Business Analytics and Time Series Forecasting:

Forecasting the Next Day's Bedtime for Better Sleep Readiness

Cindy Liu 110078701
Jamie Chan 111078515
Matthew Bobea 111078860
Wilma Chen 112078518

Instructor: Galit Shmueli

National Tsinghua University, 2024 January 2nd

One-page executive summary

Smart home devices struggle to personalize sleep features due to the absence of automatic adjustments based on individual sleep patterns. Examples of automatic adjustments include automatic blue light and temperature adjustments, falling short of user expectations for advanced health solutions. Our business seeks to provide smart home device manufacturers with personalized sleep prediction solutions to enhance user experience. The goal is next-day-ahead forecasts of individuals' bedtimes to promote sleep readiness. The client, a manufacturer of smart home devices, can utilize our forecasting model to enhance their sleep readiness product, using the forecasted bedtime to turn on/off features of their device that enhance sleep readiness. By incorporating our predictive model into their product, they aim to offer users a more intelligent and personalized sleep experience.

The data comprises 84 days of bedtimes from 4 students. The data was collected through a combination of each user's electronic devices as well as self-reported information. Additional predictors of bedtime were explored, including phone use, computer use, walking distance, holidays, and workloads, etc.

In our forecasting solution, we employ various machine learning methods while utilizing simple forecasting methods as benchmarks. The dataset includes 84 days of data, covering the period from September 25, 2023, to December 17, 2023. Roll-forward validation was used to incorporate new daily data into the analysis. We tested multiple statistical and machine learning methods for predicting bedtimes. Models were assessed by comparing forecasting performance. The selection process favored simple models with fewer variables which were shown to have better performance in the final forecasts. Final models were visualized to be evaluated through plots and checked for overall error metrics and bias towards under- or over-forecasting. Also we found forecasting bedtimes within 15 intervals did not improve forecasting accuracy, unfortunately.

Overall, it was discovered that accurate prediction of certain bedtime patterns can be challenging without external information, highlighting the importance of incorporating external factors into forecasting models. It is crucial to recognize that the optimal forecasting models can vary significantly from person to person, underscoring the need for personalized approaches. Utilizing electronic devices to automate data collection processes can improve efficiency and reduce human errors. To enhance existing forecasting models, it is recommended to gather a year-long dataset to analyze seasonal trends and patterns comprehensively. Additionally, exploring the impact of screen usage timing on sleep beyond daily totals, including late-night phone use, can provide valuable insights.

1. Problem description

Business Goal

Our Business goal is to provide smart home device manufacturers with a tailored predictive modeling solution that enhances their products by enabling personalized sleep predictions that prepare users for better sleep. By taking into account an individual's specific life habits and sleep patterns, this approach allows device settings to sync more accurately to those with irregular bedtimes than traditional fixed bedtime methods. By addressing the absence of automatic adjustments based on individual sleep patterns, we aim to boost user satisfaction by empowering manufacturers to refine sleep adjustment settings and enhance the value of their products.

Forecasting Goal

In this work, the primary objective is to conduct a forward-looking forecast of the bedtime (measured by minute) of a user for the next day. The series under consideration is their bedtime data, and the forecast horizon is set to one day. The prediction time is specifically scheduled for the next day at 7 PM for that day's bedtime. The client, who intends to utilize our forecasting model to improve their sleep readiness product.

2. Data Description

The data collection process involved collecting bedtime from a sample comprising four students: 'A', 'C', 'K', and 'M', with bedtime data being collected over a duration of 84 days, corresponding to a period of approximately three months from September 25th to December 17th (See Figure 1). Bedtime data was gathered from the health app on each individual's iPhone, wherein bedtime was defined as the commencement of an extended duration during which the absence of user motion was detected.

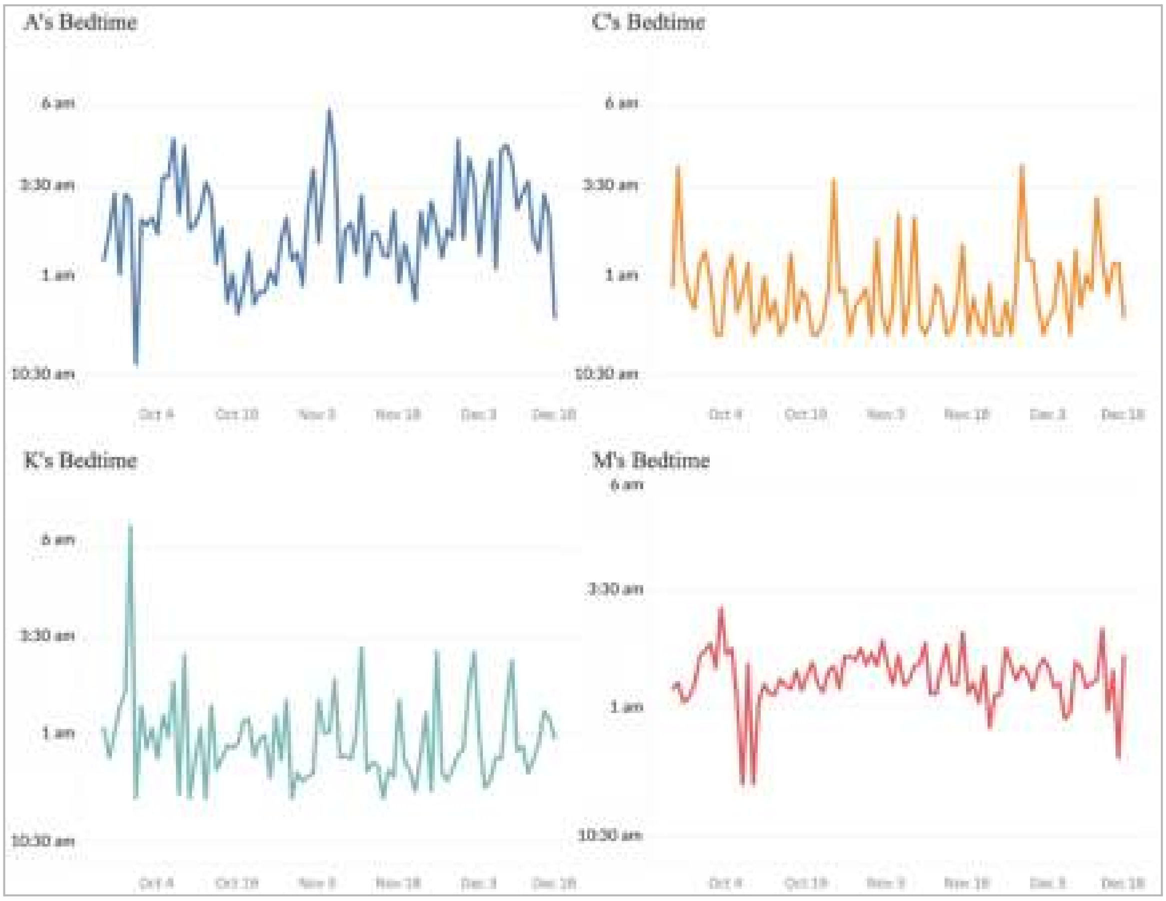


Figure 1: Daily bedtimes of four students

In order to construct an effective time series forecasting model, several external variables were considered. These variables comprised the walking distance before 7 PM, PC screen time, iPhone screen time, iPhone pickup times, as well as the presence or absence of holidays and workload factors. A snapshot of the data can be seen in Figure 2. To address the blank spaces in the figure, we wished to highlight that only user M had access to the variable 'pickups', while the variable 'PC screen time' was not available to him/her.

In order to enhance the accuracy of our forecasts, we experimented with predicting bedtimes in 15 minute windows. In this case, the dependent variable used was the "Bedtime 15 min" column in place of the “Bedtime” column as shown in Figure 2.

When the current day's data was not available at the time of forecasting, we relied on the previous day's data to forecast the bedtime for the next day. For instance, as the full day's phone usage is unknown until the end of the day, it cannot be used to predict the bedtime for that night. Appendix 1 notes these variables as "Lag1" while also providing additional information regarding each external variable and their sources.

Name id	Date	Bedtime	Bedtime 15 min	Holiday	Sleep duration (approx)	Distance before 7pm	Schedule workload	PC screen time	Phone screen time	Pickups
A	2023/9/24	25.45	25.25	1	0.27	2.07	1	6.683	3.43	
C	2023/9/24	24.75	24.75	1	0.30	2.21	4	11.033	1.40	
K	2023/9/24	25.40	25.25	1	0.34	1.64	5	2.433	7.93	
M	2023/9/24	25.18	25.00	1	0.30	0.03	2		7.47	83
A	2023/9/25	26.23	26.00	0	0.26	7.27	4	8.650	4.72	
C	2023/9/25	28.12	28.00	0	0.20	3.25	3	8.983	4.10	
K	2023/9/25	25.53	25.50	0	0.26	2.73	4	6.200	5.47	
M	2023/9/25	24.45	24.25	0	0.27	1.42	5		6.03	86
A	2023/9/26	27.33	27.25	0	0.18	3.98	5	9.117	5.77	

Figure 2: Data Snapshot

3. Data preparation

Bedtime data was preprocessed through the following steps. First, raw bedtime data was cleaned and formatted. Next, we noticed that the app recorded multiple bedtimes each day. To identify which was the real bedtime, we calculated the duration of sleep following each bedtime and selected the bedtime which has the longest duration. Third, to address challenges arising from the 24-hour clock schedule mislabeling bedtimes after midnight as the next day, times were scaled and shifted to continue under the same day for the subsequent 12 hours. For instance, 1:00 AM sleep would be represented as 25, resulting in bedtimes ranging from 12 to 35, corresponding to the period from 12 PM to 11 AM on the same day. Fourth, PC screen time and phone screen time were aggregated by day.

In order to predict an individual's bedtime based on their daily travel distance, it was necessary to establish a cut-off point before their latest preferred bedtime. This cut-off point allowed for the aggregation of travel data for the sleep forecast of that particular night. For instance, if a user typically sleeps at 10:00 PM, the travel distance data would be aggregated well before this time, such as at 7:00 PM. Therefore, 7:00 PM was selected as the cut-off point for collecting travel data and running the sleep prediction model.

Missing values (NA's) in the dataset were imputed using a naive forecasting approach due to the limitations of using centered moving averages (CMA). CMAs are unsuitable for estimating f(t) near the ends of the time series meaning that it can be affected by recent missing values. Furthermore, the CMA method assigned equal weight to all data points within the window, disregarding the importance of recent data in certain cases. Additionally, centered moving averages require a continuous series of data points, making it challenging to accurately calculate them in the presence of consecutive periods of missing data. Overall, these populated NA cells had a negligible impact, accounting for less than 1% of the total data.

4. Forecasting solution

Method applied

To forecast the bedtime for each series, we built the model for each series separately. The methods included Exponential Smoothing (ETS), Regression with and without external variables, Regression with AutoRegressive (AR) residuals, ARIMA with and without external variables, and Neural Network (NNETAR) with and without external variables. To set a benchmark, we visualized and compared Naive, Seasonal Naive, and Sample Mean methods with models. From the weekly bedtime patterns plot (see Appendix 2), it seems that M had weekly seasonality. Therefore, we also included seasonal naive as a benchmark. The external variables generally had weak correlation with one another allowing us to apply multiple combinations of external variables to each model. Our analysis revealed weak correlation between bedtime data and external variables.

Data Partitioning

We used 70 days of daily data, from 2023/9/25 to 2023/12/3, as the training set. The validation set is from 2023/12/4 to 2023/12/17. We applied roll-forward validation, integrating new data that came in every day.

Evaluation Criterion

The model selection process involves a three-step approach. Initially, in method categories like TSLM and ARIMA, where there were nearly 100 models due to various combinations of external variables, we selected the most optimal model from each category based on plots, considering RMSE, and favoring models with fewer external variables. Subsequently, these chosen models from different categories were plotted together to evaluate their performance across various scenarios (see Appendix 3 for an example). Including examining periods such as peaks and valleys, and comparing training and validation outcomes to prevent overfitting risks. Our goal is for the sleep readiness function activating one hour prior to the forecasted bedtime. Over-forecasting by more than an hour may result in users not being exposed to the sleep readiness function. Therefore, we prefer under-forecasting rather than over-forecasting. We annotated points over 1 or smaller than -1 on the residual plot, indicating instances of over- or under-forecasting. Following this, we identified the best model.

In the third step, we recompared this best model with three benchmarks through plots. (see Figure 3, Appendix 4, 5, 6). Although in a few specific periods (e.g., M's data), benchmarks showed better predictions, suggesting possibilities for ensemble investigations, overall, the best model for each individual performed better, especially in minimizing over-forecasting instances. The best models also exhibited superior RMSE compared to the benchmarks, leading to the finalization of the best model for each person (see Figure 4).

A side note: It was suspected that bedtime rounded down to the nearest 15-minute interval could provide a better prediction. Unfortunately, it did not enhance the model's performance.

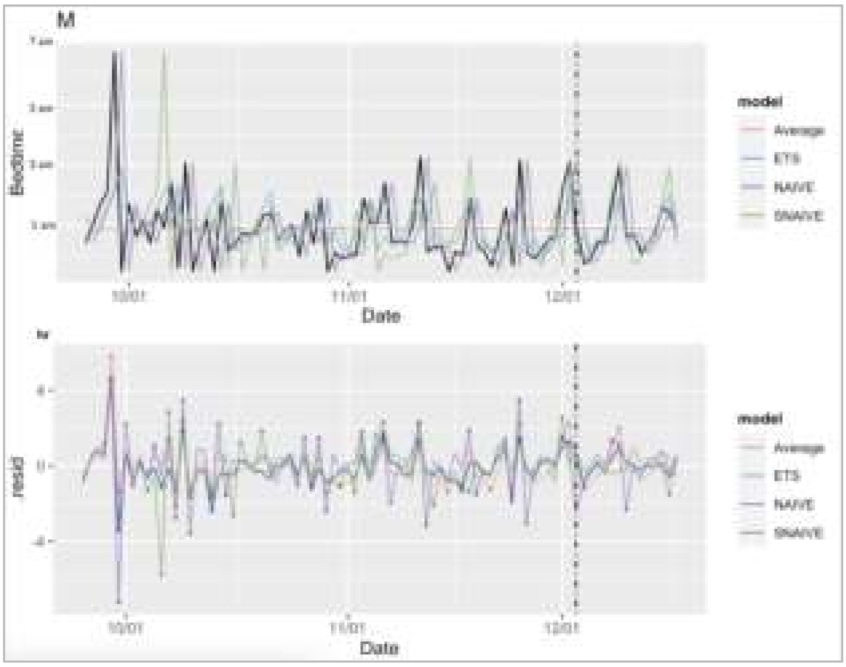


Figure 3: M's best model with benchmark

Name	Formula of the selected best model	Parameters
A	ARIMA(Bedtime ~Yesterday’s Sleep duration + Tomorrow’s workload)	<LM w/ ARIMA(0,0,3) errors>
C	NNETAR(Bedtime ~ Yesterday’s Sleep duration + Today’s walking distance before 7pm + Today’s workload)	set.seed(201), <NNAR(1,1,3)[7]>
K	NNETAR(Bedtime ~ Yesterday’s PC screen time + Yesterday’s Phone screen time)	set.seed(201), <NNAR(2,1,3)[7]>
M	ETS(Bedtime)	<ETS(M,N,A)>

Figure 4: The selected best models for each person

5. Time plot of series with future forecasts

After selecting the best models for each individual, we retrained the model using the entire dataset, maintaining the parameters identified in the previous round of training. Subsequently, we generated a one-day ahead forecast for December 18th with 95% prediction intervals (see Figure 5 for M’s time plot with forecast). For time plots of other individuals' forecasts, see Appendix 7, 8, and 9.

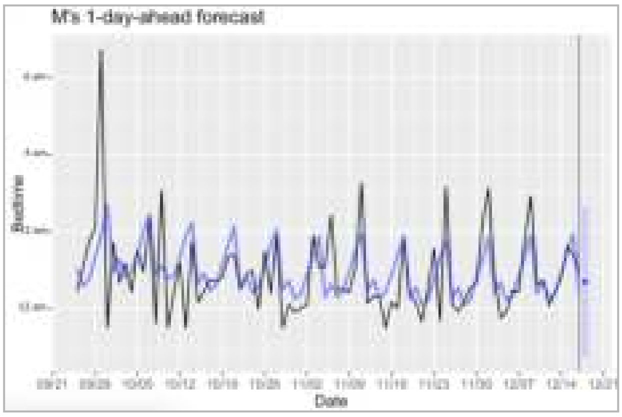


Figure 5: M’s time plot with future forecast

6. Conclusions

Recommendations

It is crucial to acknowledge that certain bedtime series exhibit inherent complexities that make them challenging to predict accurately without the incorporation of external information, making said predictors a critical factor in model prediction. Likewise, it is essential to consider that the optimal forecasting models can differ significantly among individuals, highlighting the need for personalized approaches. Lastly, the utilization of electronic predictors offers the advantage of automating data collection processes, enhancing efficiency and reducing human error.

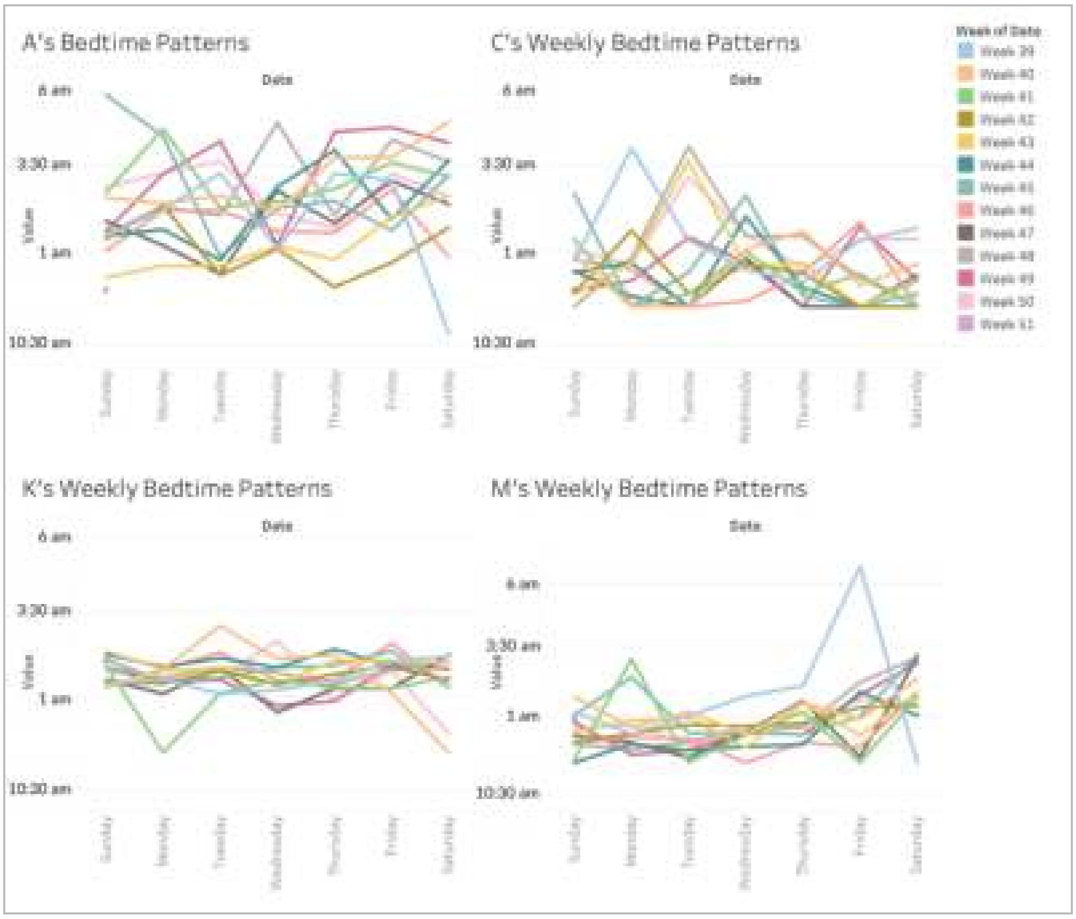
Limitations & Future works

Our methods are limited in that individuals with nocturnal work schedules struggle to fit within the 7 PM prediction timeframe and standard time scaling used during preprocessing. Furthermore, late-night phone usage disrupts accurate sleep onset detection, as bedtime is currently inferred from phone inactivity. Finally, with only one semester's worth of data, we haven't explored the potential impact of seasonal trends, such as differences between winter and summer sleep patterns. In terms of future works, the existing forecasting model should be revisited with a year-long dataset, allowing for a comprehensive analysis of potential quarterly seasonality and data patterns. Additionally, investigating the impact of screen usage timing on sleep, beyond daily totals, could be highly valuable. This could involve exploring the differential effects of late-night phone use, for instance. Next, developing a robust method for handling time zone changes within the time series data analysis is crucial for accurate forecasting. Lastly, to promote model selection during times of automation, we plan to use a more advanced method of model selection when compared to traditional RMSE metrics by utilizing a three-step approach. First, RMSE assesses overall performance. Second, a weighted penalty scheme, harsher for over-forecasting than under-forecasting, adjusts RMSE to prioritize accurate ground truth forecasts. Finally, the model with the lowest RMSE is chosen, balancing minimization and overforecasting penalty. This method prioritizes accurate forecasts while effectively addressing overestimation's potential impacts.

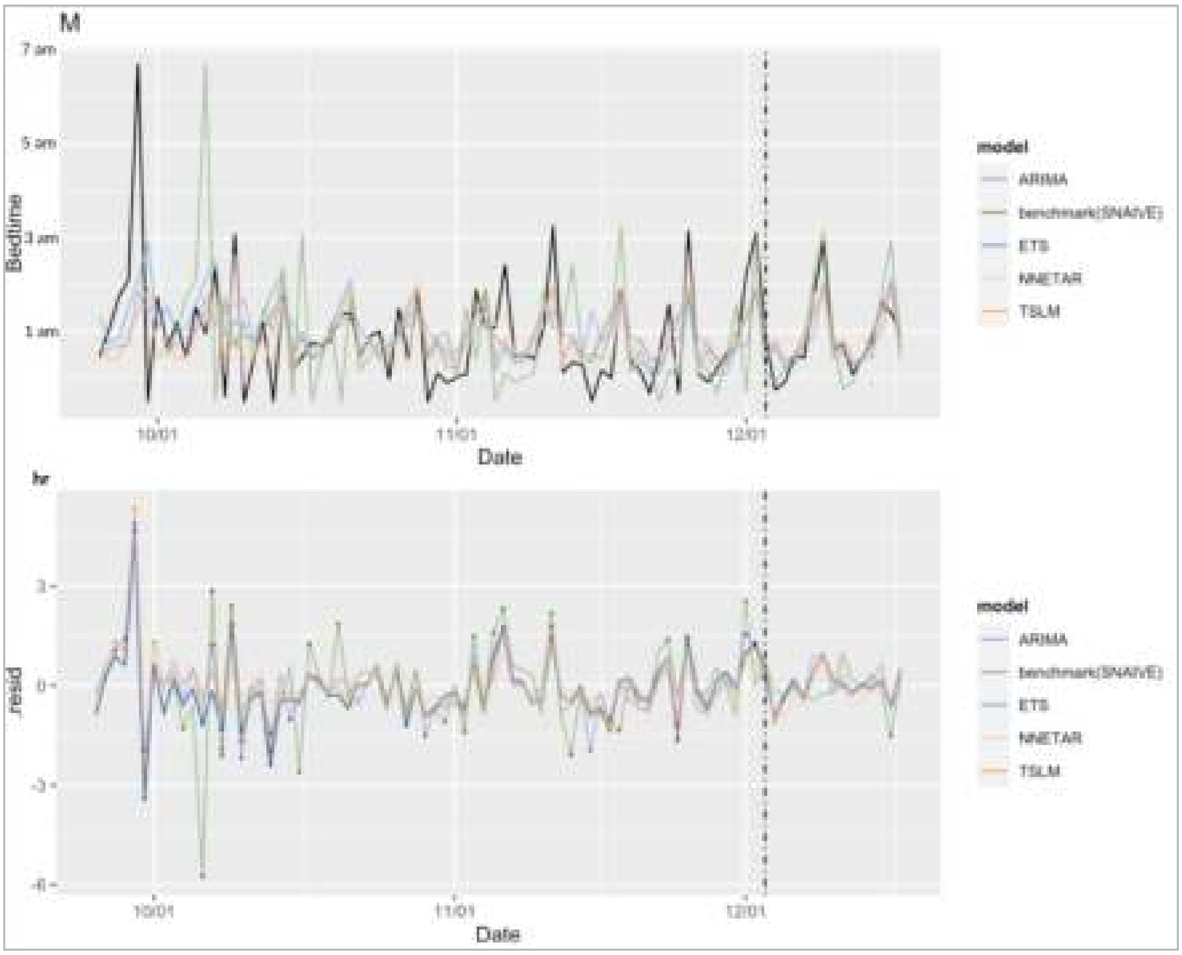
Appendix

Variable	Source	Frequency	Has Lag1 Data
Date	Phone Health App	Daily	No
Bedtime	Phone Health App	Daily	No
Bedtime 15 min	Phone Health App	Daily	No
Holiday	Calendar	Daily	No
Sleep Duration (approx)	Calculated Using the Phone's Health App	Daily	Yes
Distance Before 7pm	Calculated Using the Phone's Health App	Minute	No
Schedule Workload	User Input	Daily	No
PC Screen Time	PC app	Daily	Yes
Phone Screen Time	Phone App	Daily	Yes
Pickups	Phone App	Daily	Yes

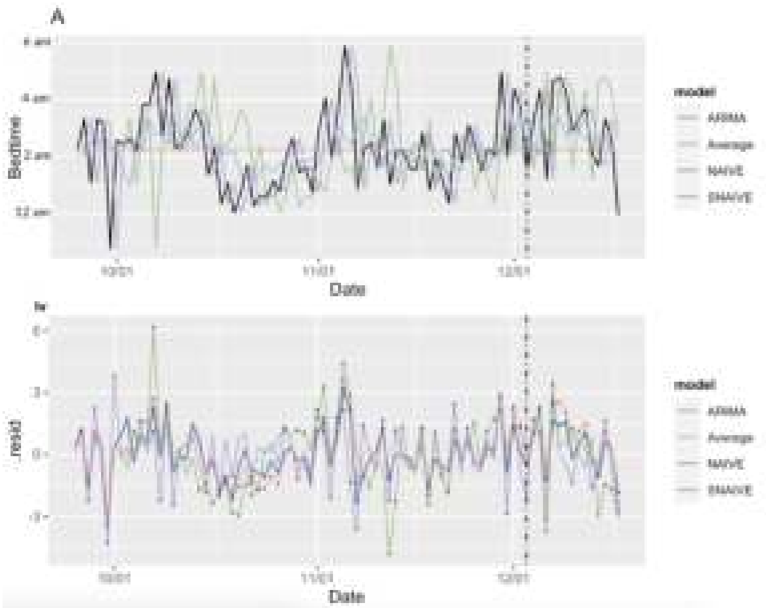
Appendix 1: Variable Information



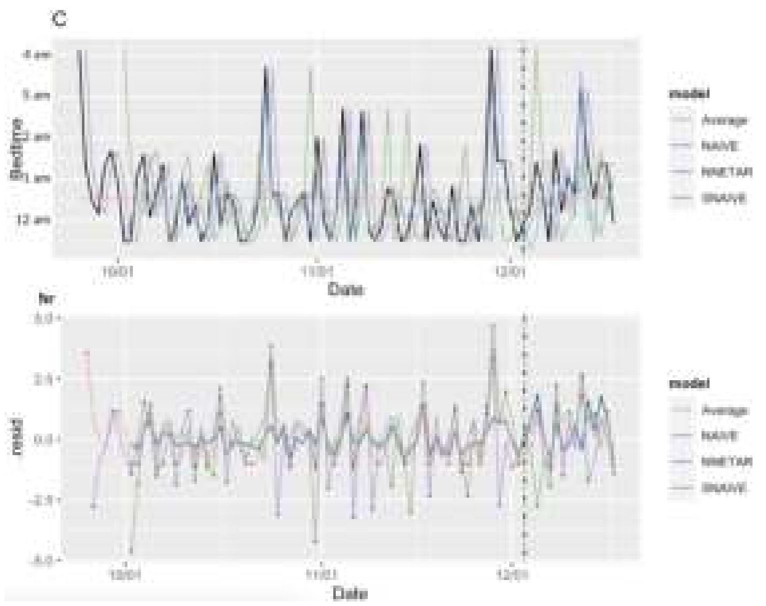
Appendix 2: A, C, K, M Weekly Bedtime Patterns



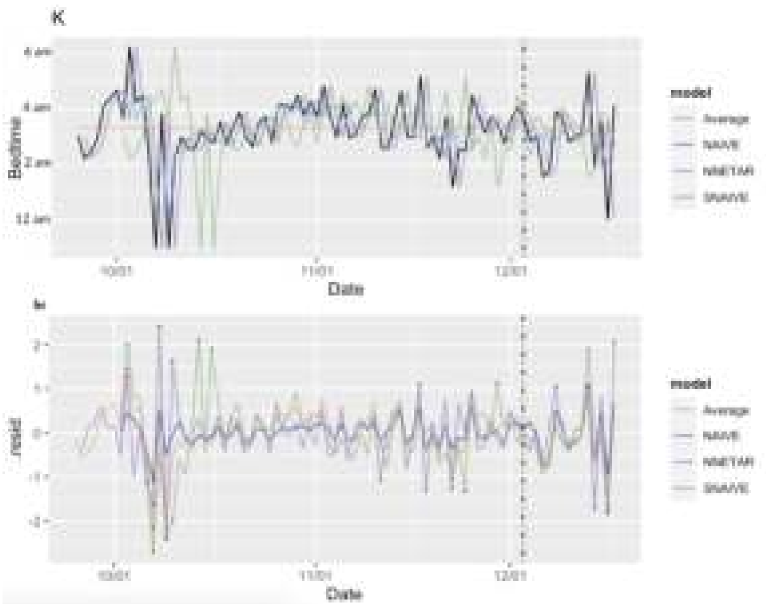
Appendix 3: M's multiple models with the lowest-RMSE-benchmark



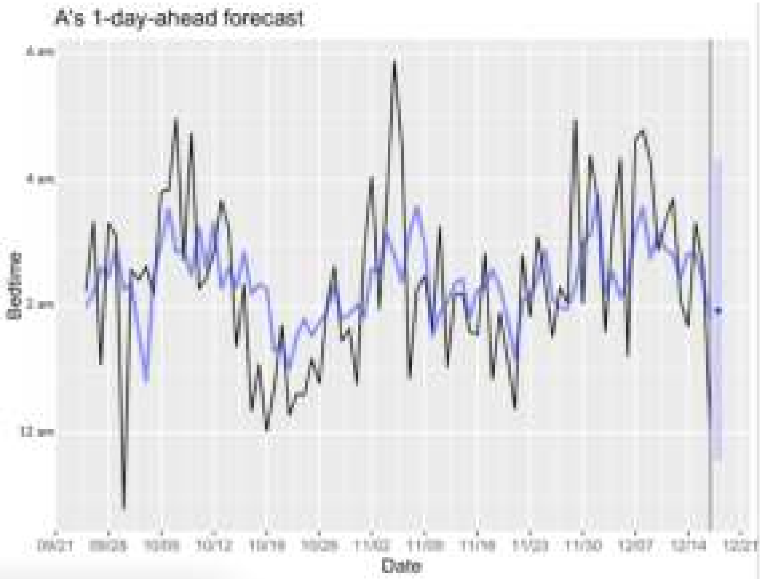
Appendix 4: A's best model with benchmarks



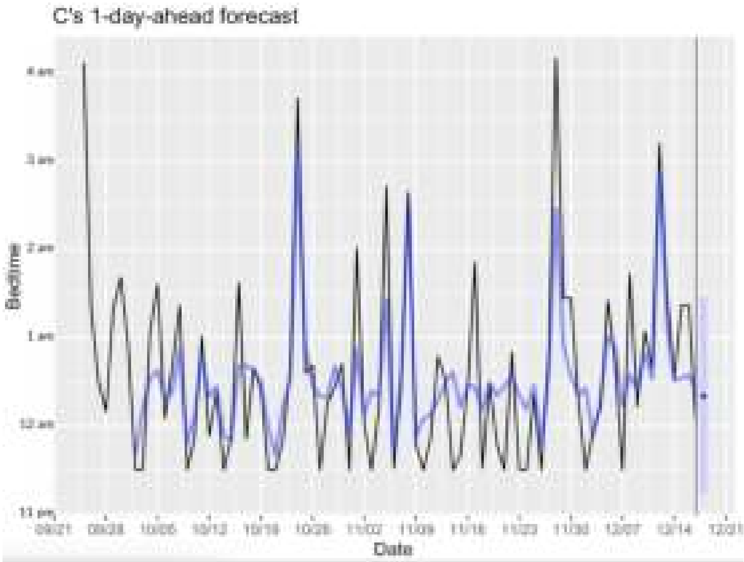
Appendix 5: C's best model with benchmarks



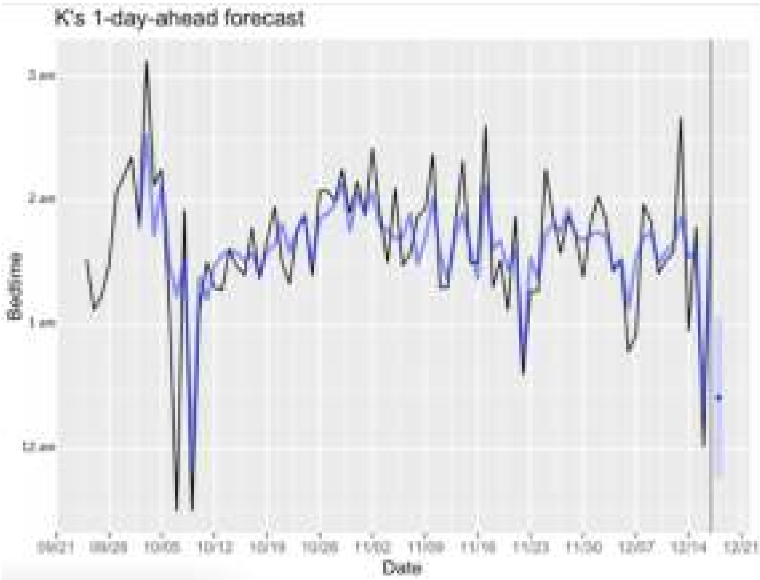
Appendix 6: K's best model with benchmarks



Appendix 7: A's time plot with future forecast



Appendix 8: C's time plot with future forecast



Appendix 9: K's time plot with future forecast