Heatwaves

Dustin Leatherman

November 15, 2020

Introduction

In this analysis, daily temperatures are analyzed to determine the number of heat waves per year in Phoenix, Denver, Las Vegas, Albuquerque, Tucson, Salt Lake City, Los Angeles, San Francisco, and San Diego.

Frich et al define a heatwave as period where the daily high temperature exceeds the average maximum high temperature by more than 9 degrees for at least 5 days. (Frich et al. 2002).

In an attempt to form a definition for a heatwave that equally applied to all biomes of the US, Robinson a heatwave as a combination of relative humidity and air temperature; however, relative humidity is not present in the dataset so the starting definition provided by the NWS as "the exceedence of a fixed percentile of all observed values." (Robinson 2001) In our case, we will assume air temperature values that exceed an overall 95th percentile for at least 2 consecutive days to be considered a heatwave. It is interesting to note that both of these definitions allow for heatwaves to occur during periods outside of annual peak temperature.

The primarily interest is to determine a model that can predict the number of heatwaves that occur in a given city and year. Let Y_{ij} represent the number of heatwaves for the ith year and the jth city. Since this is a count model the likelihood function is $Y_{ij} \sim Poisson(N\lambda)$

Model 1

$$\lambda_{ij} \sim Gamma(Y_{i-1,j}, N)$$

where $N = 365$ and $Y_{1j} \sim InvGamma(100, 100)$

Model 2

A Poisson Regression model with Number of Heatwave Days as a covariate in a linear model.

$$log(\lambda_{ij}) = \beta_{0j} + \beta_{1j} X_{ij}$$

where $\beta_{0j}, \beta_{1j} \sim N(0, 1000)$

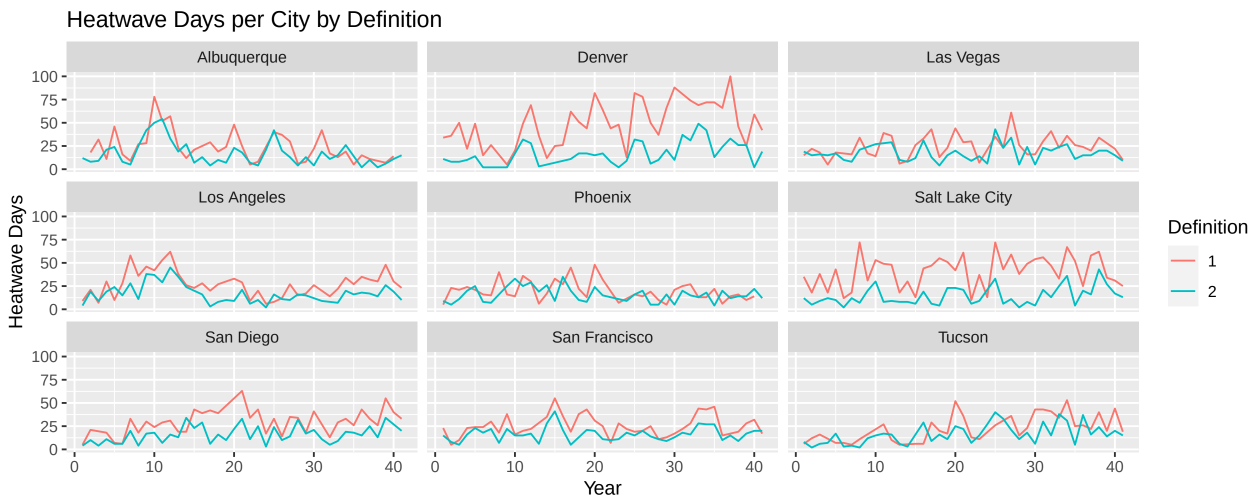
Model 3

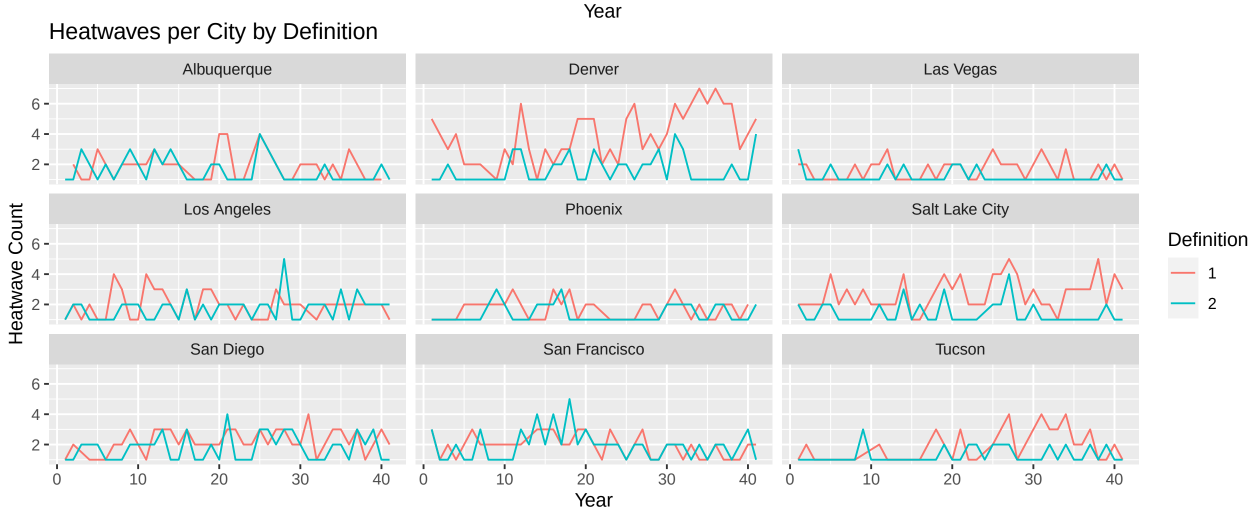
A Poisson Regression model with Number of Heatwave Days, some dependence on the previous value, and accounting for potential within-city variance.

$$log(\lambda_{ij}) = \beta_{0j} + \beta_{1j} X_{ij} + \rho Y_{i-1,j} + \alpha_j$$

where $\alpha_j, beta_{0j}, \beta_{1j} \sim N(0, 1000)$ and $\rho \sim beta(1, 1)$

Data





Model Comparisons

Table 1: Information Criteria for Models

Model	Definition	DIC	WAIC
model2	1	1039.737	1099.369
model1	1	1332.407	1322.769
model3	1	2110.194	2055.160
model2	2	969.718	991.427
model1	2	1262.221	1216.624
model3	2	2065.427	1793.347

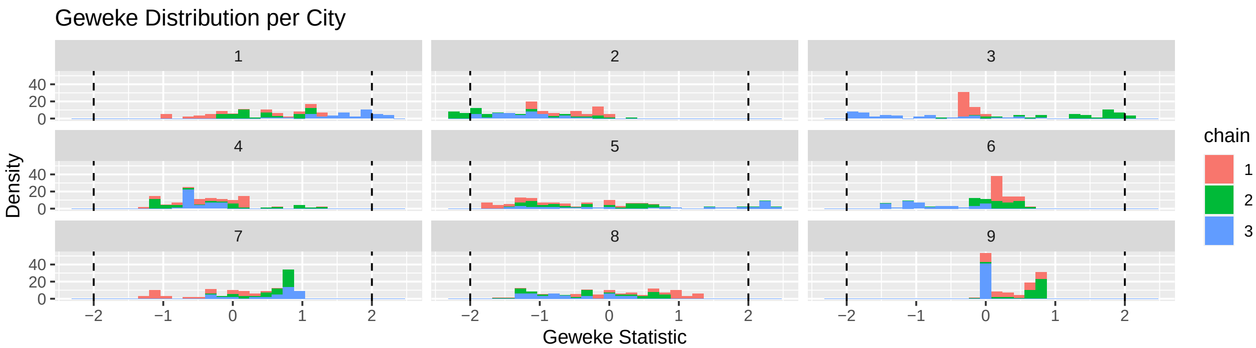
Model 2 has the lowest and WAIC and DIC so that will be used going forward. Missing values for heatwave days were imputed with the historical mean.

Convergence

Definition 1

Table 2: Gelman-Rubin Statistic Quantiles to measure convergence of chains

quantile	Point est.	Upper C.I.
min	0.9995063	0.9996456
0.25	1.0006266	1.0027274
0.5	1.0012211	1.0049579
0.75	1.0020892	1.0078415
max	1.0035161	1.0143752



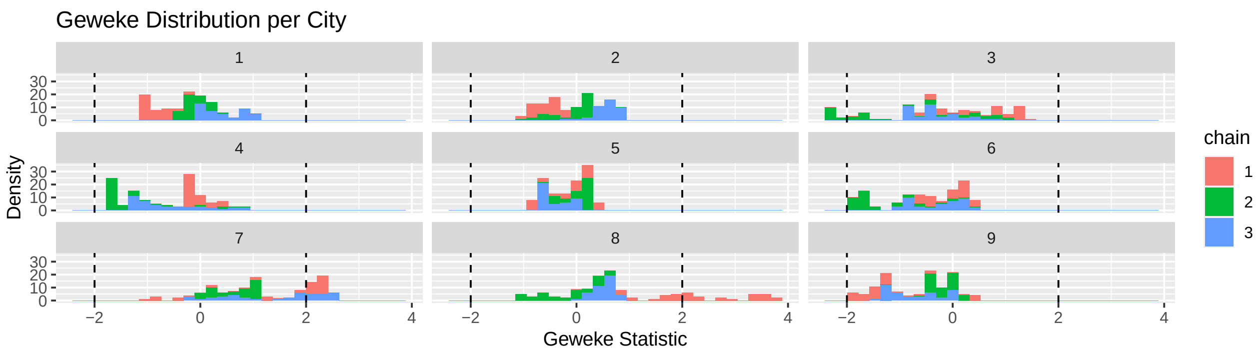
Histograms highlighting the convergence of the chains using the Geweke Statistic. It is ideal to be between the dotted lines.

The Gelman-Rubin Statistics are close to 1 which indicates convergence. Supporting this, most values of the Geweke statistic fall between (-2, 2) with maybe Tucson being the least likely to converge.

Definition 2

Table 3: Gelman-Rubin Statistic Quantiles to measure convergence of chains

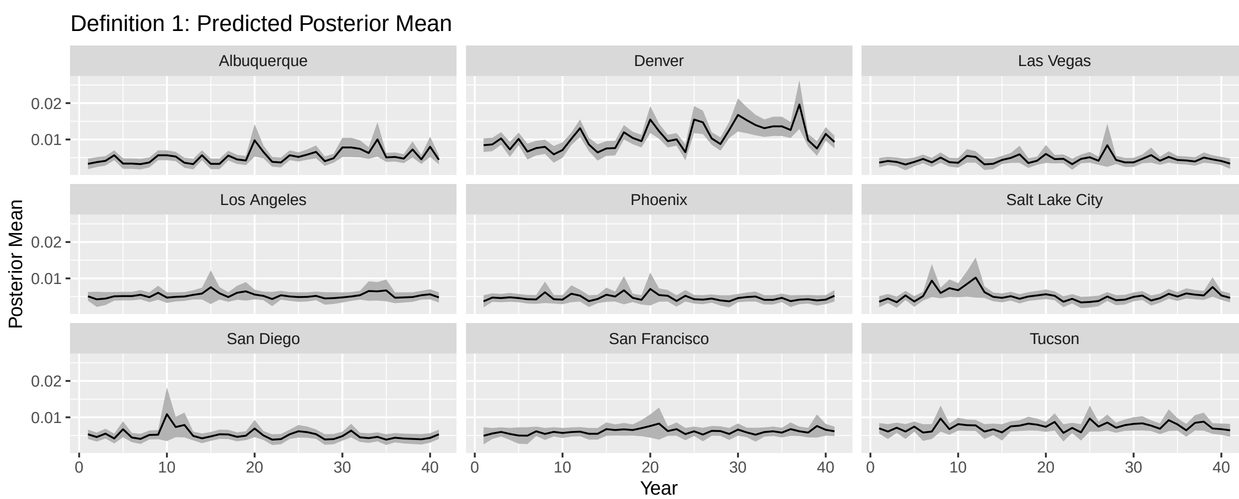
quantile	Point est.	Upper C.I.
min	0.9993617	0.9994901
0.25	0.9999548	1.0004358
0.5	1.0005230	1.0017753
0.75	1.0013289	1.0045448
max	1.0049623	1.0104827

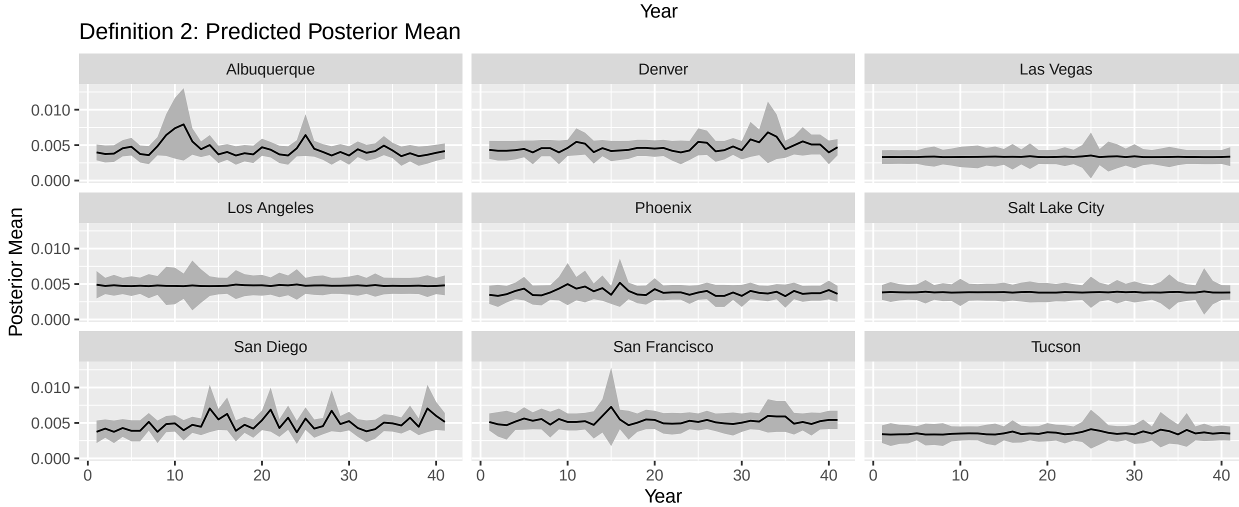


Histograms highlighting the convergence of the chains using the Geweke Statistic. It is ideal to be between the dotted lines.

Similar to Definition 1, the Gelman-Rubin statistics are close to 1 which indicate convergence. Most values of the Geweke statistic fall within the acceptable range as well.

Analysis





The KPSS test is used to test the stationarity of a univariate time series by testing for the presence of a unit root. The null hypothesis is a that a unit root exists and therefore the time series is non-stationary. The test returns diagnostics run against three different models: No Drift No Trend, Drift No Trend, and Drift with Trend. The model used in the diagnostic should be chosen based on the data. The data appears to have a

	Definition 1		Definition 2	
City	kpss	p.value	kpss	p.value
Albuquerque	0.5874306	0.0237790	0.1515276	0.1000000
Denver	0.4885967	0.0442350	0.3247026	0.1000000
Las Vegas	0.1930029	0.1000000	0.1620935	0.1000000
Los Angeles	0.0673660	0.1000000	0.2633187	0.1000000
Phoenix	0.3470366	0.0999842	0.3475203	0.0997757
Salt Lake City	0.2152194	0.1000000	0.0480312	0.1000000
San Diego	0.2739511	0.1000000	0.3591265	0.0947730
San Francisco	0.2027174	0.1000000	0.0415549	0.1000000
Tucson	0.3569821	0.0956974	0.5725897	0.0253176

non-zero mean and no trend so those respective diagnostics are used.

There is moderate evidence that the posterior means for Denver and Albuquerque are non-stationary under Definition 1 and there is moderate evidence that the posterior means for Tucson are non-stationary under Definition 2. There is not enough evidence for other cities to indicate that they are non-stationary. The and Albuquerque Denver graph for the posterior means supports this claim as it looks like it isn't straight across; however, the posterior means for Tucson appear stationary. The swings in Albuquerque for Definition 2 would also indicate some non-stationarity but the KPSS test doesn't show any indication that this is the case. This should be looked into more.

Final Thoughts

The analysis indicates that heatwave counts per year are generally constant as time goes on which bodes well for those living in the American West. Continuous monitoring should be ensured as Climate Change threatens to disrupt this. Further improvements can be made by gathering other covariates such as humidity which also plays a large role in determining whether a heatwave has occurred.

Frich, P., LV Alexander, P. Della-Marta, B. Gleason, M. Haylock, AMG Klein Tank, and T. Peterson. 2002. "Observed coherent changes in climatic extremes during the second half of the twentieth century." *Climate Research* 19 (January): 193–212. doi:10.3354/cr019193.

Robinson, Peter J. 2001. "On the Definition of a Heat Wave." *Journal of Applied Meteorology* 40 (4): 762–75. doi:10.1175/1520-0450(2001)040<0762:OTDOAH>2.0.CO;2.