About the Presenter

Baoyuan Zhou is an undergraduate student at the University of Illinois Urbana–Champaign, majoring in Computer Science and Statistics. He conducts research in the Department of Statistics, where he applies mixed-effects models to study the relationship between climate and health outcomes. His academic interests include hierarchical modeling, database systems and distributed systems.

PL-088: An analysis of Emergency Department Visits for the 10 Health Division Regions across the US during the period 2018– 2025

Baoyuan Zhou — University of Illinois Urbana–Champaign — Undergraduate Researcher
Dr. Lelys Bravo de Guenni — University of Illinois Urbana–Champaign — Associate Professor
Tyler Hecht — University of Illinois Urbana–Champaign —Graduate Researcher
Abhi Thanvi — University of Illinois Urbana–Champaign —Graduate Researcher

MWSUG 2025
Cincinnati, OH

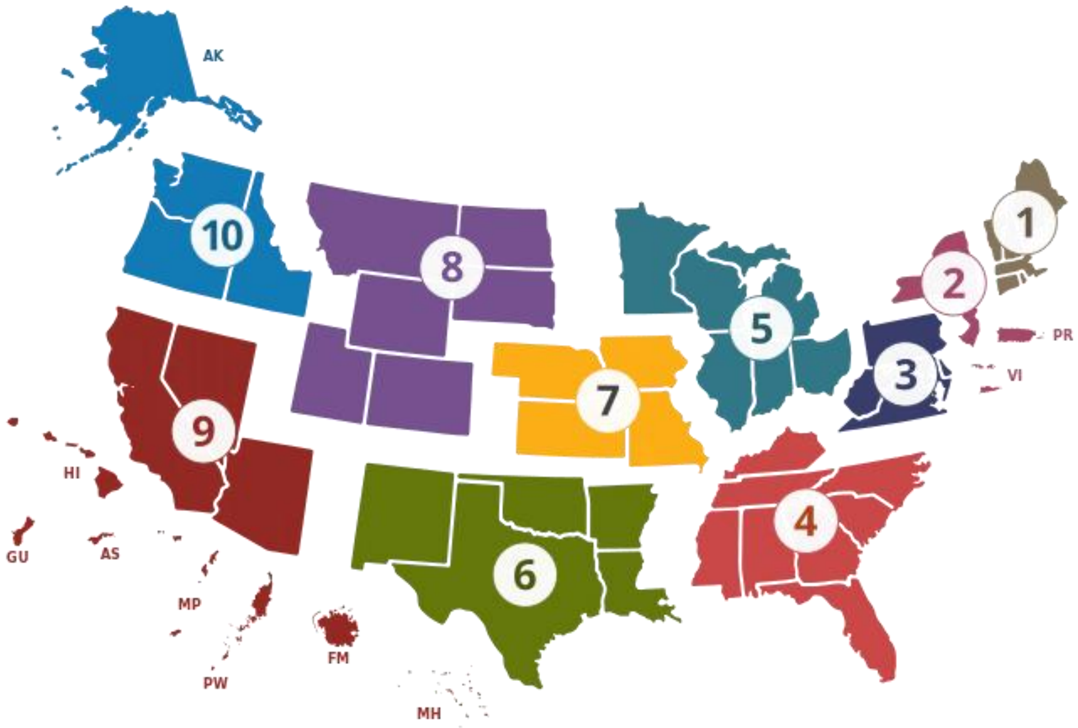
Introduction

Problem Motivation

- Extreme heat and cold events are one of the most significant causes of climate-related deaths worldwide
- Five million deaths were attributed to extreme heat and cold globally between 2000-2019 (Chen et al., 2024), and projections indicate that heat-related deaths will increase
- We aim to understand the impacts of extreme heat on healthand health services demand to provide a better estimation offuture climate-related illness burden

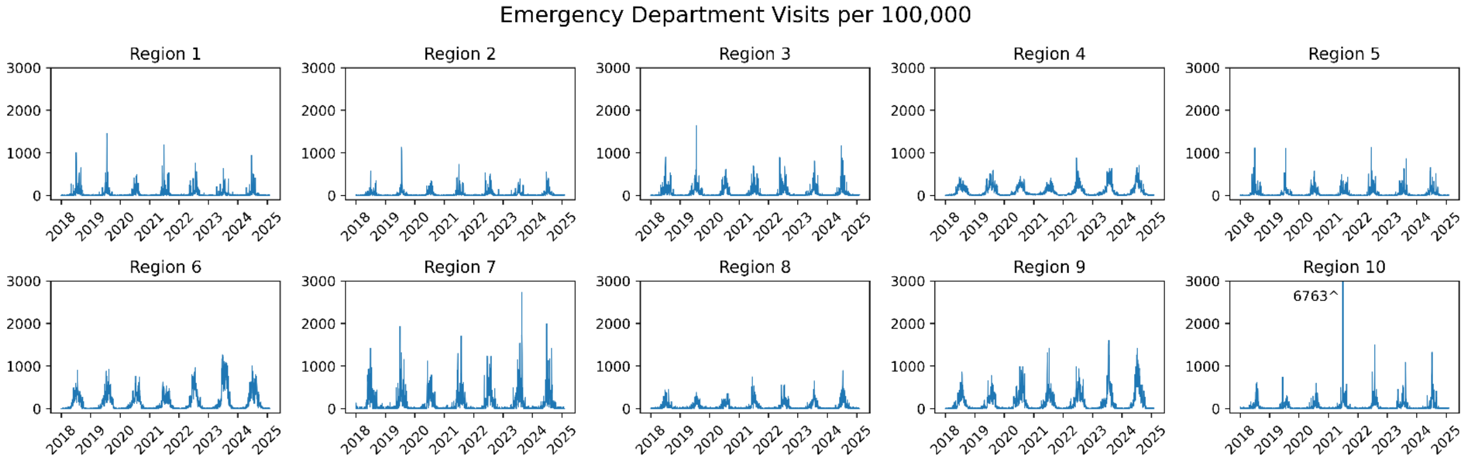
Data Granularity

- All data are aggregated into the 10 health-department region(HHS Regions) set by Office of Intergovernmental and External Affairs. Data are recorded daily and ranges from 2018 to 2025.



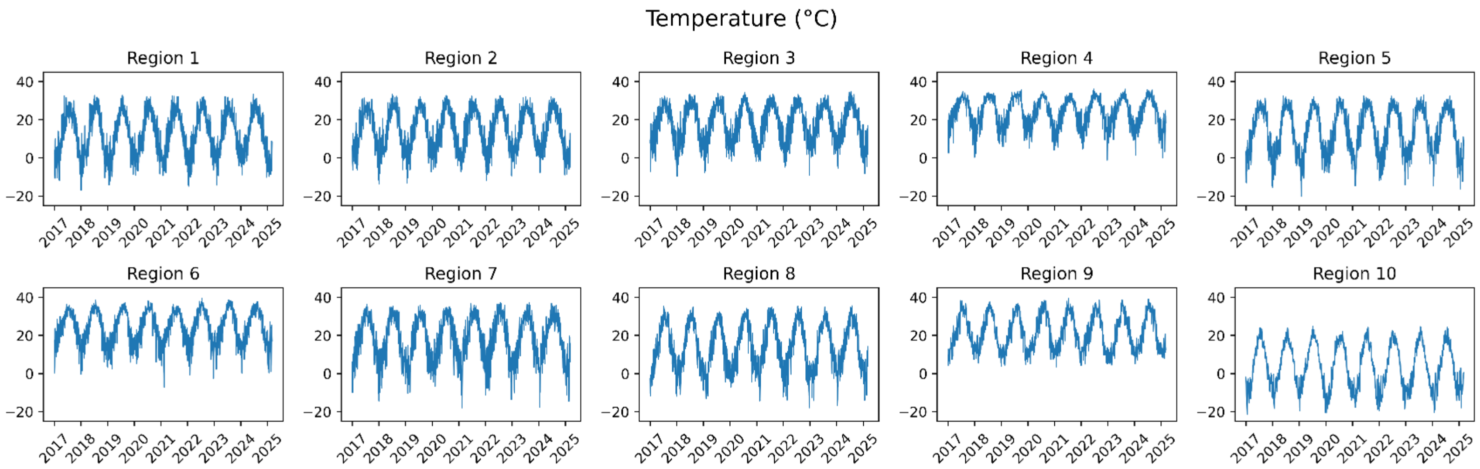
Emergency Department Visits (EDV)

- Source: Center of Disease Control and Prevention (CDC)
- Daily Heat-related illness (HRI) visits rate per 100,000 population
- Aggregated across 10 Health Department Regions (HDRs)



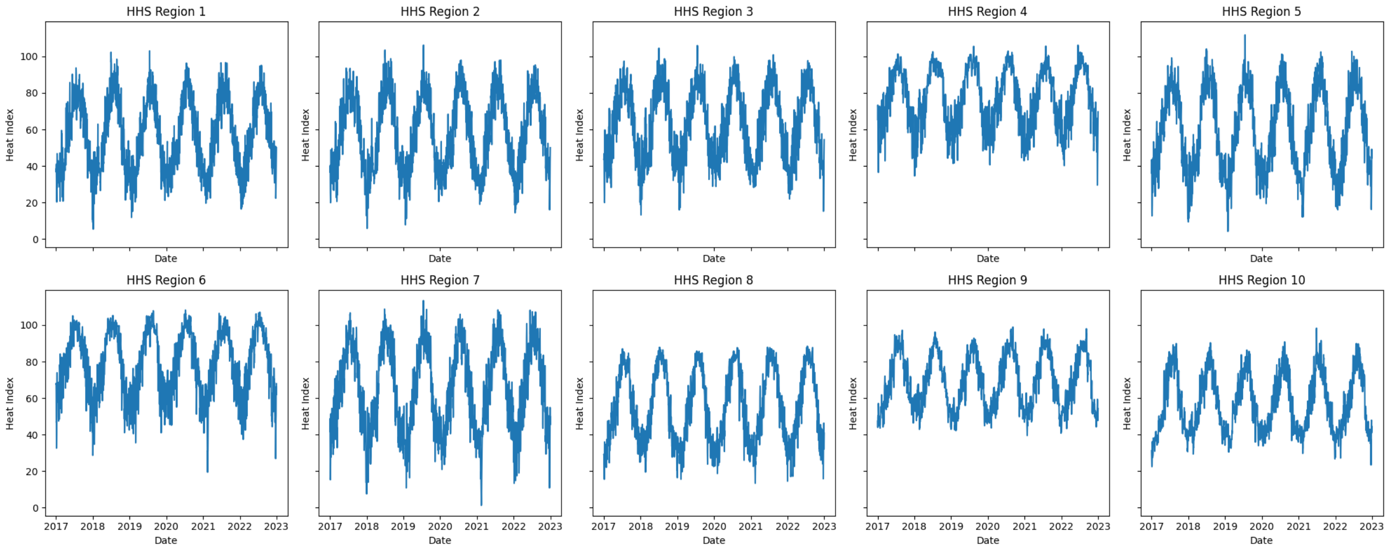
Temperature

- Source: National Oceanic and Atmospheric Administration(NOAA)
- Originally daily data gridded to 0.5x0.5 degrees lat/lon
- Aggregated to regional data with equal weighting per grid square
- Weighted proportional to area when crossing state borders



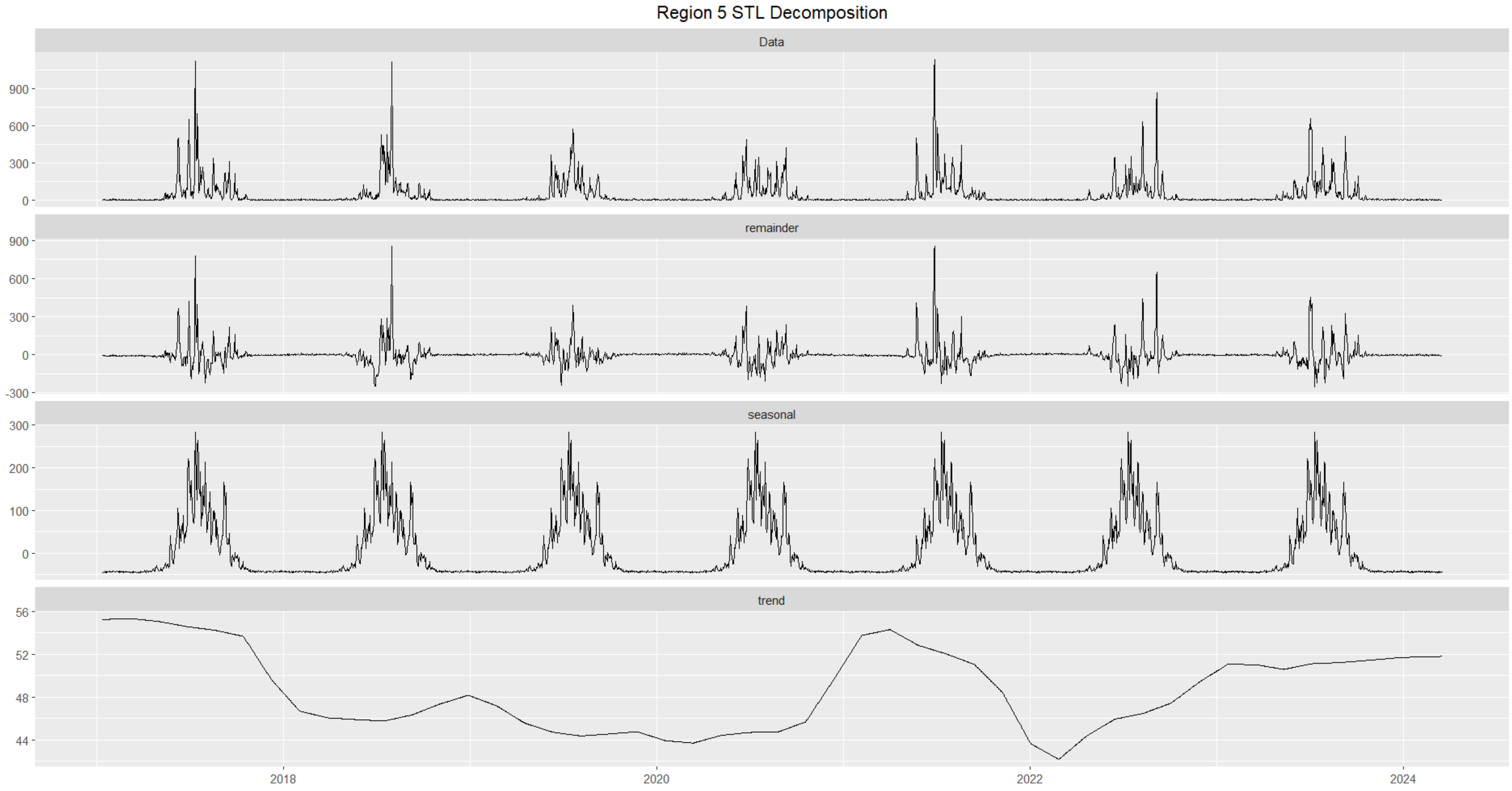
Heat Index

- Source: Center of Disease Control and Prevention (CDC)
- The heat index incorporates both the temperature and therelative humidity
- The raw data are collected at a county level and is aggregated to regional level with population weighting



Analysis

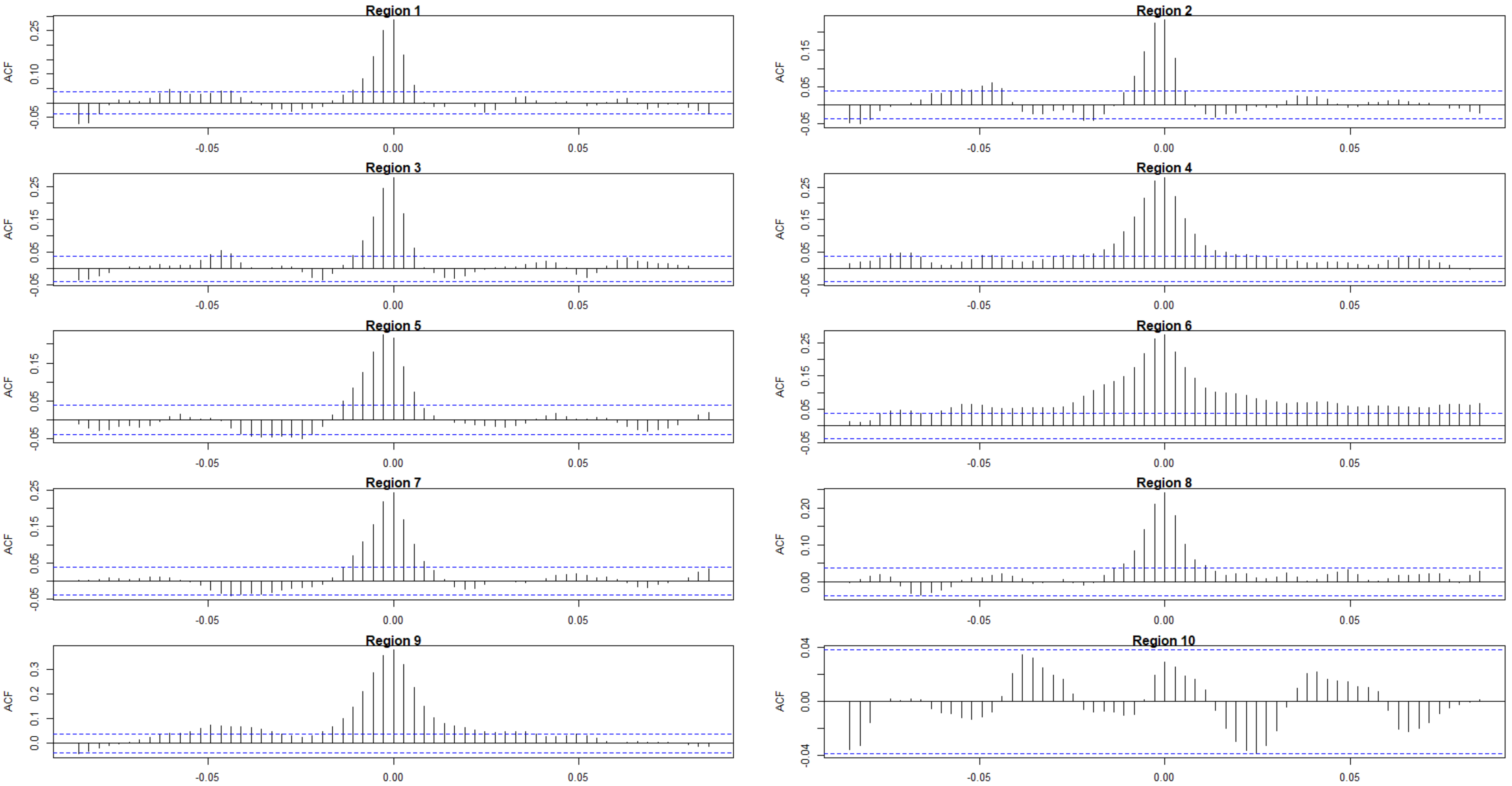
Seasonal Analysis and Lag analysis



#MWSUG2025 #PL088

MWSUG 2025
Cincinnati, OH

Cross-correlation between EDV and temperature



Mixed Effects Model

- The following mixed effects models were considered:

1. $\ln(E[EDV_{j,t}]) = \beta_0 + \beta_s \sin(\varphi t) + \beta_c \cos(\varphi t) + b_{0,j}$
2. $\ln(E[EDV_{j,t}]) = \beta_0 + \beta_s \sin(\varphi t) + \beta_c \cos(\varphi t) + b_{0,j} + b_{s,j} \sin(\varphi t) + b_{c,j} \cos(\varphi t)$
3. $\ln(E[EDV_{j,t}]) = \beta_0 + \beta_s \sin(\varphi t) + \beta_c \cos(\varphi t) + \beta_x X_{j,t} + b_{0,j}$
4. $\ln(E[EDV_{j,t}]) = \beta_0 + \beta_s \sin(\varphi t) + \beta_c \cos(\varphi t) + \beta_x X_{j,t} + b_{0,j} + b_{x,j} X_{j,t}$
5. $\ln(E[EDV_{j,t}]) =$
   $\beta_0 + \beta_s \sin(\varphi t) + \beta_c \cos(\varphi t) + \beta_{x,j} X_{j,t} + b_{0,j} + b_{s,j} \sin(\varphi t) + b_{c,j} \cos(\varphi t)$
6. $\ln(E[EDV_{j,t}]) =$
   $\beta_0 + \beta_s \sin(\varphi t) + \beta_c \cos(\varphi t) + \beta_{x,j} X_{j,t} + b_{0,j} + b_{s,j} \sin(\varphi t) + b_{c,j} \cos(\varphi t) + b_{x,j} X_{j,t}$

where $\varphi = 1/365.25$ and X is either temperature or heat index.

The poisson family was used with a log link.

Prediction Accuracy (Temperature)

- Mean squared error (MSE) was used to evaluate the models using multiple train-test splits.
- The four models that included temperature performed much better, with similar MSEs.

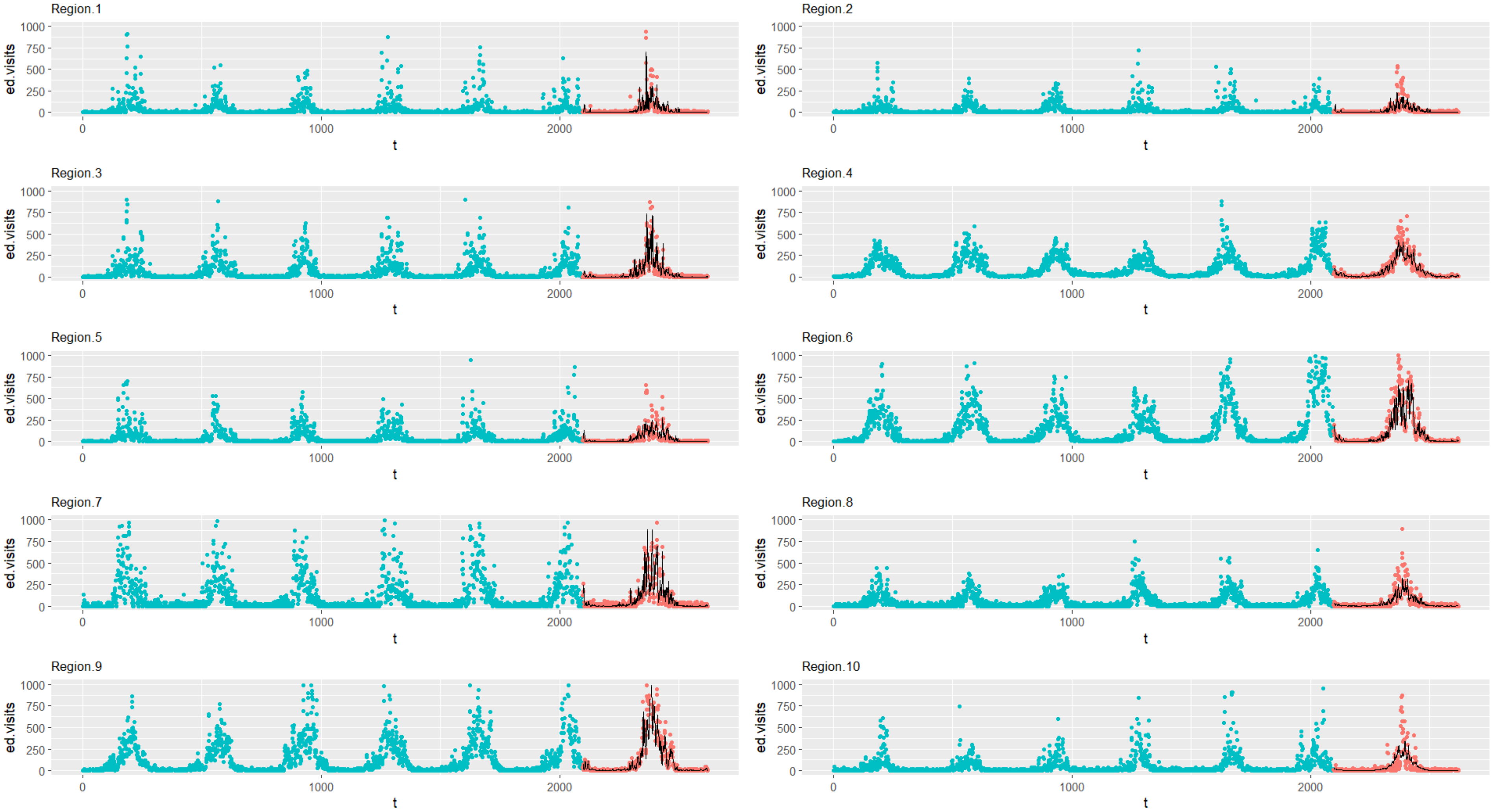
Best Fitting Model Interpretation (Temperature)

- The best-performing model is:

$$\ln(E[EDV_{j,t}]) = \beta_0 + \beta_s \sin(\varphi t) + \beta_c \cos(\varphi t) + \beta_{x,j} \mathrm{T}_{j,t} + \\ b_{0,j} + b_{s,j} \sin(\varphi t) + b_{c,j} \cos(\varphi t) + b_{x,j} \mathrm{T}_{j,t}$$

- Substantial variation across regions in baseline ED visit rates(random intercept variance = 1.59)
- Modest variation across regions in the effect of temperature(random slope variance = 0.00204)
- Strong negative correlation (−0.96) between random interceptsand temperature slopes, meaning that regions with higherbaseline ED visits tend to have a smaller effect of temperatureon ED visits.

Predictions (Temperature)

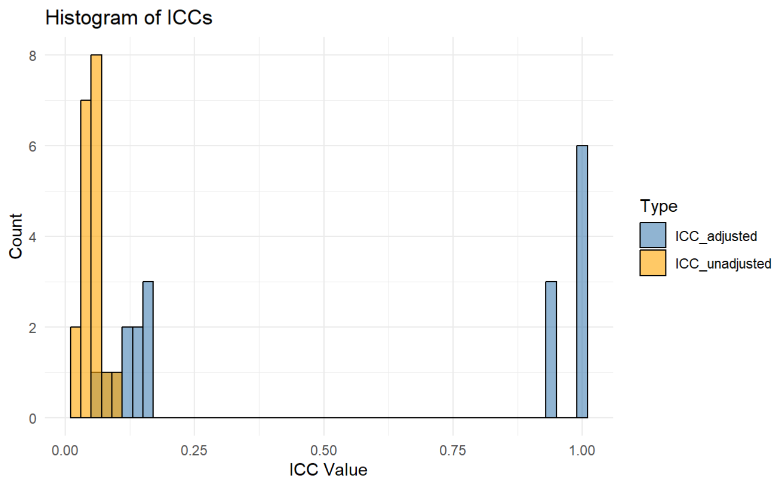


#MWSUG2025 #PL088

MWSUG 2025
Cincinnati, OH

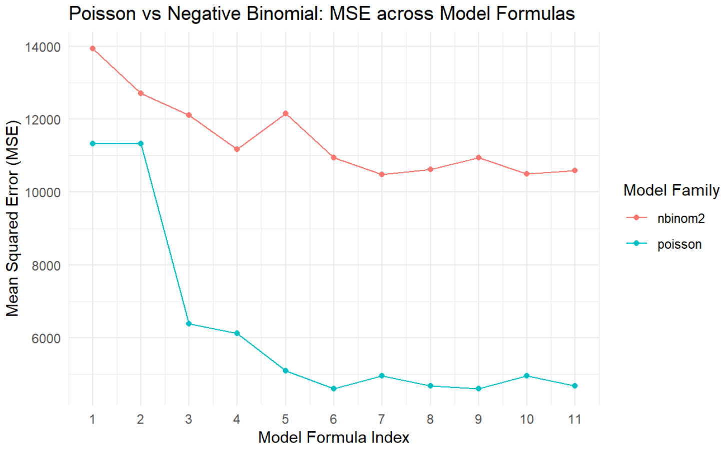
Intraclass Correlation Coefficient (Heat Index)

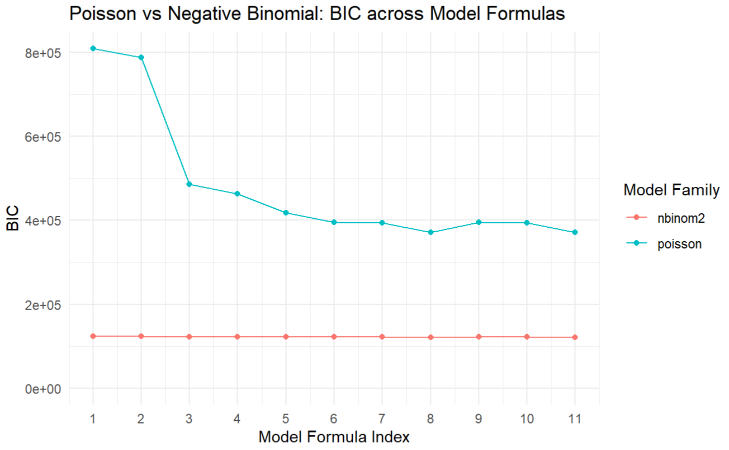
- Most unadjusted ICCs are below 0.05, meaning that the varianceexplained by regions is insignificant based only on random effects.
- Adjusted ICCs are close to 1, meaning that almost all theunexplained variance is explained by region after controlling for thecovariates.



Fitting Performance (Heat Index)

- The Poisson models fit consistently better in terms of MSE but worsein AIC/BIC. This might be an reflection of the bias-variance trade-off.The Poisson models perform better in terms of MSE due to higherbias but lower variance, which favors generalization





Best Fitting Model Interpretation (Heat Index)

- The best-performing model is:

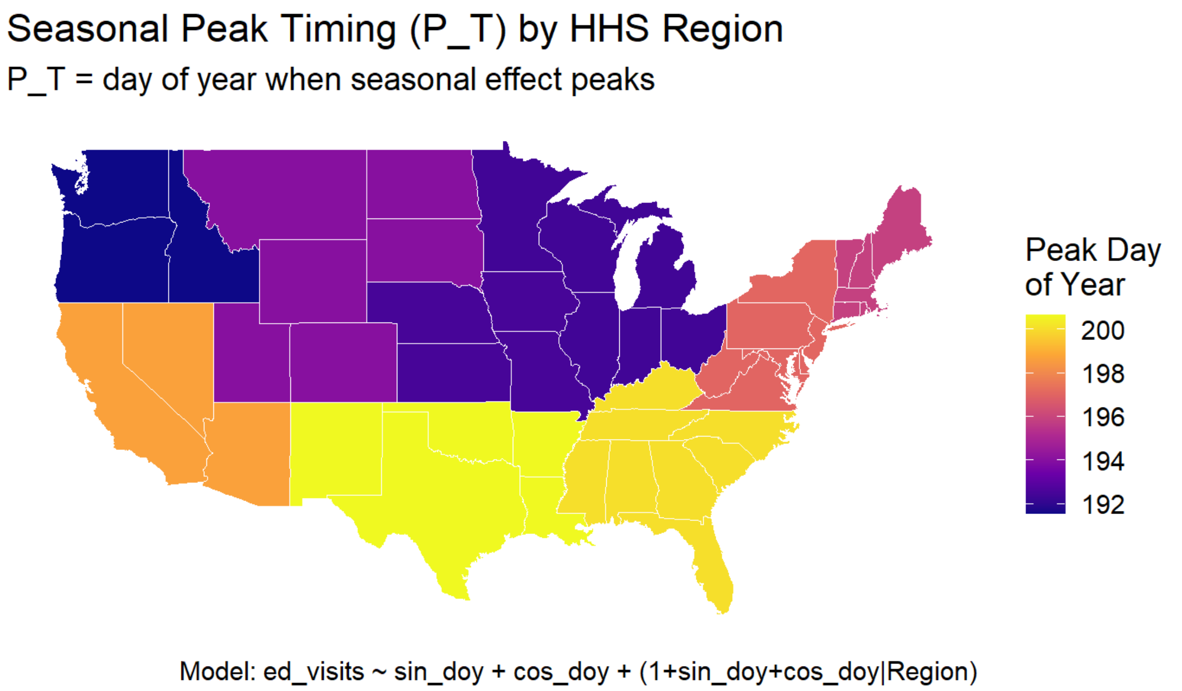
$$\ln(E[EDV_{j,t}]) = \beta_0 + \beta_s \sin(\varphi t) + \beta_c \cos(\varphi t) + \beta_x \mathtt{HI}_{j,t} + b_{0,j} + b_{x,j}\mathtt{HI}_{j,t}$$

- Substantial variation across regions in baseline ED visit rates(random intercept variance = 1.49)
- Modest variation across regions in the effect of heat index(random slope variance = 0.00029)
- Strong negative correlation (−0.94) between random intercepts and heat index slopes, meaning that regions with higher baseline ED visits tend to have a smaller effect of heat index on ED visits.
- Allowing seasonal components to vary randomly across regions slightly reduced AIC/BIC, but not enough to justify the added complexity

Conclusion

Applications

- Beyond simply predicting counts, understanding the overall trend provides more valuable public health insights
- Peak timing is measure on the time took to peak in EDV in days of a year and is derived from model coefficients.



Discussion

- Both temperature and heat index are valuable for predictingfuture emergency department visits
- The Poisson family performed better than the negative binomial family when evaluated using MSE
- Models including these variables in some way performed considerably better than models only incorporating seasonality
- Models incorporating temperature or heat index in some way performed about the same as each other
- The lack of available EDV data limited the scope of our analysis, and future research on data spanning more years may prove more conclusive

References i

- K. Chen, E. de Schrijver, S. Sivaraj, F. Sera, N. Scovronick, L. Jiang,D. Royé, E. Lavigne, J. Kyselý, A. Urban, A. Schneider, V. Huber,J. Madureira, M. Mistry, I. Cvijanovic, A. Gasparrini,A. Vicedo-Cabrera, and Y. Guo.Impact of population aging on future temperature-relatedmortality at different global warming levels.Nature Communications, 15, 02 2024.
- S. Dewi, R. Kasim, N. Sutarsa, and S. Hall Dykgraaf.A scoping review of the impact of extreme weather events onhealth outcomes and healthcare utilization in rural and remoteareas.BMC Health Services Research, 24, 11 2024

References ii

- R. Kaiser, A. Le Tertre, J. Schwartz, C. Gotway, W. Daley, andC. Rubin.The effect of the 1995 heat wave in chicago on all-cause andcause-specific mortality.American journal of public health, 97 Suppl 1:S158–62, 05 2007.
- A. Liss, R. Wu, K. Kwan, H. Chui, and E. Naumova.Heat-related hospitalizations in older adults: An amplifiedeffect of the first seasonal heatwave open.Scientific Reports, 7, 01 2017.
- E. N. Naumova, R. B. Simpson, B. Zhou, and M. A. Hartwick.Global seasonal and pandemic patterns in influenza: Anapplication of longitudinal study designs.International Statistical Review, 90(S1):S82–S95

Thank You