CIRCADIAN RHYTHM MONITORING

Naisha Shah, Raaghav Saboo, Nandini Aggarwal

THE PROBLEM

How can we use **non-invasive methods** to detect **circadian misalignment** and support **better mental and metabolic health** outcomes?

- **Circadian rhythms** are the body's internal ~24-hour biological cycles that regulate sleep, hormone release, body temperature, and metabolism.
- The most accurate way to measure biological time is **Dim Light Melatonin Onset (DLMO)** which is **expensive, invasive, and impractical** for large-scale or everyday monitoring.
- As a result, although circadian misalignment can be measured scientifically, it is **rarely monitored** in real-world settings.

THE IDEA

- Use machine learning to **predict circadian rhythm** outside controlled laboratory environments.
- Combine signals such as **light exposure, activity patterns, sleep timing, and physiological indicators.**
- Enable **continuous monitoring** of circadian alignment in real-world settings.
- Support applications such as sleep optimization, shift-work scheduling, jet lag management, and mental health monitoring.
- Ultimately **improve sleep quality, productivity, and long-term health outcomes.**

LITERATURE REVIEW #1

What They Did

- Studied 120 healthy sleepers - 60 on fixed sleep schedules, 60 on free sleep schedules - over 6 days at home
- Measured DLMO (Dim Light Melatonin Onset) via salivary melatonin every 30 minutes under dim light
- Tested whether sleep timing alone could reliably estimate DLMO

Conclusions

- DLMO was more closely linked to sleep timing in free sleepers (r = 0.70) than fixed sleepers (r = 0.44)
- Sleep timing can estimate DLMO reasonably well only in people with no work or family constraints on sleep
- For people with irregular or constrained schedules - shift workers, patients - sleep timing fails badly
- DLMO itself requires expensive, time-consuming, invasive salivary collection making it not feasible in real life
- Explicitly called for better methods to estimate DLMO in people with irregular lives

Shortcomings We Are Fixing

- Still relied on invasive saliva collection for ground truth - our project aims to eliminate this entirely
- Only tested young healthy adults - not generalizable to diverse populations
- Sleep timing as a proxy breaks down for anyone with an irregular lifestyle - our ML approach handles this
- No wearable data used at all - purely behavioral sleep logs - we use objective actigraphy signals

Burgess et al. established that DLMO is the gold standard body clock marker but is too invasive to measure in real life, and sleep timing alone is not reliable enough to replace it. Our project directly solves this by predicting DLMO non-invasively from wearable data using ML.

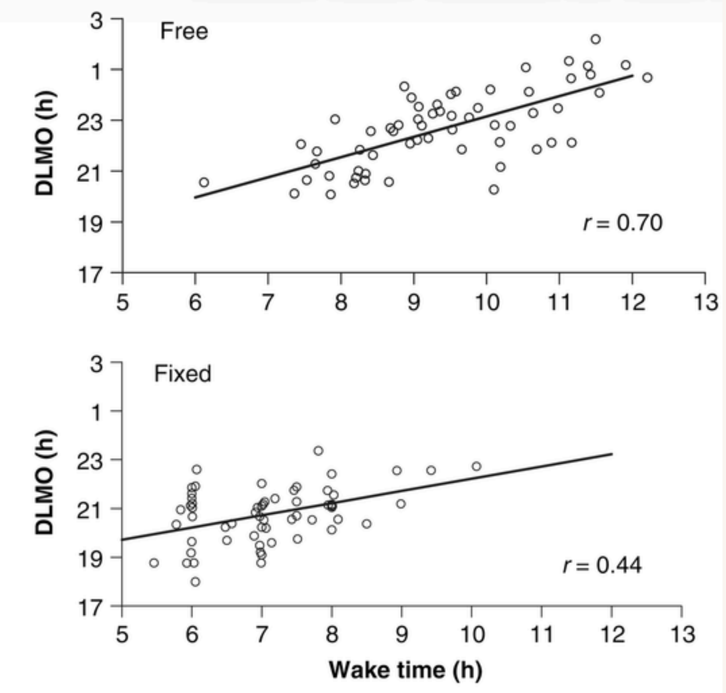


Figure 2.
The scatterplots and associated linear regressions between the dim light melatonin onset (DLMO) and wake time in the free sleepers (top) and fixed sleepers (bottom). The regression equation for the relationship between the DLMO and wake time in the free sleepers is: DLMO (decimal time) = 0.80 × wake time (decimal time) −8.83.

https://pmc.ncbi.nlm.nih.gov/articles/PMC3841975/

Burgess, Helen J., and Charmane I. Eastman. "The dim light melatonin onset following fixed and free sleep schedules." Journal of sleep research 14, no. 3 (2005): 229-237.

LITERATURE REVIEW #2

What They Did

- Built a classification ML model (neural network + regression) to estimate circadian phase directly from actigraphy and light data
- Used DLMO as ground truth - tested across normal sleepers and shift workers
- Compared 3 model types: limit cycle ODE models, regression models, and neural network models
- Specifically tested whether ML could outperform traditional math models for circadian phase prediction

Conclusions

- ML models achieved prediction errors between 0.4 and 1.1 hours - competitive with ODE models
- Neural networks performed well on normal sleepers but struggled with shift workers and irregular patterns
- All 3 model types had similar accuracy - suggesting the data and features matter more than model complexity
- Explicitly stated their biggest limitation was small sample size - not enough participants to train ML properly
- Called for future work using larger, more diverse datasets to improve ML-based circadian prediction

Shortcomings We Are Fixing

- Used only lab-collected actigraphy - not real-world population data like NHANES
- Sample size was too small to fully train ML - we use 1000+ NHANES participants
- Predicted circadian phase as a classification (early/late/normal) - we predict exact DLMO time as a continuous value
- Did not engineer rich features like morning/evening light windows or sleep midpoint - our feature pipeline is far more detailed

Mayer et al. showed ML can estimate circadian phase from actigraphy but were limited by small lab samples and treated it as classification. We directly extend this using 1000+ real-world NHANES participants with XGBoost to predict exact DLMO timing continuously.

Table 2.

Performance of Classifiers for DLMO Estimation in College Students.

The classification accuracy of whether a timepoint falls before or after experimentally observed DLMO in cross-validation (mean ± standard deviation) and on the independent test set on different subsets of the data.

Cross-Validation Results					
Model	**Full Dataset**	**Actigraphy Data**	**Activity & Skin Temperature**	**Activity & Light**	**Light & Skin Temperature**
Single Layer	89.1±3.7 (100 nodes)	89.4±3.6 (90 nodes)	88.8±3.4 (50 nodes)	89.8±3.4 (90 nodes)	87.0±3.0 (30 nodes)
Single Layer with Dropout	89.3±3.4 (p=0.4)	89.3±3.2 (p=0.3)	89.1±2.9 (p=0.5)	89.5±3.8 (p=0.5)	87.1±3.4 (p=0.2)
Double Layer	88.9±3.6 (50x50 nodes)	89.3±3.6 (30x50 nodes)	89.2±3.5 (30x40 nodes)	90.0±3.9 (40x40 nodes)	86.8±3.7 (30x40 nodes)
Double Layer with Dropout	89.3±3.3 (p=0.4)	89.5±3.4 (p=0.3)	88.7±3.2 (p=0.1)	89.6±3.8 (p=0.4)	87.1±2.9 (p=0.2)
Test Results					
Model	**Full Dataset**	**Actigraphy Data**	**Activity & Skin Temperature**	**Activity & Light**	**Light & Skin Temperature**
Single Layer	90.7	90.2	89.1	89.9	86.1
Single Layer with Dropout	89.9	90.2	89.8	89.8	85.8
Double Layer	90.1	90.9	89.4	90.5	86.4
Double Layer with Dropout	89.2	89.8	88.8	90.1	85.1

https://pmc.ncbi.nlm.nih.gov/articles/PMC8474125/

Brown, Lindsey S., Melissa A. St. Hilaire, Andrew W. McHill, Andrew JK Phillips, Laura K. Barger, Akane Sano, Charles A. Czeisler, Francis J. Doyle III, and Elizabeth B. Klerman. "A classification approach to estimating human circadian phase under circadian alignment from actigraphy and photometry data." Journal of pineal research 71, no. 1 (2021): e12745.

LITERATURE REVIEW #3

What They Did

- Tested the 3 most popular consumer wearables - Apple Watch Series 8, Fitbit Sense 2, Oura Ring Gen3 - against the gold standard polysomnography (PSG)
- 35 healthy adult participants wore all three devices simultaneously in a single overnight lab study
- Measured the accuracy of sleep stage detection for each device

Conclusions

- All devices were good at detecting sleep vs. wake (sensitivity ≥ 95%)
- Oura Ring was the most accurate for sleep stage classification - 5% better than Apple Watch, 10% better than Fitbit
- Apple Watch overestimated light sleep by 45 minutes and underestimated deep sleep
- All devices track when you sleep reasonably well - but none can tell you where your body clock is
- Millions of people already own these devices, but they are completely blind to the circadian phase

Shortcomings We Are Fixing

- Study only tested sleep stage accuracy - zero attempt at circadian phase or DLMO prediction
- Only 35 participants - very small sample
- All participants were healthy with no sleep disorders - not representative of the general population
- The exact activity and light data these devices collect is already enough for our ML model - they just never tried it

Robbins et al. confirmed that millions of people already wear accurate activity tracking devices. But these devices only track sleep - not body clock timing. Our project unlocks the circadian intelligence already hidden in this data.

Table 4. Comparison of Oura, Fitbit, and Apple Watch to PSG with respect to sleep/wake and sleep stages.

	Oura n = 35 *Mean (SD)*	PSG n = 35 *Mean (SD)*	Fitbit n = 33 *Mean (SD)*	PSG n = 33 *Mean (SD)*	Apple Watch n = 29 *Mean (SD)*	PSG n = 29 *Mean (SD)*
Total Sleep Time (min)	421 (34)	430 (41)	428 (23)	431 (40)	442 (35)	434 (40)
Wake (min)	59 (34)	50 (41)	49 (26)	49 (40)	39 (35) *	46 (40)
Light Sleep (min)	233 (28)	239 (41)	258 (37) *	240 (41)	289 (24) *	244 (38)
Deep Sleep (min)	95 (21)	95 (35)	79 (27) *	94 (36)	51 (18) *	94 (38)
REM (min)	93 (25)	96 (21)	90 (28)	97 (21)	102 (22)	96 (21)
Sleep Latency (min)	18 (25) *	13 (22)	13 (21)	13 (22)	17 (26)	15 (23)
WASO (min)	42 (26)	38 (37)	40 (15)	36 (35)	22 (23) *	32 (33)
Sleep Efficiency (%)	88% (7%)	90% (9%)	89% (5%)	90% (8%)	92% (7%)	90% (8%)

Notes. Missing data were reassigned as wake (see section on wake interpolation). Sample size for sleep latency n = 34 for PSG vs. Oura, n = 32 for PSG vs. Fitbit, and n = 29 for PSG vs. Apple Watch. * Indicates significant difference between the device and PSG using paired *t*-tests.

https://pmc.ncbi.nlm.nih.gov/articles/PMC11511193/

Robbins, Rebecca, Matthew D. Weaver, Jason P. Sullivan, Stuart F. Quan, Katherine Gilmore, Samantha Shaw, Abigail Benz et al. "Accuracy of three commercial wearable devices for sleep tracking in healthy adults." Sensors 24, no. 20 (2024): 6532.

LITERATURE REVIEW #4

What They Did

- Reviewed 4 specific methods used to extract circadian patterns from actigraphy: cosinor modelling, wavelet analysis, empirical mode decomposition, and fractal analysis
- Analysed how each method works, what features it produces, and where it breaks down
- Evaluated how well these methods have been applied in large population studies like the UK Biobank and NHANES
- Built and released an open-source tool called ezActi2 to standardise circadian feature extraction from actigraphy

Conclusions

- Every method produces different results on the same data - existing studies cannot be compared
- Cosinor modelling assumes a perfect 24h sine wave - breaks down for irregular sleepers
- Large NHANES actigraphy studies only extract basic sleep metrics - nobody has predicted DLMO from them
- Current methods describe what a rhythm looks like, but cannot tell you where your body clock actually is
- Explicitly called for work that translates actigraphy features into real circadian phase markers like DLMO

Shortcomings We Are Fixing

- They reviewed methods but built no predictive ML model - we do
- Cosinor limitation they identified is exactly why we also engineer light, activity and sleep features beyond cosinor alone
- They highlighted NHANES as underutilised - we directly build on NHANES actigraphy data

Methods	Strength	Weakness	Contexts to be used
Parametric analysis			
Cosinor model	• Mathematically simple with analytical solution • Computational flexible to allow (theoretically) infinite number of harmonics	• Only allows predefined constant cycle length(s) • Computational complexity grows exponentially with more harmonics	• Traditionally the first choice for laboratory -based studies in which rhythms are minimally interfered by environmental factors • When regular rhythms are expected (i.e., relatively constant amplitude and phase across cycles)
Wavelet analysis	• Can handle non -stationary signals	• Requires expertise and experience in choosing wavelet function • Requires expertise in interpreting the scalogram	• When rhythms are expected to change over time
Non-parametric analysis	• No assumption of sinusoidal shape	• Only allows one predefined constant cycle length	• Traditionally a compliment of parametric analysis • When the consistency, fragmentation, or the levels of activities are the primary variables of interest
Data adaptive approach			
EMD and its variants	• No assumption on the shape of base function • Can handle non -stationary signals	• The original EMD suffers from mode mixing/separation issues (while uniform-phase EMD offers a reasonable solution)	• When rhythms are expected to change over time • Can be a first choice for irregular rhythms as EMD approaches have no assumptions on base functions
Nonlinear approach			
Fractal analysis	• Captures multiscale patterns independent of rhythmicity • Minimally masked by the signal magnitude	• Not immediately obvious in translating back to rhythmicity	• Can be used to infer the functional status of the circadian network based on ambulatory data

Fig. 1.
Strengths, weaknesses, and applicable contexts of different approaches for analyzing circadian rest-activity rhythms.

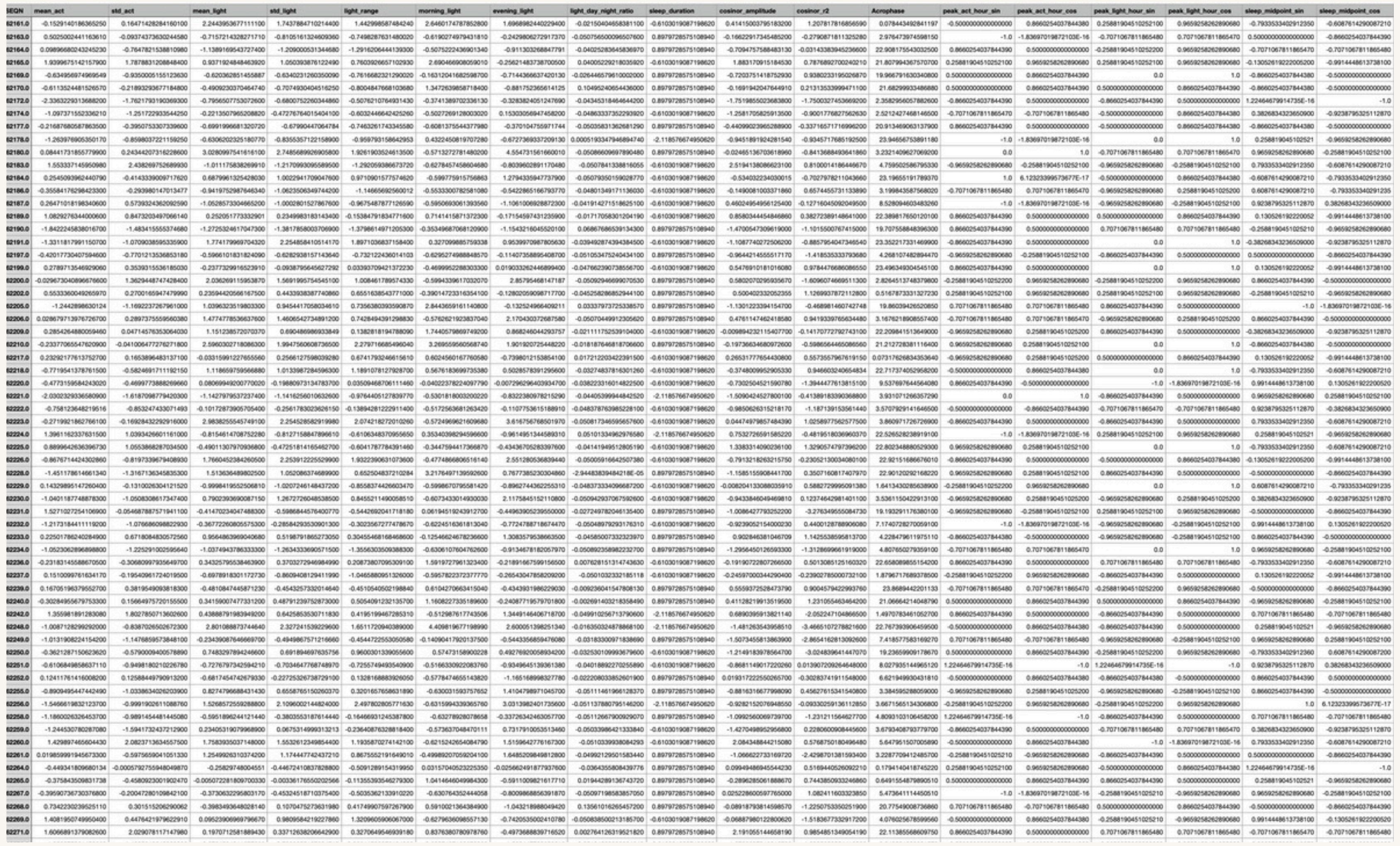
https://pmc.ncbi.nlm.nih.gov/articles/PMC12829919/

Gao, Chenlu, Shahab Haghayegh, Max Wagner, Ruixue Cai, Kun Hu, Lei Gao, and Peng Li. "Approaches for assessing circadian rest-activity patterns using actigraphy in cohort and population-based studies." Current sleep medicine reports 9, no. 4 (2023): 247-256.

Gao et al. showed that existing circadian actigraphy methods each have serious limitations and cannot predict actual body clock timing like DLMO or Acrophase. Our project directly addresses this by combining multiple feature types into a single XGBoost model trained on NHANES population data to predict acrophase.

THE GAPS

- **Gold Standard Is Too Invasive** The most accurate body clock marker requires hourly saliva samples in a controlled lab over 24 hours - completely impractical for everyday use or continuous monitoring
- **Wearables Don't Track Body Clock** Millions of people wear fitness trackers daily - but no current device can tell you where your internal body clock actually is. They track sleep. Not circadian phase.
- **Existing Models Are Too Rigid** Current prediction approaches use fixed mathematical equations - they cannot learn from data, cannot adapt to individuals, and break down for people with irregular schedules
- **ML Attempts Were Too Small Scale** Machine learning has been tried for circadian prediction but always on tiny lab samples - never enough data to train a model that generalises to the real population
- **No Population-Scale Circadian Prediction Exists** Large wearable datasets with thousands of participants already exist - but no one has built a standardised ML pipeline on them to predict actual body clock timing
- **ML Stops at Disorder Detection** The few ML models that do use circadian wearable data only classify sleep disorders - none predict the exact timing of the body clock continuously



- Dataset from the **National Health and Nutrition Examination Survey (NHANES**) conducted by the **Centers for Disease Control and Prevention (CDC)**
- Large **population-based** health dataset
- Number of **rows**: **6,040**
- Number of **columns (features)**: **19**
- Type of dataset: **Tabular numerical dataset** designed for circadian rhythm analysis using wearable sensor data.

ABOUT THE DATASET

- Target variable: **Acrophase –** the peak activity in the 24-hour cycle.
- Calculated using a cosinor model, which estimates three parameters**: Mesor** (average activity level)**, Amplitude** (strength of the rhythm)**, and Phase shift.**
- **Formula:**
- **$Y(t) = M + A \cdot \cos(2\pi t/\tau - \varphi)$**
- For individuals with irregular or noisy activity patterns, wavelet analysis can also be used to detect rhythmic patterns before estimating circadian features.

ABOUT THE DATASET

WHY THIS DATASET

- Provides **24-hour continuous activity data**, required for circadian rhythm modeling
- **Large population-scale** dataset
- Enables extraction of circadian features such as: Activity rhythms and Acrophase (peak activity time)
- **Contains key variables** needed for the project: **Light exposure, Activity** and **Sleep-related patterns**

HOW THE DATA COLLECTION

- NHANES selected participants using **representative population sampling**
- Participants wore **wrist-mounted accelerometer devices**
- Sensors continuously recorded: **Physical activity levels and Ambient light exposure**
- Data **collected continuously** over multiple days
- Sensors generated **time-stamped measurements** throughout the day

ETHICAL CONSIDERATIONS

- **Approved** by the National Center for Health Statistics Research **Ethics Review Board**
- Participants provided **informed consent**
- Participation was **voluntary**
- Dataset released publicly in fully **anonymized** form
- All **personal identifiers were removed** to protect participant privacy

FEATURES PREPROCESSING

1. **Light Sensor Noise Removal**

2. **Safe Light Ratio Calculation**
   - Recomputed day–night light ratio

3. **Removing Infinite Values**

4. **Outlier Removal (IQR Method)**

5. **Handling Missing Values**

6. **Removing Original Time Columns**
   - Original time variables were removed after encoding to avoid redundancy.

7.**Circular Encoding of Time Features**

Circadian variables are cyclic (24-hour cycle).

To preserve this structure:

- Peak activity hour
- Peak light hour
- Sleep midpoint

were converted into sine and cosine representations.

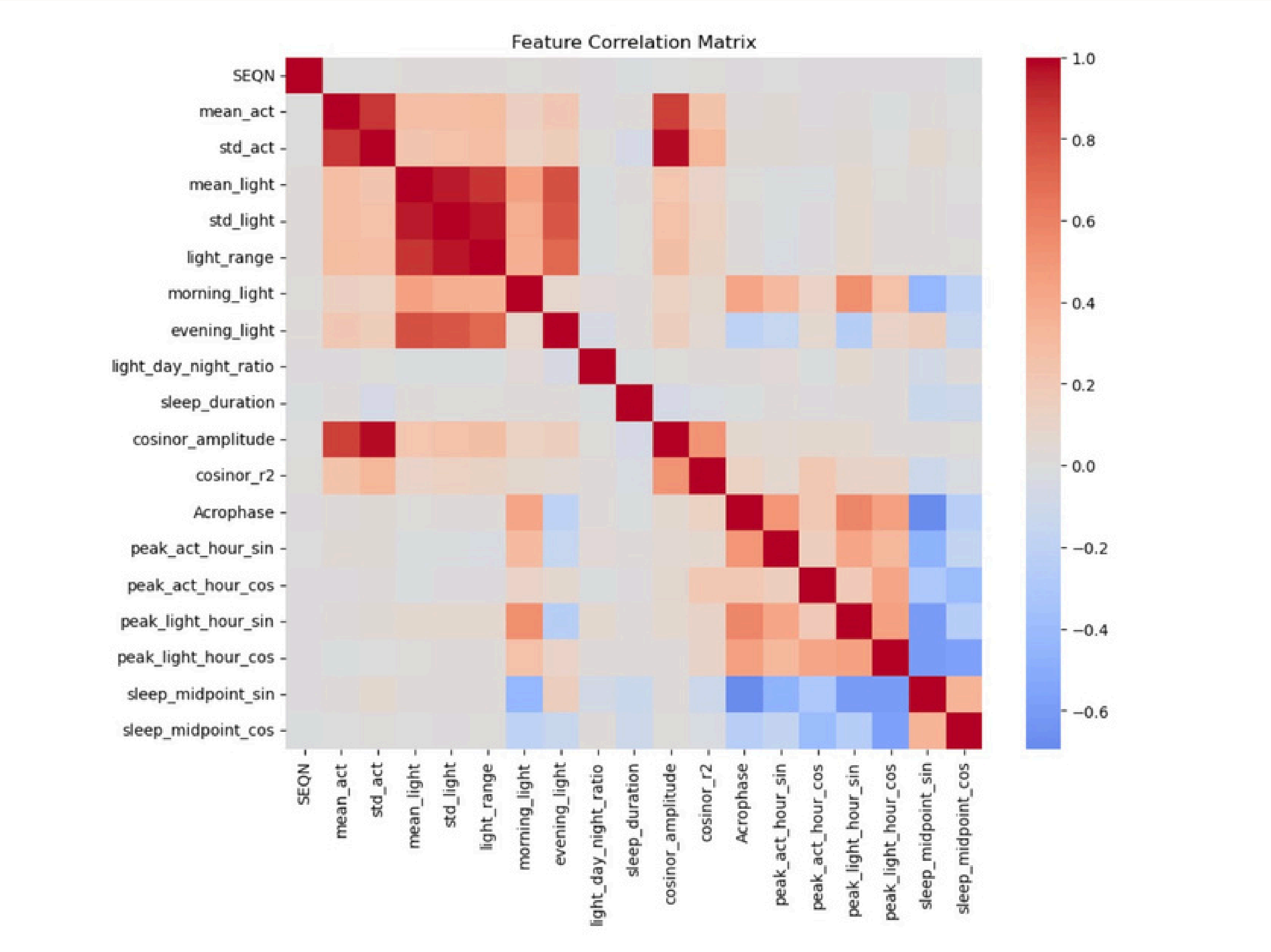
8.**Feature Standardization**

- Applied z-score normalization

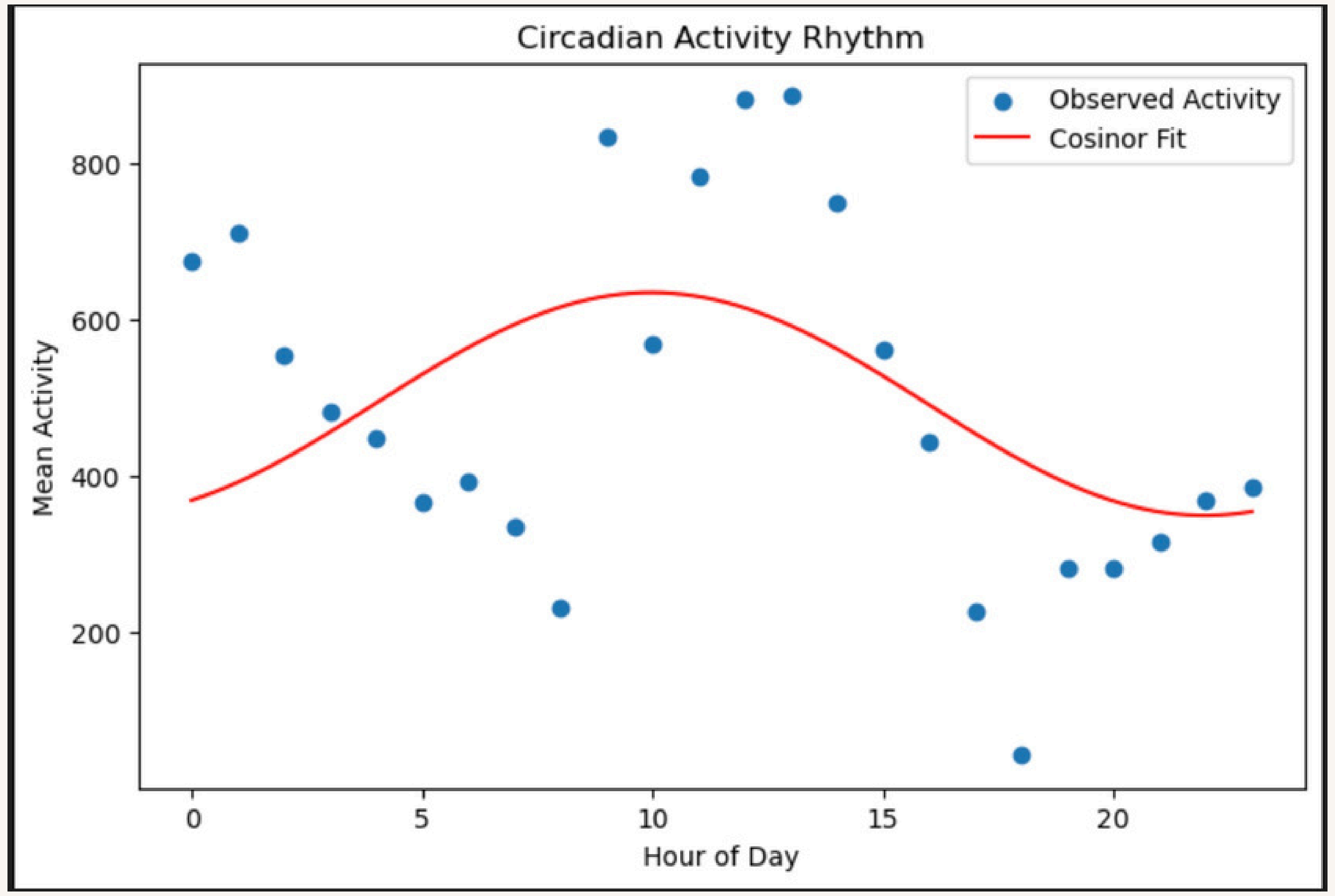
```python
dataset_final = dataset_clean.copy()

num_cols = [
    'mean_act','std_act',
    'mean_light','std_light',
    'light_range','morning_light',
    'evening_light','light_day_night_ratio',
    'sleep_duration','cosinor_amplitude',
    'cosinor_r2'
]

scaler = StandardScaler()
dataset_final[num_cols] = scaler.fit_transform(dataset_final[num_cols])
```

```python
def circular_encode(df, column, max_val=24):
    df[column + '_sin'] = np.sin(2 * np.pi * df[column] / max_val)
    df[column + '_cos'] = np.cos(2 * np.pi * df[column] / max_val)
    return df

for col in ['peak_act_hour','peak_light_hour','sleep_midpoint']:
    dataset_clean = circular_encode(dataset_clean, col)

dataset_clean.drop(columns=[
    'peak_act_hour',
    'peak_light_hour',
    'sleep_midpoint'
], inplace=True)

dataset_clean.dropna(inplace=True)
```

```python
dataset_clean = dataset.copy()

# Remove light sensor noise floor
for col in ['mean_light','morning_light','evening_light']:
    dataset_clean.loc[dataset_clean[col] < 1, col] = np.nan

# Safer light ratio
dataset_clean['light_day_night_ratio'] = (
    dataset_clean['morning_light'] /
    dataset_clean['evening_light']
)

dataset_clean.replace([np.inf, -np.inf], np.nan, inplace=True)

# IQR Outlier Removal
def iqr_filter(df, column):
    Q1 = df[column].quantile(0.25)
    Q3 = df[column].quantile(0.75)
    IQR = Q3 - Q1
    lower = Q1 - 1.5 * IQR
    upper = Q3 + 1.5 * IQR
    return df[(df[column] >= lower) & (df[column] <= upper)]

for col in [
    'mean_act','std_act',
    'mean_light','std_light',
    'cosinor_amplitude'
]:
    dataset_clean = iqr_filter(dataset_clean, col)
```



FEATURES PREPROCESSING



Correlation Analysis

We generated a correlation heatmap to: **Understand relationships between variables** and **detect potential multicollinearity**

Dimensionality Reduction

Not used in this project because:

- The dataset contains **only 19 features**
- All features capture meaningful circadian characteristics
- Dimensionality reduction techniques such as PCA are generally required when datasets contain hundreds or thousands of features.

STEP 01

Data Collection

NHANES accelerometry · sleep, activity, light, demographics

STEP 02

Feature Engineering

24h Cosinor · Mesor · Amplitude · Acrophase · Day/Night ratio

STEP 03

Preprocessing & Transformation

Normalize · scale · encode acrophase as sin(θ)/cos(θ)

STEP 04

Model Training

MLP · LightGBM · KNN · Optuna tuning · 80/20 split

STEP 05

Benchmarking

Internal cross-model · vs. circadian literature

STEP 06

Circular Validation

5-Fold CV · arctan2 · Cyclic RMSE · Cyclic R^2

STEP 07

Biological Validation

SHAP · Gain importance · SCN circadian alignment

STEP 08

Final Outcome

Best model · population-scale · biologically grounded

Implementation Pipeline

ML METHODOLOGY

KNN

- Circadian phase is influenced by a combination of sleep timing, light exposure, and activity levels - **KNN checks whether people with similar daily rhythm profiles naturally cluster around similar acrophase values**
- No assumptions needed, which suits our diverse NHANES population spanning different ages, lifestyles, and health conditions

LightGBM

- Our features (Mesor, Amplitude, Day/Night ratio, morning light) are derived actigraphy statistics - structured tabular data where gradient boosting consistently outperforms other methods
- Circadian phase shifts non-linearly with lifestyle factors, and **LightGBM captures those thresholds and interactions without us having to define them manually**

MLP

- Cosinor modeling extracts summary statistics, but the raw interplay between sleep regularity, light timing, and activity intensity may contain deeper patterns **- MLP's hidden layers can learn those latent biological representations**
- Chosen specifically to test whether a neural approach recovers circadian structure that hand-crafted features alone cannot express

Model Script	R^2 (Linear)	RMSE (Hours)	MAE (Hours)
`mlp_baseline.py`	**0.7089**	**4.9420**	**2.8683**
`lightgbm_baseline.py`	-	5.3372	2.9813
`knn_baseline.py`	0.6117	5.6685	-

Baseline Models Performance

ML METHODOLOGY

Feature Engineering and Model Tuning

Model Architecture	Model Version	R^2 (Goodness of Fit)	RMSE (Hours)	MAE (Hours)
MLP Neural Network	*Linear Baseline*	0.7089	4.9420	2.8683
MLP Neural Network	**Cyclic Optimized**	**0.9353**	**0.9855**	**0.7726**
LightGBM Regressor	*Linear Baseline*	-	5.3372	2.9813
LightGBM Regressor	**Cyclic Optimized**	**-**	**1.0650**	**0.8270**
KNN (K=20)	*Linear Baseline*	0.6117	5.6685	-
KNN (K=20)	**Cyclic Optimized**	**0.8983**	**1.2404**	**-**

Universal Changes (for all 3 models)

- Fixed circular-time issue using sin/cos encoding
- Replaced standard RMSE with circular-aware metrics
- Added interpretable Night / Morning / Afternoon / Evening bins

KNN

- Predicted sine & cosine components separately
- Reconstructed acrophase using arctan2
- Added distance-weighted neighbors
- Dynamically optimized K (1–40) using Cyclic RMSE

LightGBM

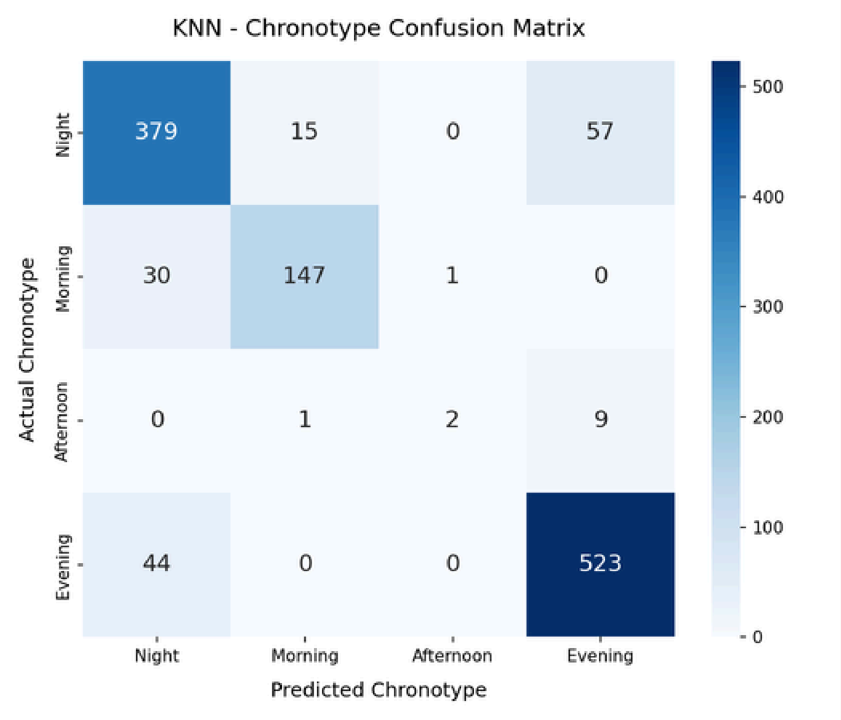
- Trained separate sine and cosine models
- Tuned depth, learning rate, and regularization
- Used 3-Fold RandomizedSearchCV to reduce overfitting
- Added ±1h and ±2h tolerance accuracy metrics

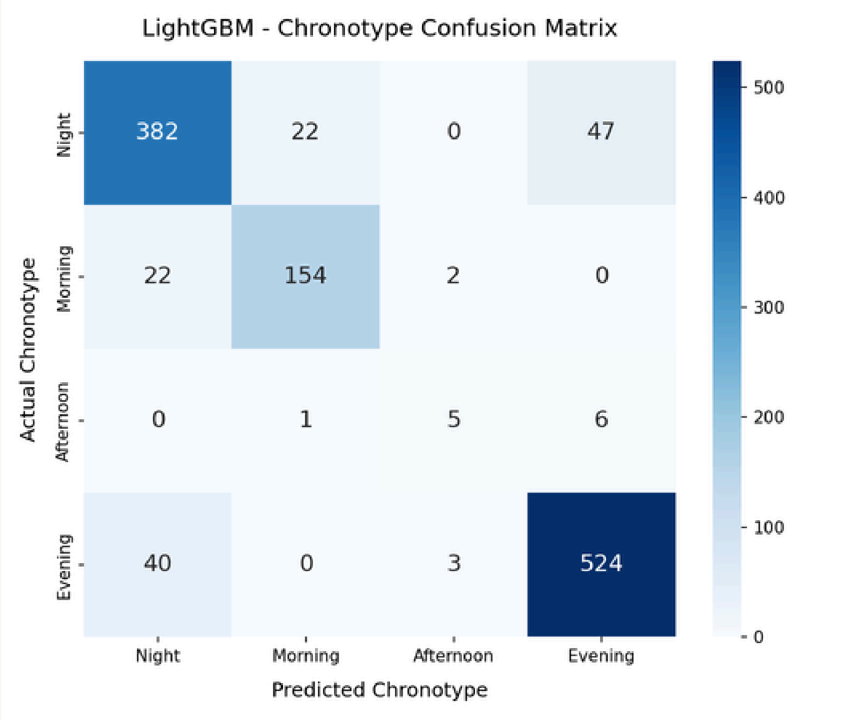
MLP

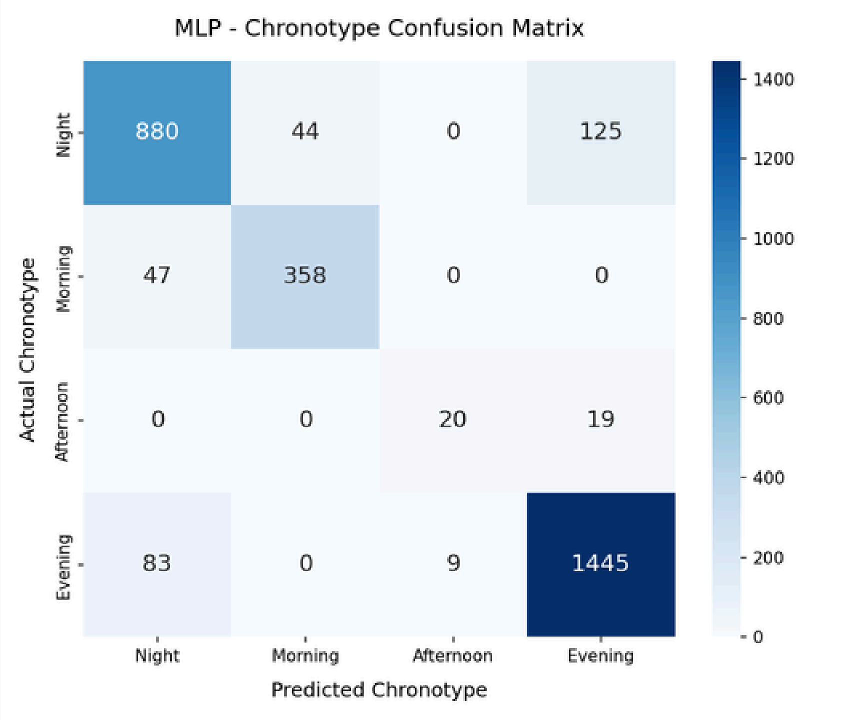
- Filtered unreliable samples using cosinor R^2 threshold
- Added circadian behavior features (light ratio, rhythm strength, sleep efficiency)
- Expanded network: 256 → 128 → 64 with GELU activations
- Added Tanh-bounded outputs for circular prediction
- Replaced MSE with Huber Loss for outlier robustness
- Evaluated using 5-Fold Cross Validation

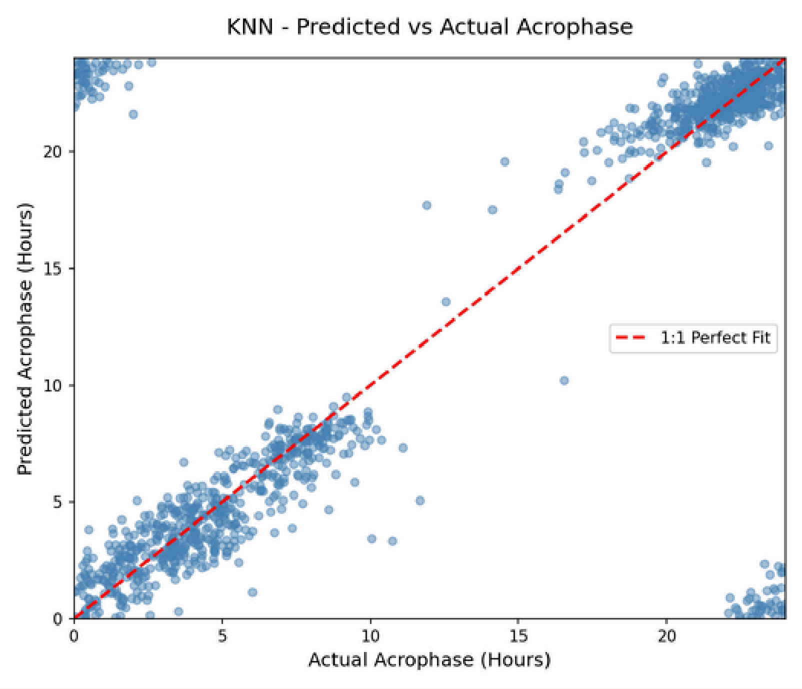
ML METHODOLOGY

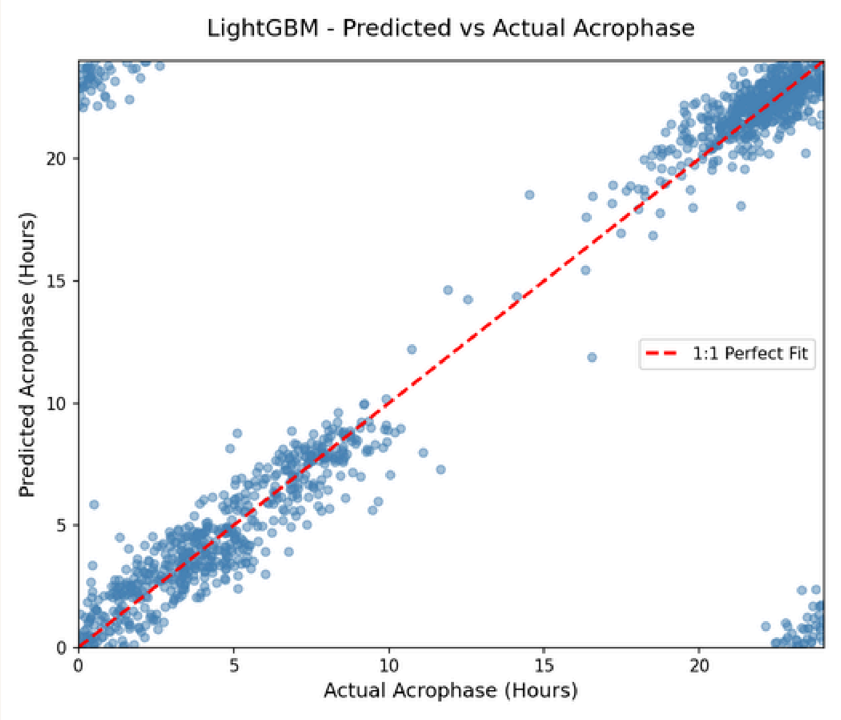
Model Performance

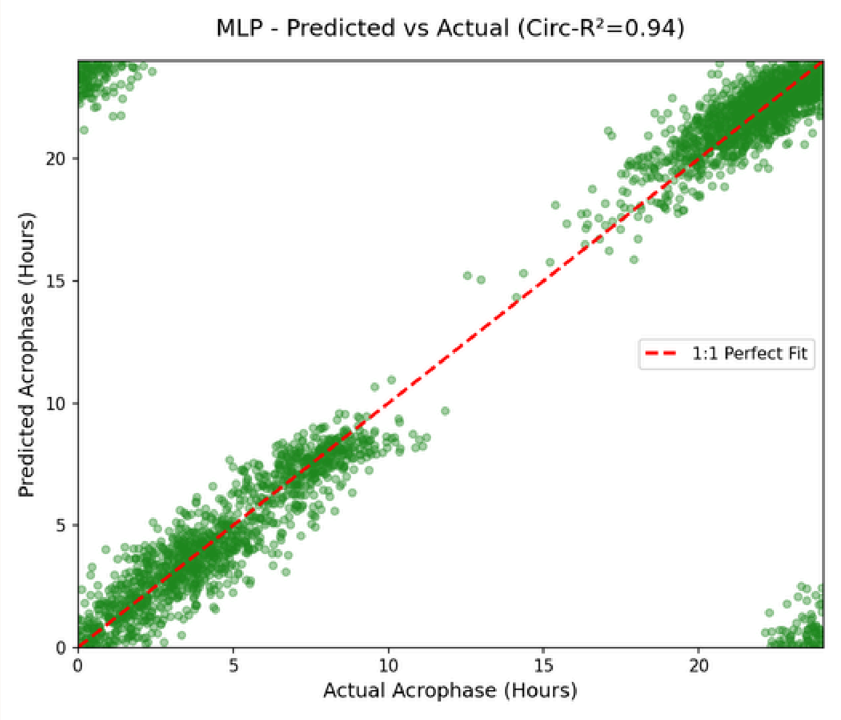




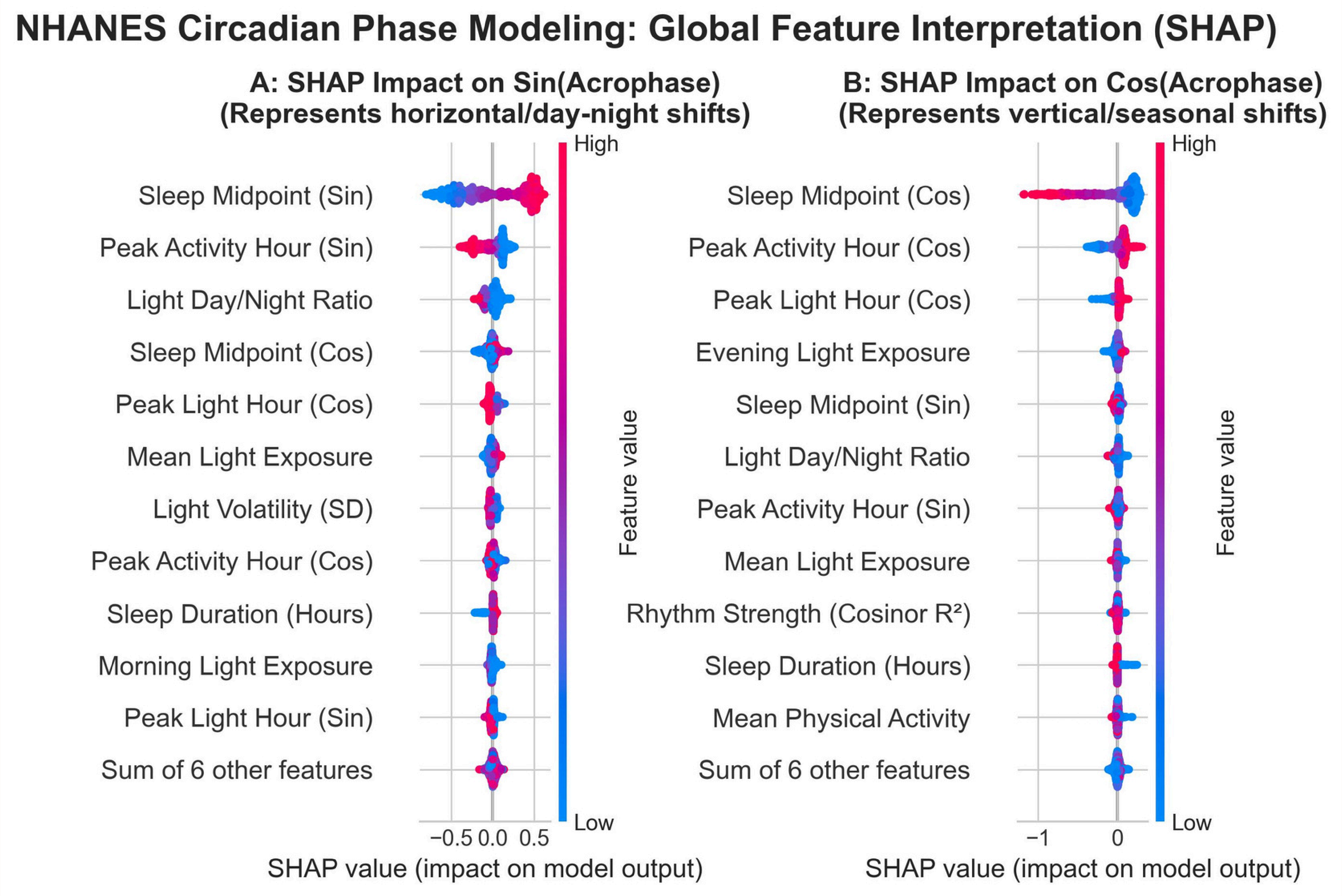








ML METHODOLOGY



Feature Evaluation

ML METHODOLOGY

Circadian Phase Estimation: Model Benchmarking & Literature Comparison

Model / Study	Inputs	Methodology	MAE (h)	RMSE (h)	Other Metrics / Details
Our Deep MLP	Wearable Actigraphy Proxy	4-layer MLP (BatchNorm+ GELU)	0.77 h	0.99 h	R^2: 0.935 (5-Fold CV)
Our LightGBM	Wearable Actigraphy Proxy	Multi-output tree ensemble	0.83 h	1.07 h	88.2% Chronotype Accuracy
Our KNN (K=15)	Wearable Actigraphy Proxy	Scaled distance-weighted KNN	—	1.24 h	87.0% Chronotype Accuracy
Daniel Stone et al. (2019)	Actigraphy + Light	Limit-cycle oscillator model	0.69 h (Day), 0.95 h (Night)	—	80% Day, 68% Night within ±1 h
Huang et al. (2021)	Wearable Activity	Mathematical physical models	~1.2–1.5 h	—	Cross-population generalizability
Xu et al. (2026)	Activity + Light	Tree-ML + Sequence models	1.19 h (Circular)	—	Real-time acrophase prediction
Walch et al. (2019)	Acc + PPG Heart Rate	Deep learning wearable pipeline	—	—	73% Accuracy (Sleep cycles), κ: 0.65-0.69

Industry Benchmarking

Daniel Stone et al. (2019)
- Used wrist actigraphy + light exposure
- Predicted circadian phase in shift workers
- Model: Limit-cycle oscillator model

Huang et al. (2021)
- Used wrist-worn wearable activity data
- Predicted biological circadian phase
- Models: Mathematical circadian models

Jennifer Walch et al. (2019)
- Used accelerometer + PPG heart-rate wearable data
- Predicted sleep stages/temporal rhythms
- Models: Deep learning wearable pipeline

Xu et al. (2026)
- Used wearable activity + light exposure data
- Predicted real-time circadian phase/acrophase
- Models: Tree-based ML + sequence models

ML METHODOLOGY

FUTURE PROSPECTS

Our solution is highly deployable in a university environment.

- Students often experience irregular sleep schedules, sleep deprivation, and circadian misalignment
- The model can help monitor biological rhythms using wearable/smartphone data
- Could improve student wellness, productivity, and academic performance

How Would It Work?

Data Collection

- Students connect smartwatches or fitness trackers
- Collect activity, heart rate, and light exposure data passively

Model Deployment

- Data processed through cloud/server pipelines
- Trained MLP/LightGBM models predict circadian acrophase

Student & University Insights

- Personalized recommendations for sleep, study, and exercise timing
- Aggregate trends can help optimize class schedules and wellness initiatives

Challenges During Large-Scale Deployment

Device Variability

- Different wearables generate inconsistent/noisy data

Model Generalization

- NHANES-trained models may not fully match Plaksha student behavior

Concept Drift

- Exams, deadlines, and festivals can suddenly disrupt sleep patterns

Privacy & Ethics

- Continuous health monitoring requires strong consent and anonymization policies

Infrastructure Costs

- Real-time data processing and storage for thousands of students is expensive

REFERENCES

Burgess, Helen J., and Charmane I. Eastman. "The dim light melatonin onset following fixed and free sleep schedules." Journal of sleep research 14, no. 3 (2005): 229-237.

Brown, Lindsey S., Melissa A. St. Hilaire, Andrew W. McHill, Andrew JK Phillips, Laura K. Barger, Akane Sano, Charles A. Czeisler, Francis J. Doyle III, and Elizabeth B. Klerman. "A classification approach to estimating human circadian phase under circadian alignment from actigraphy and photometry data." Journal of pineal research 71, no. 1 (2021): e12745.

Robbins, Rebecca, Matthew D. Weaver, Jason P. Sullivan, Stuart F. Quan, Katherine Gilmore, Samantha Shaw, Abigail Benz et al. "Accuracy of three commercial wearable devices for sleep tracking in healthy adults." Sensors 24, no. 20 (2024): 6532.

Gao, Chenlu, Shahab Haghayegh, Max Wagner, Ruixue Cai, Kun Hu, Lei Gao, and Peng Li. "Approaches for assessing circadian rest-activity patterns using actigraphy in cohort and population-based studies." Current sleep medicine reports 9, no. 4 (2023): 247-256.

Stone, Julia E., Xavier L. Aubert, Henning Maass, Andrew JK Phillips, Michelle Magee, Mark E. Howard, Steven W. Lockley, Shantha MW Rajaratnam, and Tracey L. Sletten. "Application of a limit-cycle oscillator model for prediction of circadian phase in rotating night shift workers." Scientific reports 9, no. 1 (2019): 11032.

Huang, Yitong, Caleb Mayer, Philip Cheng, Alankrita Siddula, Helen J. Burgess, Christopher Drake, Cathy Goldstein, Olivia Walch, and Daniel B. Forger. "Predicting circadian phase across populations: a comparison of mathematical models and wearable devices." Sleep 44, no. 10 (2021): zsab126.

Walch, Olivia, Yitong Huang, Daniel Forger, and Cathy Goldstein. "Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device." Sleep 42, no. 12 (2019): zsz180.

Xu, Mengzhu, Nemanja Cabrilo, Merel van Gilst, and Jean-Paul Linnartz. "Toward Real-Time Circadian Phase Estimation with Low Latency from Wearable Sensing Data." arXiv preprint arXiv:2605.00910 (2026).

THANK YOU