

NCSU ST 540

Final Exam

Bruce Campbell

Model Definitions

In this report we detail a Bayesian analysis of temperature data to test if the distribution of heatwaves in the American Southwest has temporal dependence. We investigated a number of heat wave indices and found 2 common ones used in climate studies. The World Meteorological Organization (WMO) is specific in its definition by stating that a heat wave is when the daily temperature for more than five consecutive days exceeds the average temperature by 9 degrees F. There are other definitions used in climate research - for example the TX90P and WSDI - which take a quantile approach and look for windows of days exceeding the 90th quantile of temperature. These indices and links to further definitions can be found in (Alexander 2015). In this work we use the WMO definition and a variant of the quantile method TX90P that can be found in (Kueh et al. 2017) and uses a window of five days above the quantile to define a heat wave.

From the raw temperature data we calculated heat wave counts by year for the 9 cities in the data set. Recent climate research (Lau and Nath 2012) indicates that changes in high temperature are primarily seen by a shift in mean and not in higher moments. This leads us to consider the GLM family of models. We investigated Poisson and negative binomial models for this analysis. These random variables are invariant in distribution family under taking sums so we first fit models to the 41 yearly counts for each city $Y_{t,k} \sim Pois(\lambda)$, $Y_{t,k} \sim NB(\lambda)$. After considering these fits, we proceeded to fit the Poisson and negative binomial regression models for identifying trends over time.

For the regression models, the log link functions in both models. We use i for the time index - the explanatory variable in our model. The Poisson model is given by $\log(E[Y_{t,k}|t]) = \alpha_k + \beta_k t$ where $Y_{t,k} \sim Pois(e^{\alpha_k + \beta_k t})$.

The negative binomial model is given by $\log(E[Y_{t,k}|t]) = \alpha_k + \beta_k t$. Now if we let $p_{t,k}$, r_k be the parameters of a negative binomial, $E[Y_{t,k}] = \frac{pr}{1-p}$ after some rearranging we see then that $p_{t,k} = \frac{r_k}{r_k + \log(\alpha_k + \beta_k t)}$ and $Y_{t,k} \sim NB(p_{t,k}, r_k)$. Note we used one r_k per city to account for different dispersions. The priors used for both models are below - r is only present in the negative binomial model.

- $\alpha_k \sim N(0, \tau_k)$
- $\tau_k \sim Gamma(0.1, 0.1)$
- $\beta_k \sim N(0, \nu_k)$
- $\nu_k \sim Gamma(0.1, 0.1)$
- $r_k \sim Uniform(0, 10)$

The parameterization of the binomial given above is motivated by the author of our textbook in a blog posting (John-Kruschke 2012) where it's noted that direct estimation of p and r causes correlation in the MCMC chains. We put priors on the mean and r as suggested for the negative binomial. JAGS code can be found in the Appendix.

Convergence

Once we established working JAGS models, we ran a short run of the chains and calculated the Raftery and Lewis diagnostic. This is a run length diagnostic intended to provide insight on how

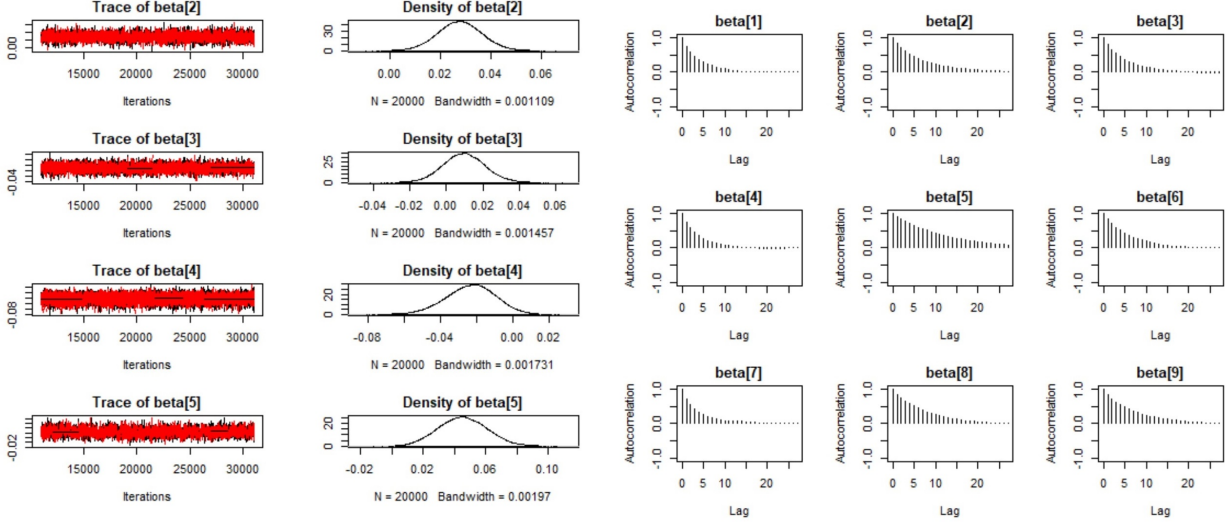


Figure 1:

long to run a chain. The diagnostic estimates the number of iterations required to estimate the quantile q to within an accuracy of $\pm r$ with probability p . We ran the chains for 1000 iteration, and the diagnostic indicated running the chains for 3746 iterations using $q = 0.25$, $r = \pm 0.005$, and $p = 0.95$. We choose to run the models for 20,000 iterations with a thinning of 2. We noted during the initial runs that one of the cities which had significant slopes in $\alpha_k + \beta_k t$ was more northerly so we added a third model for consideration which was a Poisson GLM with an added a term for latitude.

Convergence for the main run of the MCMC nodes was demonstrated through the following diagnostics; visual Inspection, the Gelman and Rubin Diagnostic, the Geweke Diagnostic, and the Heidelberg and Welch Diagnostic. We discuss these below and present convergence results for the Poisson GLM on the WHO heatwave counts. This model has the lowest penalized DIC for both definitions of heat wave; we present the DIC results below after discussing convergence. We inspected convergence plots for all of the coefficient nodes in all of the models, and spot checked other nodes. An example of the plots is given in Figure 1. Space does not permit us to display convergence diagrams for all of the models, but they can be recreated by running the code in the GitHub repository at [<https://github.com/brucebcampbell/bayesian-learning-with-R/tree/master/E3>].

The Gelman Rubin diagnostic tests whether multiple chains have converged to their stationary distribution by comparing the estimated between and within chain variances. Large differences between these variances indicate that the stationary distribution has not been reached.

The Geweke diagnostic takes two non-overlapping portions of the MCMC chain and compares the means. The test is a t-test with the standard errors adjusted for auto-correlation. In the figure below we plot the Z-scores for the t-tests.

The effective sample size (ESS) is a diagnostic describing the sample size of the MCMC chain adjusted for auto-correlation. Auto-correlation reduces the ESS. The convergence diagnostics for the Poisson GLM model fit to the WMO heatwave count data is displayed in figure 2. We did note that the ESS of some of the chains had a low value. We subsequently ran the MCMC chain for 100,000 iteration to see if we could get consistent ESS values for all the tracked nodes but we did

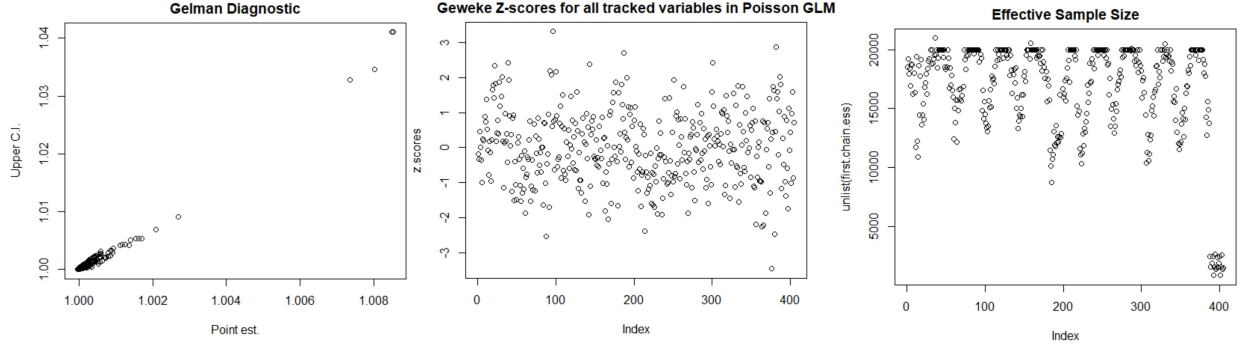


Figure 2:

not. The low ESS was observed on the intercept and slope coefficients of the GLM which do have significant auto-correlation out to lag 15-20 which could explain the low observed ESS values. We increased the thinning to 10 in order to accommodate the auto-correlation.

The Heidelberg and Welch Diagnostic uses Cramer-von-Mises statistic to test the null hypothesis that the MCMC is generating values from it's stationary distribution. See (Heidelberger and Welch 1983) for the details of this. We ran this diagnostic on all of the nodes in all of the models and kept track of the number of test that failed. There were never any runs where all of the test passed, but the number of failing tests was always less than five. This was investigated and the chain length was increased to remedy however there were still a few failed tests.

All of the models we fit showed evidence of good convergence based on the diagnostics.

Posterior Predictive Checks

We instrumented the JAGS code to perform posterior predictive checks. We simulate from the fitted model and calculate the mean and standard deviation of a sample. Then we compare to the mean and standard deviation of actual data. These are calculated by city. None of the p-values for any of the models were very close to 1 or 0, but the sd for Los Angeles was 0.095. We present the p-values for the Poisson GLM fitted to the WMO data below.

City	pval.mean	pval.sd
Phoenix	0.456	0.2665
Denver	0.4625	0.4139
Las Vegas	0.4738	0.3793
Albuquerque	0.4135	0.2149
Tucson	0.4962	0.6583
Salt Lake City	0.4418	0.5938
Los Angeles	0.4472	0.09505
San Francisco	0.5058	0.2925
San Diego	0.4985	0.2085

HPD for slopes

We present the 0.95% HPD intervals for β_k that are significant for a positive slope- i.e. do not contain zero in the 95% HPD interval and both endpoints are positive. Both the Poisson and negative binomial models gave the same significant cities for both of our heat wave indices. One of the four models we built had a negative association for Albuquerque. The Poisson model for the quantile version of the heat wave feature was significant for a negative association with heatwaves over time for that city. Tucson had the strongest relationship between heatwaves and time.

Table 2: Significant 0.95 HPD Intervals for slopes for WMO Poisson GLM

City	lower	upper
Denver	0.01086	0.04505
Tucson	0.0143	0.07484

FIT and DIC

The jags code was instrumented to provide predictions which were used calculate the MSE on the training data. The DIC was also evaluated to compare models. The Poisson is a limiting distribution for the negative Binomial. The DIC for the 2 models and the training MSE are very close. We would choose the Poisson model with linear predictor $\alpha_k + \beta_k t$ based on its simplicity (fewer model parameters) and lower DIC.

Table 3: WMO Heat Wave Data model comparison diagnostics

Model	Penalized deviance	Training MSE
Poisson GLM $\alpha_k + \beta_k t$	974.1	1.106
negative binomial GLM $\alpha_k + \beta_k t$	976.9	1.109
Poisson GLM $\alpha_k + \text{latitude}_k + \beta_k t$	975	1.108

Additional model considerations

Other models were investigated as well that could form a follow on analysis. The Zero Inflated Poisson model (ZIP) and a Bayesian change point model were experimented with. The ZIP model is used when the zero count is too high for a Poisson model. The distribution is described by a random sum of Bernoulli and Poisson random variables where a zero draw from the Bernoulli results in no contribution from the Poisson. The change point model attempts to find the location where a count process changes. $Y_{t,k} \sim \text{Pois}(\lambda_1) \forall t \in 1 : n$, $Y_{t,k} \sim \text{Pois}(\lambda_2) \forall t \in n + 1 : T$ and $n \sim \text{Uniform}(1, \dots, T)$.

Our models were checked by frequentist methods using R's `stats::glm` for the Poisson models and `MASS::glm.nb` for the negative binomial models. No discrepancies were noted.

Bibliography

Alexander, Lisa. 2015. “Introduction to Heatwave Indices.” <http://www.wmo.int/pages/prog/wcp/ccl/opace/opace4/meetings/WorkshopEnhancingClimIndices.php>.

Heidelberger, Philip, and Peter D. Welch. 1983. “Simulation Run Length Control in the Presence of an Initial Transient.” *Operations Research* 31 (6): 1109–44. doi:10.1287/opre.31.6.1109.

John-Kruschke. 2012. “Negative Binomial Reparameterization,” April. <http://www.wmo.int/pages/prog/wcp/ccl/opace/opace4/meetings/WorkshopEnhancingClimIndices.php>.

Kueh, M-T, C-Y Lin, Y-J Chuang, Y-F Sheng, and Y-Y Chien. 2017. “Climate Variability of Heat Waves and Their Associated Diurnal Temperature Range Variations in Taiwan.” *Environmental Research Letters* 12 (7): 074017. <http://stacks.iop.org/1748-9326/12/i=7/a=074017>.

Lau, Ngar-Cheung, and Mary Jo Nath. 2012. “A Model Study of Heat Waves over North America: Meteorological Aspects and Projections for the Twenty-First Century.” *Journal of Climate* 25 (14): 4761–84. doi:10.1175/JCLI-D-11-00575.1.