Modeling and Inference for Precipitation Time Series

Nicholas Vasko, Penn State, SCRiM Summer Scholars Program

Adviser: Dr. Murali Haran

The Pennsylvania State University · Department of Statistics · The Network for Sustainable Climate Risk Management

Summary

- This study analyzes monthly precipitation for Phoenix, AZ and Warren, PA
- These were chosen for their differences in climate patterns
- Normal ARMA models break down for precipitation because non-gaussian and must be positive
- A Generalized Autoregressive Moving Average model is used instead
- The GARMA model works well for capturing the mean structure of precipitation, but breaks down for the conditional variance

Background and Exploratory Data Analysis

- Climate change will affect precipitation patterns in certain areas differently
- Some areas will experience more severe drought
- Others will have more extreme rainfall

Table 1

- Precipitation amounts differ by location and season of the year
- Variability is also dependent on location

Location	Mean Annual Precipitation (in.)	Season of Max Average Precipitation	Season of Min Average Precipitation	Ratio of Max / Min
Phoenix, AZ	7.417	Winter	Spring	1.91
Warren, PA	43.684	Summer	Winter	0.79

Data and Methodology

- Precipitation data obtained from the NOAA Climate Data Online database
- The GARMA model is an extension of the more popular ARMA model
- It allows for non-gaussian error structure in the model's residuals.
- A Gamma-GARMA model is used for precipitation due to the distribution
- Parameters were estimated using Markov Chain Monte Carlo Methods
- R software was used for the analysis

GARMA Model Definition

- For a Gamma-GARMA (1,0) model we have the following
- A time series $\{Y_t\}$, for t = {1,..T}, defined by three parameters
- $Y_t \sim Gam(c\mu_t^d, c\mu_t^{d-1})$,
- It uses a link function g such that,
  - $g(\mu_t) = v + \varphi\, g(y_t)$,
  - Where $g(y_t) = \log(y_t)$

Simulation Study

- Here 500 random time series were drawn from a GARMA model
- The true parameters were chosen for similarity to actual data
- They were v = 0, c = 1, d = 0, φ = 0.3
- The rjags package in R was used to estimate the parameters
- The results show that the estimates performed significantly well for the data
- For c, d, and φ, the true value was contained in 96.6%, 94.8%, and 96.4% of respective critical intervals

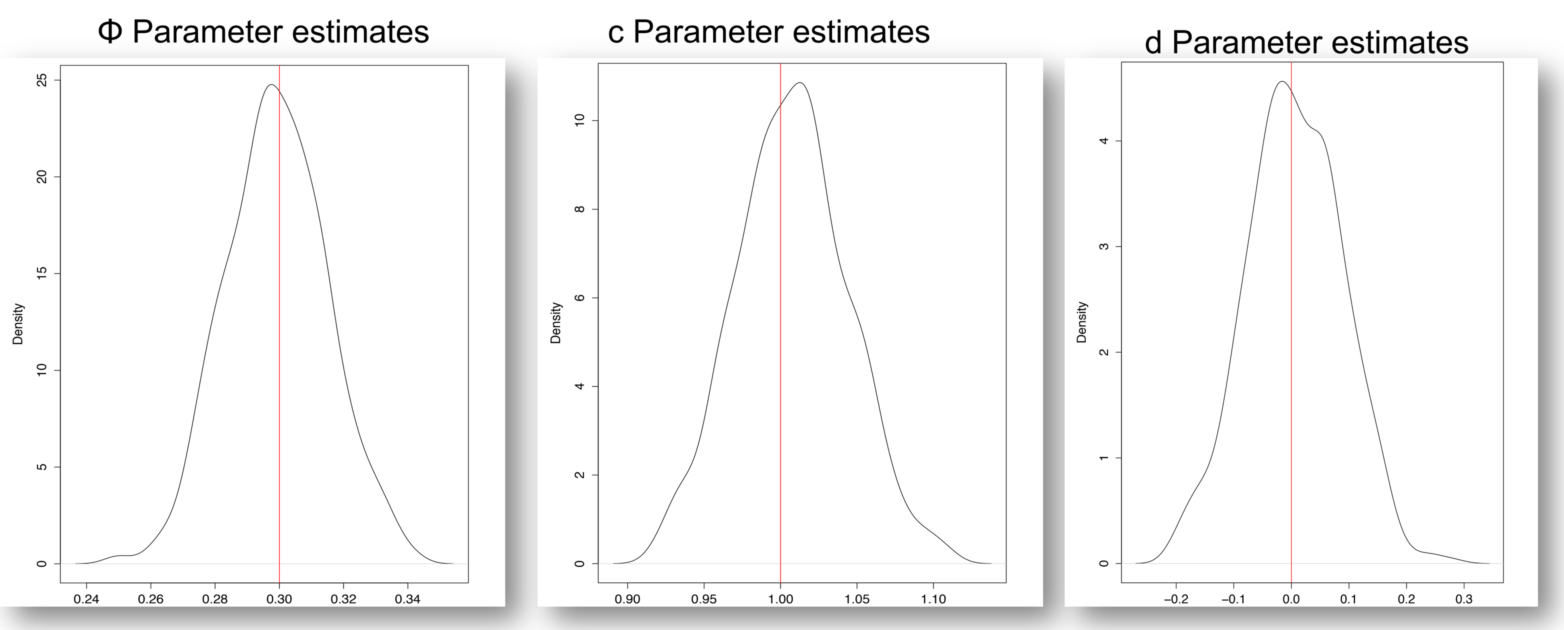

Figure 1
The three probability plots above represent the range of probabilities for each parameter estimate. The true value is in red.

Preliminarily Results

- Warren, PA and Phoenix, AZ have very different characteristics
- Monthly means regression models were considered to model seasonality
- The best regression models were determined based on the highest r^2 values
- The best model was based on whether the residuals were uncorrelated
- Also, whether one-step ahead forecasts fit the data well

Warren, PA

- Best model: Gamma-GARMA(0,0) with monthly means regression
- Very little autocorrelation between months
- The monthly means model does a good job of modeling this structure
- This model has non-gaussian error structure that is uncorrelated

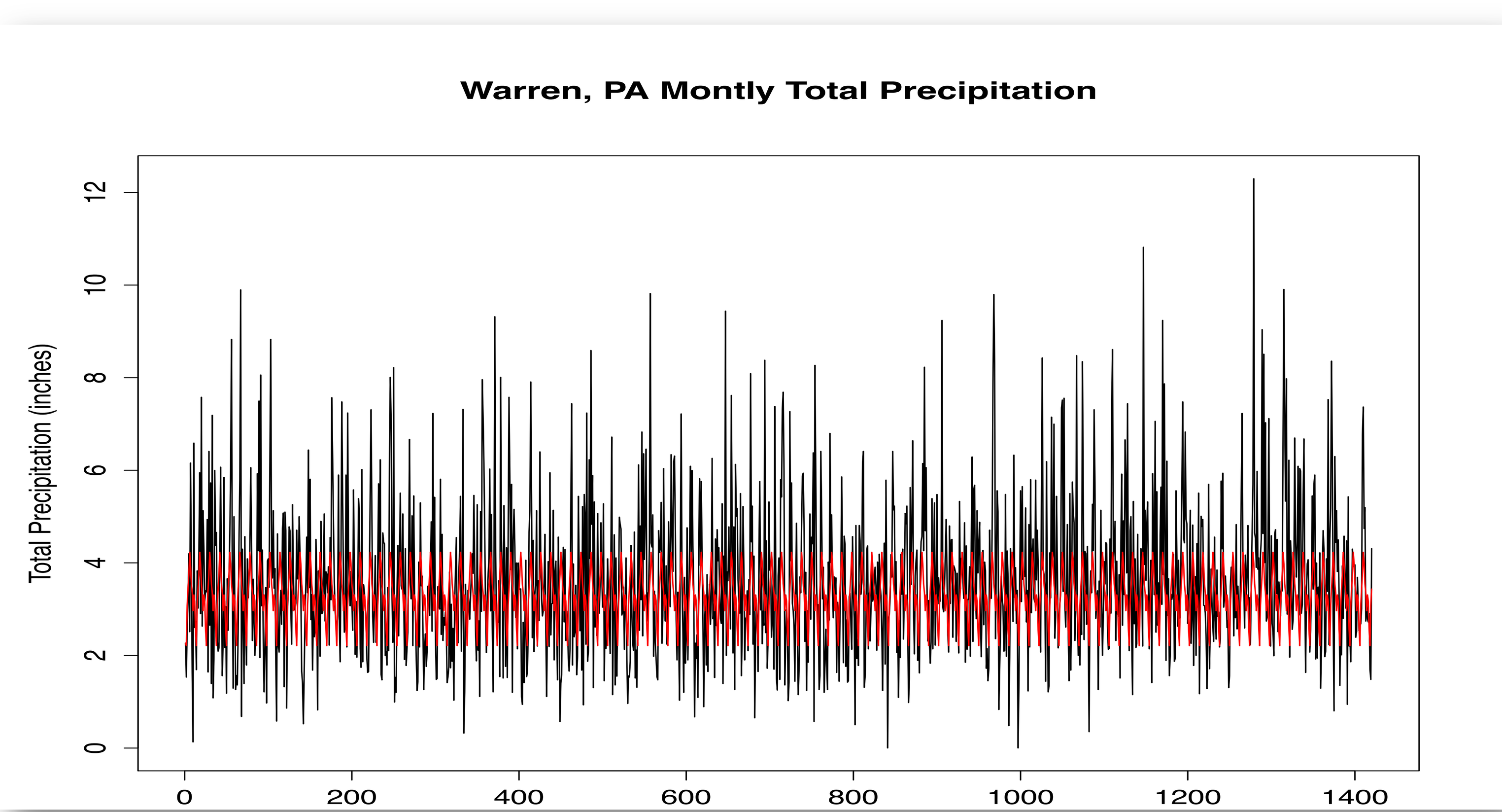

Figure 2 – Warren, PA (1897-2015)
The precipitation data is in black, with GARMA forecasts in red.

Phoenix, AZ

- A monthly means regression model still has correlated residuals
- A Gamma-GARMA(1,0) model was determined to be a best fit
- This model accounts for the autocorrelation that was seen in the data

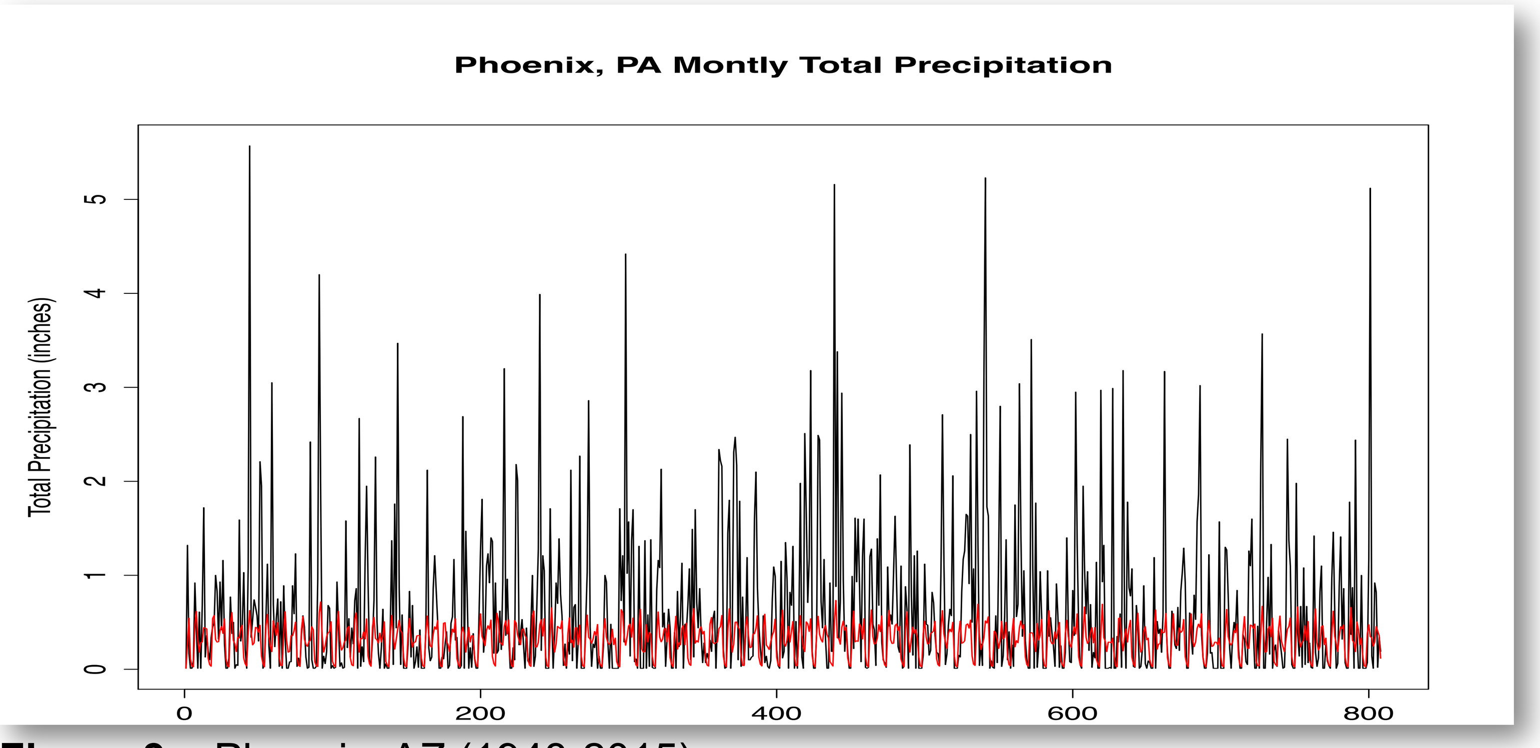

Figure 3 – Phoenix, AZ (1948-2015)
The precipitation data is in black, with GARMA forecasts in red.

Discussion/ Future Work

- The Gamma-GARMA models do a good job at accounting for autocorrelation
- Mean values for these precipitation data are modeled well
- However, the models seem to break down when it comes to modeling the conditional variance
- Future work on modeling this could involve using a GARCH model
- Looking into dry and wet spells at each of these locations, as well as others across the United States
- From this analysis, future work will focus on prediction over the next century
- Unusual dry periods are of specific importance because:
  - Changes in precipitation may cause problems with water supply availability
  - Water supplies may run out in an extended extreme drought
  - This may cause problems with people's health and growing crops

Acknowledgements

This work was supported by the National Science Foundation through the Network for Sustainable Climate Risk Management (SCRiM) under NSF cooperative agreement GEO-1240507.

References

- Zheng.T., et al., Generalized ARMA models with martingale difference errors. Journal of Econometrics (2015). http://dx.doi.org/10.1016/j.jeconom.2015.03.040
- Briet OJT, Amerasinghe PH, Vounatsou P (2013) Generalized Seasonal Autoregressive Moving Average Models for Count Data with Application to Malaria Time Series with Low Case Numbers. PLoS ONE 8(6): e65761. doi: 10.1371/journal.pone.0065761

PENNSTATE 1855

NSF

SCRiM an NSF-sponsored research network for Sustainable Climate Risk Management