## Essays on Bayesian Modeling of Power Outages

### Synopsis

It is essential to have a reliable power system because power outages can have significant negative impacts on households, businesses, and manufacturers. Outages can lead to financial losses causing damage to perishables goods, downtime, and lost production. The duration of the outage is as important as the number of individuals affected. The longer the outage duration is, the greater the incurred losses are likely to be. Campbell [4] shows that the estimated cost of only storm-related power outages on the U.S. economy is between \$20 billion and \$55 billion. The consequences of a lack of preparation for power outages can be huge for utility companies. When they do not have enough resources to deal with a sudden need for a large number of emergency repair crews and vehicles, waiting for help from state agencies and major power companies from other areas can take weeks and cause great losses [14]. Accurately estimating the number of power outages can help utility companies be better prepared and potentially save a lot of money. In order to develop prevention strategies and prepare response policies such as repair crew scheduling and positioning, it is crucial to be able to predict power outages and forecast potential damages they may cause.

The effects of power outages are not limited to financial losses mentioned above. Klinger et al. [8] discuss the impact of power outages on communication, health, and safety. Not being able to use landlines to dial 911 or charge phones, loss of functioning home oxygen supplies, carbon monoxide poisoning caused by the unsafe use of generators for electricity and gas-powered heaters for cooking and heat generation, lack of temperature and light control, and compromised drug storage are only some of the health and safety related effects of power outages. Therefore, every household affected by a power outage can be considered as a household whose safety is in jeopardy. Accurately predicting the number of households affected by the power outages or the total customer outage hours caused by a power outage would be very valuable to utility companies for crew and repair scheduling.

A power network consists of many components, including the transmission substations and lines, distribution substations, overhead and underground service lines. Each of these components are vulnerable to different types of risks. Some might be more exposed to bad weather or might be more sensitive to natural disasters because of their geographical surroundings. Thus, modeling and prediction of power outages is a challenging task. Due to its criticality for utility companies, prediction of power outages and their consequences have gained considerable attention during the last two decades. A recent review of literature on power outage prediction can be found in Kabir, Guikema and Quiring [7].

As weather being one of the major factors affecting the reliability of power distribution systems, some studies use rough classifications of adverse and normal weather to model failures of power distribution lines [13] [16] and distribution feeders [17]. Others use more specific weather related variables such as wind speed, wind gust, soil moisture and the temperature. Li et al. [9] considers weather predictions provided by Deep Thunder, a deep learning service by IBM in their hierarchical model. A major challenge with these models is reliance on accurate weather predictions. As pointed out by Kabir, Guikema and Quiring [7], the uncertainty in predictors as well as the uncertainty associated with paramter estimation may contribute to forecast errors. Effect of hurricanes on power outages in multiple locations have been studied using various models. Liu et al. [10] develops generalized linear models (GLMs) for estimating the spatial distribution of power outages during hurricanes by using indicator variables for each hurricane. Han et al. [5] develops a GLM to predict the number of power outages during hurricanes in multiple locations using different characteristics of the hurricanes. But again these models rely on hurricane predictions for predicting outages.

Nonparametric models such as Bayesian Additive Regression Trees and Random Forests are also used to model storm outages [12, 6, 15]. More recently, Kabir, Guikema and Quiring [7] used Boosting Trees, Random Forests and Support Vector Machines for predicting thuderstorm-induced power outages. As in the majority of machine learning (ML) methods, these models rely on large number of predictors many of which will not be known for predicting future outages. For example, in the Random Forest model of Kabir, Guikema and Quiring [7] majority of the important predictors are weather related and thus, rely on predictions.

In this dissertation, our objective is to develop Bayesian power outage models that enable us to predict future outages as well as their consequences in terms number of households affected or total customer outage hours in different locations. In doing so, we present a Markov modulated compound Poisson process model. This is achieved by taking the homogeneous compound Poisson process model and making its rates and jump probabilities dependent on a random environmental process. The environmental process is governed by a continuous time Markov chain. The attractive feature of this approach is its ability to implicitly capture stochastic predictors such as weather conditions via the environmental process. Furthermore, other deterministic predictors can be incorporated in the model by deterministic modulation of the rates and the jump probabilities as will be discussed in Chapter 2 (Essay 1: Modeling Number of Households Affected by Power Outages).

There are many examples of modulation in the reliability literature. For example, random environments are used to provide a tractable model of the stochastic dependence among the components of a device where the environment is an external process that depicts all physical, structural, operational, and other conditions which affect the deterioration, aging, and failure of the system. Since all components are subject to the same environmental conditions, their lifetimes are dependent on their common environmental process. Thus, the environmental process is actually a factor of variation in the failure structure of the system. These ideas were introduced by Cinlar and Ozekici [14] who propose to construct an intrinsic clock which ticks differently in different environments to measure the intrinsic age of the device. This intrinsic aging model is studied further by Cinlar et al. [15] to determine the conditions that lead to associated component lifetimes. The association of the lifetimes of components subjected to a randomly varying environment is discussed by Lefevre and Milhaud [22]. Singpurwalla [38] provides a review by discussing potential hazards in reliability modelling. Applications also include hardware reliability where a device performs a stochastic mission and its failure rate depends on the stage of the mission. Cekyay and Ozekici [11] discuss issues related to mean time to failure and availability when the mission or environmental process is semi-Markovian. The reader is referred to Cekyay and Ozekici [13, 12] for issues related to performance analysis and maintenance of such modulated reliability models. First consideration of modulation in software reliability applications was in Ozekici and Soyer [28] who assume that the failures of the software depend on its operational profile, which is now the environmental process that represents the sequence of operations that the software performs. In a recent article, Landon et al. [21] present a tractable Bayesian approach Markov modulated Poisson model for software reliability.

Use of Markov modulated compound Poisson processes have not been considered in the power outage modeling literature. The benefit of using a Markov modulated compound Poisson process to model power outages and the number of affected households is that it does not require any weather or geographical information to capture the effect of any of these factors. The model treats the environment that system operates in as a stochastic process whose different states capture different environmental characteristics. Also, the Markov modulated compound Poisson process relaxes the independent and stationary increments properties of the homogeneous compound Poisson process model.

Furthermore, Bayesian analysis of Markov modulated compound Poisson processes which we present in Chapter 2 and their extensions are new for the literature in Bayesian inference in stochastic processes. In our development in Chapter 2, we show that the Markov modulated compound Poisson process model is a special case of the doubly stochastic Markov jump processes; see Ay et al. [1]. Thus, the Bayesian methods presented in Chapter 2 can be adopted for a richer class of Markov modulated Markov jump processes. Other methodological contributions of Chapter 2 include development of a prediction algorithm for time to the next outage, number outages over a fixed time period as well as for number of households affected by each outage. Finally, Markov modulated nonhomogeous compound Poisson process model and its Bayesian analysis represents another modest methodological contribution of Chapter 2.

Geographical characteristics play an important role in developing models for power outage forecasting. Liu et al. [11] and Han et al. [5] have considered spatial models for outages due to hurricanes and ice storms. These models are spatial versions of GLMs and they rely on stochastic predictors. In Chapter 3 (Essay 2: A Random Environment Model for Multiple Location Outages) we present spatial versions of Markov modulated compound Poisson processes. We start our development by considering multiple location compound Poisson processes affected by a common random environment. Thus, the random environment model induces both temporal and spatial correlations. Bayesian analysis of the model requires modification of our Gibbs sampler of Chapter 2 to deal with with superposition of compound Poisson processes. The second extension of the model involves incorporating location based (derministic) covariates (predictors) to further modulate the rates and the jump probabilities. The introduction of the covariates requires a modification of the Bayesian analysis by use of Metropolis within Gibbs. In the final part of Chapter 3 we introduce a Markov modulated spatial model where we explicitly describe spatial correlations using spatial autoregressions in the sense of Besag [3] and Banerjee, Carlin and Gelfand [2]. To the best of our knowledge, our modeling and methodology development in Chapter 3 present contributions to both the literature in power outage modeling and Bayesian inference in stochastic processes.

# References


[1] A. Ay, J. Landon, S. Özekici, and R. Soyer, Bayesian Analysis of Doubly Stochastic Markov Processes in Reliability, *Probability in the Engineering and Informational Sciences* 2020, Advance online publication. doi: 10.1017/s0269964820000157

[2] S. C. Banerjee, B. P. Carlin and A. E. Gelfand Hierarchical Modeling and Analysis for Spatial Data, *London: Chapman and Hall.*

[3] J. E. Besag Nearest-neighbor systems and the auto-logistic model for binary data, *J. Royal Statist. Soc.*, 1972 34:75 - 83

[4] R. J. Campbell, Weather-Related Power Outages and Electric System Resiliency, *CRS Report for Congress* 2012, 1-15.

[5] S.R. Han, S. D. Guikema, S. M. Quiring, K.H. Lee, D. Rosowsky, R. A. Davidson, Estimating the spatial distribution of power outages during hurricanes in the Gulf coast region, *Reliability Engineering and System Safety*, 2009 94: 199 - 210.

[6] J. He, D. W. Wanik, B. M. Hartman, E. N. Anagnostou, M. Astitha and M. E. B. Frediani , Nonparametric Tree-Based Predictive Modeling of Storm Outages on an Electric Distribution Network, *Risk Analysis*, 2017 37: 441 - 458.

[7] E. Kabir, S.D. Guikema and S. Quiring, Predicting thunderstorm-induced power outages to support utility restoration, *IEEE Transactions on Power Systems*, 2019 34: 4370-4381.

[8] C. Klinger, W. Landeg and V. Murray, Power Outages, Extreme Events and Health: a Systematic Review of the Literature from 2011-2012 , *PLOS Currents Disasters*, 2014 1: 1-23.

[9] H. Li, L.A. Treinish, J.R.M. Hosking, A statistical model for risk management of electric outage forecasts, *IBM J. Res. Develop.*, 2010 54: 8:1-8:11.

[10] H. Liu, R. A. Davidson, D.V. Rosowsky and J.R. Stedinger, Negative Binomial Regression of Electric Power Outages in Hurricanes, *Journal of Infrastructure Systems*, 2005 11: 258-267.

[11] H. Liu, R. A. Davidson, T.V. Apanasovich, Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms., *Reliability Engineering & System Safety*, 2008 93: 897-912.

[12] S. Shashaani, S. D. Guikema, C. Zhai, J. V. Pino and S. M. Quiring Multi-Stage Prediction for Zero-Inflated Hurricane Induced Power Outages, *IEEE Access*, 2018 6: 62432 - 62449.

[13] P. Wang and R. Billinton, Reliability Cost/Worth Assessment of Distribution Systems Incorporating Time-Varying Weather Conditions and Restoration Resources, *IEEE Transactions on Power Delivery*, 2002 17: 260–265.

[14] S. Wang, B. R. Sarker, L. Mann Jr. and E. Triantaphyllou, Resource planning and a depot location model for electric power restoration, *European Journal of Operational Research*, 2004 155: 22–43.

[15] D.W. Wanik, E.N. Anagnostou, B.M. Hartman, M.E.B. Frediani and M. Astitha, Storm outage modeling for an electric distribution network in Northeastern USA, *Nat Hazards* 2015 79: 1359–1384.

[16] Y. Zhou, A. Pahwa and S. Das, Prediction of weather-related failures of overhead distribution feeders, in *Proc. Probability Methods Applications Power Systems Conf.*, Sep. 2004.

[17] Y. Zhou, A. Pahwa and S. Yang, Modeling Weather-Related Failures of Overhead Distribution Lines, *IEEE Transactions on Power Systems*, 2006 21: 1683-1690.