A dive into Bayesian Weather Modelling

Victor Lei, Charles Liu, Leonhard Spiegelberg, Vinay Subbiah

1.Introduction:

For our project we were interested in exploring different approaches of statistical nature that are used currently in the complicated topic of weather modelling. Our investigation thereby focuses on two different types of statistical scenarios:

1. Statistical Models for weather measures (temperature, precipitation, pressure)
2. Statistical Models for events (Rain, Snow, Hurricane)

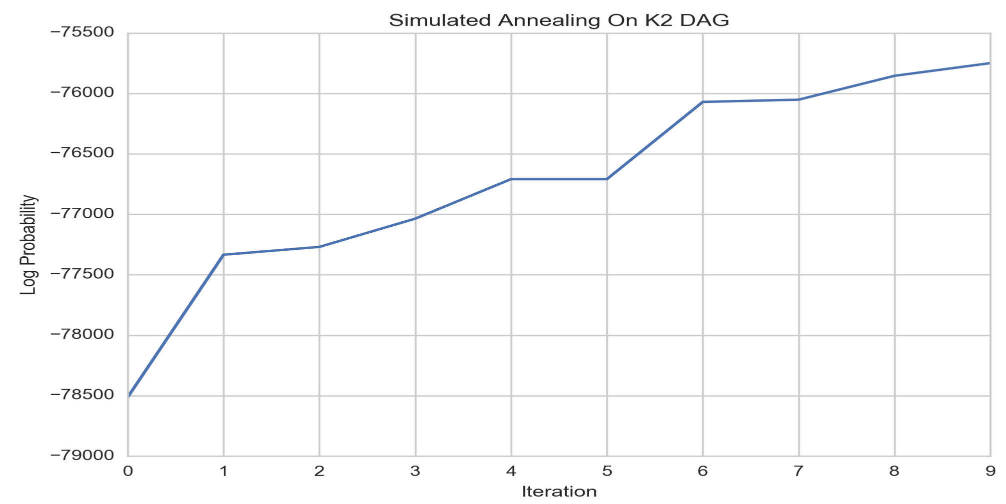
Our primary data source for this project is www.wunderground.com which provides historical data of major weather stations on a daily basis for up to 50 years.

2. Bayesian Networks

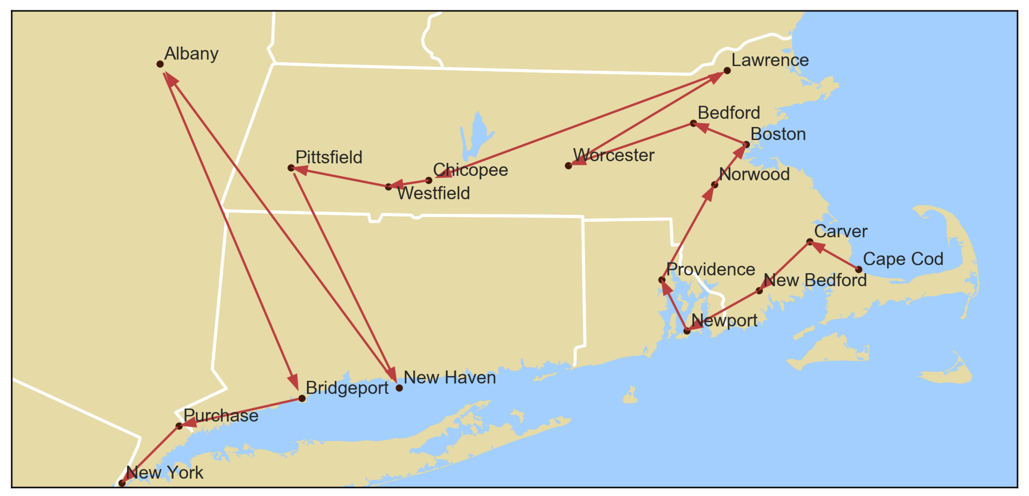
Bayesian Networks (BNs) are one of the most famous ways to model arbitrary dependencies via a directed acyclic graph (DAG). For our study we constructed a BN for precipitation that could be easily extended to other weather measures. Our network involves 18 different regions in the Northeast as shown on our maps.

3. Bayesian Network Learning:

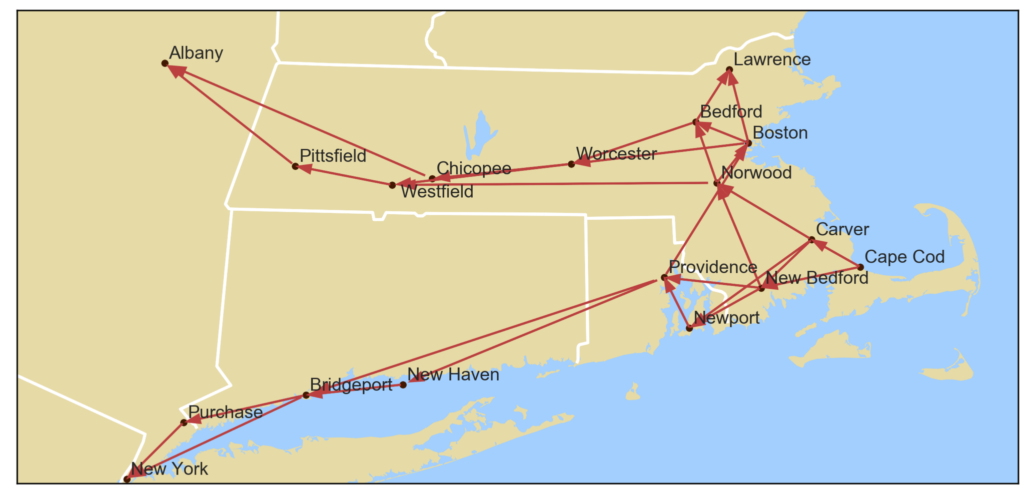
a. Minimum Spanning Tree based on Euclidean Distance



c. Simulated Annealing Starting From K2 Network



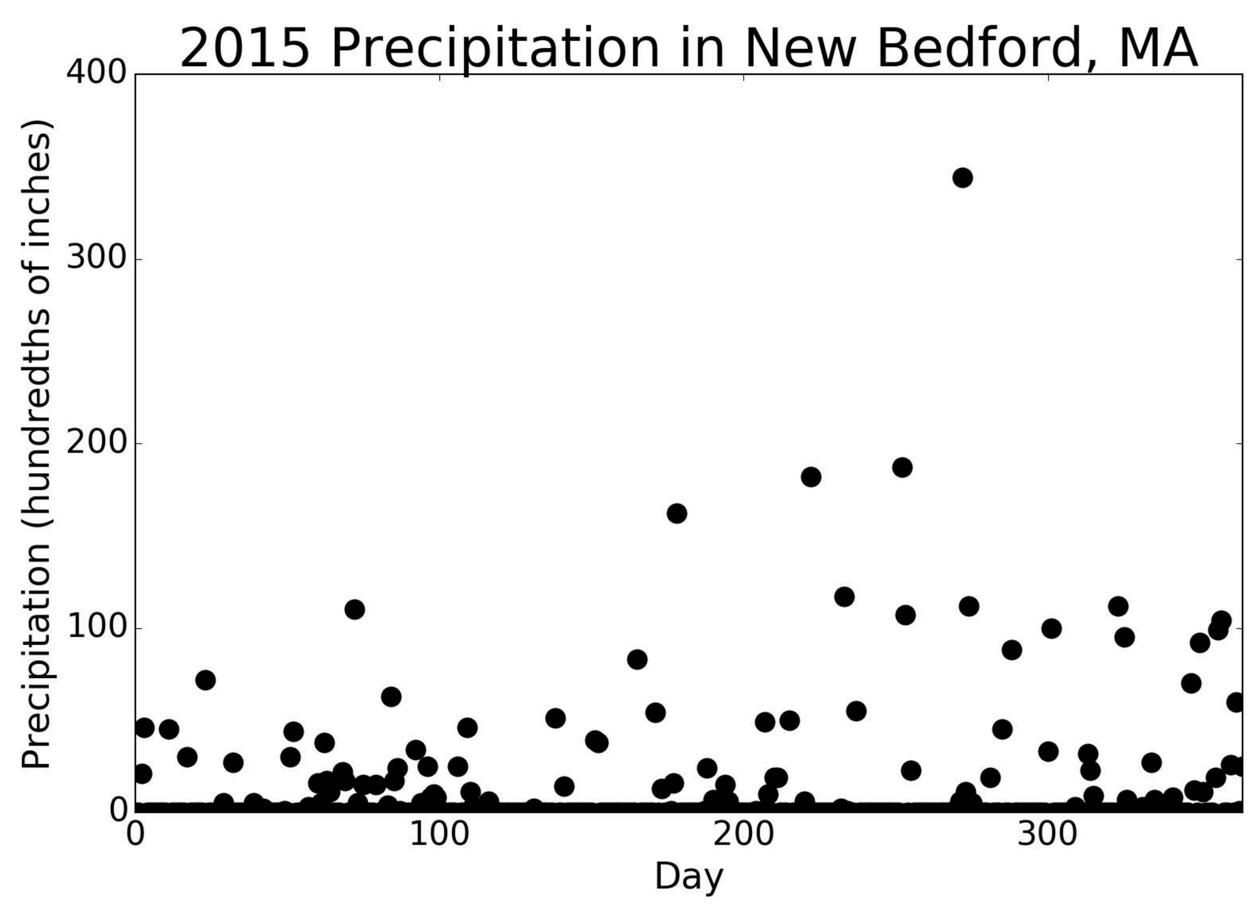
b. Topological Ordering obtained via BFS



d. Final Bayesian Network After K2+SA

4. Bayesian Hierarchical Rainfall

Early statistical models for rainfall were fairly simple due to computational constraints. We revisit and extend the classical rainfall model using MCMC, and apply it to recent rainfall data across several north-eastern cities in the US.



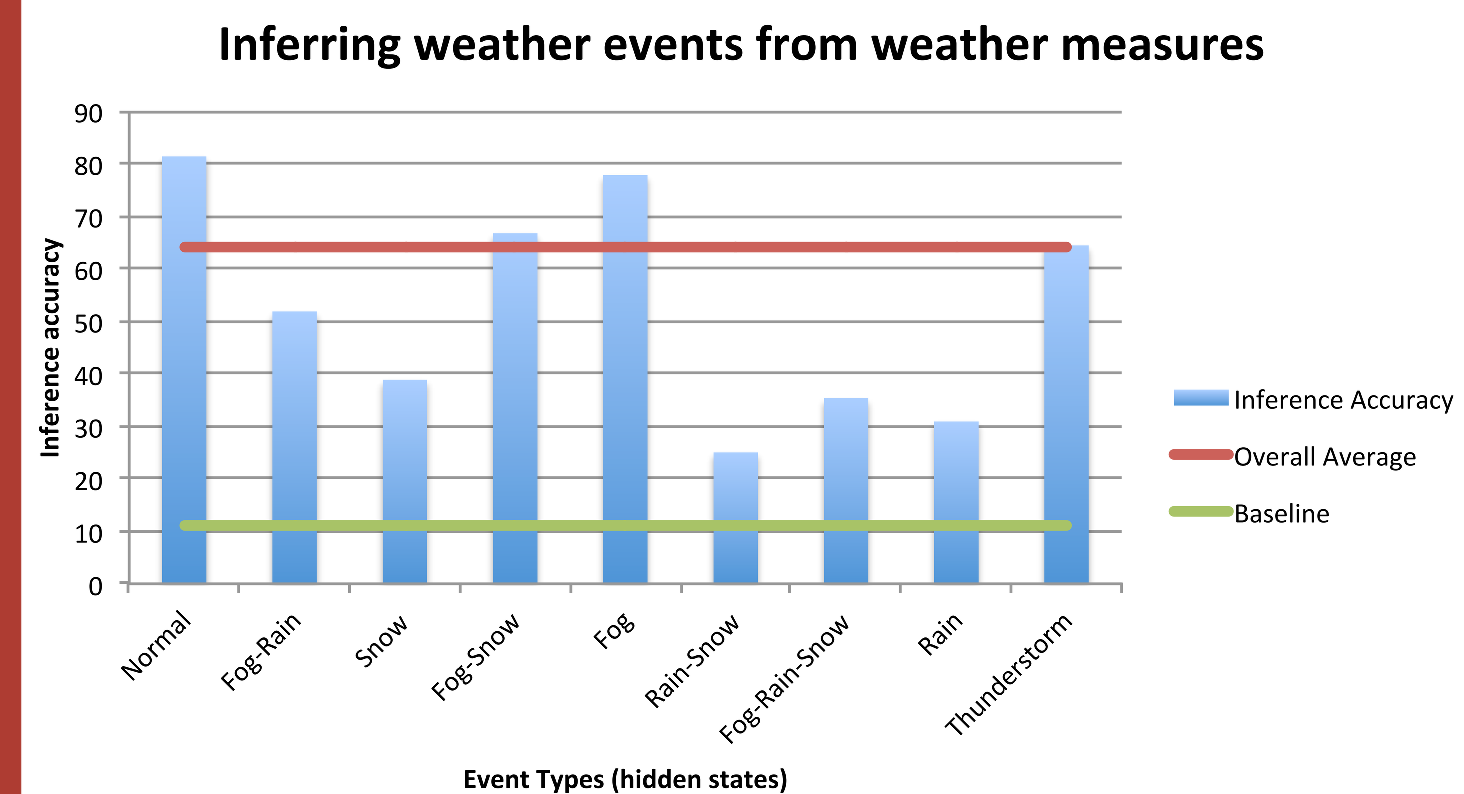
We identify approximate changing points in rainfall volume and find that there are statistically significant differences in rainfall volume during different periods of the year.

However, in accordance with empirical evidence, the shape parameters of the Gamma distribution tend not to change.

6. 2nd order Hidden Markov Models:

We want to infer weather events based on weather measures. Specifically, given measures like temperature or visibility we infer the occurrence of events like Fog-Snow, Rain, Thunderstorm etc. We use a second order Hidden Markov Model to infer these events (hidden states) and model the emission probabilities of the weather measures using a joint distribution of Gaussian and log-normal distributions.

The model was trained on 50 years of data (1960 - 2010) and the test results below are from the next five year period 2011 - 2015.



7. References:

1. Coe, R., and R. D. Stern. "Fitting models to daily rainfall data." Journal of Applied Meteorology 21.7 (1982): 1024-1031.
2. Cano, R., Sordo, C. and Gutiérrez, J.M., 2004. “Applications of Bayesian networks in meteorology”. In Advances in Bayesian networks (pp. 309-328). Springer Berlin Heidelberg.
3. Yang He, "Extended Viterbi algorithm for second order hidden Markov process," Pattern Recognition, 1988., 9th International Conference on, Rome, 1988, pp. 718-720 vol.2. doi: 10.1109/ICPR.1988.28338