

Personalized Data-Driven State Models of the Circadian Dynamics in a Biometric Signal

Chukwuemeka O. Ike^{*†a}, Yunshi Wen^{*†a}, John T. Wen^{*†}, Meeko M.K. Oishi[‡], Lee K. Brown[§], and A. Agung Julius^{*†}

Email: {ikec,weny2,juliua2}@rpi.edu

^{*}Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA

[†]Lighting Enabled Systems and Applications (LESA) Engineering Research Center, Troy, NY, USA

[‡]Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, USA

[§]Department of Internal Medicine, University of New Mexico, Albuquerque, NM, USA

^aThese two authors contributed equally.

Abstract—Circadian rhythms are endogenous 24-hour oscillations that are vital for maintaining our overall well-being. They are driven at a high level by a core circadian clock located in the brain, making their dynamics difficult to track. Various modeling approaches exist to predict the dynamics, but as the models are typically designed on population-level data, their performance is diminished on the individual level. This paper proposes a method for learning personalized latent state models, i.e., dynamical models that explicitly use latent state variables, that relate circadian input(s) to observable biometric signals. Our models combine an autoencoder with a recurrent neural network and use the pair to model the salient dynamics present in the data. We validate our method using experimental data, where the circadian input is light and the biometric data are actigraphy signals. We demonstrate that our method results in models with low-dimensional latent state that can accurately reconstruct and predict the observable biometric signals. Further, we show that the oscillation of the learned latent state agrees with the subjects' circadian clock oscillation as estimated with melatonin measurements.

Clinical relevance — This proposes a technique for personalized modeling of the circadian system with potential applications in feedback control and individualized circadian studies.

I. INTRODUCTION

Circadian (\sim 24-hour) rhythms are an ever-present phenomenon across biological organisms, and they play a vital role in the overall well-being of those organisms. In humans, these rhythms include the sleep-wake cycle, core body temperature (CBT), and the production cycles of a host of hormones. They are regulated by the core circadian clock in the brain, which is strongly driven by light exposure. When the circadian system is in sync with the environmental day-night cycle, the biological processes that depend on this synchrony run optimally. However, desynchrony or disruption of the cycle has been linked with multiple issues including reduced cognitive ability, cardiovascular disease, and various cancers [1], [2].

Circadian control refers to the process of manipulating or optimizing the circadian rhythms, for example, to mitigate circadian disruption. This is done by appropriate scheduling of circadian inputs, such as light, sleep, or pharmaceuticals. Models of human circadian rhythms have been used as tools for circadian control. These models typically describe the

circadian clock as a dynamical system whose internal states are driven by the circadian inputs. Circadian control is then performed in two steps that form a feedback loop. First, the state of the system is estimated based on some measurable signals. This step is called *state estimation*. Second, the models are used in the optimization process to find the best circadian input schedule for the control purpose, e.g., to eliminate jetlag. For a comprehensive review of modeling paradigms in circadian research, we refer the reader to [3].

Many dynamical models for circadian rhythms have been developed and used. Among them, the standout Kronauer model [4] and its variants were designed based on the circadian dynamics present in CBT. While these models have been used extensively with appreciable success, they were designed based on population-level data. As such, they often struggle to predict the full range of variations seen on the individual level [5], [6], [7]. Moreover, attempts to fine-tune the model parameters on an individual basis have been stifled by the fact that CBT is a signal that is measurable only with invasive techniques. Furthermore, state estimation typically relies on a model that relates the state of the system to the observables. Therefore, there is a need for dynamical models that are based on more easily obtainable biometric signals. We note that there are also statistical models that relate the phase of circadian rhythms with such signals, for example, in [8], [9], [10], [11]. However, these models do not contain the internal dynamics of the circadian clock and cannot be used to predict the impacts of the inputs (e.g., lighting) on the dynamics.

The main contribution of this paper is a method for learning latent state models, i.e., dynamical models that explicitly use latent state variables, that relate circadian input(s) to observable biometric signals. Our method can, therefore, construct personalized data-driven models. We validate our method using experimental data, where the circadian input is light and the biometric data are actigraphy signals. We have shown in our earlier publications that phase shifts in actigraphy signals can be used to estimate phase shifts in the circadian clock [12], [13], [14]. Our models combine an autoencoder with a recurrent neural network (RNN) and use the pair to model the salient dynamics present in the data. The encoder takes

a windowed history of past measurements of actigraphy and light and brings it to a two-dimensional space, while the RNN serves as a state space model within that latent space. We impose structure on the latent space by adding a contrastive component to the loss function used in the model training. This component uses circadian phase estimates obtained from our previously proposed filter [14].

Notable features of our models are:

- (1) Accuracy: accurate reconstruction and prediction of biometric signals,
- (2) Simplicity: low-dimensional latent state variable that is suitable for circadian control applications,
- (3) Consistency with circadian clock: the phase shifts of the latent state variables agree with the phase shifts of the circadian clocks as measured using Dim Light Melatonin Onset (DLMO) protocol.

II. METHOD

A. Dataset

We used a dataset containing the actigraphy, light exposure, and melatonin concentrations collected from a group of eight healthy young adults (5 female and 3 male) aged between 18 and 34 y (25.8 ± 6.6 y) in a clinical study conducted at the University of New Mexico (UNM). We have previously used this dataset in [12], [13], [14]. All participants gave their informed written consent, and the experiments followed the principles in the Declaration of Helsinki from the World Medical Association. The experiments were monitored by the UNM Health Sciences Center Human Research Protections office and approved by the UNM Institutional Review Board (IRB). The study's IRB number is 14-002.

The actigraphy and light data were collected at 1-minute intervals using an ActiGraph GT3X+ Monitor (Pensacola, FL) worn on each subject's non-dominant wrist over approximately 14 days, yielding $\sim 20,160$ entries. Melatonin samples were taken at 30-minute intervals on the 7th and 14th days with Salimetrics SalivaBio Oral Swabs (State College, PA) starting 5 hours before and ending 30 minutes after the average bedtime.

To pre-process, we first took the 3-hour moving average of the measured light, then further downsampled the actigraphy and light data by taking the average of each 5-minute interval. Finally, the series were transformed using z-score normalization to a mean of zero and variance of one. Figure 1 shows samples of the pre-processed data.

B. Model Architecture

We are concerned with learning a model of an unknown system using a training dataset of input-output samples $\mathcal{D} = \{u_1, y_1, \dots, u_T, y_T\}$ collected from the system where $u_t \in \mathbb{R}^{n_u}$ is a measurable exogenous input, and $y_t \in \mathbb{R}^{n_y}$ is a measurable output of the system, both taken at time t . We take light as the system input, in line with pre-existing knowledge of the circadian system, and actigraphy as the output.

This learning process amounts to fitting a parameterized model, expressed as a group of neural networks, to the

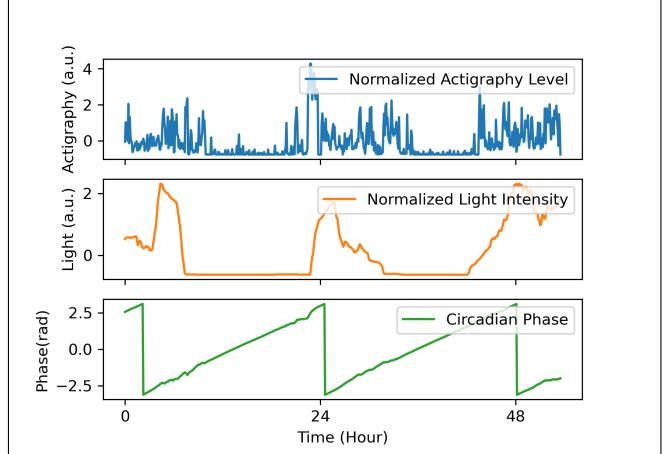


Fig. 1. Sample actigraphy, input light intensity, and circadian phase trajectories. Circadian phase is obtained by running a tuned Kalman filter [14] on the raw actigraphy data.

available data using the loss function detailed in Subsection C. The network architecture is shown in Figure 2 and detailed below, where θ is the model parameters:

$$\text{Transition model: } x_{t+1} = f_\theta(x_t, g_t, u_t), \quad (1)$$

$$\text{Encoder model: } s_t = e_\theta(y_{t-h:t}, u_{t-h:t}), \quad (2)$$

$$\text{Predictor model: } \hat{s}_t = p_\theta(x_t), \quad (3)$$

$$\text{Decoder model: } \hat{y}_t = d_\theta(\hat{s}_t, u_t), \quad (4)$$

$$\text{Update gate: } g_t = \mathbb{1}(y_t)s_t + (1 - \mathbb{1}(y_t))\hat{s}_t, \quad (5)$$

where $x_t \in \mathbb{R}^{n_x}$ is the latent state, $s_t, \hat{s}_t \in \mathbb{R}^{n_s}$ are the encoded and predicted features, respectively, $\hat{y}_t \in \mathbb{R}^{n_y}$ is the predicted output, and $\mathbb{1}(\cdot)$ is the indicator function where $\mathbb{1}(y_t) = 1$ if the external observation y_t is available. $(y_{t-h:t}, u_{t-h:t})$ is the windowed history of outputs and inputs from some time $t - h$ up to t that is passed to the encoder.

The components f_θ , e_θ , p_θ , and d_θ are the neural networks that make up the model. The encoder, decoder, and predictor are all multilayer perceptrons with ReLU activations, while the transition model is a gated recurrent unit (GRU) [15]. The overall structure is analogous to the Recurrent State-Space Model (RSSM) [16], but we remove the stochasticity in s_t to improve the model's potential ease of use in down-stream control tasks.

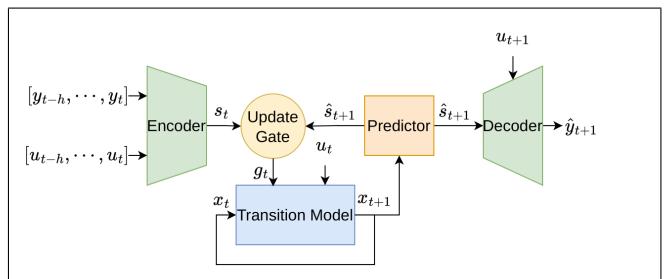


Fig. 2. Architecture of the State Space Model. Arrows illustrate the flow of information through the model.

The dimensions of the variables are $n_y = 1$, $n_u = 1$, $n_s = 16$, and $n_x = 2$, with an input window of 10 samples (50 minutes). The choice for a 2-dimensional latent space is based on its relative simplicity for control tasks and for visualization. Moreover, it has been shown in previous literature that high-dimensional circadian clock models can be reduced to 2-dimensional manifolds while still maintaining the overt 24-hour periodicity and input response characteristics [17], [18].

C. Training

The loss function used to train the model is given by

$$\mathcal{L} = \lambda_{\text{rec}} \mathcal{L}_{\text{rec}} + \lambda_{\text{feat}} \mathcal{L}_{\text{feat}} + \lambda_{\text{contrast}} \mathcal{L}_{\text{contrast}}, \quad (6)$$

where the λ 's are the weights of the components described below.

- Reconstruction loss to penalize inaccurate predicted observations:

$$\mathcal{L}_{\text{rec}} \triangleq \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2. \quad (7)$$

- Feature alignment loss to minimize the difference between features from the encoder and features from the predictor:

$$\mathcal{L}_{\text{feat}} \triangleq \frac{1}{T} \sum_{t=1}^T (s_t - \hat{s}_t)^2. \quad (8)$$

- Contrastive loss [19] to shape the state trajectory:

$$\mathcal{L}_{\text{contrast}} \triangleq \frac{1}{T^2} \sum_{i=1}^T \sum_{j=1}^T c(x_i, x_j, \phi_i, \phi_j),$$

where $c(x_i, x_j, \phi_i, \phi_j)$ (9)

$$= \begin{cases} [\|x_i - x_j\|^2 - m_{\text{pos}}]_+ & \text{if } \phi_i = \phi_j, \\ [m_{\text{neg}} - \|x_i - x_j\|^2]_+ & \text{if } \phi_i \neq \phi_j. \end{cases}$$

In Equation 9, m_{pos} and m_{neg} are hyper-parameters that define the positive and negative margins, and $[\cdot]_+$ is a function that clips the value to be positive. I.e., $[x]_+ \triangleq \max(x, 0)$. ϕ is the circadian phase estimated with a Kalman filter optimized for each individual using the procedure in [14].

With this loss function, we train a separate model for each subject using only their own data. We set the hyperparameters $\lambda_{\text{rec}} = 1$, $\lambda_{\text{feat}} = 1.5$, $\lambda_{\text{contrast}} = 2$, $m_{\text{pos}} = 0.1$, and $m_{\text{neg}} = 0.2$. These values were chosen heuristically, since there is no quantitative metric to measure the quality of a given state trajectory, making it infeasible to perform automatic tuning.

The model was implemented and trained end-to-end in PyTorch [20]. Optimization was done with the Adam algorithm [21] and a cosine annealing with warm restart learning rate schedule [22] with an initial learning rate of 10^{-3} .

III. RESULTS

In our experiments, we dropped the initial three days of data for each subject due to the Kalman filter transients present in the estimated phase ϕ [14]. In practice, the removal is only necessary the first time the filter is started on an individual's data. After dropping the transient days, we used the data from day 4 to 12 for training, and the 48 hours of data after day 12 for evaluation.

A. Contrastive Loss for Structuring the State Space

In an initial, completely unsupervised, training run without the contrastive loss component, the model struggled to produce an orbit in the latent space that is typical of periodic systems. Instead, the phase portrait often collapsed into a line or an irregular set of points as can be seen in Figure 4 even with low output prediction error as seen in Figure 3. To rectify this, we employed the filter developed in our previous work [14]. The filter tracks the phase of a selected harmonic component in an input signal - the 24-hour harmonic in this case. Running the filter on the actigraphy produces a continuous phase estimate ϕ , which we use as the similarity label in the contrastive loss detailed above to improve the overall structure of the latent trajectory. The added loss component resulted in the more structured orbit shown in Figure 6, allowing us to conclude that the contrastive loss is effective in shaping the states into a distinct periodic orbit.

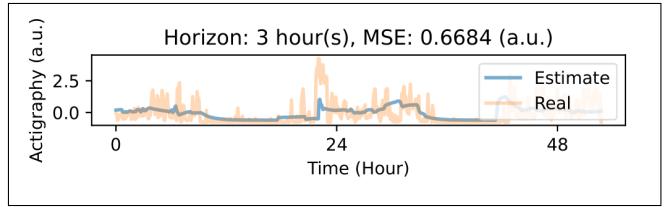


Fig. 3. Real and predicted actigraphy for subject 3 using a model trained without the contrastive loss component.

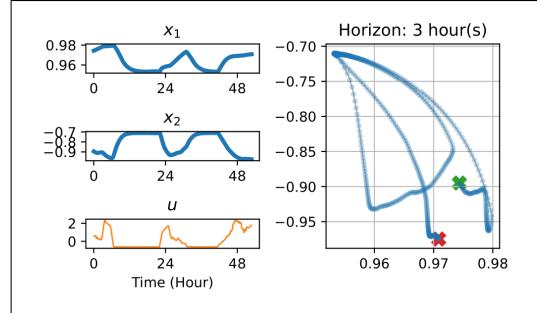


Fig. 4. Latent state trajectory for subject 3 using a model trained without the contrastive loss component. The figure contains the two latent state variables x_1 and x_2 , the input trajectory u , and the phase portrait.

B. Output Prediction

Table I presents the mean squared error of predicting the actigraphy in the test set using different prediction horizons, where prediction horizon is the frequency of reading external actigraphy and encoding it into s_t . Within the horizon, the model estimates \hat{s}_t and decodes it into predicted actigraphy.

Figures 5 and 6 compare the actigraphy prediction and state-space trajectory under different prediction horizons. We can summarize that the 15-minute horizon may have relatively low prediction error compared to longer horizons, but is less stationary in states since actigraphy measurements are noisy. Other prediction horizons result in mostly identical performance and state trajectories.

TABLE I
MEAN SQUARED ERROR IN PREDICTED ACTIGRAPHY FOR EACH SUBJECT ACROSS MULTIPLE HORIZONS.

Horizon (hour)	Subject					
	3	4	6	7	8	10
0.25	0.678	0.858	0.773	0.855	0.580	0.777
1	0.687	0.876	0.760	0.700	0.626	0.829
2	0.693	0.883	0.749	0.692	0.629	0.843
3	0.697	0.848	0.747	0.691	0.628	0.925
4	0.697	0.858	0.736	0.688	0.637	0.867
5	0.711	0.845	0.740	0.687	0.632	0.884
6	0.700	0.853	0.737	0.689	0.631	0.861
7	0.691	0.863	0.743	0.687	0.630	0.894
8	0.701	0.860	0.739	0.685	0.637	0.901
24	0.702	0.869	0.738	0.684	0.634	0.893
48	0.703	0.870	0.738	0.684	0.634	0.898

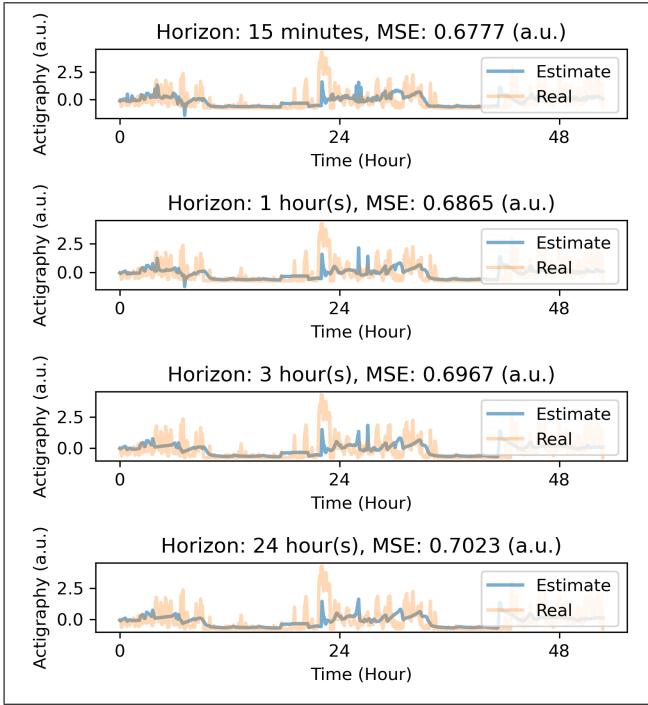


Fig. 5. Real vs predicted actigraphy for subject 3 across 4 prediction horizons.

C. Phase Shift of the Latent State Variables

To assess the structural quality of the learned state space, we measure the level of agreement between the phase shifts in the latent state variables with those of the subjects' circadian clocks. Phase is the most studied parameter of the circadian system, and as such, it provides a standardized metric of comparison to existing techniques.

The clinical standard for estimating the circadian phase is dim light melatonin onset (DLMO), which is the clock time at which melatonin concentrations in the body rise above a calculated threshold in low light conditions [23]. This DLMO time serves as a phase marker on any given day, and to estimate the phase shift between two days, we take the difference between the DLMO times on those days. DLMO calculations for subjects 5 and 9 did not yield valid estimates, so we exclude these subjects in the presented results.

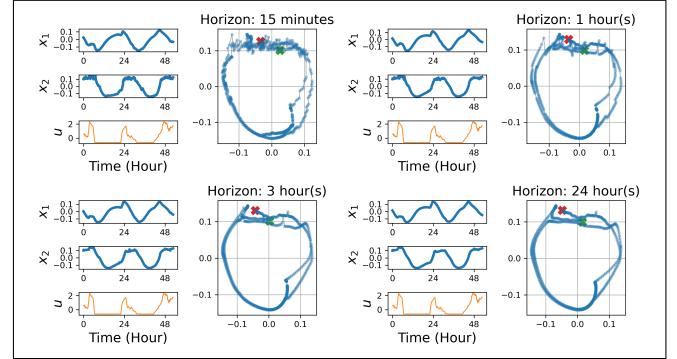


Fig. 6. Latent state trajectories for subject 3 across 4 prediction horizons. Each subfigure contains the two latent state variables x_1 and x_2 , the input trajectory u , and the phase portrait.

To calculate the phase shift of the latent state variables, we first use the trained encoder to project every data point into the latent space, producing a 14-day 2-dimensional state trajectory for each individual. We then compare the trajectories from the days DLMO was estimated - days 7 and 14. To do so, we stack the x_1 and x_2 time series to form a wide matrix, then take the 2-dimensional cross-correlation between the 24-hour window representing day 14 and that representing day 7. Lastly, we take the phase shift to be the lag value where the cross-correlation between the days is maximized.

The phase shift comparison is shown in Table II. We see that in 4 of the 6 subjects with valid DLMO readings, the model predicts the phase shift to within ~ 30 minutes of the DLMO estimate. Subjects 3 and 7 have errors of 75.75 and 66 minutes, respectively. The performance on subject 7's data can be attributed to the large amount of dropped data (~ 60 hours) toward the end of the clinical study, but we do not have a clear reasoning for the performance on subject 3. However, the overall performance achieved on the dataset suggests that the learned latent space has a structure that preserves the phase shifts present in the actigraphy signal.

Note that the phase shift comparison was used in assessing the structural quality of the latent space. The true utility of the proposed framework is in its input-state modeling. The Kalman filter from [14] provides accurate phase shifts with significantly less complexity and would thus be more appropriate in situations where only phase estimation is required. However, the filter does not model the relationship between circadian inputs and state, and is thus not as useful for control applications. The proposed framework provides such a model that can then be used more readily in control system design.

IV. CONCLUSION

In this work, we proposed a framework that is capable of learning a 2-dimensional dynamical representation of the circadian variation present in a measured biometric signal, along with the transformations from the measurement space to the latent space. When tested on real actigraphy and light data, the system performed well in predicting the cyclic nature of actigraphy readings, a common measurement in circadian

TABLE II

ESTIMATED PHASE SHIFTS PRESENT IN THE ENCODED LATENT STATE FOR EACH SUBJECT COMPARED WITH THE MEAN DLMO-ESTIMATED SHIFT.

Subject	Mean DLMO (mins)	Predicted Shift (mins)	Abs Error
3	70.75	-5	75.75
4	4.29	0	4.29
6	14	0	14
7	-66	0	66
8	85.25	55	30.25
10	-52.2	-40	12.2
Mean Absolute Error (mins)		33.75	
% within 30 min		50	
% within 60 min		67	

rhythm research. Moreover, the phase shift in the learned latent space showed agreement with DLMO-estimated shifts, serving as confirmation of the structure in the space. The mean absolute error of ~ 34 minutes across 6 subjects was well in line with the state of the art for predicting circadian phase shifts.

While the focus on personalization in this paper stemmed from previous studies linking poor generalization with population-level model development [5], [6], [7], [9], further work needs to be done to train and evaluate a subject-independent model for comparison to the work done here. It might be possible to borrow from the field of transfer learning by training a base model on a large population and then fine-tuning on an individual basis. Future work is also necessary to evaluate the framework's ability to learn the dynamics of a known system, and to test how control policies designed with it perform when deployed on such a system.

ACKNOWLEDGMENTS

This work was primarily supported by the National Science Foundation (NSF) through Grant DMS-2037357 and the Smart Lighting Engineering Research Center (EEC-0812056), the Army Research Office through Grant W911NF-17-1-0562 and W911NF-22-10039, and New York State under NYSTAR contract C130145. The authors would like to thank Dr. Jiawei Yin for his contribution to the initial portion of this research as part of his doctoral research and the staff of the University of New Mexico Hospital Sleep Disorders Center who carried out the DLMO procedures: Steven Lopez, Alex Valdez, Cielo Ortiz, and Ryan Valdez.

REFERENCES

- [1] M. H. Vitaterna, J. S. Takahashi, and F. W. Turek, "Overview of circadian rhythms," *Alcohol Research & Health*, vol. 25, no. 2, p. 85, 2001.
- [2] A. Knutsson, "Health disorders of shift workers," *Occupational medicine*, vol. 53, no. 2, pp. 103–108, 2003.
- [3] A. Asgari-Targhi and E. B. Klerman, "Mathematical modeling of circadian rhythms," *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, vol. 11, no. 2, p. e1439, 2019.
- [4] D. B. Forger, M. E. Jewett, and R. E. Kronauer, "A simpler model of the human circadian pacemaker," *Journal of biological rhythms*, vol. 14, no. 6, pp. 533–538, 1999.
- [5] J. E. Stone, S. Postnova, T. L. Sletten, S. M. Rajaratnam, and A. J. Phillips, "Computational approaches for individual circadian phase prediction in field settings," *Current Opinion in Systems Biology*, vol. 22, pp. 39–51, 2020.
- [6] M. A. St. Hilaire, H. M. Lammers-van der Holst, E. D. Chinoy, C. M. Isherwood, and J. F. Duffy, "Prediction of individual differences in circadian adaptation to night work among older adults: application of a mathematical model using individual sleep-wake and light exposure data," *Chronobiology International*, vol. 37, no. 9-10, pp. 1404–1411, 2020.
- [7] Y. Huang, C. Mayer, P. Cheng, A. Siddula, H. J. Burgess, C. Drake, C. Goldstein, O. Walch, and D. B. Forger, "Predicting circadian phase across populations: a comparison of mathematical models and wearable devices," *Sleep*, vol. 44, no. 10, p. zsab126, 2021.
- [8] V. Kolodyazhnyi, J. Späti, S. Frey, T. Götz, A. Wirz-Justice, K. Kräuchi, C. Cajochen, and F. H. Wilhelm, "An improved method for estimating human circadian phase derived from multichannel ambulatory monitoring and artificial neural networks," *Chronobiology International*, vol. 29, no. 8, pp. 1078–1097, 2012.
- [9] J. E. Stone, A. J. Phillips, S. Ftouni, M. Magee, M. Howard, S. W. Lockley, T. L. Sletten, C. Anderson, S. M. Rajaratnam, and S. Postnova, "Generalizability of a neural network model for circadian phase prediction in real-world conditions," *Scientific reports*, vol. 9, no. 1, p. 11001, 2019.
- [10] L. S. Brown, M. A. St. Hilaire, A. W. McHill, A. J. Phillips, L. K. Barger, A. Sano, C. A. Czeisler, F. J. Doyle III, and E. B. Klerman, "A classification approach to estimating human circadian phase under circadian alignment from actigraphy and photometry data," *Journal of pineal research*, vol. 71, no. 1, p. e12745, 2021.
- [11] C. Bowman, Y. Huang, O. J. Walch, Y. Fang, E. Frank, J. Tyler, C. Mayer, C. Stockbridge, C. Goldstein, S. Sen *et al.*, "A method for characterizing daily physiology from widely used wearables," *Cell reports methods*, vol. 1, no. 4, 2021.
- [12] J. Yin, A. Julius, J. T. Wen, M. M. Oishi, and L. K. Brown, "Actigraphy-based parameter tuning process for adaptive notch filter and circadian phase shift estimation," *Chronobiology International*, vol. 37, no. 11, pp. 1552–1564, 2020.
- [13] C. O. Ike, J. T. Wen, M. M. Oishi, L. K. Brown, and A. A. Julius, "Fast tuning of observer-based circadian phase estimator using biometric data," *Helijon*, vol. 8, no. 12, 2022.
- [14] ———, "Efficient estimation of the human circadian phase via kalman filtering," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2023, pp. 1–6.
- [15] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [16] D. Hafner, T. Lillicrap, I. Fischer, R. Villegas, D. Ha, H. Lee, and J. Davidson, "Learning latent dynamics for planning from pixels," in *International conference on machine learning*. PMLR, 2019, pp. 2555–2565.
- [17] D. B. Forger and R. E. Kronauer, "Reconciling mathematical models of biological clocks by averaging on approximate manifolds," *SIAM Journal on Applied Mathematics*, vol. 62, no. 4, pp. 1281–1296, 2002.
- [18] P. Indic, K. Gurdziel, R. E. Kronauer, and E. B. Klerman, "Development of a two-dimension manifold to represent high dimension mathematical models of the intracellular mammalian circadian clock," *Journal of biological rhythms*, vol. 21, no. 3, pp. 222–232, 2006.
- [19] E. Simo-Serra, E. Trulls, L. Ferraz, I. Kokkinos, P. Fua, and F. Moreno-Noguer, "Discriminative learning of deep convolutional feature point descriptors," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 118–126.
- [20] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.
- [21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [22] I. Loshchilov and F. Hutter, "Sgdr: Stochastic gradient descent with warm restarts," *arXiv preprint arXiv:1608.03983*, 2016.
- [23] A. J. Lewy and R. L. Sack, "The dim light melatonin onset as a marker for circadian phase position," *Chronobiology International*, vol. 6, no. 1, pp. 93–102, 1989.