Extreme Heat Days and Poor Mental Health Outcomes in U.S. in 2024: A Geospatial Analysis

Sanne Glastra
Columbia University, Mailman School of Public Health

EHSCP 9380: Advanced GIS and Spatial Analysis

Jeremy Porter, PhD

May 8, 2025

TABLE OF CONTENTS

INTRODUCTION

Anthropogenic climate change is driving global temperature increases, with averages rising by up to 0.51°F per decade since 1979 (*Climate Change Indicators: U.S. and Global Temperature* | *US EPA, 2025*). In the United States, warming is occurring at a rate 60% faster than the global average, making the country and its citizens particularly vulnerable to extreme heat events (Extreme Heat | US EPA, 2025). Not surprisingly, therefore, it is projected that extremely hot days will become more frequent and heat ways will intensify in the next few years (*Extreme Heat* | *US EPA*, 2025). Consequently, projections show that extremely hot days will become more frequent and heat waves will intensify in the next few years (Extreme Heat | US EPA, 2025).

At the same time, mental illness has become a major public health concern in the United States, affecting an estimated 23.1% of adults – roughly 59.2 million people (Mental Illness, n.d.). Given the growing concern around both extreme health and ill mental health, and the plausible link between the two, researchers have explored whether a relationship exists between extreme heat exposure and mental health outcomes. For example, a systematic review of over 15 studies by Thompson et al. (2018) found a strong relationship extreme heat and suicide risk in the U.S. Similarly, higher temperatures have also been associated with increased rates of hospital admissions for mental health conditions. Still, significant knowledge gaps remain in our understanding of how extreme heat affects mental health.

In previous work, I investigated the relationship between extreme heat and mental health outcomes using conventional statistical methods that did not account for spatial dependence. The analysis relied on a CDC-defined “Extreme Heat Day” exposure variable, measured as the number of days above the 95% historical average temperature; therefore, this predictor framed extreme heat as a relative, not absolute, measure. Findings from the study indicated a significant inverse relationship: each additional extreme heat day corresponded to a 6.3% decreased odds of poor mental health.

The present study builds upon that earlier analysis in two key ways:

1. Employs a spatial regression model to account for spatial dependence.
2. Utilizes an absolute temperature metric as the exposure variable, offering an alternative proxy for extreme heat.

In doing so, this study aims to generate a more nuanced understanding of the spatial dynamics linking extreme heat to mental health outcomes in the U.S.

METHODS

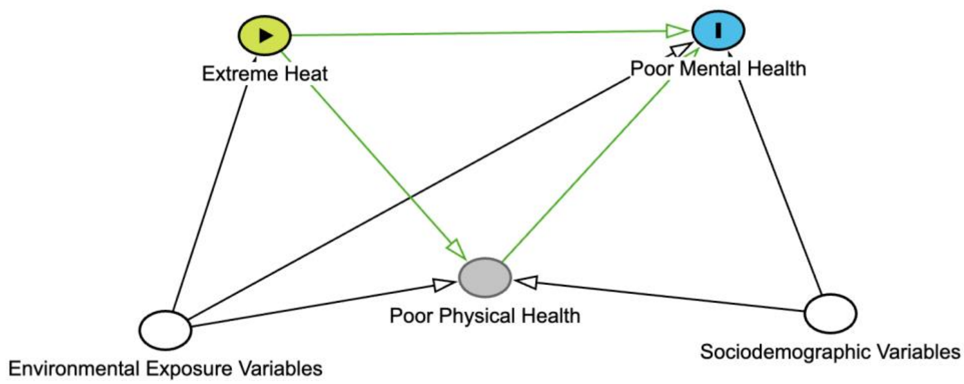
Data for this project was collected for the year 2024 from multiple sources. Data for extreme heat days as well as general control variable data was collected from the CDC HeatTracker, a tool developed by the Centers for Disease Control and Prevention to monitor heat-related risks and trends across the U.S. (Centers for Disease Control and Prevention, n.d.). The data are annual (year 2024) and provided at the Zip Code Tabulation Area (ZCTA) level.

Maximum temperature data was collected from the PRISM Climate Group, which provides high resolution gridded climate data for the U.S. using a combination of monitoring networks. Additionally, the percentage of white population per zip code (proxy for racial and ethnic variation) was obtained from the U.S. Census Bureau's American Community Survey (ACS) 5-Year Estimates, using the variable B02001_002 for the white population and B02001_001 for the total population (U.S. Census Bureau, 2024). Maximum temperature raster data was then converted to zip code level data using zonal statistics with this U.S. Census Bureau shapefile.

The outcome variable for this study is mental health status, a continuous outcome variable for mental health status was used: crude prevalence of persons (age 18+) reporting “not good” mental health for at least 14 days at the zip code level. There are two main predictor for this study. The first is a relative measure of extreme heat: number of extreme heat days, a continuous variable defined as the number of days above the 95% historical average temperature per ZCTA. The second is an absolute measure of extreme heat: maximum temperature.

Covariates were selected using a Directed Acrylic Graph (DAG) to account for confounding and causal relationships (**Figure 1**); green arrows in the figure indicate the hypothesized causal pathway. Poor physical health (gray) was identified as a mediator and therefore was dropped as a covariate from the model. Covariates kept in the model included sociodemographic score, which is a continuous summary indicator averaging the percentile ranks of various specific sociodemographic indicators, with higher scores indicating worse sociodemographic conditions for each ZCTA. Other continuous covariates in the model include the % of ZCTA with impervious surface coverage, % of ZCTA with tree canopy coverage, % of ZCTA with mobile homes, % of ZCTA with households without vehicle access, annual mean days above the PM2.5 regulatory standard (3-year average per ZCTA), and % white population in each ZCTA.

Figure 1: Directed Acrylic Graph (DAG) of Predictor, Outcome, and Covariates



To assess global relationship between extreme heat and poor mental health, global linear regression models were applied using either 1.) maximum temperature (absolute predictor) and 2.) extreme heat days (relative predictor). Next, to understand whether we needed to account for spatial dependence, a Global Moran's I was calculated and LISA Cluster map was created to assess spatial clustering of poor mental health outcomes. When spatial dependence was confirmed, we ran two spatial lag models; the first one with maximum temperature as the main predictor, and the second one with extreme heat days as a predictor. All analyses were conducted in Python.

RESULTS

Figure 2 shows highest prevalence of poor mental health in the southeastern U.S. as well as in areas that overlap almost perfectly with Native American reservations. The indicated is supported by **Figure 3**, which shows that areas with low percentage of white populations coincide with poorer mental health outcomes.

Figure 4 and **Figure 5** illustrate the spatial distribution of extreme heat variables. Absolute measures of heat, measured through maximum temperature, has the highest values in the southern portion of the U.S. Relative measures of heat, as measured through extreme heat days, shows the highest amount of extreme heat days in the West, with low amount of extreme heat days in the Southeast.

Figure 2: Spatial Distribution of Poor Mental Health Prevalence in U.S. in 2024 (%)

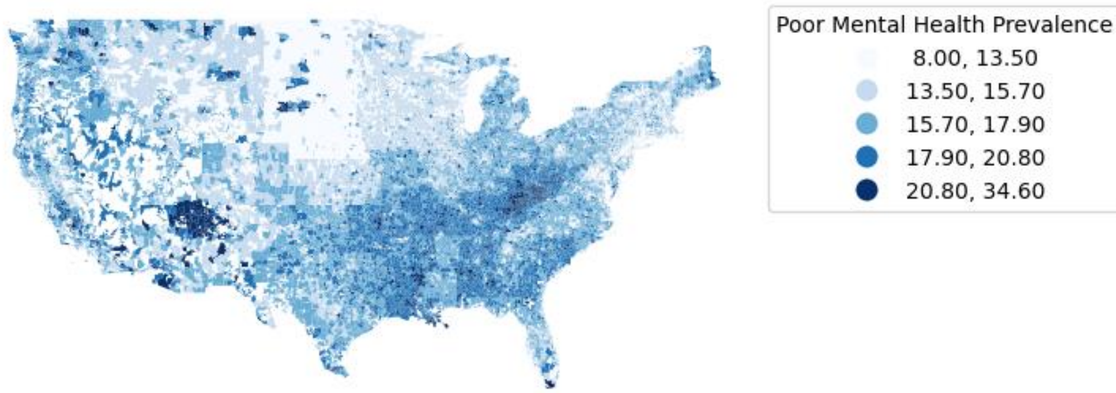


Figure 3: Spatial Distribution of Percent White Population in U.S. in 2024

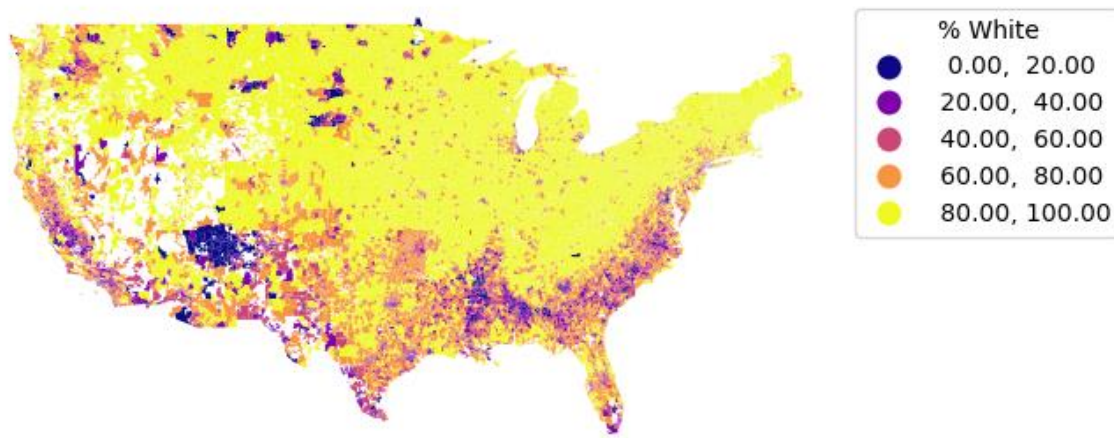


Figure 4: Spatial Distribution of Maximum Temperature in U.S. in 2024 (°C)

Figure 5: Spatial Distribution of Extreme Heat Days in U.S. in 2024

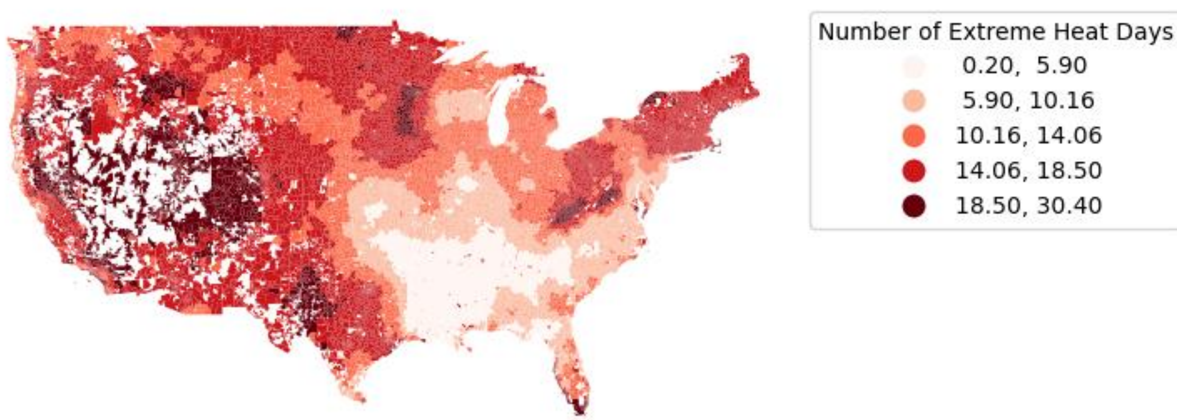


Figure 6 illustrates a Global Moran's I value of 0.71 (p = 0.001), indicating strong positive spatial autocorrelation of poor mental health outcomes in the United States. Examining the local clustering pattens in **Figure 7** reveals an abundance of high-high clustering in the Southeast, Navajo Nation, and some other sporadic areas in the western portion of the U.S. (some of which overlap with other Native American reservations). There is a significant amount of low-low clustering also in the Midwest as well as parts of the Northeast. It is important to note that within these low-low clustered areas, there are pockets of high-high

clustering that directly overlap with Native American reservations that have low prevalence of white populations.

Figure 6: Moran's I Scatterplot: Spatial Autocorrelation of Poor Mental Health

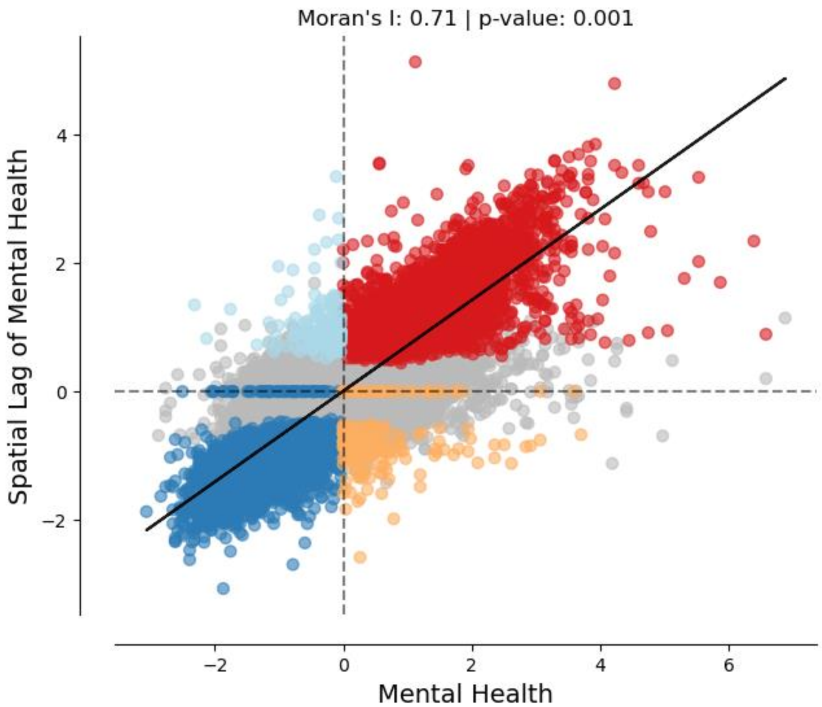


Figure 7: LISA Cluster Map: Spatial Autocorrelation of Poor Mental Health

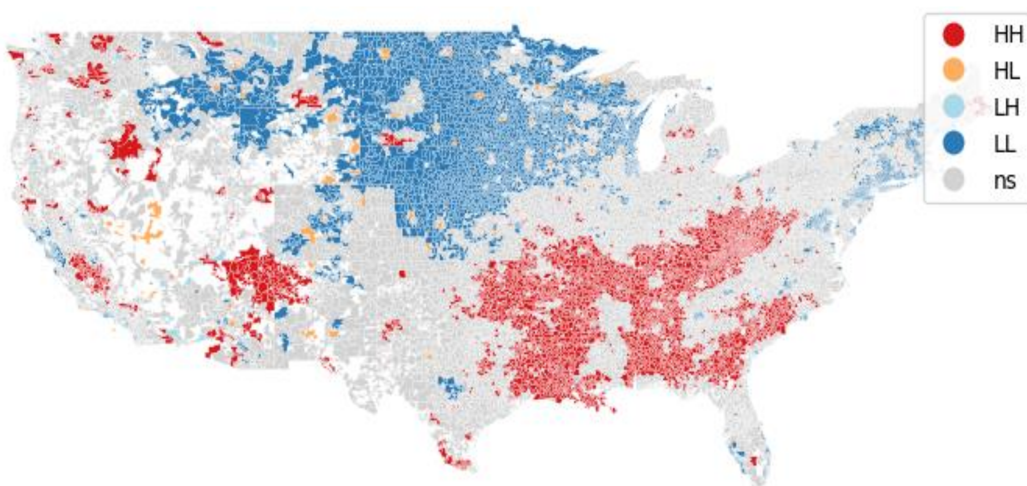


Table 1 and **Table 2** show results from the global linear regression model and spatial lag model using maximum temperature as an extreme heat proxy. The global linear regression model shows a statistically significant relationship between all predictors and poor mental health outcomes. Of all predictors, maximum temperature has the largest effect size, with every 1 degree increase in maximum temperature coinciding with a 0.16% increase in poor mental health prevalence, controlling for all other predictors. Even when accounting for spatial dependence in the spatial lag model (in addition to all other predictors), maximum temperature still has a significant positive relationship with poor mental health outcomes with an effect size of 0.11.

Table 1: Global Linear Regression Model with Max Temperature as Main Predictor

Variable	**Coefficient**	**P-Value**
Constant	14.48	0.00
Maximum Temperature (°C)	0.16	0.00
Socioeconomic Score	0.00	0.00
Impervious Surface Coverage (%)	-0.01	0.00
Tree Canopy Coverage (%)	0.03	0.00
Mobile Homes (%)	0.00	0.00
Households Without Vehicle (%)	0.00	0.00
Days Above PM2.5 Regulatory Threshold	0.01	0.02
White Population (%)	-0.03	0.00

Table 2: Spatial Lag Model with Max Temperature as Main Predictor

Variable	**Coefficient**	**P-Value**
Constant	11.20	0.00
Maximum Temperature (°C)	0.11	0.00
Socioeconomic Score	0.00	0.01
Impervious Surface Coverage (%)	0.00	0.00
Tree Canopy Coverage (%)	0.02	0.00

Mobile Homes (%)	0.00	0.00
Households Without Vehicle (%)	0.00	0.00
Days Above PM2.5 Regulatory Threshold	0.00	0.31
White Population (%)	-0.03	0.00
Spatial Lag of P_MNTHL	0.27	0.00

Table 3 and **Table 4** also highlight global linear regression model and spatial lag model results, this time using extreme heat days as an extreme heat proxy. All predictors in the global model have a statistically significant relationship with poor mental health outcomes, with the exception of extreme heat days and days above PM2.5 regulatory threshold. All predictors have a very low effect size, however, with coefficients close to 0. When account for spatial dependence on the spatial lag model, we see that percent tree canopy, days above PM2.5 regulatory threshold, percent white population, and the spatial lag term have a significant relationship with poor mental health outcomes. However, effect size for all indicated predictors are very low at around zero, with the exception of the spatial lag term with the highest coefficient value of 0.74. This suggests poor mental health outcomes are heavily influenced by neighboring areas, more so than local characteristics. In both model, extreme heat days does not have a significant relationship with poor mental health outcome when controlling for relevant controls and spatial dependence.

Table 3: Global Linear Regression Model with Relative Extreme Heat as Main Predictor

Variable	**Coefficient**	**P-Value**
Constant	18.90	0.00
Extreme Heat Days[1]	0.00	0.22
Socioeconomic Score	0.00	0.00
Impervious Surface Coverage (%)	0.01	0.00
Tree Canopy Coverage (%)	0.02	0.00
Mobile Homes (%)	0.01	0.00
Households Without Vehicle (%)	0.01	0.00

Days Above PM2.5 Regulatory Threshold	0.01	0.13
White Population (%)	0.04	0.00

[1] Defined as the number of days above the 95% historical average temperature per ZCTA.

Table 4: Spatial Lag Model with Relative Extreme Heat as Main Predictor

Variable	**Coefficient**	**P-Value**
Constant	5.81	0.00
Extreme Heat Days[1]	0.00	0.12
Socioeconomic Score	0.00	0.63
Impervious Surface Coverage (%)	0.00	0.21
Tree Canopy Coverage (%)	0.01	0.00
Mobile Homes (%)	0.00	0.76
Households Without Vehicle (%)	0.00	0.12
Days Above PM2.5 Regulatory Threshold	-0.01	0.01
White Population (%)	-0.02	0.00
Spatial Lag of P_MNTHL	0.74	0.00

[1] Defined as the number of days above the 95% historical average temperature per ZCTA.

DISCUSSION

At baseline, our results indicate clear geographic disparities in poor mental health outcomes in the U.S. in 2016. There is a high amount of clustering of poor mental health outcomes in the Southeastern U.S. and around Native American reservations, suggesting that underlying social and structural stressors may be influencing these outcomes in these areas.

Our results furthermore indicated clear strong spatial autocorrelation and therefore spatial dependence in poor mental health outcomes, confirming that location and context matter for mental health outcomes. A significant spatial lag term in both our models furthermore indicates that neighboring counties influence each other's mental health outcomes, potentially due to shared environmental or socioeconomic conditions. These findings underscore the importance of accounts for spatial effects when modeling mental health outcomes. Ignoring spatial dependence – as conventional regression methods do – may miss true associations between extreme heat and poor mental health. Notably, my prior analysis did not account for spatial dependence and yielded substantially different results.

In comparing our findings for different extreme heat indicators, notably maximum temperature (absolute measure) and extreme heat days (relative measure), it becomes clear that the absolute heat indicator is a much stronger predictor of poor mental health than the relative heat indicator. Specifically, extreme heat days (relative indicator) did not have a significant relationship with poor mental health outcomes in 2024. In contrast, the absolute measure of maximum temperature has a very strong positive relationship with poor mental health outcomes, even when controlling for relevant control variables and spatial dependence. These outcomes draw attention to the importance of indicator selection, emphasizing that the intensity of heat experienced is a much stronger predictor of poor mental health than whether temperatures are considered "extreme" based on historical norms.

Overall, our results indicate that as maximum temperature increases, so does poor mental health. It is important to note that effect of maximum temperature on poor mental health is particularly strong, having more effect than any other predictor in our model (with the exception of the spatial lag term). The fact that maximum temperature is a stronger predictor of poor mental health than factors like percent white population or socioeconomic status, even when controlling for spatial dependence, is a major finding and warrants further investigation.

It is important to acknowledge limitations in our study due to constraints in our methodology. First, we did not thoroughly verify whether our model assumptions were met, which means multicollinearity, for example, could have influenced the coefficient estimates in our models. Another key limitation is our selected absolute heat indicator. Though maximum temperature captures the highest recorded temperatures of a particular day, it does not fully reflect the heat experienced by populations. To better capture this, wet bulb temperature or heat index that includes humidity would provide a more accurate measure (Writer, 2024). A final note is that our data had a significant number of missing values to due to the U.S. Census not having data for all zip codes. Future studies should address these limitations by verifying model assumptions and using more comprehensive extreme heat indicators.

Ultimately, our study highlights the significant spatial relationship between maximum temperature and poor mental health. As climate change continues to drive up temperatures, understanding its impact on mental health is crucial, particularly in the U.S. where poor mental health is already a growing concern (Mental Illness, n.d.).

REFERENCES

Centers for Disease Control and Prevention. (n.d.). *Heat & Health Tracker*. U.S. Department of Health & Human Services. Retrieved April 9, 2025, from https://ephtracking.cdc.gov/Applications/heatTracker/

Climate change indicators: U.S. and global temperature | US EPA. (2025, February 4). US EPA. https://www.epa.gov/climate-indicators/climate-change-indicators-us-and-global-temperature

Extreme Heat | US EPA. (2025, March 27). US EPA. https://www.epa.gov/climatechange-science/extreme-heat

Fernandez-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems. *Journal of Machine Learning Research*, *15*. https://doi.org/10.5555/2627435.2697065

López, O. a. M., López, A. M., & Crossa, J. (2022). Random forest for genomic prediction. In *Springer eBooks* (pp. 633–681). https://doi.org/10.1007/978-3-030-89010-0_15

Mental illness. (n.d.). National Institute of Mental Health (NIMH). https://www.nimh.nih.gov/health/statistics/mental-illness

PRISM Climate Group. (n.d.). *Annual maximum temperature data*. PRISM Climate Group, Oregon State University. https://prism.oregonstate.edu

Thompson, R., Hornigold, R., Page, L., & Waite, T. (2018). Associations between high ambient temperatures and heat waves with mental health outcomes: a systematic review. *Public Health*, *161*, 171–191. https://doi.org/10.1016/j.puhe.2018.06.008

Writer, S. (2024, November 22). What System is Best for Measuring Extreme Heat? | Harvard T.H. Chan School of Public Health. *Harvard T.H. Chan School of Public Health*. https://hsph.harvard.edu/environmental-health/news/what-system-is-best-for-measuring-extreme-heat/

U.S. Census Bureau. (2023). *American Community Survey 5-Year Estimates: Detailed tables, race and ethnicity data* [Data set]. U.S. Department of Commerce. https://www.census.gov/data.html