $\tilde{I}$ is the light that reaches the retina and n is the fraction of activated photoreceptors. The factor $(1-S)$ in equation (5) models the fact that eyes are closed during sleep, so that sleep has a gating effect on light exposure. Equation (6) states that when light reaches the retina, the rate of saturation of photoreceptors is given by $\alpha_0 \left(\frac{\tilde{I}(t)}{I_0}\right)^p$ and the rate of decay of activated receptors is given by β. The values of α_0, I_0, p and β have been determined from experimental data collected during phase response experiments [26, 27]. Of particular note is that the response to light of intensity $\tilde{I}$ is 'compressed' since it appears in the form $(\tilde{I})^p$ and $p <$1.

The pacemaker P is modelled as a van der Pol oscillator, here, as in [27], given by,

$$\kappa \frac{dx}{dt} = \gamma \left(x - \frac{4x^3}{3} \right) - y \left(\left(\frac{24}{f\tau_c} \right)^2 + kB \right), \tag{8}$$

$$\kappa \frac{dy}{dt} = x + B, \tag{9}$$

where $\kappa = \frac{24 \times 3600}{2\pi}$ is a time scaling factor and

$$B(t) = (1 - bx)(1 - by)\,\hat{B}. \tag{10}$$

A typical form for $B(t)$ for one specific light profile is shown in Fig. 1 and shows that it incorporates both so-called 'phasic' and 'tonic' elements [38]. Phasic effects induce abrupt changes in the position of the oscillator, as occurs as a result of the spike in $B(t)$ at the transition from dark to light. Tonic effects change the speed of the clock through modulating the period of the pacemaker and occur via the fact that $B(t)$ settles to an approximately constant value in constant light.

In equations (8), (9), $y(t)$ represents the circadian pacemaker activity, and $x(t)$ is an auxiliary variable. The physiologically motivated parameters result in the van der Pol oscillator operating in its weakly nonlinear regime where $x(t)$ and $y(t)$ oscillate in a close to sinusoidal manner, see Fig. 1. In the absence of light, these oscillations have a natural period, the 'intrinsic circadian period', of τ_c hours. The parameter γ is the stiffness of the oscillator which describes how easily the system is perturbed from its natural oscillatory state and the speed of recovery following perturbation. Consistent with observations that to induce large changes in phase in humans repeated bouts of high intensity light are required [39], the value of γ is taken to be relatively large. This means that the periodic solution (limit cycle) is strongly attracting,

The effect of light on the van der Pol oscillator comes from the light drive, $\hat{B}$ modulated by a sensitivity function $(1 - bx)(1 - by)$ that describes how the strength of the drive varies with the position of the oscillator.

In the HCL model, the function $C(t)$ is taken to be

$$C(t) = c_{20} + \alpha_{21}x + \alpha_{22}y + \beta_{21}x^2 + \beta_{22}xy + \beta_{23}y^2 \tag{11}$$

where the coefficients c_{20}, α_{ij} and β_{ij} were found by fitting to the circadian drive for wakefulness reported in [36], see Fig. 1. Further details of the fitting process are given in the Supplementary Material Section S2, along with the values of the coefficients.

Hence the complete HCL model consists of four differential equations: one for the sleep homeostat, (2); two for the van der Pol oscillator, (8), (9), and one for the photoreceptors, (6). In addition, there are the two switching conditions: switching from wake to sleep occurs when $H(t) = H^{+}(t)$ and switching from sleep to wake when $H(t) = H^{-}(t)$, where $H^{\pm}(t)$ are given by equations (3), (4) and (11). Finally, the light exposure pattern $I(t)$ is needed as an input, see equation 5. In this paper, $I(t)$ will come from light intensity data measured in individuals. For scenario testing and intervention design, synthetic profiles could also be used.

2.1.1 Parameter values in the HCL model

The HCL model contains a total of 14 parameters, five directly linked to the sleep-wake regulation aspects of the model ($\mu, \chi, \overline{H}_0$, Δ, c_a) and nine related to the interaction of light with the pacemaker ($f\tau_c$, G, p, k, B, γ, α_0, β, I_0).

Default parameter values for $\chi, \overline{H}_0$ and Δ are motivated by considering the HCL model as a slow manifold reduction of the neuronal model in [14]. Hence the gap between the thresholds Δ is much smaller, and typical values for the circadian amplitude c_a are much larger, than those used in the original two-process model. Furthermore, consistent with neuronal models, we take the homeostatic time constant χ to be 45 hours whereas the two-process model has different time constants for the rate of rise of sleep homeostasis during wake and rate of fall of sleep homeostasis during sleep. These time constants have been measured from slow wave activity in sleep electroencephalography (EEG) studies to be approximately 18 hours and four hours respectively [23]. These time constants appear very different, but since the actual rise / decay rates of the homeostat depend not only on the homeostatic time constants but also on the values of the upper (and lower) asymptotes it is not easy to make a direct comparison. Further comparison between the HCL model parameter regime and that of the original two-process model is given in the Supplementary Material, Section S3 and Fig. S2. In Supplementary Section S3 we also demonstrate that the ratio of the average rate of rise of sleep homeostasis during wake compared with the average fall during sleep is given by the ratio of sleep to wake duration). Thus for an average sleeper who is asleep for 8 hours and awake for 16 hours, the ratio is one half.

For the purposes of this paper, default values for μ and c_a were chosen such that, with the same default light profile used in [14] and shown gated by sleep in Fig. 1, mid-

sleep time was 03:16 hh:mm. This matches mid-sleep time data reported in [40, 41] for people aged approximately 65 y. Default values for the ten light / circadian parameters were fixed to the standard parameters used in [28, 14, 30].

All parameters, their meanings and their default values are listed in Table 1.

Sleep / wake regulation parameters		
Parameter	Value	Meaning
μ	17.87	Upper asymptote giving the upper bound for homeostatic sleep pressure and controlling the rate of rise of homeostatic sleep pressure during wake.
χ	$45 \times 60 \times 60$ s	Time constant for rise and decay of homeostatic sleep pressure.
$\overline{H}_0$	13	Mean level of the wake propensity rhythm.
Δ	1	Separation between the thresholds.
c_a	1.72	Circadian amplitude.
Light / circadian parameters		
Parameter	Value	Meaning
τ_c	24.2 h	Circadian period. Natural period of the van der Pol oscillator.
f	0.99669	Correction factor so that in the absence of light the van der Pol oscillator gives oscillations with a period of τ_c.
G	19.9	Gain factor determining the magnitude of the effect of light on the driving force $\hat{B}$ on the circadian pacemaker.
p	0.6	Light sensitivity. Determines the steepness of the dose response curve to light.
k	0.55	Determines the relative effect of light on the oscillator variables x and y. It changes the shape of the velocity response curve.
b	0.4	Sensitivity modulation factor that determines the degree to which the strength of the drive on pacemaker P depends on the position of the pacemaker. Shifts the time of maximum sensitivity to light.
γ	0.23	Stiffness of the van der Pol oscillator determining how quickly oscillations return to the limit cycle when perturbed.
α_0	$0.16/60$ s^{-1}	Factor determining the magnitude of the effect of light on the fraction of activated photoreceptors and on the drive.
β	$0.013/60$ s^{-1}	Decay rate of fraction of activated photo receptors.
I_0	9500 lux	Scaling factor for light.

Table 1: **Default parameter values for the HCL model and their meanings**.

2.1.2 Differential equation solver

Simulations of the HCL model were carried out in MATLAB [42] using the stiff solver ode15s with relative and absolute tolerances set to 10^{-4}. The event solver was used to locate switching points from wake and sleep and from sleep and wake.

2.2 Data collected in everyday life

The primary purpose of developing the HCL mathematical model was to construct a simpler, more transparent model that was computationally less complex than existing models and could be used to fit longitudinal data for deeper understanding of causal factors driving sleep phenotypes in close to real time.

In order to test our approach, we used light, actigraphy and sleep diary data collected for 7-14 days at home from 35 participants (age (mean $\pm$ SD) 70.9 $\pm$ 5 years, range 65–83 years, 14 women). Participants also undertook an overnight polysomnography (PSG) recording at the Surrey Sleep Research Centre, Guildford, UK. The study received a favourable opinion from the University of Surrey Ethics Committee and was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from participants before any procedures were performed.

To be deemed eligible for the study participants had to be between the age of 65 and 85 years, be self-declared mentally stable and physically healthy but could have stable controlled medical conditions including Type 2 diabetes, hypertension, arthritis. They needed to be able to perform daily living activities independently, drink less than 29 units of alcohol per week and be non-smokers. Eligibility for the study was assessed via a telephone interview and an in-person screening visit which included provision of demographic information, measurement of height, weight and vital signs, self-reported medical history, symptom directed physical examination and completion of a series questionnaires including the Epworth Sleepiness Scale [43] and the Pittsburgh Sleep Quality Index [44]. Data were collected from the first 18 participants from January to March 2020. Following a temporary halt in data collection due to COVID, data from the remaining 17 participants were collected from July to November 2021.

Subjective sleep-wake patterns were assessed via daily completion of the Consensus Sleep Diary [45]. Parameters collected through the sleep diary included the time the

participant went to bed, when they tried to go to sleep, how long it took them to fall asleep (sleep latency), number of awakenings, and time of final awakening. For the purposes of this paper, sleep onset times for each participant were calculated as the sleep diary time when participants started to try to go to sleep plus the sleep latency (i.e. answers to the questions: What time did you try to go to sleep? How long did it take you to fall asleep?). Sleep offset times were considered to be the sleep diary final awakening time (i.e. answer to the question: What time was your final awakening?).

Light exposure patterns were assessed by continuous wearing of a wrist worn actiwatch (Actiwatch Spectrum, Philips Respironics, Murrysville, PA, USA). Environmental white light levels, recorded in lux at one minute intervals, were extracted from the Actiwatch datasets. One participant was excluded from the fitting because the Actiwatch stopped recording after two days. The second group of participants in addition wore a second light measuring device (HOBO Pendant MX2202 Onset Data Logger, Bourne, MA, USA) that was worn as a pendant and clipped to clothing near the shoulder. Results were calculated both using data recorded at the wrist and data recorded at the shoulder to test the sensitivity of results to the position of the sensor.

2.2.1 Light imputation

Participants were encouraged to keep all light monitors uncovered. Nevertheless, there were some periods in which the participants reported being awake but the recorded lux values were zero. For values that were zero and occurred during wake, imputation was carried out and zero values were replaced by the median value for the half an hour before and half an hour after, calculated across all available days of data for the participant concerned. Fittings were carried out with both raw and imputed light data to test the sensitivity of the method to imputation.

Light was recorded at one minute intervals. When the ordinary differential equation solver used to solve the HCL equations required values of light at intermediate intervals, light data were linearly interpolated.

2.2.2 Slow wave activity

Full PSG data were collected using the SomnoHD system (SOMNOmedics GmbHTM, Germany) according to the American Academy of Sleep Medicine (AASM) guidelines during the overnight in-lab recording which consisted of an in-bed period of ten hours. The Domino software associated with the SomnoHD system was used to score sleep at 30 s intervals [46]. A consensus hypnogram was generated using the manual sleep scoring performed by a registered polysomnographic technologist and an experienced scorer. The six channels of electroencephalography (EEG) data collected in the PSG included frontal (F3-M2; F3-M1), central (C3-M2; C4-M1) and occipital (O1-M2; O2-M1) derivations. The artefacts in the EEG were manually scored by an expert and the labels were verified by a second expert to ensure accurate removal of the artefacts. Following the artefact removal, slow wave activity (SWA) was calculated during NREM sleep in the 0.75-4.5 Hz range from the artefact free EEG power spectrum [47]. The SWA from the two frontal and central derivations were averaged to produce a single measure of SWA for each participant.

2.3 Parameter fitting

For each participant, the light data were fed into the HCL model and then parameters were found that minimized the difference between the average diary sleep duration and the average modelled predicted sleep duration and the average diary mid-sleep timing and the average model predicted mid-sleep timing. Care was taken to remove transients in the calculation of sleep duration and mid-sleep time by starting from default settings for the initial conditions for H, x, y and n and then integrating repeatedly for the duration of the light data (truncated to an integer number of days) until the model predicted average sleep duration and mid-sleep time had converged.

In order to test the robustness of our modelling approach, sleep data from each participant were fitted four times for the first cohort in which light data were collected at the wrist, and eight times for the second cohort in which light data were collected from both the wrist and shoulder. First, the impact of light imputation was tested by fitting using both raw and imputed light data. Next, the impact of the position of the sensor was tested by fitting using both light data collected from the shoulder and from the wrist. In all cases, fitting was carried out for the pair of parameters (μ, τ_c) and separately for

the pair of parameters (μ, c_a) to investigate different possible physiological interpretations. The motivations for choosing these two pairs of parameters are discussed in the results.

The fittings were implemented in MATLAB. For reasons of computational speed, different minimization methods were chosen depending on the fitting parameters. Specifically, when fitting for the pair of parameters (μ, c_a) the sum of squared residuals between the observed and model predicted averages of sleep duration and mid-sleep time was minimized using a Nelder-Mead search algorithm (fminsearch). The optimization routine was considered to have successfully converged to a solution if the residuals were less than 0.0005 and the parameters were physiologically realistic. Specifically, we required all parameters to be positive and $c_a + \overline{H}_0 + \frac{1}{2}\Delta < \mu$. The reason for imposing the latter constraint is that when $c_a + \overline{H}_0 + \frac{1}{2}\Delta > \mu$, there are some circadian phases at which it is impossible for homeostatic sleep pressure to reach the upper threshold. Consequently, the model would predict that, however long someone had been awake, it would not be possible for the model to spontaneously fall asleep during circadian phases close to the maximum wake propensity [37] (which in the literature is known as the wake maintenance zone or forbidden zone for sleep).

Fitting for the pair of parameters (μ, τ_c) was carried out iteratively using knowledge of model behaviour, as reported in the results. Specifically, sleep duration is independent of τ_c and depends monotonically on μ, whereas sleep timing depends monotonically on τ_c, see Section 3.1. Hence, first the value of μ that gave a model predicted average sleep duration that matched the average diary sleep duration was found using a bisection-based method (fzero). Second, the value of τ_c that matched the average diary mid-sleep time was found using a bisection-based method (fzero). The process was repeated if the residual did not meet the tolerances specified. In both cases, an interval containing the solution was first found based on the known monotonic dependence of the outcome measure on the parameters. Fitting iteratively was faster than using fminsearch.

2.4 Metrics for light exposure patterns

Light exposure patterns can be quantified in different ways, and it remains unclear which metric is most relevant for light's effect on the circadian system. Here, five stan-

dard metrics were calculated for each participant: the mean daily number of hours of bright light exposure (light $>$ 500 lux); the mean lux value per minute; the geometric mean light exposure; the mean time of day at which half the daily light exposure occurred by considering the cumulative lux; the mean time of day at which half the daily light exposure occurred by considering the cumulative log(lux+1).

2.4.1 Novel light metric quantifying the biological effect of light

In addition, a novel measure of the biological effect of light was constructed by feeding each day of light data into the light-circadian model of Forger et al [27] (essentially equations (5 - 10) for default parameters and $S = 0$) and integrating for one day. Initial conditions were set to be those for the state of the light-clock system (given by x, y and n) at midnight for someone with a mid-sleep of 04:30, a sleep duration of 8.22 h and the synthetic light profile used in [14]. These values were chosen to be consistent with the 'average' person considered in [14]. The light-circadian model is designed to capture the interaction of light on the human velocity response curve, including the fact that at some times of day light speeds up the clock and at others it slows down the clock. Integrating for one day and considering whether the clock has covered more than or less than one cycle (here expressed in minutes) quantifies the extent to which that day of light data has a net effect of speeding up or slowing down the clock. Positive values mean that the clock has undergone more than one cycle in 24 hours, negative numbers mean that the clock has undergone less than one cycle in 24 hours.

2.5 Statistics

General linear and / or mixed effect models were used to quantify the contribution of different factors to sleep timing.

Intraclass correlations (ICCs) were calculated for daily measures of mid-sleep time, sleep onset, sleep offset, sleep duration and light metrics in order to assess the within versus the between subject variance. The ICCs were calculated by first fitting a linear mixed effect model using the MATLAB fitlme function with participant as a random variable, i.e.

$$y_{ij} = \beta_0 + b_{0j} + \epsilon_{ij}, \tag{12}$$

for observation y_{ij} and participant j, with the assumptions that b_{0j} and ϵ_{ij} are normally

distributed, $b_{0j} \sim \mathcal{N}(0, \sigma_b^2)$ and $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. Hence, σ_b^2 is the between participant variance and σ^2 is the residual variance (within participant variance). The ICC is then given by

$$ICC = \frac{\sigma_b^2}{\sigma_b^2 + \sigma^2} \tag{13}$$

Pairwise Spearman's rank correlations were computed using the MATLAB function corr. Significance levels were corrected for false discovery rate (FDR) using the Benjamini-Hochberg method [48] with a FDR of 5%.

3 Results

3.1 Behaviour of the HCL model: sensitivity of sleep duration and timing to model parameters and light availability patterns

Before fitting, we perform a sensitivity analysis of the nonlinear HCL model to understand the underlying mathematical structure, deduce relevant parameters and assess which parameters are identifiable from available data.

We find parameters largely separate into those that primarily affect sleep duration and those that affect sleep timing. Sleep timing is also affected by light exposure. Results are summarised in Fig. 2 and Supplementary Figs. S3 and S4 and discussed further below.

3.1.1 Parameters that affect sleep duration

For the five sleep-wake regulation parameters listed in Table 1, sleep duration is primarily controlled by μ (the upper asymptote) and $\overline{H}_0$ (the mean level of the wake propensity rhythm). Both of these parameters affect sleep duration through a similar mechanism. Either increasing $\overline{H}_0$ or decreasing μ result in a net reduction in the rate of increase of sleep homeostasis during wake. The slower rise in sleep homeostasis during wake results in longer wake periods and consequently shorter sleep periods.

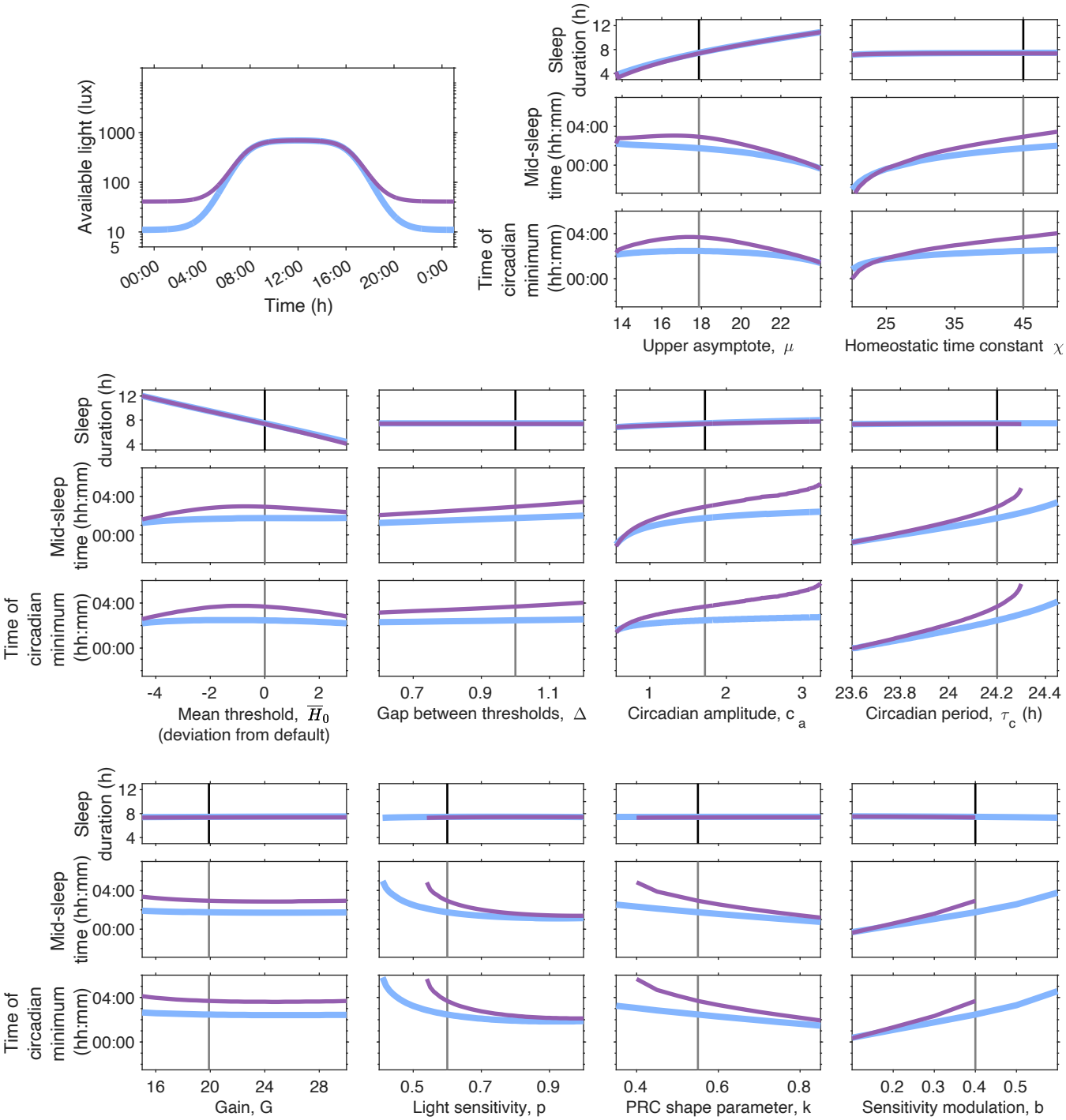

Figure 2: **Dependence of model predicted sleep duration and timing on model parameters and light exposure.** The top left panel shows two different light availability profiles. In all other groups of panels, the dependence of sleep duration (top panels), mid-sleep time (middle panels) and time of the circadian minimum (bottom panels) for the two light profiles are shown. The black vertical line indicates the default parameter values. Results are shown for all five sleep-wake parameters and five of the circadian-light parameters. The equivalent graphs for the remaining four circadian-light parameters are shown in Supplementary Fig. S3.

The parameter χ, which simultaneously affects the rate of decay of sleep homeostasis during sleep and the rate of rise of sleep homeostasis during wake, has little impact on sleep duration over the range considered. Similarly, sleep duration is insensitive to the gap between the thresholds, Δ over the range considered. Circadian amplitude c_a, has a small effect on sleep duration.

There are ten light / circadian parameters listed in Table 1, although only nine of them are independent since f and τ_c only appear in the combination $f\tau_c$. All of these nine parameters act to change the timing of the circadian rhythm relative the 24-hour day with barely observable changes to the amplitude or shape of $C(t)$ (graphs for five of these nine parameters are shown in Fig. 2, the remaining four are shown in Supplementary Fig. S3. There is therefore no mechanism for any of these parameters to change sleep duration.

3.1.2 Parameters that affect sleep timing

Each of the five sleep-wake regulation parameters affects sleep timing to a greater or lesser extent. The change in sleep timing occurs through two mechanisms. First, where there are changes to sleep duration there are consequently also changes to the duration of the light period. Second, all five parameters alter the internal relationship between sleep and circadian rhythmicity to some extent (see Supplementary Fig. S4). The fixed light availability profile used here combined with the gating effect of sleep, mean that a shift in the circadian minimum towards the end of the sleep episode results in more bright light earlier in the circadian day. More bright light earlier in the circadian day means more light at times when light acts to speed up the circadian clock, leading to earlier sleep timing.

Each of the nine independent circadian parameters adjust sleep timing by changing the magnitude or the effect of light on the circadian pacemaker, consistent with their definitions in Table 1. The parameter τ_c, giving the intrinsic period of the circadian clock. Consistent with entrainment theory [49], sleep timing is later for longer τ_c.

The parameters α_0, p, I_0 simultaneously alter the magnitude of the light signal, the rate of saturation of photoreceptors and the overall receptor saturation level. It is not necessarily obvious what the net effect will be, since in general if, for example, the signal to the SCN, $B(t)$ is scaled up, then it will 'speed up' the clock more at some phases of

the circadian cycle and 'slow down' the clock more at others. The parameter β alters the rate of decay of activated photoreceptor. Larger values of β will in general lead to smaller values for n and thus larger values for $\hat{B}$ and B. k alters the shape of the velocity response curve that describes the effect that light has on the velocity of the clock as a function of the phase of the clock. b alters the timing when the clock is most sensitive to light. The parameter γ alters the speed at which the van der Pol oscillator returns to the limit cycle. Under the entrained conditions here with a smooth, regular periodic input it has little impact on mid-sleep timing.

3.1.3 The effect of light on sleep timing

The sensitivity results presented in Fig. 2 and Fig. S3 are shown for two different example light profiles. These two profiles have a similar shape, with light switching smoothly between a high level during daylight hours and a low level at all other times, representing 24-hour access to electric light. The duration of the daylight hours was selected to approximate the natural photoperiod around the equinoxes. We refer to the low level as 'evening light' because the alignment of sleep with respect to daylight hours means that the primary effect of the low level of light is after dusk. The two profiles differ in the amount of available evening light. Note that the model assumes that light is available during wake, but is turned off when the switch to sleep occurs and turned back on when the switch to wake occurs. The consequent 'self-selection' means that sleep timing controls the timing of light, resulting in feedback between sleep and circadian rhythmicity [13, 14]. This feedback has a 'destabilising' effect since even if the input light signal is periodic the signal reaching the retina will not be periodic unless sleep is periodic.

In all cases, increasing evening light shifts sleep time later with no visible impact on sleep duration. However, the magnitude of the delay is dependent on the parameter values. For the default parameters, the increase in evening light results in a delay of approximately one hour because increasing light during the evening results in more light exposure during times when light acts to slow down the circadian clock. For the intrinsic period τ_c, longer circadian periods result in greater sensitivity to evening light, consistent with [14].

In general, sleep timing is more sensitive to the values of the parameters for evening light of 40 lux than of 10 lux because of the gating effect of sleep on light exposure.

3.1.4 The same phenotype may have different driving factors

Fig. 2 highlights that multiple model parameters lead to a particular sleep phenotype. For example, an individual may have a late sleep timing preference because of their physiology (e.g. a long intrinsic circadian period) or because of their pattern of light exposure. Different phenotypes and associated model parameters are summarised in Table 2.

For given sleep timing and duration data, there may therefore be multiple ways to explain the same data. With access to individual light exposure data, it should however, be possible to separate endogenous physiological factors from exogenous environmental factors.

Sleep phenotype	**Physiological / environmental interpretation**	**Associated model parameter**
Long sleep	Rapid increase of homeostatic sleep pressure during wake. High sleep need.	High value of μ.
	Low level of wake propensity.	Low mean level of thresholds $\overline{H}_0$
Short sleep	Slow increase of homeostatic sleep pressure during wake. Low sleep need.	Low value of μ.
	High level of wake propensity.	High mean level of thresholds $\overline{H}_0$.
Early sleep	Low circadian amplitude.	Small circadian amplitude c_a.
	Intrinsic period shorter than 24 h giving a tendency to phase advance in the absence of zeitgebers of sufficient strength.	Short τ_c.
	Insensitive to the effect of light in the evening and / or sensitive to bright light during the day.	High value of p, high value of k.
		High levels of bright light during the day and / or low levels of light in the evening.
Late sleep	High circadian amplitude.	High c_a.
	Intrinsic period longer than 24 hours giving a tendency to phase delay in the absence of zeitgebers of sufficient strength.	Long τ_c.
	Sensitive to the effects of light in the evening and / or insensitive to the effects of bright light during the day.	Low value of p or k or high value of b.
	Low levels of light during the day or too much light in the evening.	
Irregular sleep	Low circadian amplitude.	Low c_a.
	Low zeitgeber strength	Low levels of light during the day or too much light in the evening.

Table 2: **Different sleep phenotypes, associated model parameters and the physiological and environmental interpretation.**

3.2 Identifying endogenous factors driving sleep phenotypes: combining the HCL model with data

Next we demonstrate that we can match individual average sleep duration and timing by fitting for two pairs of parameters namely (μ, τ_c) and (μ, c_a). The choice of these pairs is motivated by the sensitivity analysis and aims to provide sufficient flexibility to fit to a wide range of sleep durations and mid-sleep times. Essentially, for each pair, one parameter primarily determines sleep duration (μ), whilst the other primarily determines sleep timing (τ_c or c_a). We fit two pairs of parameters to illustrate some of the pros and cons of parameter selection.

3.2.1 Observed sleep and circadian phenotypes

In order to test the HCL model, it was important to consider individuals who exhibit a range of different phenotypes. In our cohort, sleep timing (mid-sleep ICC = 0.65, sleep onset ICC = 0.56, sleep offset ICC = 0.61), duration (ICC = 0.47) and day-to-day variability differed between individuals. Five examples are shown in Fig. 3 and include: an early sleeper (mid-sleep time $\pm$ standard deviation 02:11 hh:mm $\pm$ 0:26 h:mm; sleep onset 22:54 hh:mm $\pm$ 0:30 h:mm; sleep offset 05:28 $\pm$ 0:49 h:mm); a late sleeper (mid-sleep time 04:18 hh:mm $\pm$ 0:42 h:mm; sleep onset 00:52 $\pm$ 0:47 h:mm; sleep offset 07:43 hh:mm $\pm$ 0:55 h:mm); a long sleeper (sleep duration $\pm$ standard deviation 8.92 h $\pm$ 0.44 h); a short sleeper (sleep duration 4.67 h $\pm$ 0.74 h) and an irregular sleeper (sleep duration 5.37 h $\pm$ 1.85 h). Also shown are boxplots for sleep duration and mid-sleep time for each participant and the distributions of the participant average and standard deviation in mid-sleep time and sleep duration for the whole cohort. Equivalent boxplots and distributions for sleep onset and offset are shown in Supplementary Fig. S5.

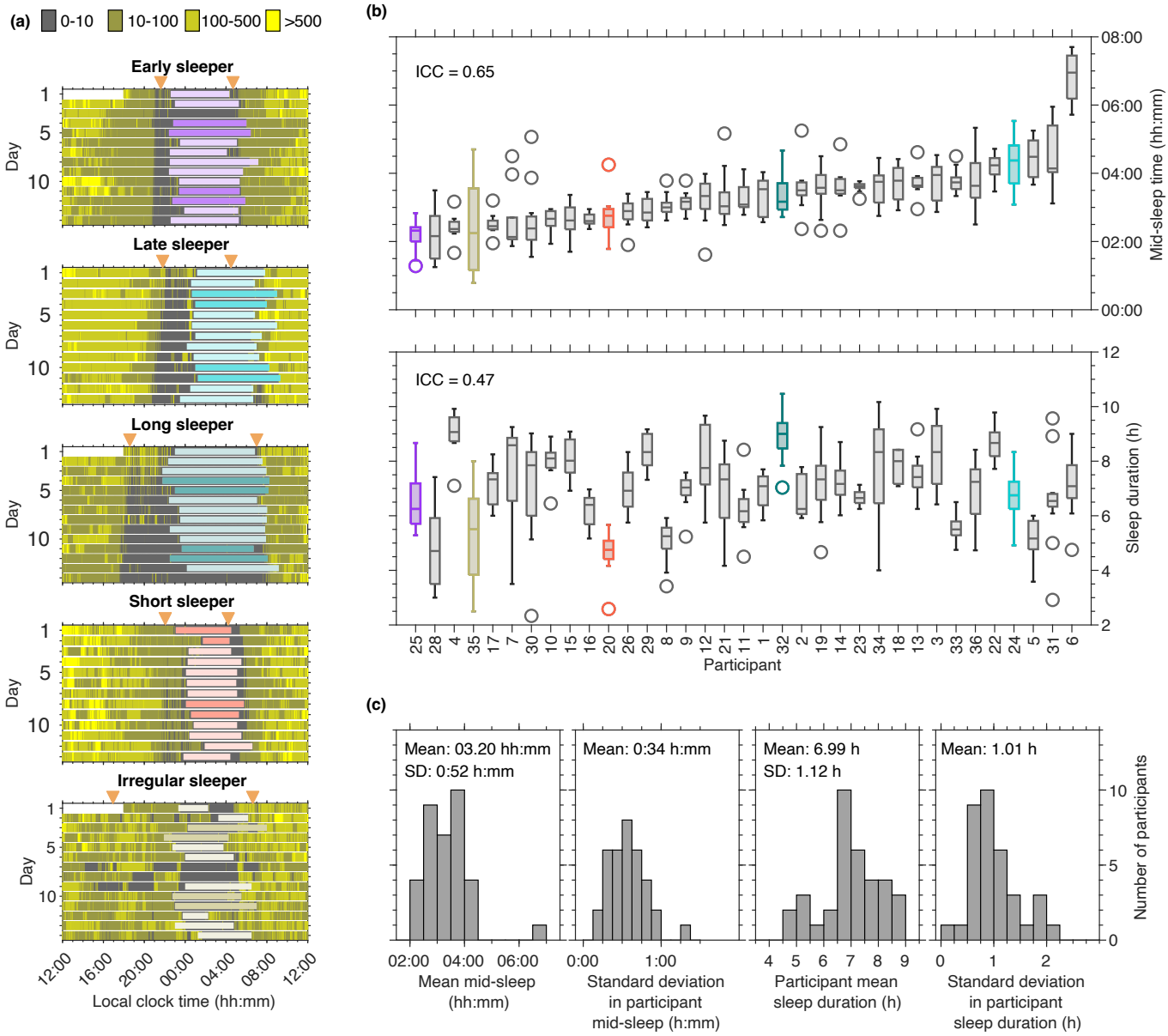

Figure 3: **Example of field data for five different sleep phenotypes and summary information for sleep timing and duration for each participant and the cohort.** Panels (a) show daily sleep timing (coloured horizontal bars, darker colouring shows sleep on Friday and Saturday nights). Light exposure is shown in the background for four different intensity bands. Dusk and dawn are indicated by the orange triangles. Panels (b) show box plots for mid-sleep timing and sleep duration respectively for each participant. Participants have been ordered according to their mean mid-sleep time. Panels (c) show the distributions of mean participant mid-sleep, standard deviation of participant mid-sleep, mean participant sleep duration, and standard deviation of participant sleep duration respectively.

3.2.2 Personalised model parameters were retrieved for each participant

Model parameters (μ, τ_c) that matched average sleep timing and duration for every individual were found for all 34 participants. Fitting for (μ, c_a) successfully retrieved parameters for 27 of the participants when raw light data were used and 30 of the 34 participants when imputed light data were used. Example fits and the corresponding personalised HCL models are shown in Fig. 4. Scatter plots showing the values of all parameters can be seen in Fig. 5.

For both pairs of parameters (μ, τ_c) and (μ, c_a), the fitted values accurately predict average sleep duration and mid-sleep timing within individuals to within the tolerances set for the optimization algorithm i.e that the residual calculated as the square root of the sum of the squares of the error in the average sleep duration and the average mid-sleep was less than one minute, Supplementary Fig. S6.

There is a very little association between the homeostatic parameter and the circadian period τ_c, supporting the fact that they are capturing different phenotypic dimensions, see Fig. 5. There was a weak association between the homeostatic parameter and the circadian amplitude, c_a.

The method was robust in that fitted parameters calculated using data collected from the wrist were strongly associated with those calculated using data collected from the shoulder (Spearman's rho for the homeostatic parameters $\mu \geq 0.96$, and those for the timing parameters of c_a or $\tau_c \geq 0.79$), see Supplementary Figs. S7 and S8. Likewise, parameters fitted using imputed data were strongly associated with those calculated using raw data (Spearman's rho values for the homeostatic parameters ≥ 0.96, and for the timing parameters of τ_c or $c_a \geq 0.90$).

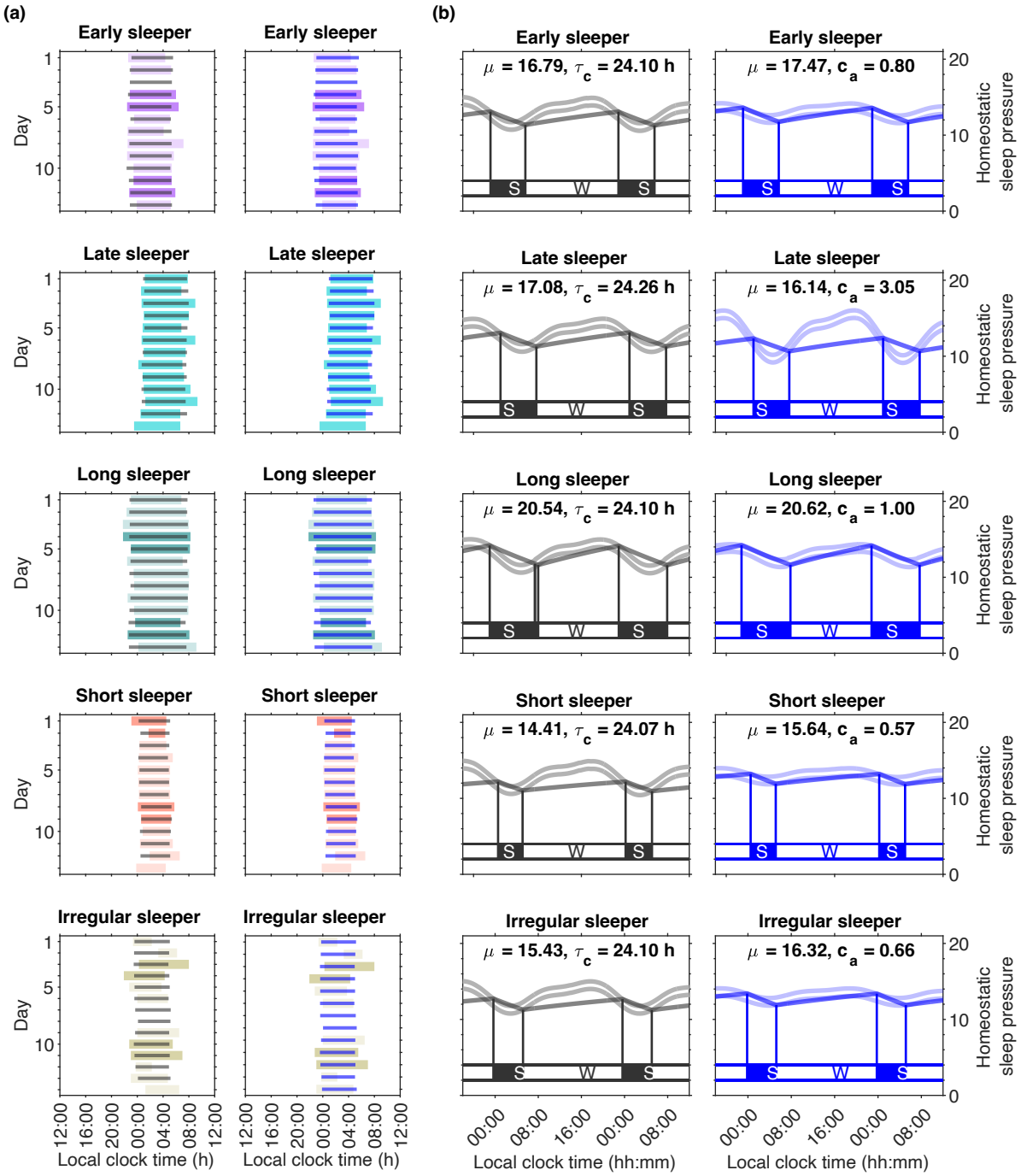

Figure 4: **Example fits to mid-sleep and sleep duration.** Panels (a) show simulated raster plots superimposed on the sleep timings (background bars coloured according to phenotype) as a result of fitting for the homeostatic parameter μ with the two alternative circadian timing parameters, i.e. the intrinsic circadian period, τ_c (grey, left hand column), and the circadian amplitude c_a (blue, right hand column). Panels (b) show the corresponding Homeostatic-Circadian-Light (HCL) models from fitting for (μ, τ_c) (grey, left hand column, c_a fixed at the default value of 1.72) and for (μ, c_a) (blue, right hand column, τ_c fixed at the default value of 24.20 h).

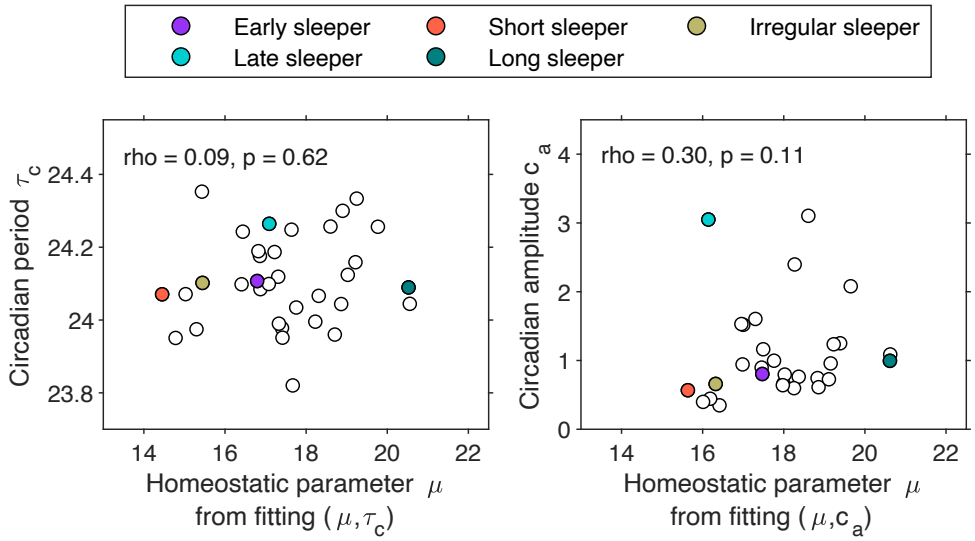

Figure 5: **Associations between the fitted homeostatic parameter μ with the two alternative circadian timing parameters, i.e. the intrinsic circadian period, τ_c and the circadian amplitude c_a.**

3.2.3 Relation of fitted parameters to phenotypes and physiology

Consistent with the sensitivity analysis, we find that for each pair of parameters, one parameter (μ) is strongly associated with sleep duration, with little association with mid-sleep time, see Table 3. Larger μ is associated with longer sleep duration, shifting sleep onset earlier and sleep offset later by approximately the same amount.

Also consistent with the sensitivity analysis, τ_c or c_a are (weakly) associated with mid-sleep time, with larger values of τ_c and larger values of c_a indicative of later mid-sleep time. We note that although we consider c_a as a 'timing' parameter, as found in the sensitivity analysis, it does also affect sleep timing and indeed there is a moderate association between c_a and sleep duration.

The two homeostatic parameters μ are strongly associated with each other (Spearman's rho 0.96), as are the timing parameters τ_c and c_a, (Spearman's rho 0.84), see Supplementary Fig. S6(c).

The HCL model is designed to capture the gross features of key physiological processes that regulate sleep. Since SWA activity is viewed as a marker of homeostatic sleep pressure it is interesting to note that the homeostatic parameter μ derived from

	Spearman correlations (p-values)				
	Mid-sleep time	Sleep onset	Sleep offset	Sleep duration	Slow wave activity
Homeostatic parameter μ (μ, τ_c fitting)	0.05 (0.794)	-0.60 **($<$0.001*)**	0.59 **($<$0.001*)**	0.99 **($<$0.001*)**	0.39 **(0.022)**
Homeostatic parameter μ (μ, c_a fitting)	0.01 (0.960)	-0.29 (0.121)	0.24 (0.196)	0.95 **($<$0.001*)**	0.31 (0.101)
Circadian timing parameter τ_c	0.33 (0.058)	0.12 (0.501)	0.33 (0.055)	0.13 (0.460)	0.06 (0.735)
Circadian timing parameter c_a	0.27 (0.151)	-0.24 (0.196)	-0.25 (0.181)	0.51 **(0.004*)**	0.20 (0.296)

Table 3: **Correlations between fitted parameters and sleep duration and timing measures and slow wave activity** Values in bold are significant (p $<$ 0.05). Those with an asterisk survive FDR correction (p $\leq$ 0.0125).

data collected at home is positively associated with SWA derived from data collected in the laboratory, see Table 3.

The width of the distribution of model intrinsic circadian periods τ_c (standard deviation 0.13-0.15 h) is similar to the width of the distribution of intrinsic circadian periods found in forced desynchrony experiments (standard deviation 0.2 h, [7]), see Fig. 6(a). However, since both circadian amplitude c_a and τ_c determine sleep timing, in the HCL model these interact, consequently the mean value for the distribution of fitted τ_c depends on the circadian amplitude. For example, when fitting for (μ, τ_c) a default value of c_a = 1.72 was used. Using a larger circadian amplitude results in later sleep timing. So fixing the default circadian amplitude at a higher value will tend to shift the distribution of τ_c to smaller values (earlier sleep timing) to compensate. Similarly, a smaller circadian amplitude shifts the distribution of τ_c to higher values, see Fig. 6(a). Hence, one cannot make inferences on the actual intrinsic circadian period as measured in a forced desynchrony protocol. We have kept the notation τ_c consistent with prior literature (e.g. [28]), but it is perhaps more appropriate to think of it as an endogenous timing parameter.

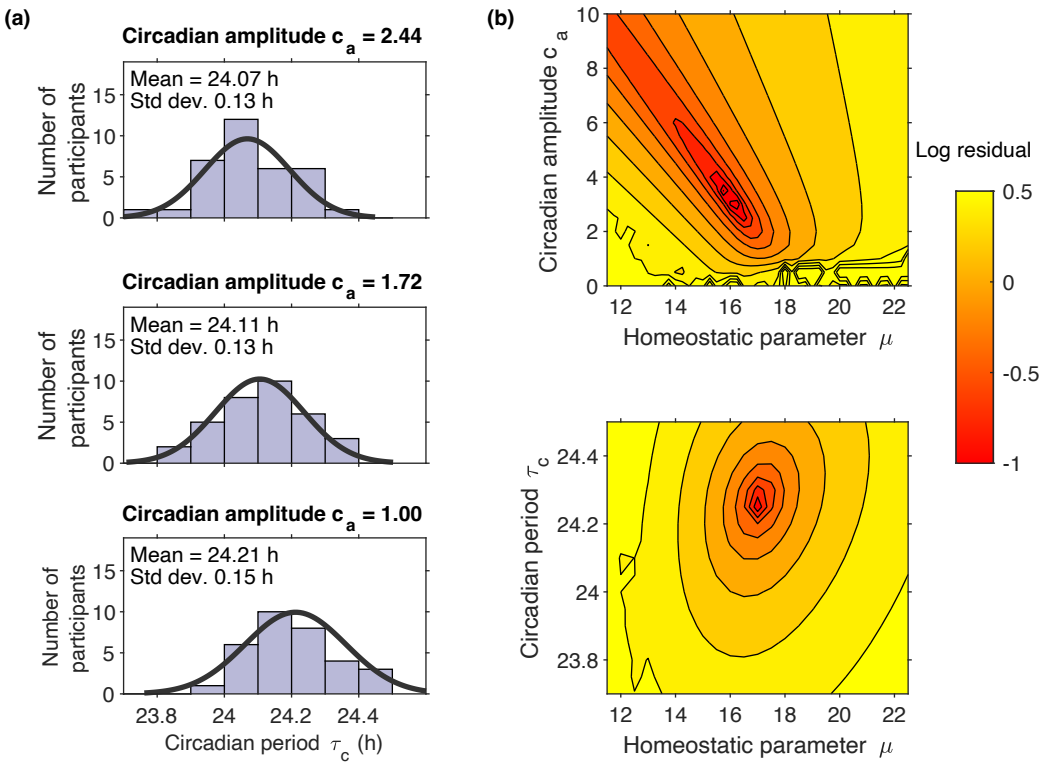

Figure 6: **Identifiability of model parameters** Panels (a) show histograms of fitted intrinsic circadian period τ_c for three different choices of the circadian amplitude c_a, highlighting the inter-dependence of these two parameters. Panels (b) show the residual for fitting simultaneously to mean sleep duration and mean mid-sleep timing for one participant as a function of (μ, τ_c) and (μ, c_a) i.e. the sum of the squared differences between mean sleep duration and mean simulated sleep duration and mean mid-sleep time and mean simulated mid-sleep time.

3.2.4 Day-to-day variability

Average sleep duration and mid-sleep timing are accurately matched, but Fig. 4 highlights that there are differences between reported and model predicted day-to-day mid-sleep timing. Specifically, average standard deviation in participant reported mid-sleep time was 0:34 h:mm (see Fig. 3), whereas model predictions were 0:09 h:mm and 0:08 h:mm for (μ, τ_c) and (μ, c_a) respectively. Similarly, average participant standard deviation in reported sleep duration was 1.01 h (see Fig. 3, whereas model predictions were 0.07 h and 0.05 h for (μ, τ_c) and (μ, c_a) respectively. Considering the day-to-day differences between predicted and reported mid-sleep time, sleep onset time, sleep offset time and sleep duration, there were no significant differences between the goodness of

fit for (μ, τ_c) and (μ, c_a).

3.2.5 Which parameters to fit?

Overall, it was faster and more people were successfully fitted with (μ, τ_c) than (μ, c_a). Plotting the residual error as a function of the parameters, see Fig. 6(b) for one example, illustrates the reasons. For (μ, τ_c) there is a well-defined minimum value since sleep duration is independent of τ_c and mid-sleep is only weakly dependent on μ. For (μ, c_a) the minimum value is much less well-defined with a line along which reducing μ and increasing c_a give similar (low) values for the residuals. The inability to fit for (μ, c_a) for a few people is because for some light exposure patterns, the model is unable to simulate late enough sleep timing by only varying amplitude (and (μ)). This is because for large amplitudes, sleep timing is very insensitive to further changes in amplitude. This can be seen in the sensitivity analysis shown in Fig. 2 and understood with reference to the HCL model for the late sleeper shown in Fig. 6(b). In the HCL model, once the amplitude is large, at the points on the thresholds where sleep onset and offset occur, the gradient is very steep. Increasing the amplitude further makes the gradient steeper, making the time of sleep onset or sleep offset relatively insensitive to the upward or downward trajectory of the sleep homeostat. Hence, from this point we focus on the pair of parameters (μ, τ_c).

3.3 Exogenous factors determining sleep phenotypes

The endogenous parameters found by fitting the HCL model were strongly associated to sleep duration, sleep onset and sleep offset, but the association with mid-sleep timing was weak. Given the acknowledged importance of light in determining the timing of the biological clock and sleep, it is interesting to examine if individual patterns of light exposure correlate with sleep phenotypes.

3.3.1 Observed light exposure patterns

Light exposure patterns varied between individuals, see Fig. 7, where light exposure patterns for the five example sleep and circadian phenotypes are shown. Also shown are boxplots indicating the daily hours of bright light and daily measure of the effect

of light on the biological clock for each participant and distributions of the participant average and standard deviation in both light measures. An average of less than one hour of bright light (light $>$ 500 lux) a day was recorded at the wrist for most people (24 individuals i.e. 71%). Equivalent boxplots and distributions for the remaining four light metrics which further quantify overall light exposure and timing of light exposure are shown in the Supplementary Figs S9 and S10.

3.3.2 Associations between sleep timing and duration and light exposure

Mid-sleep time was not associated with any of the three metrics which quantify the amount of light received, i.e. the mean lux value, the geometric mean value or the mean number of daily hours of bright light (light $>$ 500 lux) (Spearman's rho all of magnitude less than or equal to 0.11), see Supplementary Table S2. Mid-sleep was weakly associated with the two measures of the timing of light i.e. the time of day at which half the total daily amount of light (lux) has been received (Spearman's rho 0.29) and the time of day at which half the total daily amount of light (log (lux$+$1) has been received (Spearman's rho 0.32). The associations are such that later light exposure is associated with later mid-sleep time. The novel measure of the biological effect of light was also weakly assciated with mid-sleep time (Spearman's rho -0.25).

All measures that quantify the amount of light are significantly associated with each other ($p<0.001$), as are the measures that quantify the timing of light ($p<0.001$), see Table S2. The novel measure of the biological effect of light was found to be moderate to strongly associated with all measures of the amount of light (Spearman's rho from 0.56 to 0.76) and with one of the timing measures (time of half-light measured in lux, Spearman's rho -0.55).

Sleep duration was not found to associate with any of the six light metrics (Spearman's rho all of magnitude less than or equal to 0.16), see Supplementary Table S2.

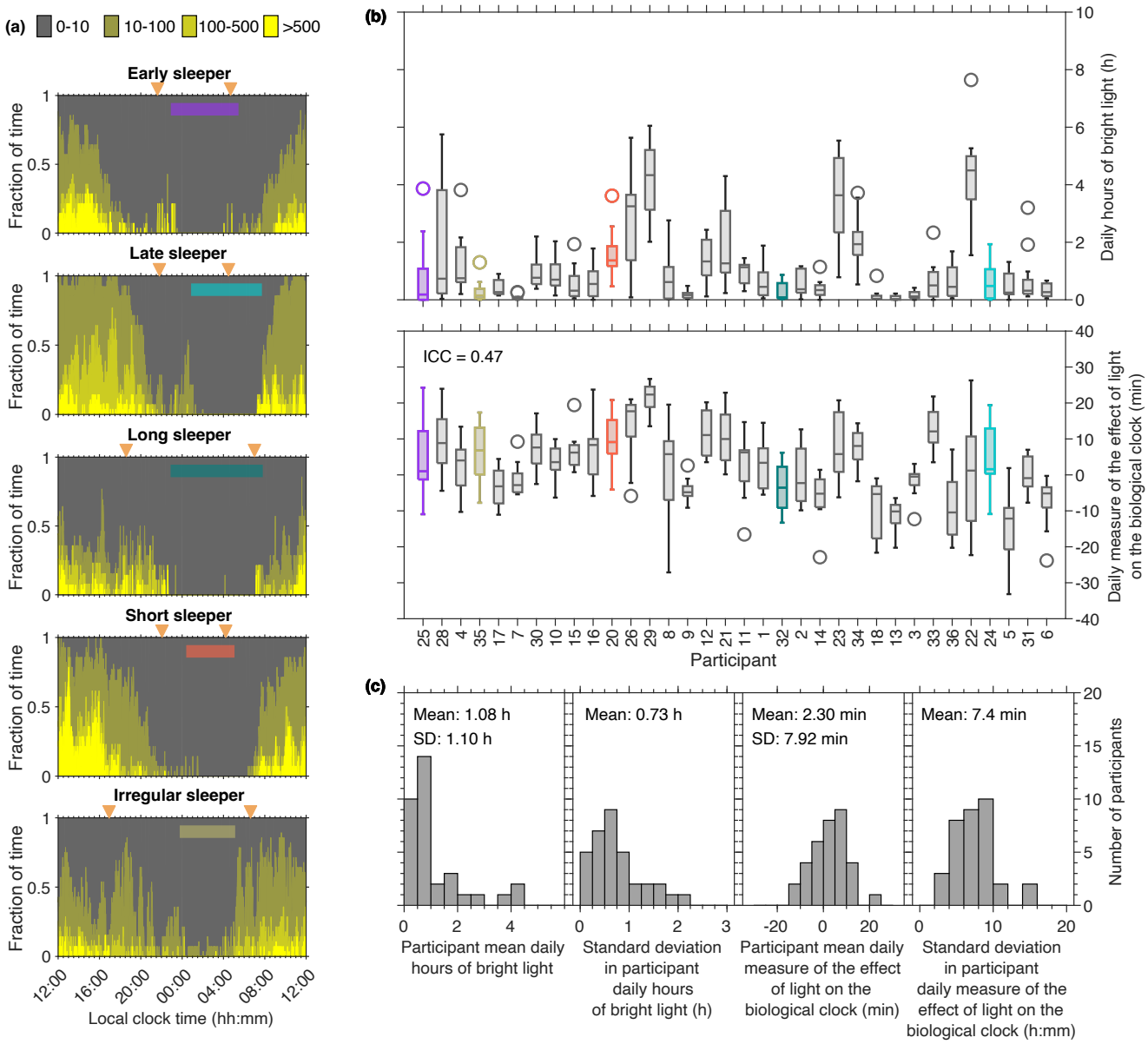

Figure 7: **Example light data for five different sleep phenotypes and summary information for each participant and the cohort.** Panels (a) show examples of light exposure patterns for five different participants. The regions show the fraction of time spent in each of four intensity bands. Average sleep timing is marked by the coloured horizontal bars. Dusk and dawn are indicated by the orange triangles. Panels (b) show box plots for the daily hours of bright light and the daily measure of the effect of light on the circadian clock respectively for each participant. Participants are ordered by their mean mid-sleep time, as in Fig. 3. Panels (c) show the distributions of the mean and standard deviation of participant daily hours of bright light and the mean and standard deviation of the participant daily measure of the effect of light on the biological clock.

3.4 Relative contribution of endogenous physiological and exogenous environmental drivers of sleep phenotypes

To summarise, individual differences in sleep timing and duration were successfully fitted by the HCL model. Individual differences in sleep duration were found to be primarily driven by the modelled homeostatic sleep process (see Table 3). However, for individual differences in mid-sleep timing the explanation is more complex. In a first approach to answer this question, we have not only quantified the contributions of parameters describing endogenous physiological factors but also considered measures of exogenous light exposure.

Physiological factors driving sleep timing are in principle captured by the fitted model circadian timing parameter, τ_c. But we found that τ_c is only weakly associated with sleep timing (see Table 3 Indeed, τ_c was only found to explain a small amount of the variance in mid-sleep time, see in Fig. 8(a)). For example, although the late sleeper labelled as 'late phenotype (physiology)' in Fig. 8(a)) has a greater value of τ_c than the 'early phenotype', there are those with a relatively small value of τ_c who have a similarly late mid-sleep time, for example the individual labelled as 'late phenotype (environmental light)'.

Similarly, mid-sleep timing was not associated with measures of total environmental light exposure and was only weakly associated with measures of the timing of light (see Table 3). As shown in Fig. 8(b), our novel metric of the biological effect of light also only explains a small amount of the variance in mid-sleep time, with the early phenotype and late phenotype (physiology) examples having similar values for the biological effect of light.

However, taken together, the fitted parameter τ_c and our novel metric of light timing explain a large amount of the variance (adjusted Rsquared, 0.62), see Fig. 8(c). Here we see that our modelling approach suggests that of the two highlighted late sleep examples, one has late phenotype as a result of their physiology and one has a late phenotype as a result of their light exposure pattern.

In Fig. 8(c) along with the data the fitted general linear model,

$$\text{Mid-sleep time} = 4.06 - 5.94 \times \mathrm{L}_{\mathrm{time}} + 5.37 \times (\tau_c - 24.2), \tag{14}$$

where L_{time} is the measure of the biological effect of light in hours, is shown. This

model estimates the deviation from a mid-sleep time of 4.06 h that occurs either as a result of a deviation of the fitted intrinsic period from 24.2 hours or as a result of an environmental light signal that for the average person would advance / delay the biological clock relative to a 24-hour day. Including higher order interaction terms does not improve the fit.

The novel light timing metric combined with τ_c explains more of the variance than combining either of the other two light timing metrics with τ_c. (Fitting for τ_c with the time of half light gives an adjusted rsquared value of 0.51, and with the time of half-loglight gives an adjusted rsquared value of 0.09.)

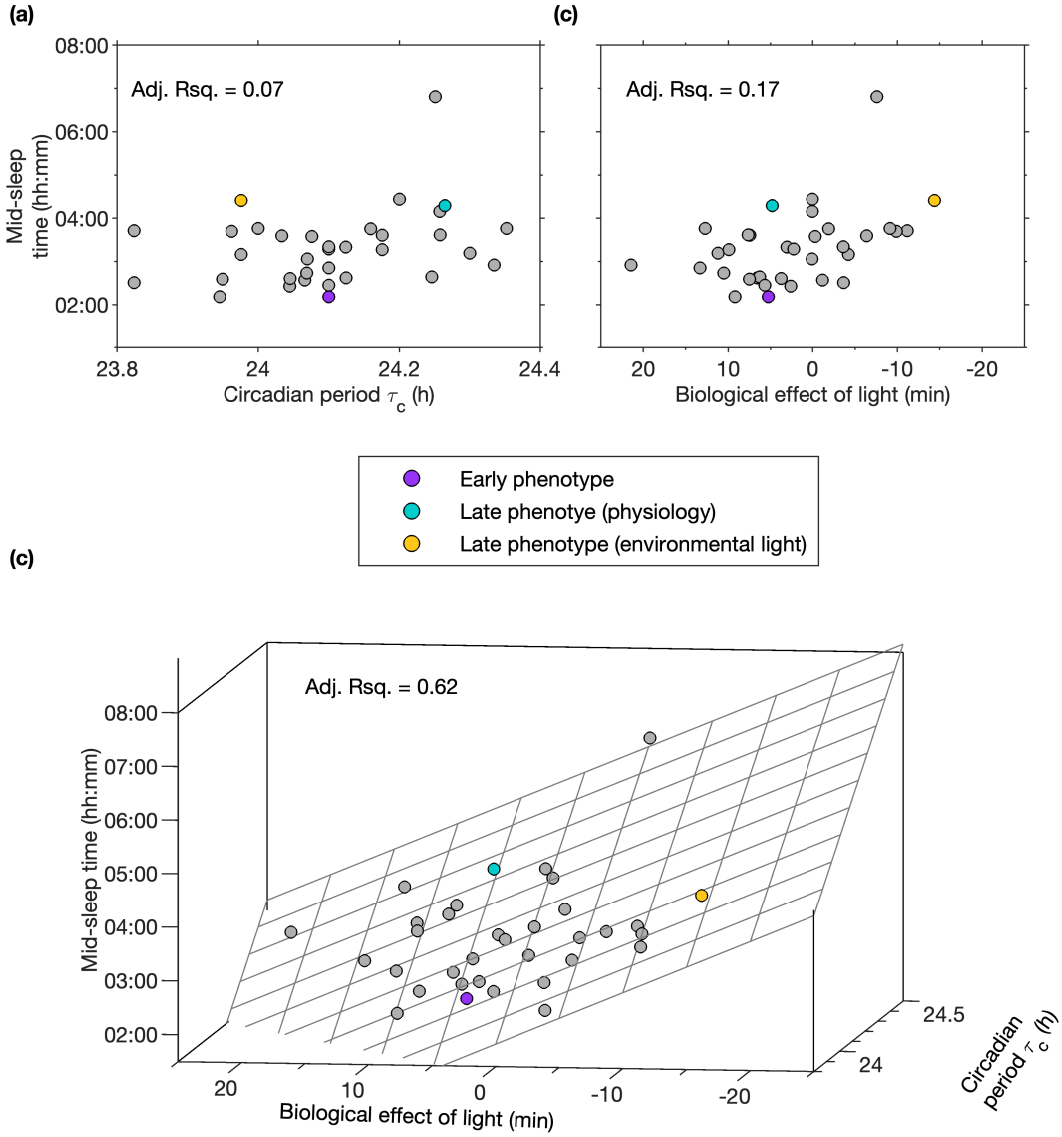

Figure 8: **Quantifying the effect of environment versus physiology in the determination of mid-sleep time.** Panel (a) shows that the fitted physiologically-motivated timing parameter, the intrinsic circadian period τ_c, also only explains a small amount of the between participant variance. Similarly, panel (b) shows that considering the novel metric of the biological effect of light also explains only a small amount of the between participant mid-sleep time. In contrast, combining environmental and physiological variables (panel (c)), explains a large amount of the variance.

4 Discussion

Individuals may have preferred sleep timing that conflicts with their work schedules or personal desired sleep timing. But preferred sleep timing is a result of both endogenous physiological factors and exogenous light exposure patterns. In order to design personalised interventions to either advance or delay sleep timing an important step is to be able to dissect the relative contributions of these separate drivers to sleep phenotype. While changing endogenous physiology drivers is challenging, exogenous light exposure patterns can readily be altered and offer a practical solution for implementation of targeted interventions at scale.

Here we have developed a new simplified model for the Homeostatic-Circadian-Light (HCL) regulation of sleep timing and duration. Critically, we have demonstrated that the model can take 7-10 days of measured light exposure patterns and accurately and robustly find personalised parameters that match individual mean sleep-timing and duration. The identification of personalised model parameters along with the use of a novel measure of the effect of light on the biological clock enabled us to suggest, for the first time, the relative contributions of endogenous physiological versus exogenous environmental light drivers for mid-sleep timing.

The HCL model could readily be combined with wearables to provide individual advice on appropriate light availability patterns to achieve a target preferred sleep timing, as envisaged in the interventional framework proposed in [30].

Relation of the HCL model to other models

The HCL model was motivated by elements of the original two-process model [23] along with elements of recent neuronal models [28, 14]. Critically, unlike the two-process model it includes light to allow for the modelling of sleep-timing relative to the 24-hour day. Compared with recent neuronal models, the HCL model is simpler (four instead of six first order differential equations, 14 instead of 23 parameters), and computationally faster. The relative simplicity of the HCL model makes it straightforward to understand the role of each parameter, see Fig. 2 and Table 2. In addition the form of the coupling between sleep and circadian rhythmicity has been updated to better reflect data on the circadian variation of wakefulness, [6]. The model can also easily be extended to

include the role of social constraints (see the Supplementary Material).

Nine of the parameter relate to light-circadian aspect of the model. In spite of the number of parameters, we elected to retain the circadian-light model as specified by Kronauer and colleagues [26, 27]. A recent systematic study suggests that simplifying the circadian-light model results in a poorer fit to circadian phase data [50]. Another popular version of the Kronauer model is that produced by St Hilaire and co-workers [51]. Given the similarities between the different models, we would expect results to be similar whichever version is used. Although we note that if a different light-circadian model was used then the form of the sleep-circadian coupling function $C(t)$ given in equation (11) would also need to be updated.

Identifying factors underlying sleep phenotypes: challenges

Based on our sensitivity analysis, we fitted for one parameter encoding sleep duration (e.g. μ) and one encoding sleep timing (e.g. c_a or τ_c). Over several weeks, with sufficiently variable light from one week to the next, it may be possible to capture the nuanced different dynamic responses that may occur with different parameter combinations. Further types of data could enable the identification of more parameters. For example, knowledge of individual alignment of circadian phase with sleep could enable fitting of one additional sleep-wake parameter. Day-to-day variation in circadian phase could enable both a light sensitivity parameter and intrinsic circadian period to be determined as suggested for light-circadian models in [52].

Limitations

The current work is based on only one to two weeks of data. As longer timeseries become available, the same approaches can be used to assess whether model parameters vary with time. The current model captures individual differences in the average timing and duration of sleep but not the night-to-night variation in sleep timing and duration. Night-to-night variation could be modelled by the inclusion of stochasticity (e.g. see [53]) or further knowledge on the reasons that determine when people go to bed. Including circadian amplitude dependence on light [54] may also be important and enable reported dependence of sleep duration on light [55] to be captured. Finally, the current work uses lux as the measure for the effects of light on the circadian clock. The

discovery of the contribution of the photopigment melanopsin to the regulation of circadian rhythmicity implies that melanopic lux rather than photopic lux may be the relevant measure [56], although we do not expect this to substantially affect our results.

Outlook

Our modelling approach suggests that it is possible to separate endogenous physiological drivers from exogenous environmental drivers for different sleep phenotypes. This implies that there is considerable scope for appropriate models to be incorporated in clinical guidance and intervention design for sleep disorders. Models could even to be implemented as standard in smart watch software for general guidance for day-to-day living or coupled to human centric lighting systems to automatically tune the light environment to optimise the alignment of sleep and circadian rhythmicity to a desired time.

Acknowledgements

The authors would like to thank the study participants in the sleep study and all others responsible for collecting and collating the data including Dr Hana Hassanin, Damion Lambert, Surrey Clinical Research Facility staff, and Daniel Barrett. We thank Giuseppe Atzori for scoring the PSG recordings.

This work is part of the sleep and circadian disturbance project (PI DJD) supported by the UK Dementia Research Institute which receives its funding from DRI Ltd, funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK.

Author contributions

Conceptualization (ACS, DJD); data curation (TRG, KKGR, VLR, CdM); formal analysis (ACS, TRG, SC); funding acquisition (DJD); investigation (ACS, TRG, SC, VLR, KKGR, CdM); methodology (ACS, DJD); software (ACS, TRG, SC); visualization (ACS, DJD, TRG); writing – original draft preparation (ACS, DJD); writing – review & editing (ACS, DJD, TRG, SC, KKGR, VLR, CdM).

Disclosure statement

DJD and ACS are consultants to F. Hoffmann-La Roche, Ltd.

Data availability statement

All software and data required for validation and replication of the results in this paper will be made available on acceptance of this manuscript.

References

[1] The Lancet. Waking up to the importance of sleep. *The Lancet*, 400:973, 2022.

[2] Meyer N, Harvey AG, Lockley SW, and Dijk D-J. Circadian rhythms and disorders of the timing of sleep. *The Lancet*, 400:1061–1078, 2022.

[3] Walker WH 2nd, Walton JC, DeVries AC, and Nelson RJ. Circadian rhythm disruption and mental health. *Transl Psychiatry*, 10:28, 2020.

[4] Logan RW and McClung CA. Rhythms of life: circadian disruption and brain disorders across the lifespan. *Nat Rev Neurosci*, 20:49–65, 2019.

[5] Katori M, Shi S, Ode KL, Tomita Y, and Ueda HR. The 103,200-arm acceleration dataset in the uk biobank revealed a landscape of human sleep phenotypes. *Proc Natl Acad Sci*, 119:e2116729119, 2022.

[6] Dijk DJ and Czeisler CA. Contribution of the circadian pacemaker and the sleep homeostat to sleep propensity, sleep structure, electroencephalographic slow waves, and sleep spindle activity in humans. *J Neurosci*, 15:3526–3538, 1995.

[7] Duffy JF, Cain SW, Chang AM, Phillips AJ, Münch MY, Gronfier C, and et al. Sex difference in the near-24-hour intrinsic period of the human circadian timing system. *Proc Natl Acad Sci*, 108:15602–8, 2011.

[8] Minors DS, Waterhouse JM, and Wirz-Justice A. A human phase-response curve to light. *Neurosci Lett*, 133:36–40, 1991.

[9] Duffy JF and Czeisler CA. Effect of light on human circadian physiology. *Sleep Med Clin*, 4:165–177, 2009.

[10] Klepeis NE, Nelson WC, Ott WR, Robinson JP, Tsang AM, Switzer P, and et al. The national human activity pattern survey (nhaps): a resource for assessing exposure to environmental pollutants. *Journal of Exposure Science and Environmental Epidemiology*, 11:231–252, 2001.

[11] Yetish G, Kaplan H, Gurven M, Wood B, Pontzer H, Manger PR, and et al. Natural sleep and its seasonal variations in three pre-industrial societies. *Curr Biol*, 25:2862–2868, 2015.

[12] Stothard ER, McHill AW, Depner CM, Birks BR, Moehlman TM, Ritchie HK, and et al. Circadian entrainment to the natural light-dark cycle across seasons and the weekend. *Curr Biol*, 27:508–513, 2017.

[13] Klerman EB, Dijk D-J, Kronauer RE, and Czeisler CA. Simulations of light effects on the human circadian pacemaker: implications for assessment of intrinsic period. *Am J Physiol*, 270:R271–R282, 1996.

[14] Skeldon AC, Phillips AJK, and Dijk D-J. The effects of self-selected light-dark cycles and social constraints on human sleep and circadian timing: a modeling approach. *Sci Rep*, 7:45158, 2017.

[15] Wittmann M, Dinich J, Merrow M, and Roenneberg T. Social jetlag: Misalignment of biological and social time. *Chronobiol Int*, 23:497–509, 2006.

[16] Hsu PK, Ptáček LJ, and Fu YH. Genetics of human sleep behavioral phenotypes. *Methods Enzymol*, 552:309–24, 2015.

[17] Jones SE, van Hees VT, Mazzotti DR, Marques-Vidal P, Sabia S, van der Spek A, and et al. Genetic studies of accelerometer-based sleep measures yield new insights into human sleep behaviour. *Nature Comm.*, 10:1585, 2019.

[18] Aeschbach D, Cajochen C, Landolt H, and Borbély AA. Homeostatic sleep regulation in habitual short sleepers and long sleepers. *Am J Physiol*, 270:R41–53, 1996.

[19] Duffy JF, Rimmer DW, and Czeisler CA. Association of intrinsic circadian period with morningness-eveningness, usual wake time, and circadian phase. *Behav Neurosci*, 115:895–9, 2001.

[20] Lazar AS, Santhi N, Hasan S, Lo JC, Johnston JD, Von Schantz M, Archer SN, and Dijk DJ. Circadian period and the timing of melatonin onset in men and women: predictors of sleep during the weekend and in the laboratory. *J Sleep Res*, 22:155–9, 2013.

[21] Kronauer RE, Czeisler CA, Pilato SF, Moore-Ede MC, and Weitzman ED. Mathematical model of the human circadian system with two interacting oscillators. *Am J Physiol*, 242:R3–R17, 1982.

[22] Borbély AA. A two process model of sleep regulation. *Hum Neurobiol*, 1:195–204, 1982.

[23] Daan S, Beersma DGM, and Borbély AA. Timing of human sleep: Recovery process gated by a circadian pacemaker. *Am J Physiol*, 246:R161–R183, 1984.

[24] Saper CB, Scammell TE, and Lu J. Hypothalamic regulation of sleep and circadian rhythms. *Nature*, 437:1257–63, 2005.

[25] Phillips AJK and Robinson PA. A quantitative model of sleep-wake dynamics based on the physiology of the brainstem ascending arousal system. *J Biol Rhythms*, 22:167–179, 2007.

[26] Kronauer RE, Forger DB, and Jewett ME. Quantifying human circadian pacemaker response to brief, extended and repeated light stimuli over the phototopic range. *J Biol Rhythms*, 14:501–515, 1999.

[27] Forger DB, Jewett ME, and Kronauer RE. A simpler model of the human circadian pacemaker. *J Biol Rhythms*, 14:533–538, 1999.

[28] Phillips AJK, Chen PY, and Robinson PA. Probing the mechanisms of chronotype using quantitative modeling. *J Biol Rhythms*, 25:217–227, 2010.

[29] Skeldon AC, Derks G, and Dijk D-J. Modelling changes in sleep timing and duration across the lifespan: Changes in circadian rhythmicity or sleep homeostasis? *Sleep Med Rev*, 28:96 – 107, 2016.

[30] Skeldon AC, Dijk D-J, Meyer N, and Wulff K. Extracting Circadian and Sleep Parameters from Longitudinal Data in Schizophrenia for the Design of Pragmatic Light Interventions. *Schiz Bull*, 48:447–456, 2021.

[31] Postnova S, Postnov DD, Seneviratne M, and Robinson PA. Effects of rotation interval on sleepiness and circadian dynamics on forward rotating 3-shift systems. *J Biol Rhythms*, 29:60–70, 2014.

[32] Postnova S, Lockley SW, and Robinson PA. Prediction of cognitive performance and subjective sleepiness using a model of arousal dynamics. *J. Biol. Rhythms*, 33:203–218, 2018.

[33] Hong J, Choi SJ, Park SH, Hong H, Booth V, Joo EY, and Kim JK. Personalized sleep-wake patterns aligned with circadian rhythm relieve daytime sleepiness. *iScience*, 24:103129, 2021.

[34] Diniz Behn CG and Booth V. A fast-slow analysis of the dynamics of rem sleep. *SIAM Journal on Applied Dynamical Systems*, 11:212–242, 2012.

[35] Piltz SH, Diniz Behn CG, and Booth V. Habitual sleep duration affects recovery from acute sleep deprivation: A modeling study. *J Theor Biol*, 504:110401, 2020.

[36] Dijk D-J, Duffy JF, Riel E, Shanahan TL, and Czeisler CA. Ageing and the circadian and homeostatic regulation of human sleep during forced desynchrony of rest, melatonin and temperature rhythms. *J Physiol*, 516:611–627, 1999.

[37] Dijk DJ and Czeisler CA. Paradoxical timing of the circadian rhythm of sleep propensity serves to consolidate sleep and wakefulness in humans. *Neurosci Lett*, 166:63–8, 1994.

[38] Daan S. Tonic and phasic effects of light in the entrainment of circadian rhythms. *Ann NY Acad Sci*, 290:51–59, 1977.

[39] Czeisler CA, Kronauer RE, Allan JS, Duffy JF, Jewett ME, Brown EN, and Ronda JM. Bright light induction of strong (type 0) resetting of the human circadian pacemaker. *Science*, 244:1328–1333, 1989.

[40] Roenneberg T, Kuehnle T, Pramstaller PP, Ricken J, Havel M, Guth A, and Merrow M. A marker for the end of adolescence. *Curr Biol*, 14:R1038–R1039, 2004.

[41] Roenneberg T, Allebrandt KV, Merrow M, and Vetter C. Social jetlag and obesity. *Curr Biol*, 22:939–943, 2012.

[42] The MathWorks Inc. Matlab version: 9.13.0 (r2022b), 2022.

[43] Johns MW. A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. *Sleep*, 14:540–545, 1991.

[44] Buysse DJ, Reynolds CF, Monk TH, Berman SR, and Kupfer DJ. The pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28:193 – 213, 1989.

[45] Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, and Morin CM. The Consensus Sleep Diary: Standardizing Prospective Sleep Self-Monitoring. *Sleep*, 35:287–302, 2012.

[46] Vaughn BV, Berry RB, Brooks R, Gamaldo C, Harding SM, Lloyd RM, and Quan SF add Troester MT. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. American Academy of Sleep Medicine, version 2. edition, 2018.

[47] Dijk D-J, Groeger JA, Stanley N, and Deacon S. Age-related reduction in daytime sleep propensity and nocturnal slow wave sleep. *Sleep*, 33:211–223, 2010.

[48] Benjamini Y and Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc B*, 57:289–300, 1995.

[49] Adler R. A study of locking phenomena in oscillators. *Proceedings of the IRE*, 34:351–357, 1946.

[50] Rea MS, Nagare R, Bierman A, and Figueiro MG. The circadian stimulus-oscillator model: Improvements to kronauer's model of the human circadian pacemaker. *Front Neurosci*, 16, 2022.

[51] St Hilaire MA, Klerman EB, Khalsa Sat Bir S, Wright KP, Czeisler CA, and Kronauer RE. Addition of a non-photic component to a light-based mathematical model of the human circadian pacemaker. *J Theor Biol*, 247:583 – 599, 2007.

[52] Stone JE, McGlashan EM, Quin N, Skinner K, Stephenson JJ, Cain SW, and Phillips AJK. The role of light sensitivity and intrinsic circadian period in predicting individual circadian timing. *J Biol Rhythms*, 35:628–640, 2020.

[53] Shochat T, Santhi N, Herer P, Dijk D-J, and Skeldon AC. Sleepiness is a signal to go to bed: data and model simulations. *Sleep*, 44:zsab123, 2021.

[54] Bano-Otalora B, Martial F, Harding C, Bechtold DA, Allen AE, Brown TM, and et al. Bright daytime light enhances circadian amplitude in a diurnal mammal. *Proc Natl Acad Sci*, 118:e2100094118, 2021.

[55] Brockmann PE, Gozal D, Villarroel L, Damiani F, Nuñez F, and Cajochen C. Geographic latitude and sleep duration: A population-based survey from the tropic of capricorn to the antarctic circle. *Chronobiol Int*, 34:373–381, 2017.

[56] Tekieh T, Lockley SW, Robinson PA, McCloskey S, Zobaer MS, and Postnova S. Modelling melanopsin-mediated effects of light on circadian phase, melatonin suppression and subjective sleepiness. *J Pineal Res*, 69:e12681, 2020.