Bayesian Inference from Continuously Arriving Informant Reports, with Application to Crisis Response*

Carter T. Butts† Fabio Leite‡

January 15, 2004; v0.3

Note: This is a draft document. Please do not cite without permission.

Abstract

Effective decision-making for crisis response depends upon the rapid integration of limited information from (possibly unreliable) human sources. Here, a Bayesian modeling framework is developed for inference from informant reports. Reports are assumed to arrive via a Poisson-like process, whose rates are dependent upon the (unknown) state of the world in addition to assorted covariates. A hierarchical modeling structure is used to represent error processes which vary based on informants' group memberships, with the possibility of multiple, overlapping memberships for each informant. Procedures are shown for sampling from the joint posterior distribution of the parameters, and for obtaining posterior predictive quantities.

Keywords: informant accuracy, hierarchical Bayesian models, event history analysis, change-point models, crisis response

1 Introduction

While sensor systems (USGS, 1999) and remote sensing (Kaiser et al., 2003; Kerle and Oppenheimer, 2002) can play a significant role in early warning,

*This material is based upon work supported by the National Science Foundation under award NSF ITR #0331707. We also acknowledge the input of the UCI RESCUE team into this research.

†Department of Sociology and Institute for Mathematical Behavioral Sciences, University of California-Irvine; SSPA 2145, Irvine, CA 92697; buttsc@uci.edu

‡Institue for Mathetmatical Behavioral Sciences, University of California-Irvine

situation assessment, and loss assessment for certain hazards (particularly those of a large-scale geological or weather-related nature), a great deal of information involving hazards such as tornadoes, fires, and severe traffic accidents is derived from human sources. When such an event occurs, informant reports from persons on the scene (e.g.,via 911 calls) may be the first indicators of trouble. To prevent the initial event from escalating into a disaster – and, generally, to enhance post-impact response – effective use of informant reports is a must.

Unfortunately for this effort, the accuracy of informant reports for a wide range of phenomena is generally acknowledged to be poor (Bernard et al., 1984). Biases of perception and recall lead to errors of both omission (false negatives) and commission (false positives), with an overall tendency towards altering reports in the direction of expected or familiar patterns (Freeman et al., 1987). (Such cognitive biases may in part account for the "apathy effect" noted both by practitioners (Auf der Heide, 1989) and by researchers in the field (Drabek, 1986).) To overcome this difficulty, researchers have turned to inferential systems which seek to model and/or compensate for errors in informant reports (Batchelder and Romney, 1988; Butts, 2003). Such modeling is of particular importance in the context of crisis events, where respondant behaviors may be especially sensitive to the anomalous nature of the situation (Drabek, 1986), and where it may be necessary to make decisions based on very limited data. By explicitly incorporating behavioral processes and background knowledge into a formal modeling framework, we can make the most of the information we have.

Here, we provide a general Bayesian model for informant reports which arrive in continuous time via a "spontaneous" reporting process (i.e., without direct inquiry on the part of the researcher). The objective of the modeling process is simultaneous inference for an event history (e.g., the unfolding of a multi-stage crisis event) and for informant accuracy. The approach used here can be considered a special case of the multiple changepoint model (e.g., Carlin et al., 1992; Raftery and Akman, 1986; Smith, 1975), in which unobservable alterations in the background environment (here assumed to be synonymous with the evolution of the crisis event) are inferred based on the sequence of incoming reports. The resulting posterior distributions can be used to characterize the current state of knowledge about an ongoing crisis (including the degree of uncertainty regarding said knowledge), thereby facilitating decision-making for post-impact response.

1.1 Notation

In the text which follows, the pdf (or pmf, where no danger of confusion exists) for a random variable x is denoted $p(x)$. The conditional density of x given y is denoted $p(x|y)$; distinctions between prior, likelihood, and posterior densities are made using conditional notation. The symbol $\sim$ (read: "is distributed as") is also used in some cases to indicate equivalence between distributions. Conventional distributions are denoted by abbreviated names, in Roman type, and will be defined where first introduced. Thus, $x \sim \mathrm{N}(\mu, \sigma^2)$ should be read as stating that x is distributed normally with parameters μ and σ^2.

2 States of the World, and Basic Modeling Assumptions

We conceive of the state of the world, $\mathcal{S}$, in terms of a series of n propositions which are either true (1) or false (0) at any given instant in time. Thus, at time t, we assume that $\mathcal{S}(t) \in \{0,1\}^n$. In general, the state of the world is unknown, although we will usually assume that we can identify some (possibly imaginary) time point t_0 at which $\mathcal{S}$ is known with certainty. Over time, the state of the world may change; that is, certain propositions may change state from 0 to 1, or 1 to 0. It will be assumed throughout the treatment which follows that all changes are irreversible for the time interval under study: thus, a proposition may move from 0 to 1, but it cannot then return to 0. Furthermore, we will generally assume that $\mathcal{S}$ is defined in such a way as to make 0 the base state, and 1 the transition state. This is not strictly necessary, but greatly simplifies exposition.

Given that $\mathcal{S}$ hypothetically begins in some known state – represented by the zero vector – at time zero, our primary objective is to infer the vector of transition times for the elements of $\mathcal{S}$ away from the ground state. We represent this transition vector by θ, where $\theta \in (0, \infty)^n$. For the moment, we defer any assumptions regarding the prior distribution of θ, as these will be developed subsequently. Of more immediate importance is the notion that the state of the world in any given instant – parameterized by θ – has a strong impact on informant reports. It is this issue to which we now turn.

3 Likelihood for Informant Reports

Let us consider some known population of m individuals, each of whom has the potential to observe $\mathcal{S}$. For the fixed time interval from t_α to t_ω (called the *reporting interval*), we allow each informant the option of reporting on the state of the world; each such report occurs as a discrete event, and for the time being we take reports to represent statements asserting a change in one of the n elements of $\mathcal{S}$. Each informant is assumed to report at most one change per element within the reporting interval.[1] The data thereby generated can be expressed by two matrices. First, we have a *report matrix*, $Y^R \in \{0,1\}^{m \times n}$, which is coded $Y^R_{ij} = 1$ iff the ith informant reported a change in element j. Second, we have a *timing matrix*, $Y^t \in (t_\alpha, t_\omega)^{m\times n}$, in which $Y^t_{ij} = t$ if $Y^R_{ij} = 1$ and the report occurred at time t. (For (i,j) such that $Y^R_{ij} = 0$, Y^t_{ij} is undefined.)

During the reporting interval, reports by informants are assumed to arrive via a Poisson-like process whose rate parameter varies as a function of the informant/element pair, the transition vector θ, and some arbitrary covariate set X. Conditional on these rates, all reporting events are taken to be independent of one another (that is, we presume that dependencies among informant reports can be modeled via the latent rate structure). For clarity in exposition, we divide the rate function into two regimes: λ^-, which represents the latent rate for the interval $0 \le t < \theta_j$; and λ^+, which represents the latent rate for the interval $\theta_j \le t \le t_\omega$. More formally:

$$\lambda^-(\theta, X, i, j) \equiv \lambda^-_{ij} \tag{1}$$

$$\lambda^+(\theta, X, i, j) \equiv \lambda^+_{ij} \tag{2}$$

where $\lambda^-_{ij}, \lambda^+_{ij} \in (0, \infty)$. While we will discuss the structure of λ in more detail below, let us note for the moment that it will generally be reasonable to presume that $\lambda^-(\theta, X, i, j) < \lambda^+(\theta, X, i, j)$ for any given parameter values. Where the two rates differ greatly, the timing of informant reports will contain a great deal of information about θ – this lies at the heart of the model which follows.

Given the above assumptions, the form for the likelihood of a single informant/element pair can be recognized as a fairly standard event history model (Blossfeld and Rohwer, 1995). Specifically, we have:

[1]That is, our data should be interepreted in terms of first reporting times – multiple reports from the same party are not considered here.

$$
\begin{aligned}
p\left(Y_{ij}^R, Y_{ij}^t \middle| \theta, \lambda^-, \lambda^+, t_\alpha, t_\omega, X\right) = & \left[S\big(\theta_j \big| \lambda_{ij}^-\big) S\big(t_\omega - \theta_j \big| \lambda_{ij}^+\big)\right]^{1-Y_{ij}^R} \\
& \times \left[\Big(h\big(Y_{ij}^t - t_\alpha \big| \lambda_{ij}^-\big) S\big(Y_{ij}^t - t_\alpha \big| \lambda_{ij}^-\big)\Big)^{I\left(Y_{ij}^t < \theta_j\right)}\right. \\
& \times \left.\Big(S\big(\theta_j \big| \lambda_{ij}^-\big) h\big(Y_{ij}^t - \theta_j \big| \lambda_{ij}^+\big) S\big(Y_{ij}^t - \theta_j \big| \lambda_{ij}^+\big)\Big)^{1-I\left(Y_{ij}^t < \theta_j\right)}\right]^{Y_{ij}^R}
\end{aligned} \tag{3}
$$

where I is the standard indicator function, h is the *hazard function* for the (i, j) pair, and S is the pair's *survival function*.[2] Utilizing the assumption of piecewise constant rates embodied in Equation 1 and 2, we observe that the waiting time distribution for informant reports must be (piecewise) exponential. For a given rate λ, the exponential density is given by

$$p(t|\lambda) = \lambda e^{-\lambda t}; \tag{4}$$

thus, the associated cumulative distribution function of the waiting time is

$$F(t|\lambda) = \int_0^t p(x|\lambda)\mathrm{d}x = 1 - e^{-\lambda t}. \tag{5}$$

To obtain the survival function, we simply apply the definition,

$$S(t|\lambda) = 1 - F(t|\lambda) = e^{-\lambda t}, \tag{6}$$

which then gives us the hazard function:

$$h(t|\lambda) = \frac{p(t|\lambda)}{S(t|\lambda)} = \lambda. \tag{7}$$

Substituting the above into Equation 3, we can then obtain the likelihood for a single informant/element pair under the piecewise constant rate model:

$$
\begin{aligned}
p\left(Y_{ij}^R, Y_{ij}^t \middle| \theta, \lambda^-, \lambda^+, t_\alpha, t_\omega, X\right) = & \left[e^{-\lambda_{ij}^+(t_\omega - \theta_j) - \lambda_{ij}^- \theta_j}\right]^{1-Y_{ij}^R} \\
\times & \left[\left(\lambda_{ij}^- e^{-\lambda_{ij}^-(Y_{ij}^t - t_\alpha)}\right)^{I(Y_{ij}^t < \theta_j)} \left(\lambda_{ij}^+ e^{-\lambda_{ij}^+(Y_{ij}^t - \theta_j) - \lambda_{ij}^- \theta_j}\right)^{1-I(Y_{ij}^t < \theta_j)}\right]^{Y_{ij}^R}
\end{aligned} \tag{8}
$$

[2] The hazard function reflects the limiting probability of a first (or, in this case, unique) state transition for an interval beginning at a specified time point, as the interval length approaches zero. Relatedly, the survival function represents the probability that no transition has occurred by a given point in time.

The joint likelihood of Y^R and Y^t then follows trivially from the conditional independence of informant/element pairs:

$$p\left(Y^R, Y^t \middle| \theta, \lambda^-, \lambda^+, t_\alpha, t_\omega, X\right) = \prod_{i=1}^{m} \prod_{j=1}^{n} p\left(Y^R_{ij}, Y^t_{ij} \middle| \theta, \lambda^-, \lambda^+, t_\alpha, t_\omega, X\right). \tag{9}$$

3.1 Reporting Rate Parameterization

While the event structure model of Equation 8 is reasonably straightforward, much of its substantive content is dependent on the structure of the latent rate functions. Although these functions could be chosen in any of a number of ways, our approach is to employ a log-linear framework. Specifically, we assume vectors of weights, β, and sufficient statistics, s, such that

$$\lambda^-\left(\theta, X, i, j\right) = e^{\sum_{k=1}^{\ell^-} \beta^-_k s^-(\theta,X,i,j)} \tag{10}$$

and

$$\lambda^+\left(\theta, X, i, j\right) = e^{\sum_{k=1}^{\ell^+} \beta^+_k s^+(\theta,X,i,j)}, \tag{11}$$

where $\beta^- \in \mathbb{R}^{\ell^+}$, $\beta^- \in \mathbb{R}^{\ell^-}$ $s^- : (\theta, X, i, j) \mapsto \mathbb{R}^{\ell^-}$, and $s^+ : (\theta, X, i, j) \mapsto \mathbb{R}^{\ell^+}$. Since the effects of the sufficient statistics are multiplicative in the reporting rates, the above implies that the associated weights can be interpreted as per a *proportional hazards model* (Cox, 1972) with the proviso that proportional effects hold only within transition regimes.[3] Thus, a dichotomous statistic, s_i, with a weight of β_i will multiply the base reporting rate by a factor of e^{β_i} where $s_i = 1$, having no effect otherwise.

The specific choice of statistics for the rate functions is a substantive matter, and will obviously depend on the problem at hand. It is notable, however, that many of the types of statistics which would normally be used in a fixed effects model would be appropriate here.[4] For instance, if informants can be labeled as belonging to a series of (possibly non-exclusive) classes, dichotomously coded variables for class membership could easily be employed as rate modifiers. Similarly, effects for particular elements (or classes of elements) of S may be appropriate, e.g., if it is believed that some

[3] I.e., the two respective intervals $[0, \theta_j)$ and $[\theta_j, t_\omega]$.

[4] Although these effects are considered to be random variables from a Bayesian point of view, they are nevertheless "fixed" in the sense that they are not modeled as arising from a population distribution. A hierarchical alternative is developed below.

elements are more easily noticed (or more subject to false reports) than others. Indeed, for $n > 1$, it is even possible to include fixed effects for each informant (reflecting individual differences in reporting rates); this is unlikely to be practical where m is large, however.

3.1.1 Hierarchical Forms for Rate Parameters

In many circumstances, it is natural to presume the some or all of the effects which contribute to the λ functions arise via a hierarchical process of some form or other. This is particularly likely to be the case where certain informant properties (e.g., group memberships) can be interpreted in terms of a series of subpopulations, such that each such population is in its turn drawn from some larger or more general category. In addition to being plausible on a priori grounds, imposition of hierarchical structure onto rate parameters can be used to concentrate the distribution of probability across models, thereby potentially increasing data efficiency.

The general approach for the use of such hierarchical forms is straightforward. For some set of weights, β^+, β^-, we express the joint density of the betas as

$$p(\beta^+, \beta^-) = p(\beta^+, \beta^- | \phi) p(\phi | \phi') p(\phi' | \phi'') \ldots, \qquad (12)$$

where $\phi, \phi', \phi'', \ldots$ reflect vectors of parameters at each respective hierarchical level. Note that – since this is a model for β rather than Y – the use of such a form does not alter the model likelihood, and indeed is more properly considered to be a constraint on the prior structure. (We treat the topic here, however, due to its substantive role.) Where successive levels reflect random population samples, a hierarchical normal model tends to suggest itself; other forms are possible, however, depending on what is known regarding the structure of the covariates.

4 Prior Structure

We have already touched on the issue of prior structure for the informant reporting model, vis a vis the use of hierarchical structure for reporting rate parameters. Here, we consider this matter more broadly. While the particular choice of priors for the informant reporting model must naturally be driven by the substantive knowledge of the problem at hand, some general guidelines can nevertheless be suggested. The treatment presented here thus focuses on basic heuristics for cases which are most likely to be encountered in practice.

Our first issue with respect to choice of parameter priors for the informant reporting model is the identification of a priori independence among parameters. Obviously where such independence may be safely assumed, it is useful to deploy it; we do not, however, wish to impose it where inappropriate. Where, then, can we generally presume independence to be a reasonable assumption? In many cases, we would suggest, such a division can be made between the elements of θ (which reflect the evolving state of the world) and the elements of β (which reflect the underlying mechanisms which govern the reporting process). Irrespective of what is being reported *on*, the mechanisms of the reporting *process* (the "laws of nature," so to speak) should remain the same. Note that such an assumption does not imply that the realized reporting *rates* (λ^+, λ^-) are a priori independent of θ, since these are indeed functions of θ – rather, it implies that whatever role θ plays in the reporting process, that role is independent of θ itself. (Note also that this assumption does not imply the *a posteriori* independence of θ and β, which is another matter entirely!) Given this, then, we may express the joint posterior with the decomposition

$$p(\theta, \beta^-, \beta^+) = p(\theta)p(\beta^-, \beta^+), \tag{13}$$

and then consider each factor in turn.

For the elementwise transition points (θ), the degree of dependence in the prior structure is predicated on the nature of the elements themselves. If, for instance, the ith element of S is present only if the jth element is present, then $p(\theta)$ must be such that $p(\theta_i \geq \theta_j) = 1$. Alternately, S may represent a series of disconnected propositions, in which case it may be reasonable to assume transition times for its elements to be independent. In this case, a natural prior model for θ would be a product of exponential densities, i.e.

$$p(\theta) = \prod_{i=1}^{n} \rho_j e^{-\rho_i \theta_i}. \tag{14}$$

Such a model includes the assumption that the a priori hazard of transition is constant and known, for each element of S. Where the constant hazard assumption is reasonable but it is difficult to identify a plausible rate, a hierarchical mixture over ρ is a logical choice. For instance, if the a priori transition rate for each element could be specified up to a log-normal distribution, a model such as

$$p(\theta) = \prod_{i=1}^{n} \left(\rho_i e^{-\rho_i \theta_i}\right) \text{Lnorm}\left(\rho_i \middle| \mu_i, \sigma_i^2\right) \tag{15}$$

would be appropriate. Alternately, a non-constant transition hazard may suggest a model based on Weibull or Gamma densities. The former, in particular, is a versatile density which is widely utilized in failure-rate research. Where independence of transitions cannot be assumed, but where dependencies between transitions are correlative rather than logical, the multivariate log-normal model may provide a reasonable approximating density. Again, the specific density to be used must be chosen based on substantive considerations.

Turning to the rate effect parameters, it should be emphasized that a priori independence cannot generally be assumed here. If nothing else, we should expect that covariates which enhance reporting accuracy will have opposing signs in β^- and β^+, which implies an a priori negative correlation between parameters for these effects. (Negative correlations may also be reasonable for some covariates which *reduce* reporting accuracy, e.g. by inflating λ^- while simultaneously reducing λ^+.) Covariates whose general tendencies are to increase or decrease overall reporting will, on the contrary, be associated with positive correlations between β^- and β^+. Similarly, complex correlation structure may be present within each β vector; the hierarchical models already considered provide one example of this. Given these considerations, a simple but fairly general model for the joint prior of the β parameters takes the form

$$p(\beta^-, \beta^+) \propto \left(1 + \frac{\left(\begin{bmatrix}\beta^- \\ \beta^+\end{bmatrix} - \mu\right)^T \Sigma^{-1} \left(\begin{bmatrix}\beta^- \\ \beta^+\end{bmatrix} - \mu\right)}{\nu}\right)^{-\frac{\nu+1}{2}} \tag{16}$$

$$\propto t_\nu \left(\begin{bmatrix}\beta^- \\ \beta^+\end{bmatrix} \middle| \mu, \Sigma \right), \tag{17}$$

i.e. the multivariate t density with degrees of freedom parameter ν, location vector μ, and scale matrix Σ. While this model allows for intuitive setting of location and scale, like the normal density, it also allows for the presence of heavy tails (an important consideration for robustness). Where a hierarchical form for effect parameter priors is desired, the above can be extended by the addition of hyperprior distributions for ν, μ, and/or Σ.

Whatever the final distributional form used, the importance of checking the joint parameter priors before model fitting should be stressed. In particular it is important that the prior predictive distributions of λ^+ and λ^- be examined, for various reporting scenarios. Since – for most applications – $\lambda^-_{ij} < \lambda^+_{ij}$ in the vast majority of cases, it should likewise be true that

$p(\lambda^-_{ij} \geq \lambda^+_{ij}) \ll 0.5$ for most models. If prior predictives for a given choice of prior are inconsistent with this expectation, they should be reexamined (and likely adjusted). Similarly, it is important to establish that the mean reporting rates produced by the prior model reasonably reflect the actual number of reports likely to be made by informants over the reporting interval. If, for instance, past experience has shown that only 5% of a given population will produce a report within a similar time frame to that of the present study, then priors which imply that something on the order of 50% of the current population will produce reports are immediately suspect. While it may be possible in some cases to spot such problems by examination of the prior parameters, the complexity of the reporting model suggests that this will not prove efficacious in most instances: where doubt remains, inspection of prior predictive draws (produced via statistical simulation) is the most straightforward way to identify potential difficulties.

5 Posterior Inference

Having constructed a likelihood for the incoming informant reports, and having chosen priors which are appropriate for the problem at hand, we are now ready to consider posterior inference. By Bayes' Theorem, the joint posterior must satisfy the relation

$$p\left(\theta, \beta^-, \beta^+ \mid Y^R, Y^t, X, t_\alpha, t_\omega\right) \propto p\left(Y^R, Y^t \middle| \theta, \beta^-, \beta^+, X, t_\alpha, t_\omega\right) p\left(\theta\right) p(\beta^-, \beta^+) \tag{18}$$

$$\propto q\left(\theta, \beta^-, \beta^+ \mid Y^R, Y^t, X, t_\alpha, t_\omega\right), \tag{19}$$

which (with appropriate substitutions) gives us the posterior density up to a constant of proportionality. Because of the difficulty of working with this density, we adopt the standard practice of conducting inference via the analysis of posterior draws; although we cannot simulate the distribution in question directly, we may closely approximate it by means of Markov Chain Monte Carlo (MCMC) methods (Gamerman, 1997).

A simple Metropolis algorithm for posterior simulation of the informant reporting model is shown in Algorithm 1. Note that this procedure begins by sampling θ and β^-, β^+ from their respective prior densities; while this is a not-unreasonable choice of starting point where the prior distribution is fairly informative, alternatives may be preferable in the event that the prior does not admit direct simulation. (So long as the starting point is within the posterior support, convergence is guaranteed in the limit.) After the initial point is chosen, simulation proceeds by drawing candidate moves which are accepted with probability $\min\left(1, \frac{q\left(\theta^{(i)},\beta^{-(i)},\beta^{+(i)}|Y^R,Y^t,X,t_\alpha,t_\omega\right)}{q\left(\theta^{(i-1)},\beta^{-(i-1)},\beta^{+(i-1)}|Y^R,Y^t,X,t_\alpha,t_\omega\right)}\right)$, where q is the unnormalized posterior density and $\cdot^{(i)}$ represents the ith simulation draw. Candidates for the simple model are drawn by perturbing each current draw by a lognormal factor (in the θ case) or a normal term (in the β case), with standard deviation ϵ set on a per-item basis. (Thus, the resulting Markov Chain is symmetric, as is required for the Metropolis algorithm.) In the limit, as the number of iterations grows, the set of draws thus obtained will converge to the joint posterior for the reporting model parameters.

Algorithm 1 Posterior Simulation for the Informant Reporting Model

procedure Draw from $\theta, \beta^-, \beta^+|Y^R, Y^t, X, t_\alpha, t_\omega$
Draw $\theta^{(1)}$ from $p(\theta)$
Draw $\beta^{-(1)}, \beta^{+(1)}$ from $p(\beta^-, \beta^+)$
$i := 2$
repeat
    for $j \in 1, \ldots, n$ **do**
        Draw $\theta_j^{(i)} \sim \text{Lnorm}(\ln \theta_j^{(i-1)}, \epsilon^2_{\theta_j})$
    end for
    for $j \in 1, \ldots, \ell$ **do**
        Draw $\beta_j^{-(i)} \sim \text{N}(\beta_j^{-(i)}, \epsilon^2_{\beta_j^-})$
        Draw $\beta_j^{+(i)} \sim \text{N}(\beta_j^{+(i)}, \epsilon^2_{\beta_j^+})$
    end for
    Draw $u \sim \text{U}(0, 1)$
    if $u > \frac{q\left(\theta^{(i)},\beta^{-(i)},\beta^{+(i)}|Y^R,Y^t,X,t_\alpha,t_\omega\right)}{q\left(\theta^{(i-1)},\beta^{-(i-1)},\beta^{+(i-1)}|Y^R,Y^t,X,t_\alpha,t_\omega\right)}$ **then**
        $\theta^{(i)} := \theta^{(i-1)}$
        $\beta^{-(i)} := \beta^{-(i-1)}$
        $\beta^{+(i)} := \beta^{+(i-1)}$
    end if
    $i := i + 1$
until $\theta^{(\cdot)}, \beta^{-(\cdot)}, \beta^{+(\cdot)} \sim \theta, \beta^-, \beta^+|Y^R, Y^t, X, t_\alpha, t_\omega$
return $\theta^{(\cdot)}, \beta^{-(\cdot)}, \beta^{+(\cdot)}$

5.1 Convergence Acceleration for the MCMC Algorithm

Although convergence of the MCMC is guaranteed by the irreducibility of the chain, this property holds only in the limit of infinite sample size – for realistically obtainable samples, slow convergence may result in very poor approximations to the posterior. In addition to the use of convergence diagnostics for the assessment of posterior draws (see Gamerman (1997) for a review), it may be desirable to attempt to accelerate chain convergence by various means. Perhaps the simplest and most obvious method of convergence for a chain such as that shown in Algorithm 1 would be the replacement of independent normal/lognormal candidate draws by joint draws with scale matrix chosen to approximate the posterior correlation structure of the parameters. While this last is not known, it can be estimated using either the Hessian matrix of the joint posterior, or a smaller set of approximate posterior draws taken using the initial (unaccelerated) algorithm. (Note that an overall scale reduction is recommended prior to using the estimated correlation structure, so as to avoid extreme moves.) Careful monitoring of convergence, together with judicious tuning of the proposal densities, can greatly improve model performance.

6 Conclusion

The process of drawing inferences from potentially error-prone informant reports is a difficult one, and one whose success depends critically upon one's ability to adequately model the report generation process. At the same time, such inferences are of central importance to crisis response. Here, a modeling framework has been shown for reports which arrive continuously at rates which vary as a function of the state of the world as well as exogenous covariates. Methods for conducting posterior inference via MCMC methods were shown, and suggestions have been provided vis a vis the setting of priors based on (pre-event) background knowledge. It is hoped that this work will serve to provide a useful "first step" towards model-based integration of human-derived information in crisis response settings, while further extending the larger literature on the problem of informant-based inference.

7 References

Auf der Heide, E. (1989). *Disaster Response: Principles of Preparation and Coordination*. Mosby, St. Louis, MI.

Batchelder, W. H. and Romney, A. K. (1988). Test theory without an answer key. *Psychometrika*, 53(1):71–92.

Bernard, H. R., Killworth, P., Kronenfeld, D., and Sailer, L. (1984). The problem of informant accuracy: The validity of retrospective data. *Annual Review of Anthropology*, 13:495–517.

Blossfeld, H. P. and Rohwer, G. (1995). *Techniques of Event History Modeling: New Approaches to Causal Analysis*. Lawrence Erlbaum and Associates, Mahwah, NJ.

Butts, C. T. (2003). Network inference, error, and informant (in)accuracy: A Bayesian approach. *Social Networks*, 25(2):103–140.

Carlin, H. P., Gefland, A. E., and Smith, A. F. M. (1992). Hierarchical Bayesian analysis of change-point problems. *Applied Statistics*, 41(2):389–405.

Cox, D. R. (1972). Regression models in life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 34:187–220.

Drabek, T. E. (1986). *Human System Responses to Disaster: An Inventory of Sociological Findings*. Springer-Verlag, New York.

Freeman, L. C., Romney, A. K., and Freeman, S. C. (1987). Cognitive structure and informant accuracy. *American Anthropologist*, 89:310–325.

Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall, London.

Kaiser, R., Spiegel, P. B., Henderson, A. K., and Gerber, M. L. (2003). The application of geographic information systems and global positioning systems in humanitarian emergencies: Lessons learned, programme implications, and future research. *Disasters*, 27(2):127–140.

Kerle, N. and Oppenheimer, C. (2002). Satellite remote sensing as a tool in lahar disaster management. *Disasters*, 26(2):140–160.

Raftery, A. E. and Akman, V. E. (1986). Bayesian analysis of a poisson process with a change-point. *Biometrika*, 73(1):85–89.

Smith, A. F. M. (1975). A Bayesian approach to inference about a change-point in a sequence of random variables. *Biometrika*, 62(2):407–416.

USGS (1999). *An assessment of seismic monitoring in the United States; requirement for an Advanced National Seismic System.* U.S. Geological Survey (USGS), Washington, D.C.