Predicting the Lifespan of Drosophila Melanogaster : A Novel Application of CNN and Zero-inflated ACP Models

Yi (Ann) Zhang, V.A. Samaranayake, Gayla Olbricht, Matthew Thimgan

Missouri University of Science and Technology

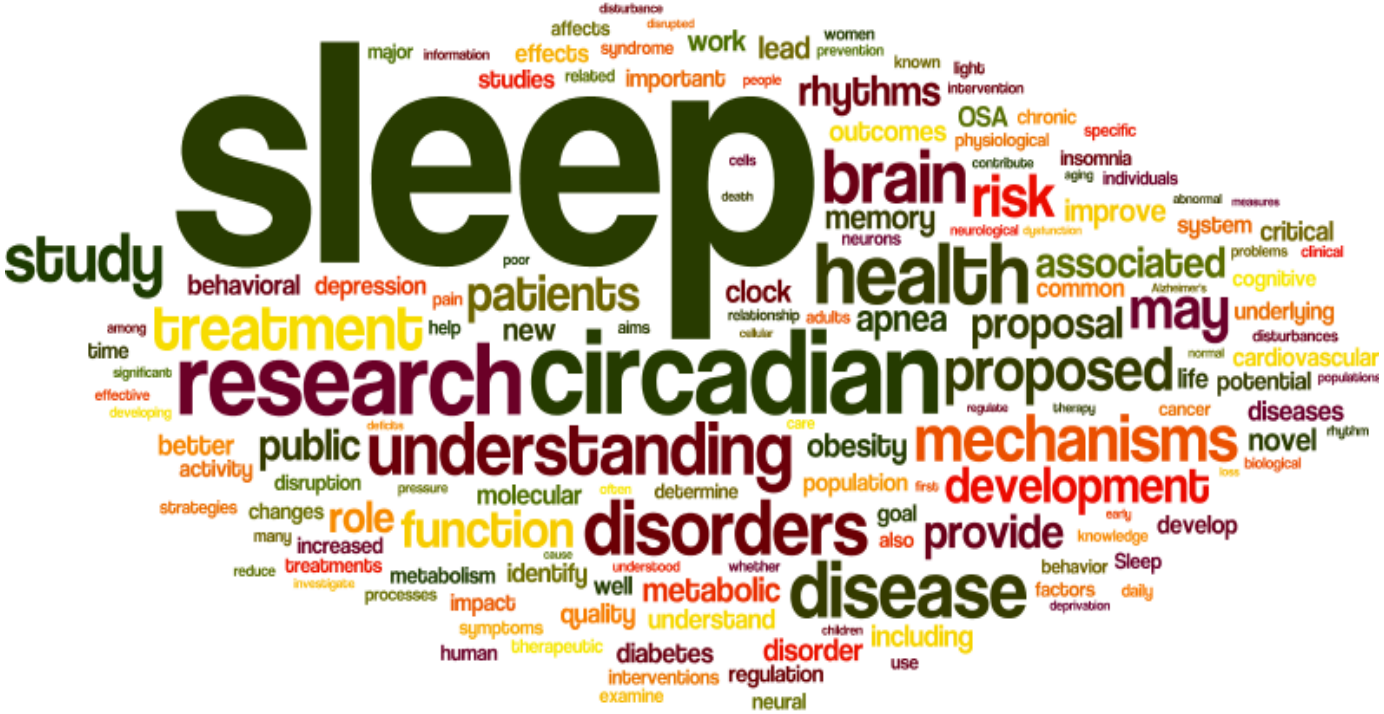
SDSS, 06/05/2020

Outline

- ❑ Background and Motivation
- ❑ Data
- ❑ Methodology and Results
- ❑ Conclusions

Background and Motivation

- Consolidated sleep benefits the animal, while inadequate sleep has deleterious effects on health and mortality.
- Fundamental research in genetically tractable model organisms helps to understand the complexity of sleep regulation and its impact.
- The objective of our study is to investigate the relationship between sleep and ageing in such a model organism.



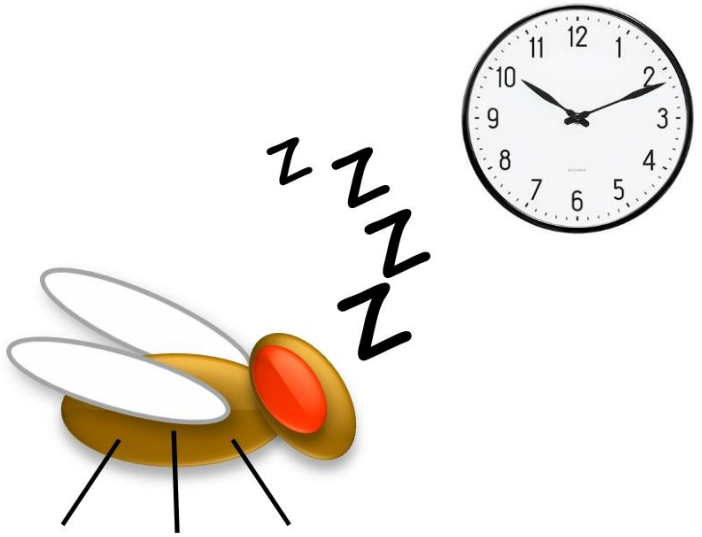
Word Cloud from abstracts of funded NIH sleep/circadian research grants.

Source: Twery, Michael, Opportunities for BSM research in Agencies Across the Federal Government, Webinar, December 10, 2018, National Heart, Lung, and Blood Institute.

Background and Motivation

- Drosophila melanogaster has emerged as an ideal model organism to study sleep.
- Features of sleep architecture associated with lifespan in flies also apply to our understanding of sleep and health. (Wallace, Stone et al. 2018)
- Flies have their primary sleep period at night. And same neurotransmitters and proteins observed in both humans and flies regulate sleep and wakefulness (Andretic and Shaw 2005, Allada, Cirelli et al. 2017)
- Fruit flies have relatively short lifespans which enables monitoring their sleep over their whole lifespans

Why fruit fly?

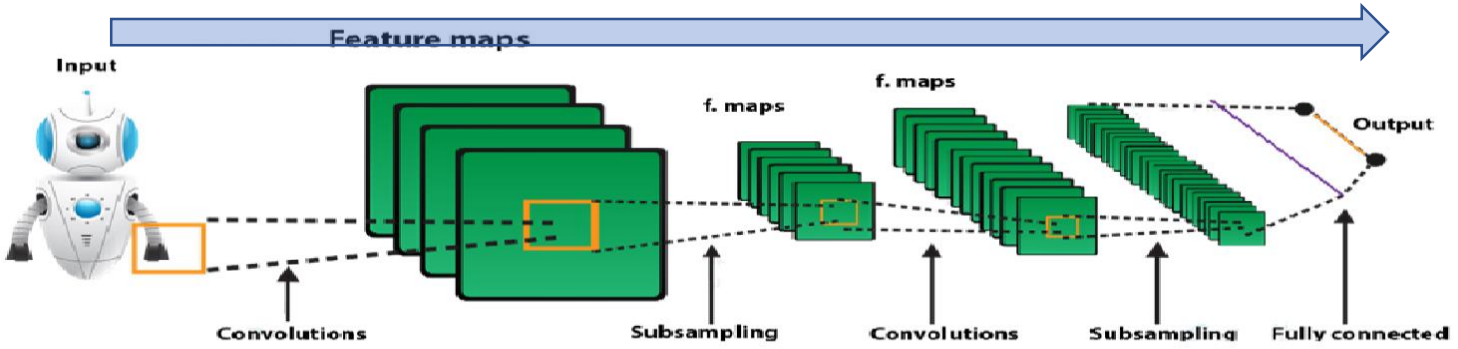


Background and Motivation

- With the development of movement-detecting monitors, high-frequency time series of count data have become common in areas such as biology.
- A model that accounts for both the dependent structure and time varying zero-inflation has been introduced recently, as a generalization of the Poisson process.
- Convolutional neural networks (CNN), have become highly useful in modeling complex data sets.

Sleep

?



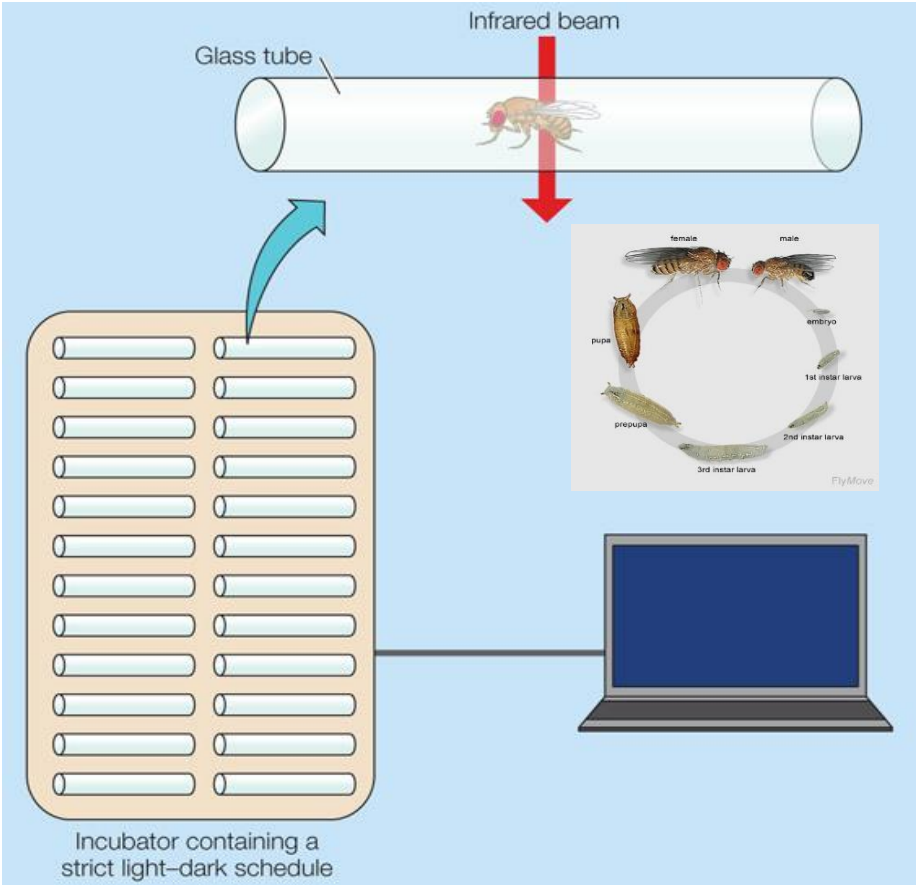
Lifespan

Data

- Minute activity counts were collected from 531 Canton S male flies over their lifespans.
- Minimum Lifespan : 32 days

  Maximum Lifespan: 74 days
- The flies were categorized into three groups based on length of lifespan, with the middle group eliminated to provide greater contrast.

Group	Short Life Group 0	Long Life Group 1
Lifespan (days)	32-51	60-74
Total	151	106

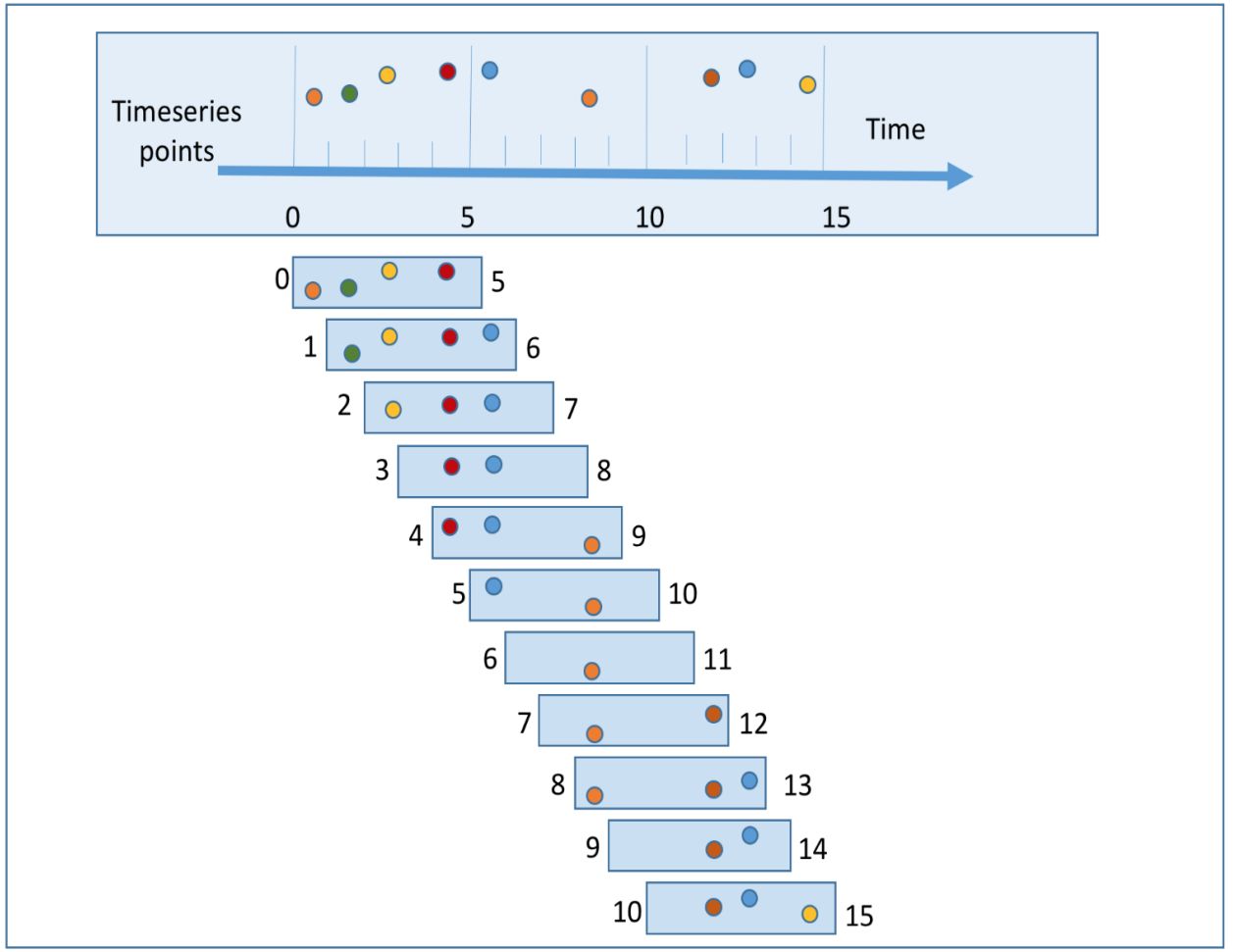


Methodology

The model for the prediction of lifespan was built in two stages

Stage1

- Model data using a zero-inflated autoregressive conditional Poisson (ACP) model.
- The zero-inflation probability was allowed to vary from hour-to-hour over the 24 hours
- A moving window of five-days was used to allow for changes in model parameters over time



Stage 1 -- Zero Inflated ACP Model Fitting

Ratnayake and Samaranayake (2017) extended an existing zero inflated ACP model to allow for zero inflation to vary cyclically.

Let X_t be a time series of count data satisfying

$$X_t|F_{t-1} \sim Poisson(\omega_t, \lambda_t)$$
$$\lambda_t = \mu + \beta X_{t-1} + \alpha \lambda_{t-1}$$
$$P(X_t = k) = \omega_t \delta_{k,0} + (1 - \omega_t) \frac{\lambda_t^k e^{-\lambda_t}}{k!},$$

where,

$$\delta_{k,0} = \begin{cases} 1 & if\ k = 0 \\ 0 & if\ k \neq 0 \end{cases},$$

with

$\omega_t = \omega_i$ when t falls into the i^{th} hour, I = 1,2,…, 24.

Stage 1 -- Results from Zero-inflated ACP Modeling

Results from the Zero-inflated ACP model fitting are illustrated using the zero inflation probabilities for two flies, one short-lived and one long-lived.

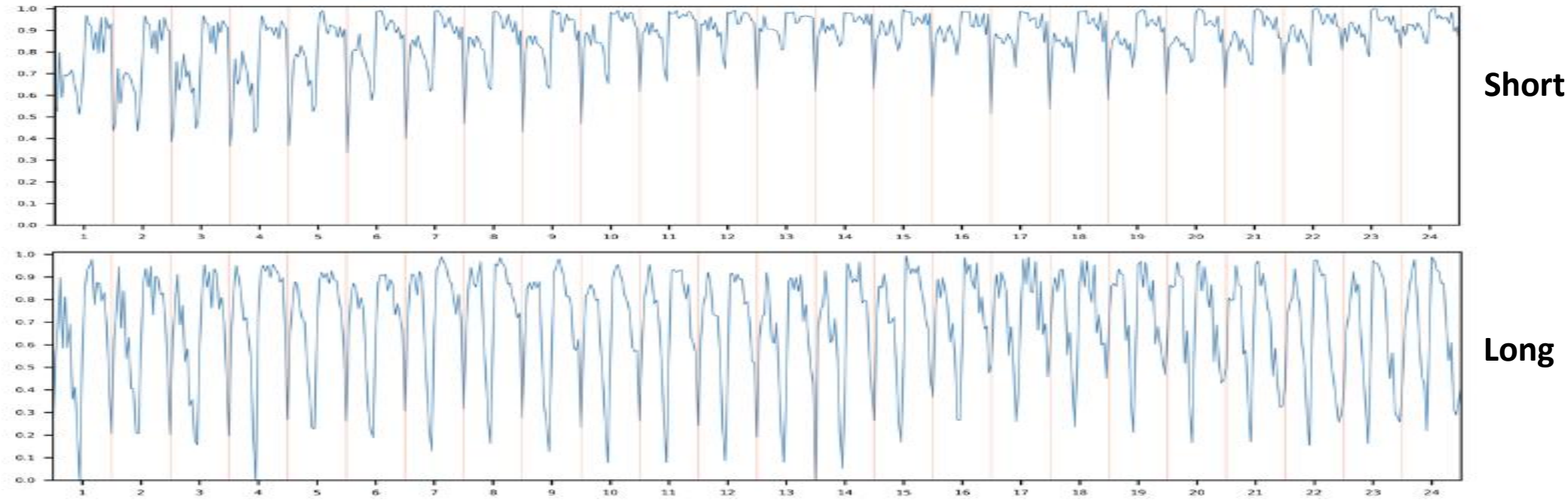


Figure 1. Estimated Zero-inflation probabilities over 24 days. The top figure is for a fly that lived for 42 days and the bottom figure is for a fly that lived 64 days.

Stage 1 -- Results from Zero-inflated ACP Modeling

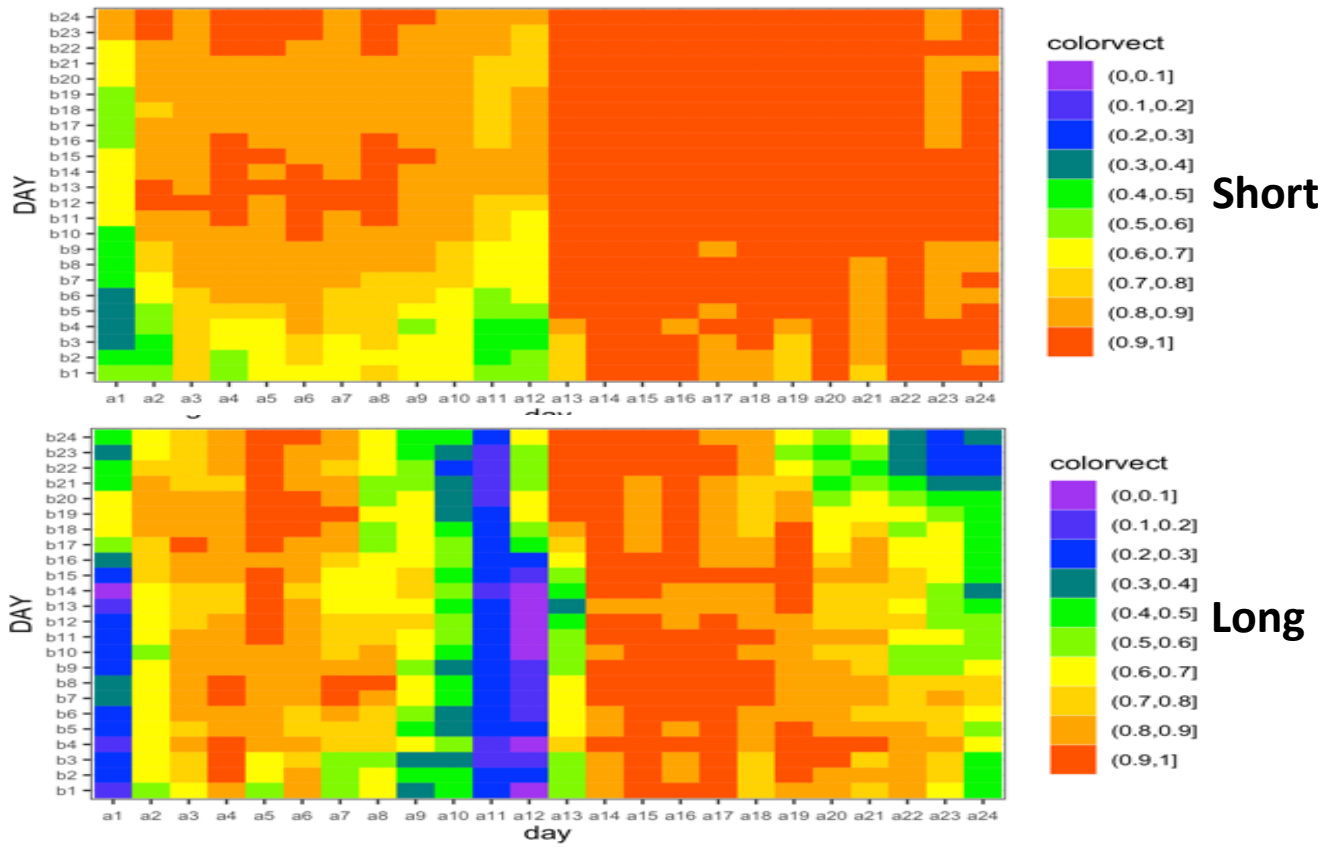


Figure 2. Heat Map of Zero-inflation Probabilities

Heat maps were generated using the 24-hour zero-inflation probabilities of each fly across the 24-day test period.

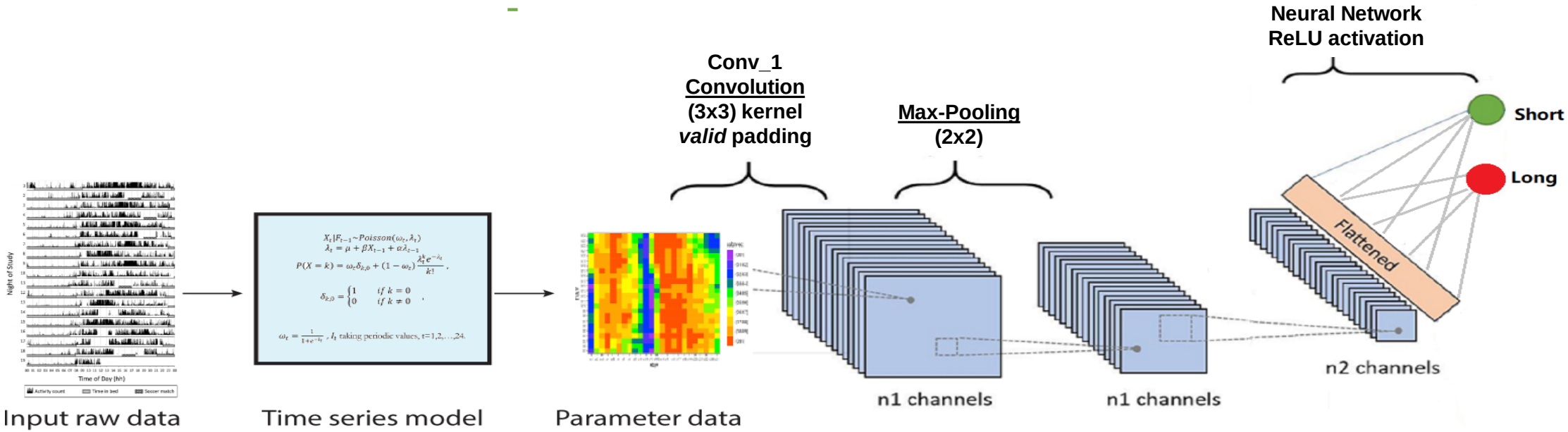
Each row of the map represents the zero-inflation probability values for one day

Each column represents these probabilities for a specific hour across the 24 days over which the data is analyzed.

Methodology

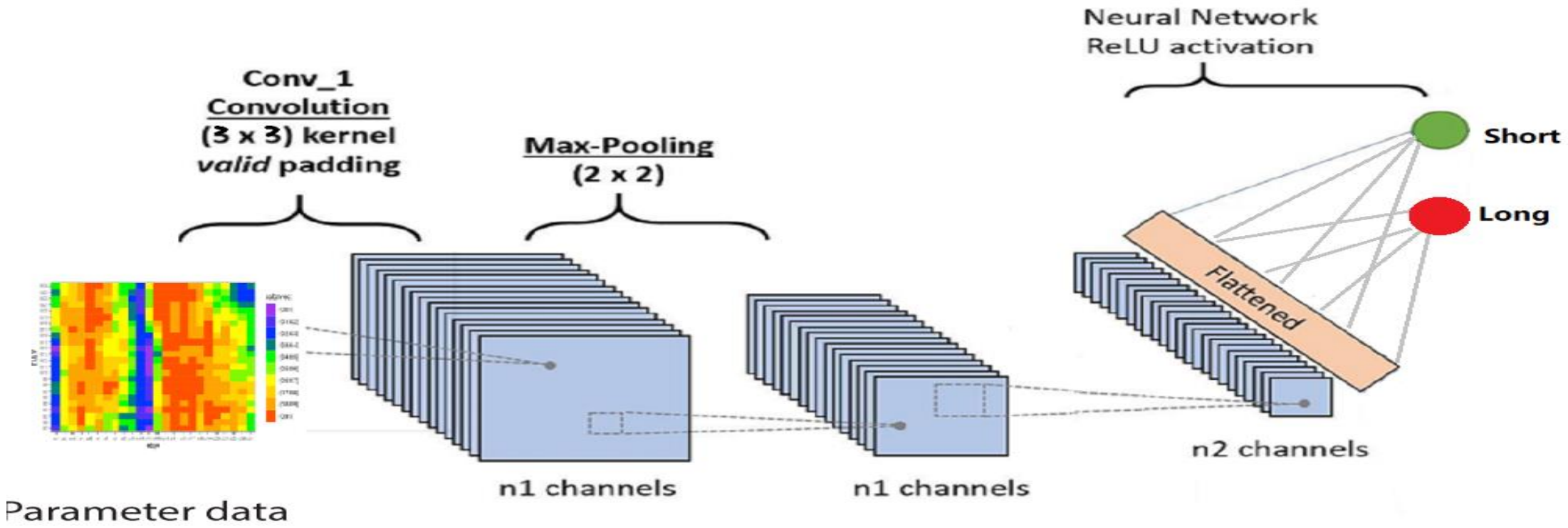
Stage 2

- The heat maps for each fly were used as input to a convolutional neural network (CNN) to build a predictive model to classify flies into the two groups.



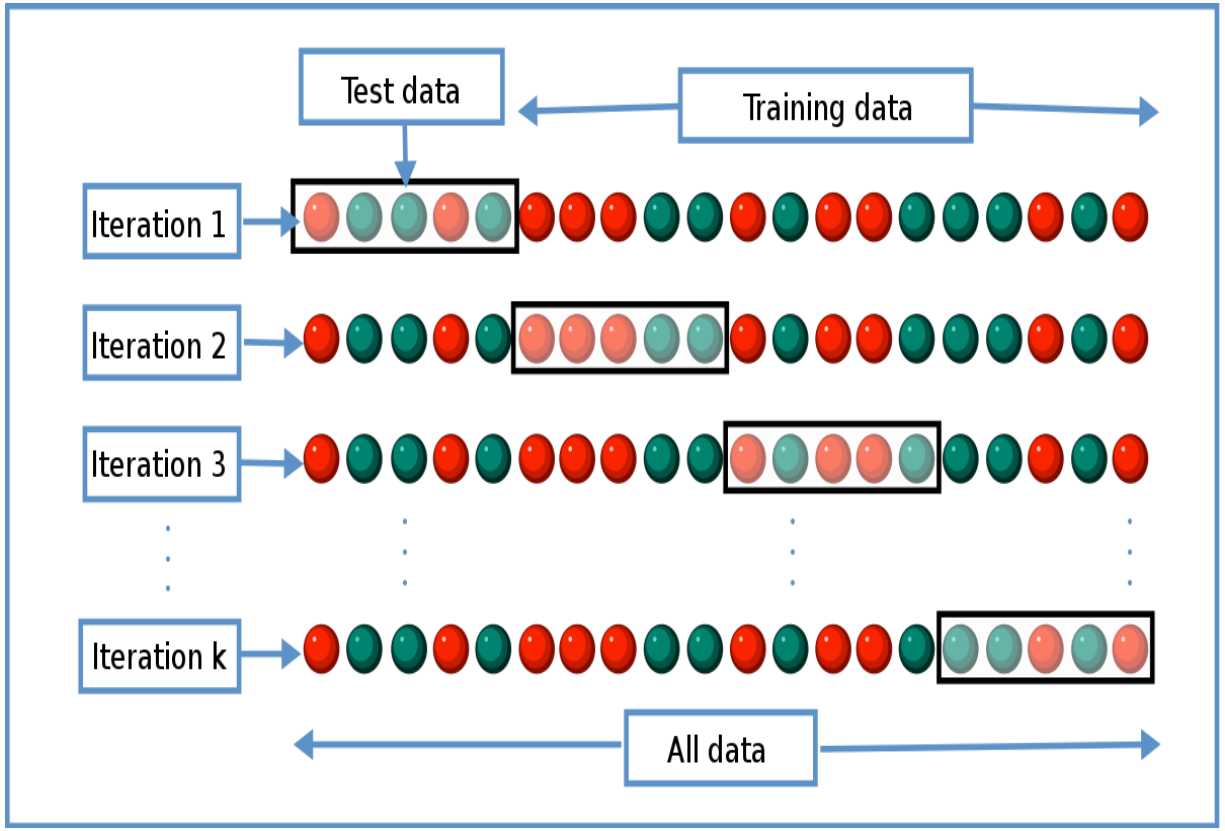
Stage 2 -- Convolutional Neural Network (CNN) Modeling

- A convolutional layer with 16 kernels of size of 3x3.
- A max pooling of size 2x2
- A fully connected neural network with RELU activation function



Stage 2 -- K-fold Cross Validation

- Four disjoint subsets of data were created using stratified random sampling from the two lifespan groups (K=4).
- Three subsets were used as training data and remaining set as the test set.
- Process was repeated until flies in each subset received a prediction.



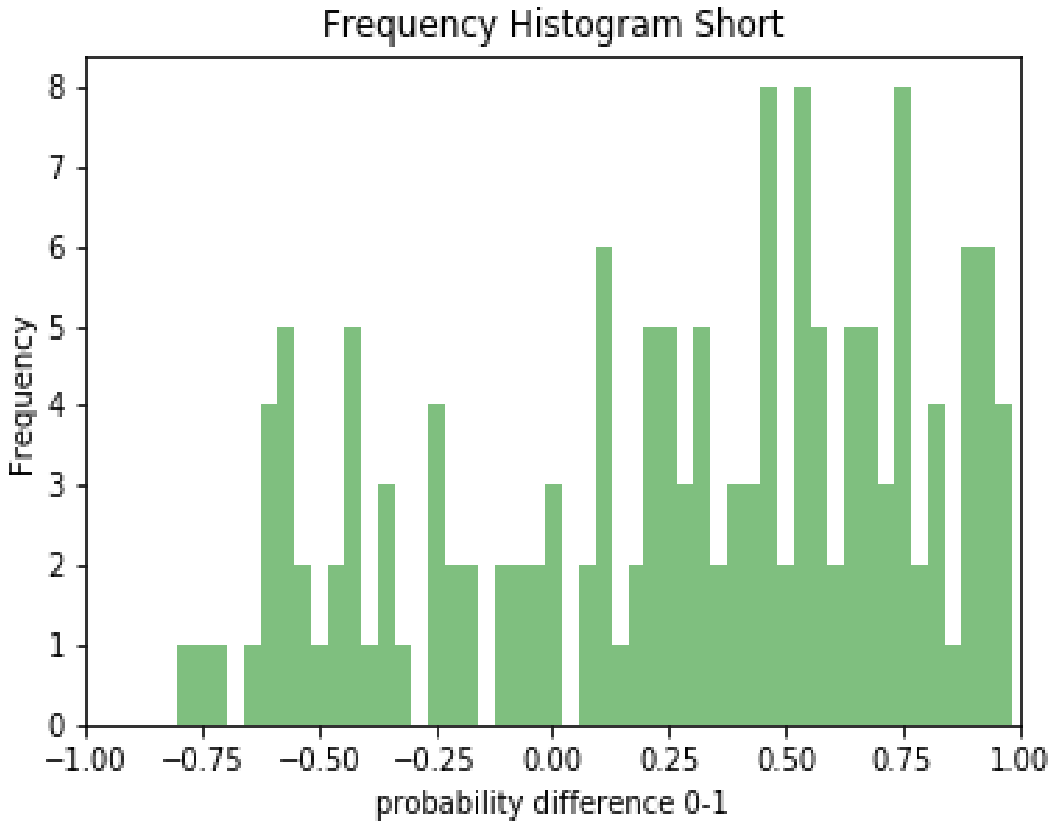
Stage 2 -- Prediction Accuracies

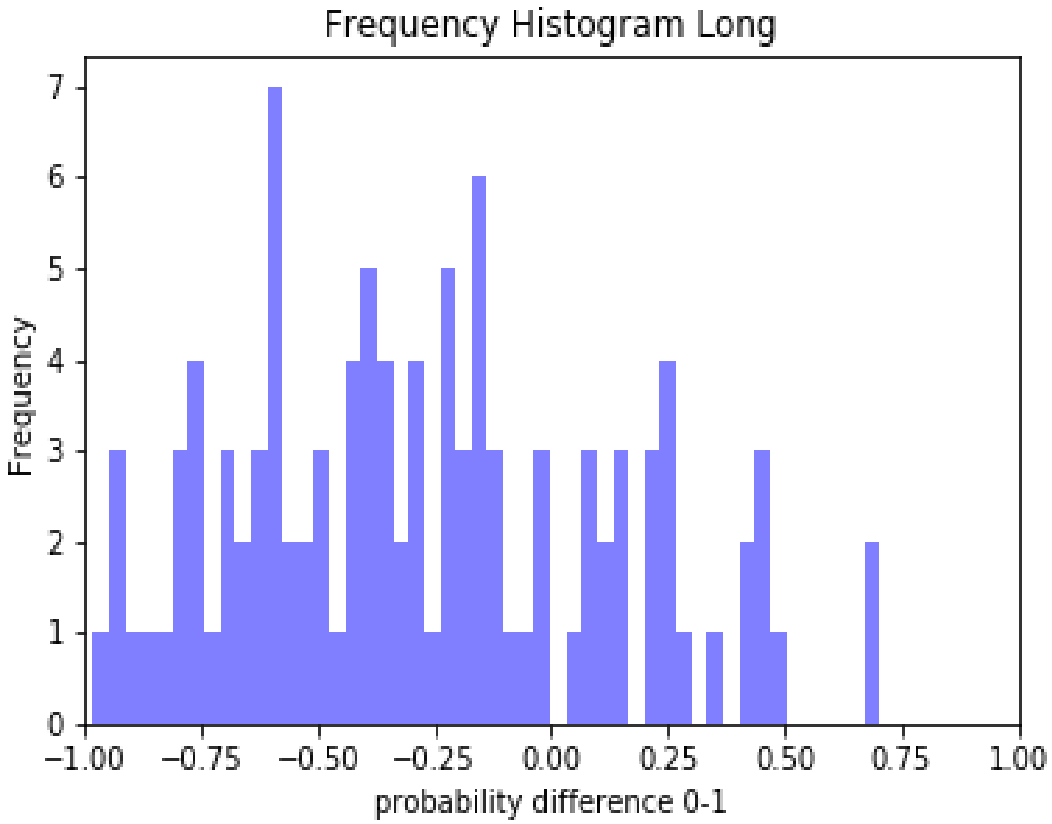
Table 1. Confusion Table Based on Test Samples

	Actual Group		Column Total
Predicted Group	0	1	
0	109	26	135 (80.74%)
1	42	80	122(65.57%)
Row Total	151 (72.18%)	106 (75.47%)	257

The prediction model yields reasonable accuracy, especially in light of the fact that sleep is only one factor that affects the lifespan of the fruit fly.

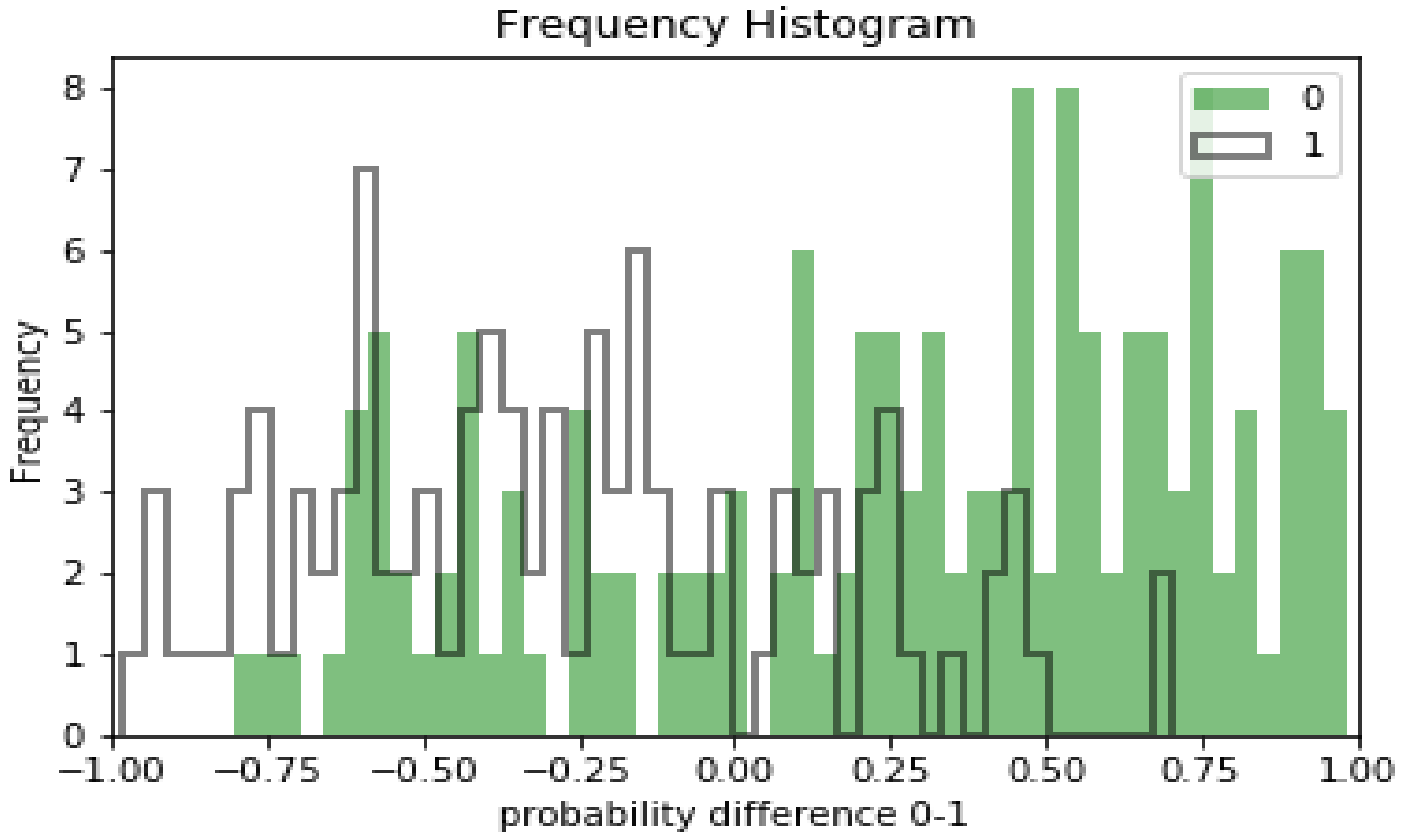
Stage 2 – Graphical View of Group Separation





Probability difference 0-1 = probability of falling into Group 0 – probability of falling into Group1

Stage 2 – Graphical View of Group Separation



Overlapping the previous two histograms provides a better view of the separation of two lifespan categories of flies

Conclusions

- A multi-stage modeling approach was employed to build a prediction model to identify fruit flies with short and long lifespans.
- Results show that the prediction model provides a reasonably accurate way (~70% accuracy) to group flies into extreme lifespan categories.
- Further improvements to the model may be possible.

Acknowledgements

We thank Josh Lisse, Elizabeth Park, and Lauren Francis for their contributions in data collection and Luyang Wang for assistance with raw data formatting. We also like to thank Austin Vandegriffe for his assistance.

This research was solely supported by the National Institute of General Medical Sciences of the National Institute of Health under award number R15GM117507.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Donlea, J.M., et al., Inducing sleep by remote control facilitates memory consolidation in Drosophila. Science, 2011. 332(6037): p. 1571-6.
- Dissel, S., et al., Sleep restores behavioral plasticity to Drosophila mutants. Curr Biol, 2015. 25(10): p. 1270-81.
- Mednick, S., K. Nakayama, and R. Stickgold, Sleep-dependent learning: a nap is as good as a night. Nat Neurosci, 2003. 6(7): p. 697-8.
- Cappuccio, F.P., et al., Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep, 2010. 33(5): p. 585-92.
- Kripke, D.F., et al., Mortality associated with sleep duration and insomnia. Arch Gen Psychiatry, 2002. 59(2): p. 131-6.
- Gallicchio, L. and B. Kalesan, Sleep duration and mortality: a systematic review and meta-analysis. J Sleep Res, 2009. 18(2): p. 148-58.
- Allada, R., C. Cirelli and A. Sehgal (2017). "Molecular Mechanisms of Sleep Homeostasis in Flies and Mammals." Cold Spring Harb Perspect Biol 9(8).
- Andretic, R. and P. J. Shaw (2005). "Essentials of sleep recordings in Drosophila: moving beyond sleep time." Methods Enzymol 393: 759-772.
- Ratnayake, I., Samaranayake VA, A GARCH Type Poisson Model for Time Series of Counts with Cyclically Varying Zero Inflation. In JSM Proceedings, Business and Economic Statistics Section, Section, Alexandria, VA, American Statistical Association. 2017; :1760-1770.
- Wallace, M. L., K. Stone, S. F. Smagula, M. H. Hall, B. Simsek, D. M. Kado, S. Redline, T. N. Vo, D. J. Buysse and G. Osteoporotic Fractures in Men Study Research (2018). "Which Sleep Health Characteristics Predict All-Cause Mortality in Older Men? An Application of Flexible Multivariable Approaches." Sleep 41(1).

Questions?