Atmospheric Research 102 (2011) 243–262

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmos

Bayesian analysis for extreme climatic events: A review

Pao-Shin Chu [a,*], Xin Zhao [b]

[a] Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, Hawaii 96822, United States
[b] Sanjole Inc., Honolulu, Hawaii 96822, United States

A R T I C L E I N F O

Article history:
Received 8 October 2010
Received in revised form 4 July 2011
Accepted 11 July 2011

Keywords:
Bayesian analysis
Extreme events
Change-point
Tropical cyclone path clustering
Tropical cyclone prediction

A B S T R A C T

This article reviews Bayesian analysis methods applied to extreme climatic data. We particularly focus on applications to three different problems related to extreme climatic events including detection of abrupt regime shifts, clustering tropical cyclone tracks, and statistical forecasting for seasonal tropical cyclone activity. For identifying potential change points in an extreme event count series, a hierarchical Bayesian framework involving three layers – data, parameter, and hypothesis – is formulated to demonstrate the posterior probability of the shifts throughout the time. For the data layer, a Poisson process with a gamma distributed rate is presumed. For the hypothesis layer, multiple candidate hypotheses with different change-points are considered. To calculate the posterior probability for each hypothesis and its associated parameters we developed an exact analytical formula, a Markov Chain Monte Carlo (MCMC) algorithm, and a more sophisticated reversible jump Markov Chain Monte Carlo (RJMCMC) algorithm. The algorithms are applied to several rare event series: the annual tropical cyclone or typhoon counts over the central, eastern, and western North Pacific; the annual extremely heavy rainfall event counts at Manoa, Hawaii; and the annual heat wave frequency in France.

Using an Expectation-Maximization (EM) algorithm, a Bayesian clustering method built on a mixture Gaussian model is applied to objectively classify historical, spaghetti-like tropical cyclone tracks (1945–2007) over the western North Pacific and the South China Sea into eight distinct track types. A regression based approach to forecasting seasonal tropical cyclone frequency in a region is developed. Specifically, by adopting large-scale environmental conditions prior to the tropical cyclone season, a Poisson regression model is built for predicting seasonal tropical cyclone counts, and a probit regression model is alternatively developed toward a binary classification problem. With a non-informative prior assumption for the model parameters, a Bayesian inference for the Poisson regression model and the probit regression model are derived in parallel. A Gibbs sampler is further designed to integrate the posterior predictive distribution. The resulting Bayesian Poisson regression algorithm is applied to predicting the seasonal tropical cyclone activity.

Contents

1. Introduction . . . 244
   1.1. Concept of Bayesian inference . . . 244
   1.2. Bayesian analysis applied to climate data . . . 245
   1.3. Bayesian analysis applied to extreme climate events . . . 246

* Corresponding author.
E-mail address: chu@hawaii.edu (P.-S. Chu).

0169-8095/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.atmosres.2011.07.001

2. Data . . . 247
3. Bayesian change-point analysis for extreme climate events . . . 248
3.1. Hierarchical Bayesian model for extreme event count series . . . 248
3.2. Exact Bayesian inference under the hypothesis with only one change-point . . . 249
3.3. Bayesian inference under the hypothesis H_1 using MCMC method . . . 249
3.4. Bayesian change-point analysis using RJMCMC method . . . 249
3.5. Prior specification . . . 250
3.6. Case study for the Bayesian change-point analysis . . . 251
3.6.1. Exact Bayesian inference . . . 251
3.6.2. Bayesian inference using MCMC method . . . 251
3.6.3. Bayesian inference via using RJMCMC method . . . 252
4. Bayesian tropical cyclone track pattern Clustering . . . 254
4.1. Finite mixture Gaussian model for tropical cyclone track pattern clustering . . . 254
4.2. Bayesian inference for tropical cyclone track pattern clustering . . . 255
4.3. Typhoon track types over the WNP via using the Bayesian track path clustering . . . 256
5. Bayesian prediction for seasonal tropical cyclone activity . . . 257
5.1. Bayesian regression via the generalized Poisson regression model . . . 257
5.2. Bayesian classification via the probit regression model . . . 258
5.3. Case study . . . 259
6. Summary . . . 260
References . . . 261

1. Introduction

1.1. Concept of Bayesian inference

In principle, Bayesian analysis is grounded on a probabilistic generative model of a process. With the generative model, Bayes' theorem provides an approach to inferring one or more parameters in a process from the observed data, where the parameters are supposed to characterize the process of interest. In the Bayesian viewpoint, probability can be used to quantify degrees of belief of inference with given assumptions. Thus, Bayesian inference deals with uncertainty of unknown parameters or hypotheses of interest in probabilistic forms. Under a Bayesian framework, the unknown quantities are modeled as random variables, instead of constants or fixed values. This feature fits well with climate research because climate should not be considered stationary but rather as something that is always changing. When new information is obtained, prior knowledge about the unknown quantities of interest is revised accordingly. In this regard, Bayes' theorem provides a formal mechanism to revise or update prior beliefs in light of new data to yield posterior probability statements about the unknown parameters or hypotheses.

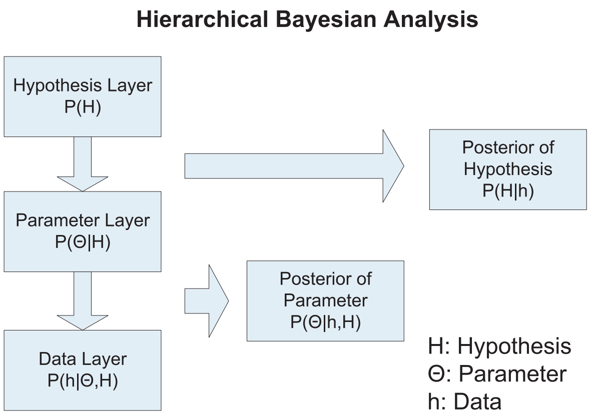

Fig. 1. A 3-layer hierarchical Bayesian analysis model.

The number of inference problems that can be tracked by Bayesian analysis is enormous and across a lot of research fields (e.g., Berger, 1985; MacKay, 2003; Trotta, 2008). The general paradigm of a Bayesian inference analysis can be sketched as the hierarchical flow chart displayed in Fig. 1. On the top of the Bayesian network is the hypothesis or model layer, which defines a hypothesis or model with its associated parameter set. Presumably, the observed data is sampled from this generative model. Before observing the data, there is some subjective belief of the hypothesis or model along with its associated parameter set, which is termed as "prior". Through the Bayes' theorem, one can update the degree of belief of a hypothesis and its relative parameters, yielding the "posterior" probability of the hypothesis and parameter set of interest. Assuming a hypothesis H is given and we denote the observed data by $\mathbf{h}$, the Bayesian formula to infer the parameter set $\boldsymbol{\theta}$ is given by:

$$P(\boldsymbol{\theta}|\mathbf{h},H) = \frac{P(\mathbf{h}|\boldsymbol{\theta},H)P(\boldsymbol{\theta}|H)}{P(\mathbf{h}|H)} = \frac{P(\mathbf{h}|\boldsymbol{\theta},H)P(\boldsymbol{\theta}|H)}{\int_{\boldsymbol{\theta}} P(\mathbf{h}|\boldsymbol{\theta},H)P(\boldsymbol{\theta}|H)d\boldsymbol{\theta}} \propto P(\mathbf{h}|\boldsymbol{\theta},H)P(\boldsymbol{\theta}|H). \quad (1)$$

If the hypothesis or model H is unanimously accepted, Eq. (1) provides the full solution for the inference problem. Note that, as the denominator $P(\mathbf{h}|H)$ does not contain any information on parameter $\boldsymbol{\theta}$, the likelihood term $P(\mathbf{h}|\boldsymbol{\theta},H)$ in Eq. (1) conveys all the "new" information obtained from the data.

Table 1
Raftery's guideline for interpreting the Bayes factor.

2lnB	Evidence for Bayesian model
0–2	Not worth more than a bare mention
2–6	Positive
6–10	Strong