JOURNAL OF TIME SERIES ANALYSIS
J. Time. Ser. Anal. **37**: 837–850 (2016)
Published online 19 May 2016 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12197

ORIGINAL ARTICLE

BAYESIAN DECONVOLUTION OF SIGNALS OBSERVED ON ARRAYS

MING LIN,[a] ERIC A. SUESS,[b] ROBERT H. SHUMWAY[c] AND RONG CHEN[d*]

[a] *Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen, China*
[b] *Department of Statistics and Biostatistics, California State University, East Bay, Hayward, CA, USA*
[c] *Department of Statistics, University of California, Davis, Davis, CA, USA*
[d] *Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ, USA*

Time series data collected from arrays of seismometers are traditionally used to solve the core problems of detecting and estimating the waveform of a nuclear explosion or earthquake signal that propagates across the array. We consider here a parametric exponentially modulated autoregressive model. The signal is assumed to be convolved with random amplitudes following a Bernoulli normal mixture. It is shown to be potentially superior to the usual combination of narrow band filtering and beam forming. The approach is applied to analyzing series observed from an earthquake from Yunnan Province in China received by a seismic array in Kazakhstan.

Received 1 October 2014; Revised 25 March 2016; Accepted 2 April 2016

Keywords: Nonlinear models, Bayes, Markov chain Monte Carlo, deconvolution, seismic arrays, nuclear monitoring.

1. INTRODUCTION

One important aspect of the general problem of monitoring nuclear tests such as those that have occurred in North Korea, Pakistan and India pertains to inferences drawn directly from arrays of sensors or seismometers that detect seismic events such as earthquakes or mining or nuclear test explosions. Amplitudes and power spectra of the signal estimated from an array of sensors can be used directly to obtain an estimate for the magnitude of an event that relates directly to yield if the event happens to be an explosion. The signal arrival times are also critical because they are used to locate the origin of the event, another important parameter of interest. Furthermore, there are a number of shorter signals called ‘phases’ that arrive with different delays and correspond to different paths taken to the recording receivers by the different phases. Generally, there might be two arrivals associated with the body wave (P) and possibly two arrivals associated with surface waves (S). For further analysis, see Gibbons *et al.* (2011). The parametric formulation allows for this kind of behaviour in the model for the convolving functions. Hence, it is critical to develop the best estimators possible for the waveform of the underlying signal on the array.

To illustrate, consider Figure 1 that shows four of the nine channels recording an earthquake in Yunnan Province in China at the Makanche Array in Kazakhstan. The series are recorded at 20 points per second, leading to a folding frequency of 10 cycles per second (Hz). The primary energy in this signal is in the 0–3 Hz range. The signals arrive at different times $T_i, i = 1, 2, \ldots, N$, relative to some arbitrary start point, with the time delays $T_i = \boldsymbol{r}_i'\boldsymbol{\theta}$ giving the relation of the array coordinate vector $\boldsymbol{r}_i = (r_{i1}, r_{i2})'$ in km to a slowness parameter $\boldsymbol{\theta}' = (\theta_1, \theta_2)$ in second/km. The slowness parameter is directly related to the velocity and azimuth of a propagating plane wave (see Shumway *et al.* (2008) for details). The coordinates for the nine vertical recording channels of the Makanche are shown in Figure 2.

* Correspondence to: Rong Chen, Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ, USA. E-mail: rongchen@stat.rutgers.edu

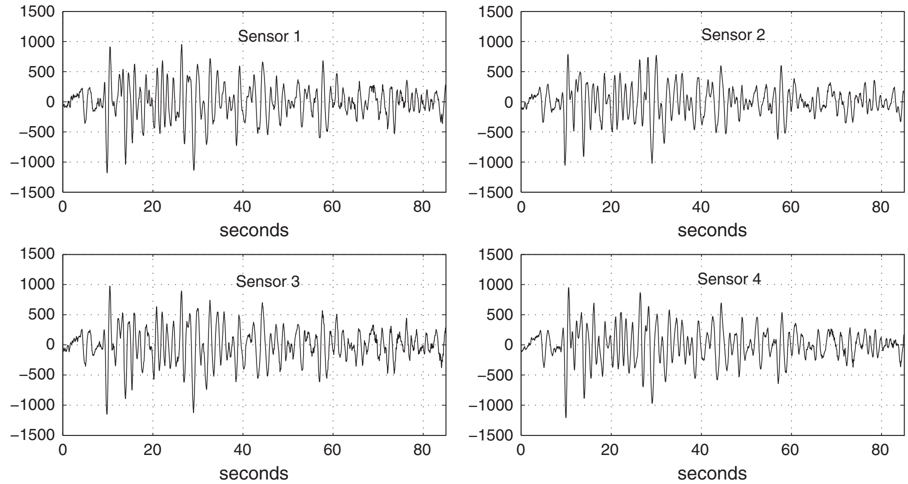

Figure 1. First four (of nine) vertical channels recording an earthquake in Yunnan Province in China, February 9, 2004 at the Makanche Array in Kazakhstan

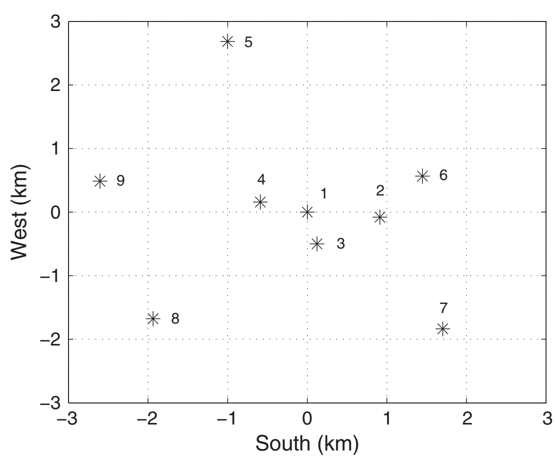

Figure 2. Array coordinates $\boldsymbol{r} = (r_1, r_2)'$ in km for the nine channels at the Makanche Array in Kazakhstan

Given an underlying signal $s(t)$, assumed to be fixed and unknown, we model the arrival at sensor i by

$$y_i(t) = s\left(t - \boldsymbol{r}_i'\boldsymbol{\theta}\right) + \varepsilon_i(t), \tag{1}$$

$i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, n$, where $\varepsilon_i(t)$ is stationarily correlated over time but uncorrelated between sensors. In the frequency domain, one can obtain an F-statistic as in Shumway (1971) for each $\boldsymbol{\theta}$. Finding the maximum as a function of slowness $\boldsymbol{\theta}$, we convert to velocity and azimuth to obtain direction. The maximum likelihood estimators for the fixed velocity and azimuth also lead to the best linear unbiased estimates (Shumway and Dean, 1968) for the signal. In this case, the best estimator for the signal can be approximated: first by filtering the data into the signal band, and then by delaying and averaging, that is, we compute the estimator

$$\hat{s}(t) = N^{-1} \sum_{i=1}^{N} y_i\left(t + \boldsymbol{r}_i'\boldsymbol{\theta}\right). \tag{2}$$

  J. Time. Ser. Anal. **37**: 837–850 (2016)
DOI: 10.1111/jtsa.12197

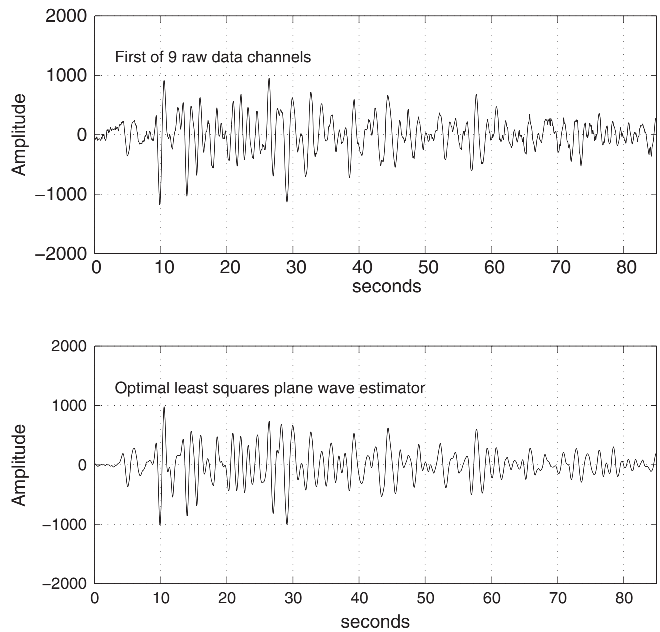

Figure 3. The estimated signal of the Yunnan earthquake data computed by equation (2) compared with the waveform at the first sensor

In the geophysical literature, this is called 'beam-forming,' and it depends critically on the assumption that the signal is a plane wave propagating across the array with a fixed velocity and azimuth (direction in degrees measured clockwise from north).

For details on this computation and extensions to multiple signal and noise sources, see Shumway *et al.* (2008). Figure 3 shows the estimated signal compared with the raw signal observed at the first sensor, and we find improvements in signal-to-noise ratio, particularly when the signal arrives; the first arrival is enhanced, and a clear representation of the first two cycles can be important for discrimination between earthquakes and nuclear explosions.

The aforementioned procedure can be relatively effective, but it still depends on the special plane wave model. Furthermore, it depends on being able to assume an exact replica for the signal at each element of the array. It might be effective to generalize it to a model that does not depend on plane wave propagation and would lead to alternatives to beam-forming. The rest of the article is organized as follows. Section 2 proposes a generalized model, in which the observations on each sensor consists of several waves of the underlying signal $s(t)$ that arrive at different time points and have different amplitudes. The generalization is achieved at the cost of more detailed processing. In Sections 3 and 4, we employ a Bayesian paradigm to extract the signal and estimate the unknown parameters. The approach is implemented via Markov chain Monte Carlo (MCMC) (Gilks *et al.*, 1995; Liu, 2001; Robert and Casella, 1999) and provides a powerful means for dealing with high dimensional nonlinear problems. A similar approach has been used in Cheng *et al.* (1996). We adopt their framework and propose some modifications to improve its performance. Section 5 applies the new model to analysing the earthquake signal from Yunnan Province in China, observed by a seismic array in Kazakhstan. Section 6 concludes.

2. A BAYESIAN FORMULATION OF THE MODEL

It is clear that the plane wave model currently in use may not be the best one for producing an estimator for the underlying signal on the array. There can be nonlinear perturbations in the idealized linear model that will cause significant changes in the estimated underlying signal. Therefore, we propose extending the model in equation (1) to one that allows departures from the idealized signal model. Consider the generalization

DOI: 10.1111/jtsa.12197

$$y_i(t) = \sum_{j=-v}^{m} a_{i,j} s(t-j) + \varepsilon_i(t), \tag{3}$$

where $a_{i,j}$ is the amplitude multiplier for each delay on each sensor and $\varepsilon_i(t)$ are independent zero-mean Gaussian white noise processes with variance σ_y^2. The model allows for $y_i(t)$ to be expressed as a sum consisting of a convolution of unknown scale amplitudes a_{ij} and the signal with additive noise. It is clear that the plane wave model (1) can be obtained by letting $a_{i,j} = \delta\left(j - \boldsymbol{r}_i'\boldsymbol{\theta}\right)$, where $\delta(\cdot)$ denotes Kronecker's delta function. The advantage of allowing more general amplitude factors in equation (3) is that it allows the signal contributions at each lag to be modified in the model, which potentially can adjust for perturbations because of scattering and other local variations in wave propagation. Many observed series have reflections because of mixtures of phases, that is, similar echoes distribute over the range of the primary signal. Such a model can also be used to describe acoustic echoes in telecommunication (Murano *et al.*, 1990; Benesty *et al.*, 2001; Vaseghi, 2006), in which $y_i(t)$ is the echo, $s(t)$ is the source of the echo and $a_{i,j}$ represents the echo path. In the problem of echo cancellation, $s(t)$ is observable, the task is to estimate the echo path then subtract the echo from the received signal. Suess *et al.* (1998) used a similar model for analysing rippled-fired mining explosion signals and used it for discriminating mining explosions and earthquakes. In the mobile communication literature for recovering convolutively mixed sources, Andrieu *et al.* (1998) studied a similar model with $\{s(t)\}$ being a sequence of (multichannel) discrete signals. Godsill and Andrieu (1999) proposed a general model with no restriction on the mixing structure. In our model, a strong restriction is imposed on the mixing coefficients $a_{i,j}$ specifically to model the signal reflection structure of seismic signals travelling through earth medium.

In order to produce a sensible parametric representation for the signal process, it is reasonable to suspect from Figure 1 that the roughly periodic behaviour might be fit well by an autoregressive (AR) model except for the indication that there is a decay in the amplitudes. This can be even more pronounced in more impulsive earthquakes and in explosions. As a simplified model, we allowed for an exponentially modulated AR model of order p, written as

$$s(t) = e^{-dt} x(t),$$

where

$$x(t) = \sum_{k=1}^{p} \phi_k x(t-k) + \epsilon(t) \tag{4}$$

is an AR(p) model and d denotes the decay parameter associated with the exponential modulations. For the Chinese earthquake signal and other earthquake events we have analysed, the roots of the characteristic polynomial is close to the unit circle, and this will be noted later on. We assume that $\epsilon(t)$ is a zero-mean Gaussian white noise process with variance σ_x^2.

The amplitudes $a_{i,j}, i = 1, 2, \ldots, N,\ j = -v, -v+1, \ldots, m$, are critical to the modelling process because they clearly control the extent to which the signal loads on each sensor. The analogy between this signal model and what might be termed a random coefficient factor model is suggested where the signal process plays the role of factors. One could consider a set of coefficients that generates a linear time invariant filter and transforms to a frequency domain factor analysis (Shumway and Der, 1985). However, we choose here to keep the model strictly in the time domain and construct a probability distribution for the amplitudes. The main idea is that the amplitudes should be zero when a particular delay is not important in determining the output at an observed sensor. The restriction essentially forces a pattern that one essentially looks for when doing factor analysis.

Specifically, we assume that the amplitude is zero with probability η and otherwise follows a normal distribution with mean μ_a and variance σ_a^2. To avoid ambiguity, we let the delay of the first arriving signal on the first sensor

DOI: 10.1111/jtsa.12197

be zero and the corresponding amplitude multiplier be 1, that is, $a_{1,j} = 0$ for all $j < 0$ and $a_{1,0} = 1$. In order to prevent catching too weak signal as the first arriving signal on the first sensor, we also restrict $|a_{1,j}| < 1.5$ for $j > 0$.

3. THE PRIOR AND POSTERIOR DISTRIBUTIONS

In this model, the variance σ_y^2 of the observation noise $\varepsilon_i(t)$ can be estimated from the background noise received by the sensors before the signal arrives. In addition, the decay parameter d can be estimated from the decay of the observed series. Hence, we will assume these two parameters are known in the following discussion. In the earthquake data presented in Figure 1, we let $\tau_y = 1/\sigma_y^2 = 1/62^2$ and $d = 0.00070$.

The rest of unknown parameters in the model are $\Theta = \left\{\tau_x = 1/\sigma_x^2, \boldsymbol{\phi}, A, \eta\right\}$, where $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_p)'$, the AR parameters in equation (4), $A = \{a_{i,j}, i = 1, \cdots, N, j = -v, -v+1, \cdots, m\}$, the set of amplitude, and $\eta = P(a_{i,j} = 0)$.

We use the following prior distributions for the unknown parameters in the model.

$$
\begin{aligned}
&P(\tau_x) \sim \text{Gamma}(\alpha_x, \lambda_x), \quad P(\boldsymbol{\phi}) \sim N(\mu_\phi, \Sigma_\phi), \quad P(\eta) \sim \text{Beta}(\beta_1, \beta_2), \\
&a_{1,j} = 0 \ \text{ for } j < 0\,, \quad a_{1,0} = 1, \\
&P(a_{1,j} \mid \eta) = \eta\,\delta(a_{1,j}) + (1-\eta)c\ \Phi\left(a_{1,j}; \mu_a, \sigma_a^2\right) I(|a_{1,j}| < 1.5) \ \text{ for } j > 0\,, \\
&P(a_{i,j} \mid \eta) = \eta\,\delta(a_{i,j}) + (1-\eta)\ \Phi\left(a_{i,j}; \mu_a, \sigma_a^2\right) \ \text{ for } i \neq 1,
\end{aligned}
$$

where $\Phi(a; \mu, \Sigma)$ denotes the normal density function with mean μ and variance Σ, evaluated at a. $I(\cdot)$ is the indicator function, c is the normalizing constant. For the signal $\boldsymbol{X} = \{x(t), t = -m+1, -m+2, \cdots, n+v\}$, we assume

$$
P(x(-m+1), \cdots, x(-m+p)) \sim N(0, \Sigma_0),
$$

and $x(t)$ follows equation (4) for $t > -m + p$.

Let $\boldsymbol{Y} = \{y_i(t), i = 1, \cdots, N, t = 1, \cdots, n\}$ be the observations, the posterior distribution of the parameters, and the signal can be written as

$$
\begin{aligned}
P(\boldsymbol{X}, \Theta \mid \boldsymbol{Y}) &\propto P(\boldsymbol{Y}, \boldsymbol{X}, \Theta) \\
&= P(\tau_x) P(\boldsymbol{\phi}) P(\boldsymbol{X} \mid \tau_x, \boldsymbol{\phi})\, P(\eta) P(A \mid \eta)\, P(\boldsymbol{Y} \mid \boldsymbol{X}, A) \\
&= P(\tau_x) P(\boldsymbol{\phi}) P\left(x(-m+1), \cdots, x(-m+p)\right) \\
&\quad \times \prod_{t=-m+p+1}^{v+n} \sqrt{\frac{\tau_x}{2\pi}} \exp\left\{-0.5\tau_x \left(x(t) - \sum_{j=1}^{p} \phi_j x(t-j)\right)^2\right\} \\
&\quad \times P(\eta) \prod_{j=1,\cdots,m} P(a_{1,j} \mid \eta) \times \prod_{i=2,\cdots,N, j=-v,\cdots,m} P(a_{i,j} \mid \eta) \\
&\quad \times \prod_{i=1,\cdots,N, t=1,\cdots,n} \sqrt{\frac{\tau_y}{2\pi}} \exp\left\{-0.5\tau_y \left(y_i(t) - \sum_{j=-v}^{m} a_{i,j} s(t-j)\right)^2\right\},
\end{aligned} \tag{5}
$$

where $s(t) = e^{-dt} x(t)$.

DOI: 10.1111/jtsa.12197

The prior distributions used here are quite standard. The hyper-parameters (α_x, λ_x) control the prior knowledge of the underlying signal strength, and (β_1, β_2) reflects prior knowledge on the probability $\eta = P(a_{i,j} = 0)$, the 'sparseness' of reflection and composition of the delayed versions of the underlying signal. As posterior distributions are often sensitive to the selection of prior distributions, especially for nonlinear models with a relatively small sample size, caution should be exercised and prior sensitivity should be tested. With no strong prior knowledge, one would choose the prior as non-informative as possible.

4. BAYESIAN INFERENCE WITH MARKOV CHAIN MONTE CARLO METHOD

Markov chain Monte Carlo is a powerful tool for solving problems in high-dimensional space. It has been applied successfully in various problems in statistics (Gelfand and Smith, 1990; Liu and Sabatti, 2000), physics (Goodman and Sokal, 1989; Marinari and Parisi, 1992), bioinformatics (Lawrence *et al.*, 1993; Liu, 1994), signal processing (Chen *et al.*, 2002; Lee *et al.*, 1995; Winkler, 1995), economics (Chib *et al.*, 2006; Chib and Ergashev, 2009; Verhofen, 2005) and other fields. In MCMC, a Markov chain $\left(\boldsymbol{X}^{(l)}, \Theta^{(l)}\right)$, $l = 1, \cdots, L$, is generated from a transition kernel whose stationary distribution is $P(\boldsymbol{X}, \Theta \mid \boldsymbol{Y})$. Then, for any function $h(\boldsymbol{X}, \Theta)$, under regularity conditions, $L^{-1} \sum_{l=1}^{L} h\left(\boldsymbol{X}^{(l)}, \Theta^{(l)}\right)$ converges to $E(h(\boldsymbol{X}, \Theta)|\boldsymbol{Y})$ as L tends to infinity.

In the following, we provide detailed construction of the Markov chain that updates $\left(\boldsymbol{X}^{(l)}, \Theta^{(l)}\right)$ from $\left(\boldsymbol{X}^{(l-1)}, \Theta^{(l-1)}\right)$. The general construction is a Gibbs sampler (Casella and George, 1992) with several modifications. The updating (sampling) is designed to cycle through the parameters one group at a time, conditional on the rest of the most recently sampled parameters. A complete cycle of updating results in a new sample of $\left(\boldsymbol{X}^{(l)}, \Theta^{(l)}\right)$. To simplify notation, we always use $P\left(\theta_i|\boldsymbol{Y}, \boldsymbol{X}^{(l)}, \Theta_{-}^{(l)}\right)$ to denote the distribution of the parameter θ_i conditional on all the rest of the most recent samples of parameters in the l-th iteration; some of them have been updated in this iteration; some of them are the samples from the previous iteration and yet to be updated.

4.1. Updating the Parameters τ_x, $\boldsymbol{\phi}$, η and A

We use the standard conditional distribution to update these parameters.

(1) Draw the precision parameter $\tau_x^{(l)}$ from $P\left(\tau_x|\boldsymbol{Y}, \boldsymbol{X}^{(l-1)}, \Theta_{-}^{(l)}\right) \sim \text{Gamma}\left(\alpha_x^*, \lambda_x^*\right)$, where $\alpha_x^* = \alpha_x + (m+n+v-p)/2$ and

$$\lambda_x^* = \lambda_x + \frac{1}{2} \sum_{t=-m+p+1}^{v+n} \left[x^{(l)}(t) - \sum_{j=1}^{p} \phi_j^{(l)} x^{(l)}(t-j) \right]^2 .$$

(2) Draw the AR model coefficients $\boldsymbol{\phi}^{(l)}$ from $P\left(\boldsymbol{\phi}|\boldsymbol{Y}, \boldsymbol{X}^{(l-1)}, \Theta_{-}^{(l)}\right) \sim N\left(\mu_\phi^*, \Sigma_\phi^*\right)$, where $\mu_\phi^* = \Sigma_\phi^*\left(\boldsymbol{b} + \Sigma_\phi^{-1}\mu_\phi\right)$ and $\Sigma_\phi^* = (C + \Sigma_\phi)^{-1}$. Here, $\boldsymbol{b}$ is a $p \times 1$ vector with components

$$b(i) = \tau_x^{(l)} \sum_{k=-m+p+1}^{n+v} x^{(l)}(k) x^{(l)}(k-j),$$

and $\boldsymbol{C}$ is a $p \times p$ matrix with elements given by

$$C(i,j) = \tau_x^{(l)} \sum_{k=-m+p+1}^{n+v} x^{(l)}(k-i) x^{(l)}(k-j).$$

  J. Time. Ser. Anal. **37**: 837–850 (2016)
DOI: 10.1111/jtsa.12197

(3) Draw the multipliers $a_{i,j}^{(l)}$, $i = 2, \cdots, q$, $j = -v, \cdots, m$, from

$$P\left(a_{i,j}|\boldsymbol{Y}, \boldsymbol{X}^{(l-1)}, \Theta_{-}^{(l)}\right) = \eta^* \delta(a_{i,j}) + (1-\eta^*)\Phi\left(a_{i,j}; \mu_a^*, \sigma_a^{*2}\right),$$

where

$$\mu_a^* = \sigma_a^{*2}\left(\frac{\mu_a}{\sigma_a^2} + \tau_y \sum_{t=1}^{n} \tilde{\varepsilon}_i^{(l)}(t) s^{(l)}(t-j)\right),$$
$$\sigma_a^{*2} = \left[\frac{1}{\sigma_a^2} + \tau_y \sum_{t=1}^{n} \left(s^{(l)}(t-j)\right)^2\right]^{-1},$$

and

$$\eta^* = \frac{\eta}{\eta + (1-\eta) d_{i,j}}, \quad \text{where } d_{i,j} = \frac{\sigma_a^*}{\sigma_a} \exp\left\{\frac{\mu_a^{*2}}{2\sigma_a^{*2}} - \frac{\mu_a^2}{2\sigma_a^2}\right\}.$$

Here, $\tilde{\varepsilon}_i(t) = y_i(t) - \sum_{k \neq j} a_{i,k}^{(l)} s^{(l)}(t-k)$. For $i = 1$, $a_{1,j}^{(l)}$, $j = 1, \cdots, m$, is drawn from

$$P\left(a_{1,j}|\boldsymbol{Y}, \boldsymbol{X}^{(l-1)}, \Theta_{-}^{(l)}\right) = \bar{\eta}^* \delta(a_{1,j}) + \left(1 - \bar{\eta}^*\right) c^* \,\Phi\left(a_{1,j}; \mu_a^*, \sigma_a^{*2}\right) I(|a_{1,j}| < 1.5),$$

where c^* is the normalizing constant for the truncated normal distribution, and

$$\bar{\eta}^* = \frac{\eta}{\eta + (1-\eta)\bar{d}_{i,j}} \text{ with } \bar{d}_{i,j} = \frac{c\,\sigma_a^*}{c^* \sigma_a} \exp\left\{\frac{\mu_a^{*2}}{2\sigma_a^{*2}} - \frac{\mu_a^2}{2\sigma_a^2}\right\}.$$

In this step, each $a_{i,j}^{(l)}$ is updated one by one sequentially.

(4) Draw $\eta^{(l)}$ from $P\left(\eta|\boldsymbol{Y}, \boldsymbol{X}^{(l-1)}, \Theta_{-}^{(l)}\right) \sim \text{Beta}\left(\eta_1^*, \eta_2^*\right)$, where $\eta_1^* = \eta_1 + (m+v+1)N - (m+1) - n_a$, $\eta_2^* = \eta_2 + n_a$. Here, n_a is the number of nonzero unfixed $a_{i,j}^{(l)}$.

The aforementioned steps complete the update for τ_x, $\boldsymbol{\phi}$, η and A. It is a standard implementation of the Gibbs sampler.

4.2. Updating the Signal X

Although one can use the standard Gibbs sampler to update each individual $x(t)$ conditioned on the rest of $\boldsymbol{X}$ and the other parameters, its mixing rate is slow. In fact, given the other parameters, the signal $\boldsymbol{X}$ can be embedded in the dynamic linear model with Gaussian innovations. Hence, the forward-filtering backward-sampling algorithm of Früwirth-Schnatter (1994) and Carter and Kohn (1994) can be used to update $\boldsymbol{X}$ as a block from $P\left(\boldsymbol{X} \mid \boldsymbol{Y}, \Theta^{(l)}\right)$. Such a procedure has a higher mixing rate and is more efficient.

Specifically, given parameters τ_x, $\boldsymbol{\phi}$, η and A, we rewrite the model in the form of dynamic linear model (Chen and Liu, 2000). Stacking the signal in the vector $\boldsymbol{x}(t) = (x(t+v), x(t+v-1), \cdots, x(t-m))'$, the data in the vector $\boldsymbol{y}(t) = (y_1(t), \cdots, y_N(t))'$ and the observe noise in the vector $\boldsymbol{\varepsilon}(t) = (\varepsilon_1(t), \cdots, \varepsilon_N(t))'$, we have

$$\begin{aligned} \boldsymbol{x}(t) &= \boldsymbol{F}\boldsymbol{x}(t-1) + \boldsymbol{B}\,\epsilon(t+v), \\ \boldsymbol{y}(t) &= \boldsymbol{H}(t)\boldsymbol{x}(t) + \boldsymbol{\varepsilon}(t), \end{aligned} \tag{6}$$

where $\boldsymbol{F}$ is a $(m+v+1) \times (m+v+1)$ matrix, with the first row being $(\phi_1, \cdots, \phi_p, 0, \cdots, 0)$, the first lower diagonal elements all being 1's and the rest elements are all 0's. The vector $\boldsymbol{B} = (1, 0, \ldots, 0)'$ is a $(m+v+1) \times 1$

DOI: 10.1111/jtsa.12197

vector and $\boldsymbol{H}(t) = \left\{\boldsymbol{e}^{-d(t-j)} a_{i,j}\right\}_{N\times(m+v+1)}$, $i = 1, \ldots, N$, $j = -v, \ldots, m$. Let $\boldsymbol{Y}(t) = \{\boldsymbol{y}(l),\ l = 1, \cdots, t\}$ if $t > 0$ and let $\boldsymbol{Y}(t) = \varnothing$ if $t \le 0$, then $\boldsymbol{X}^{(l)}$ can be sampled from $P\left(\boldsymbol{X} \mid \boldsymbol{Y}, \Theta^{(l)}\right)$ as follows.

(1) Use Kalman recursions to calculate $P\left(\boldsymbol{x}(t)|\boldsymbol{Y}(t), \Theta^{(l)}\right) \sim N(\mu_x(t), \Sigma_x(t))$ for $t = 1, \ldots, n$.

(i) Let $x(t) \equiv 0$ for $t \le -m$. When $t = -m - v + p$, $\boldsymbol{x}(t) = (x(-m+p), \cdots, x(-m+1), 0, \cdots, 0)'$. Since $-m - v + p < 0$, we have

$$
\begin{aligned}
P(\boldsymbol{x}(-m-v+p) \mid \boldsymbol{Y}\left(-m-v+p), \Theta^{(l)}\right) &= P\left(\boldsymbol{x}(-m-v+p) \mid \Theta^{(l)}\right) \\
&\sim N(\mu_x(-m-v+p), \Sigma_x(-m-v+p)),
\end{aligned}
$$

where $\Sigma_x(-m-v+p)$ is a $(m+v+1)\times(m+v+1)$ matrix with the prior Σ_0 on the top left corner block and zeros for all other elements, $\mu_x(-m-v+p) = (0, \ldots, 0)'$ is a $(m+v+1)\times 1$ zero vector.

(ii) For $t = -m - v + p + 1, \cdots, 0$, we have

$$
P\left(\boldsymbol{x}(t) \mid \boldsymbol{Y}(t), \Theta^{(l)}\right) = P\left(\boldsymbol{x}(t) \mid \Theta^{(l)}\right) \sim N(\mu_x(t), \Sigma_x(t)),
$$

where $\mu_x(t) = \boldsymbol{F}\mu_x(t-1)$ and $\Sigma_x(t) = \boldsymbol{F}\Sigma_x(t-1)\boldsymbol{F}' + \boldsymbol{B}\boldsymbol{B}'/\tau_x^{(l)}$.

(iii) For $t = 1, \cdots, n$, according to model (6) and the Kalman filter, we have

$$
P\left(\boldsymbol{x}(t) \mid \boldsymbol{Y}(t), \Theta^{(l)}\right) \sim N(\mu_x(t), \Sigma_x(t))
$$

with

$$
\begin{aligned}
\mu_x(t) &= \boldsymbol{F}\mu_x(t-1) + \Sigma_{12}(t)[\Sigma_{22}(t)]^{-1}(\boldsymbol{y}(t) - \boldsymbol{H}(t)\boldsymbol{F}\mu_x(t-1)), \\
\Sigma_x(t) &= \Sigma_{11}(t) - \Sigma_{12}(t)[\Sigma_{22}(t)]^{-1}\Sigma_{12}'(t).
\end{aligned}
$$

where $\Sigma_{11}(t) = \boldsymbol{F}\Sigma_x(t-1)\boldsymbol{F}' + \boldsymbol{B}\boldsymbol{B}'/\tau_x^{(l)}$, $\Sigma_{12}(t) = \Sigma_{21}'(t) = \Sigma_{11}(t)\boldsymbol{H}'(t)$, and $\Sigma_{22} = \boldsymbol{H}(t)\Sigma_{12} + I_N/\tau_y$, I_N is the $N \times N$ identity matrix.

(2) Draw $\boldsymbol{x}^{(l)}(n)$ from $P\left(\boldsymbol{x}(n)|\boldsymbol{Y}(n), \Theta^{(l)}\right)$, then generate $\boldsymbol{x}^{(l)}(n-1), \cdots, \boldsymbol{x}^{(l)}(1)$ recursively from

$$
\begin{aligned}
&P\left(\boldsymbol{x}(t)|\boldsymbol{x}^{(l)}(t+1), \ldots, \boldsymbol{x}^{(l)}(n), \boldsymbol{Y}(n), \Theta^{(l)}\right) \\
&\qquad \propto P\left(\boldsymbol{x}(t) \mid \boldsymbol{Y}(t), \Theta^{(l)}\right) P\left(\boldsymbol{x}^{(l)}(t+1) \mid \boldsymbol{x}(t), \Theta^{(l)}\right).
\end{aligned}
$$

Noting that only one element in $\boldsymbol{x}(t)$ is not in $\boldsymbol{x}(t+1)$, we only need to draw $x^{(l)}(t-m)$ to complete $\boldsymbol{x}^{(l)}(t)$. Partition vector $\mu_x(t) = (\boldsymbol{g}_1(t), g_2(t))'$ into an $(m+v)\times 1$ vector $\boldsymbol{g}_1(t)$ and a scalar $g_2(t)$ and denote the corresponding partition of $\Sigma_x(t)$ as

$$
\Sigma_x(t) = \begin{pmatrix} G_{11}(t) & G_{12}(t) \\ G_{21}(t) & G_{22}(t) \end{pmatrix}.
$$

Because $P\left(\boldsymbol{x}^{(l)}(t+1) \mid \boldsymbol{x}(t), \Theta^{(l)}\right)$ does not depend on $x(t-m)$ (usually $m + v > p$), we draw $x^{(l)}(t-m)$ from

$$
P\left(x(t-m) \mid x^{(l)}(t-m+1), \cdots, x^{(l)}(t+v), \boldsymbol{Y}(t), \Theta^{(l)}\right) \sim N\left(\mu, \sigma^2\right),
$$

where

$$
\begin{aligned}
\mu &= g_2(t) + G_{21}(t)[G_{11}(t)]^{-1}[(x(t+v), \cdots, x(t-m+1))' - \boldsymbol{\gamma}_2(t)], \\
\sigma^2 &= G_{22}(t) - G_{21}(t)[G_{11}(t)]^{-1}G_{12}(t).
\end{aligned}
$$

DOI: 10.1111/jtsa.12197

4.3. A Metropolis–Hasting Step

The aforementioned steps complete a Gibbs sampler implementation for updating $\left(\boldsymbol{X}^{(l)}, \Theta^{(l)}\right)$. Here, we add a Metropolis–Hasting step to improve the performance and try to avoid the Markov chain getting stuck in local mode.

(1) Move the amplitude set A to a new amplitude set $\tilde{A}$ in the following manner: randomly choose a non-zero $a_{i,j}$ and interchange the values of $a_{i,j}$ and $a_{i,J}$, where $J = j - 1$ or $J = j + 1$ with probability 0.5. Here, $a_{i,j}$ and $a_{i,J}$ need to be those unfixed amplitudes.
(2) Use the same method as in Section 4.2 to draw a new signal set $\tilde{\boldsymbol{X}}$.
(3) Decide whether to accept $\left(\tilde{\boldsymbol{X}}, \tilde{A}\right)$ as new $(\boldsymbol{X}, A)$. Let $\tilde{\Theta} = \left\{\eta, \tau_x, \boldsymbol{\phi}, \tilde{A}\right\}$, then $P\left(\tilde{\Theta}\right) = P(\Theta)$. We accept $\left(\tilde{X}, \tilde{\Theta}\right)$ with probability

$$\begin{aligned}\min\left\{1, \frac{P\left(\tilde{\Theta}, \tilde{\boldsymbol{X}} \mid \boldsymbol{Y}\right) P(\boldsymbol{X} \mid \boldsymbol{Y}, \Theta)}{P(\Theta, \boldsymbol{X} \mid \boldsymbol{Y}) P\left(\tilde{\boldsymbol{X}} \mid \boldsymbol{Y}, \tilde{\Theta}\right)}\right\} &= \min\left\{1, \frac{P\left(\boldsymbol{Y}, \tilde{\Theta}\right)}{P(\boldsymbol{Y}, \Theta)}\right\} \\ &= \min\left\{1, \frac{\prod_{t=1}^{n} P\left(\boldsymbol{y}_t \mid \boldsymbol{Y}_{t-1}, \tilde{\Theta}\right)}{\prod_{t=1}^{n} P(\boldsymbol{y}_t \mid \boldsymbol{Y}_{t-1}, \Theta)}\right\},\end{aligned}$$

where $P(\boldsymbol{y}_t \mid \boldsymbol{Y}_{t-1}, \Theta)$ and $P\left(\boldsymbol{y}_t \mid \boldsymbol{Y}_{t-1}, \tilde{\Theta}\right)$ can be calculated by the Kalman filter when generating $\boldsymbol{X}$ and $\tilde{\boldsymbol{X}}$. This acceptance rate comes from the Metropolis–Hastings method (Hastings, 1970).

5. REVISITING THE SEISMIC ARRAY DATA

We continue the analysis of the earthquake signal from Yunnan Province in China as observed at a nine-element array in Kazakhstan. In this case, we have $N = 9$ time series observed. There are $n = 1701$ points taken at a sampling rate of 20 points per second.

The choice of the prior distribution will affect inference with the posterior distribution. Here, we use a relative flat prior in the analysis of the Yunnan earthquake data. Specifically, we let $m = 40$, $v = 20$, $\beta_1 = 1$, $\beta_2 = 1$, $\mu_a = 0.7, \sigma_a = 0.15$ and $\Sigma_0 = 25 I_p$, where I_p is the $p \times p$ identity matrix. To set the hyper-parameters (μ_ϕ, Σ_ϕ) of the AR coefficient $\boldsymbol{\phi}$, we fit $\left\{y_i(t) e^{d\,t}, i = 1, \cdots, N, t = 1, \cdots, n\right\}$ by an AR(p) model, then let μ_ϕ be the estimated AR coefficients and let $\Sigma_\phi = I_p$. The hyper-parameters (α_x, λ_x) of τ_x control the prior knowledge of the strength of the underlying signal. We use the method proposed in Chib and Ergashev (2009) to set (α_x, λ_x). For a given prior distribution, we sample $\boldsymbol{Y}$ repeatedly from the joint distribution (5); then, the summary statistics of the simulated data is compared with that of the true observations. We use $\alpha_x = 25$ and $\lambda_x = 1000$ when $p = 3$. Figure 4 shows that the simulated data have similar quantiles to the true observations in different time periods.

In order to determine the AR model p in model (4), we use the deviance information criterion (DIC, Spiegelhalter *et al.* (2002)). When the time series is directly observed, Akaike information criterion (AIC) or partial autocorrelation function are often used for order determination. In the Bayesian framework, Troughton and Godsill (1997) used an MCMC approach to determine the AR model. In our case, the time series $x(t)$ in equation (4) is not directly observed and we found that DIC is a simple and effective approach for determining its AR order. Specifically, DIC is defined as

$$\mathrm{DIC}_p = E[D(\Theta) \mid \boldsymbol{Y}] + \left\{E[D(\Theta) \mid \boldsymbol{Y}] - D\left(\overline{\Theta}\right)\right\},$$

where $D(\Theta) = -2 \log P(\boldsymbol{Y} \mid \Theta)$ and $\overline{\Theta} = E(\Theta \mid \boldsymbol{Y})$. Here, $E[D(\Theta) \mid \boldsymbol{Y}]$ and $D\left(\overline{\Theta}\right)$ can be obtained by the MCMC procedure presented in Section 4. For each possible value of p, we implement the MCMC procedure and calculate its DIC value. The order with the smallest DIC value is selected. Figure 5 reports the DIC values for different p's. The result suggests that we use $p = 3$.

DOI: 10.1111/jtsa.12197

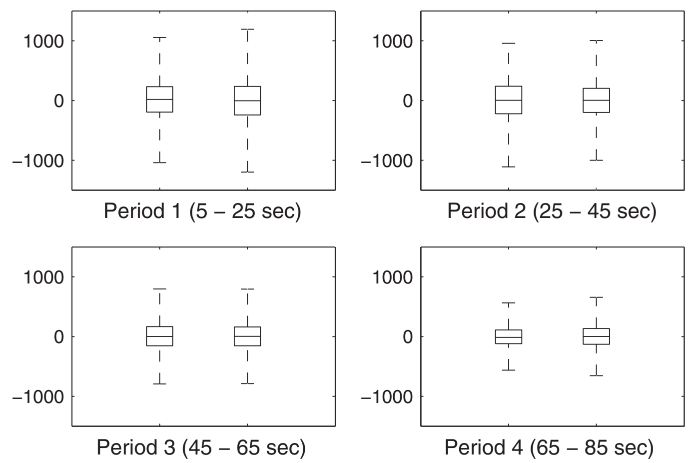

Figure 4. Boxplots of the values of the true observations (left boxes) and the simulated observations (right boxes) in different time periods

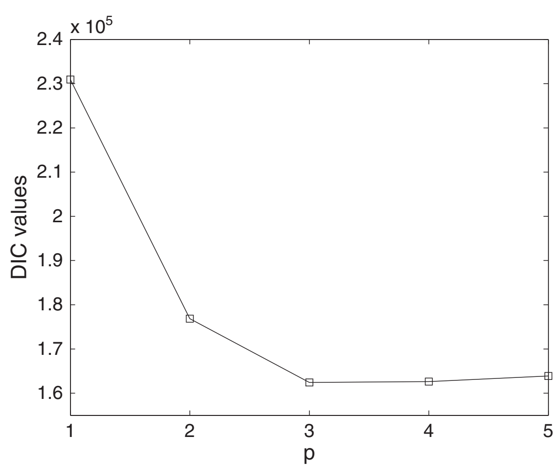

Figure 5. Deviance information criterion (DIC) values for different order p

Table I. Bayesian inference for model parameters

Parameters	Mean	Standard deviation	Credible interval (95%)
η	0.8740	0.0149	(0.8435, 0.9016)
σ_x^2	59.4198	6.1880	(47.7273, 71.8618)
ϕ_1	2.7947	0.0155	(2.7636, 2.8243)
ϕ_2	−2.6699	0.0300	(−2.7272, −2.6097)
ϕ_3	0.8673	0.0155	(0.8360, 0.8971)

We ran the MCMC procedure specified in the previous section for 15,000 iterations. The results of the last 10,000 iterations are used for the estimation of the model parameters. The inefficiency factor as defined in Chib and Ramamurthy (2010) is about 50 for most of the parameters and 300 for σ_x^2. Table I reports the estimation results.

We obtained $\widehat{\sigma}_x^2 = 59.4198$ and $\widehat{\boldsymbol{\phi}} = (2.7947, -2.6699, 0.8673)'$ for the AR parameters in model (4). The roots of the AR(3) model are quite close to the unit circle with one real root 1.1271 and two complex roots with

DOI: 10.1111/jtsa.12197

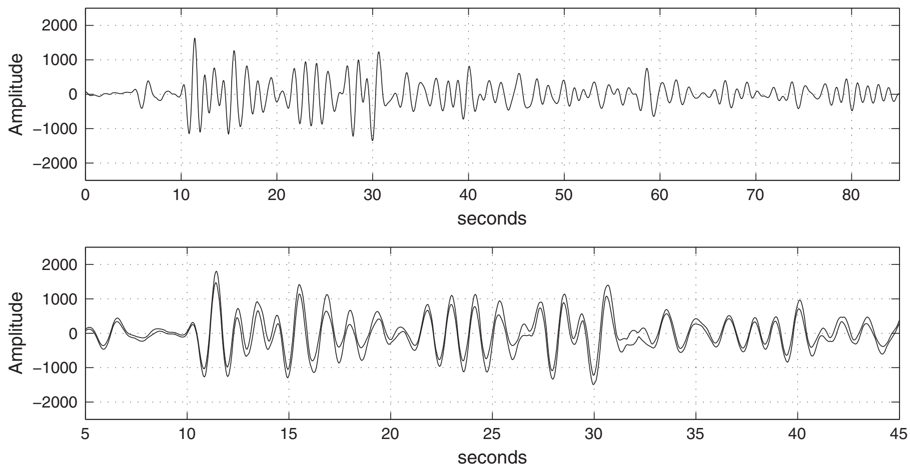

Figure 6. Deconvolved signal $s(t)$ for the Chinese earthquake data in the top panel and its 95% credible interval (from the 5th second to the 45th second) in the bottom panel

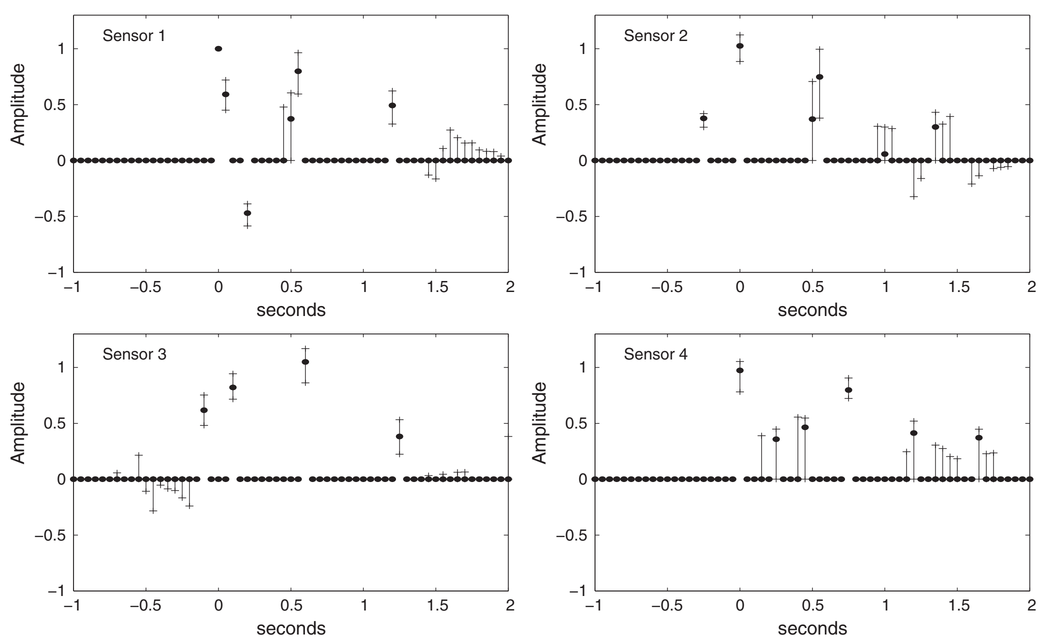

Figure 7. The estimated amplitudes $a_{i,j}$ (the black dots) and its 95% credible interval ('+'), of the first four channels for the Chinese earthquake data

magnitude 1.0114. This produces a highly periodic estimated waveform for the signal $\boldsymbol{X}$, as shown in Figure 6. This is not entirely unexpected as the idealized waveforms of earthquakes and explosions typically have a strong periodic component that is used for input into procedures for discriminating between the two classes of events.

The estimated parameter $\hat{\eta} = 0.8740$ shows that the fraction of amplitudes $a_{i,j}$ that are zero is quite high. The high value implies that many of the lags do not contain much information in determining the observed data in model (3), while about 12% of the lags are significant. By contrast, the usual model (1) only produces a unit spike at the delay predicted by the plane wave velocity model.

DOI: 10.1111/jtsa.12197

Since the posterior distribution of the amplitude $a_{i,j}$ is a mixture distribution with positive probability mass at zero, the amplitude is estimated by its posterior median, that is,

$$\widehat{a}_{i,j} = \text{median}\left\{a_{i,j}^{(l)},\, l = 1, \cdots, L\right\}.$$

Figure 7 shows the estimated amplitudes for the first four sensors. We can estimate the velocity and azimuth of the signal using the first arrival times on different channels. With each sampled $A^{(l)} = \left\{a_{i,j}^{(l)},\, i = 1, \cdots, N,\, j = -v, -v+1, \cdots, m\right\}$, we can obtain the time delays of each channel relative to the first

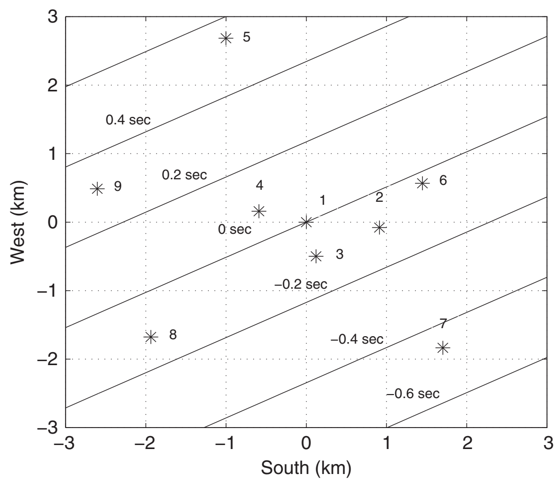

Figure 8. Contour map of the first arrival times

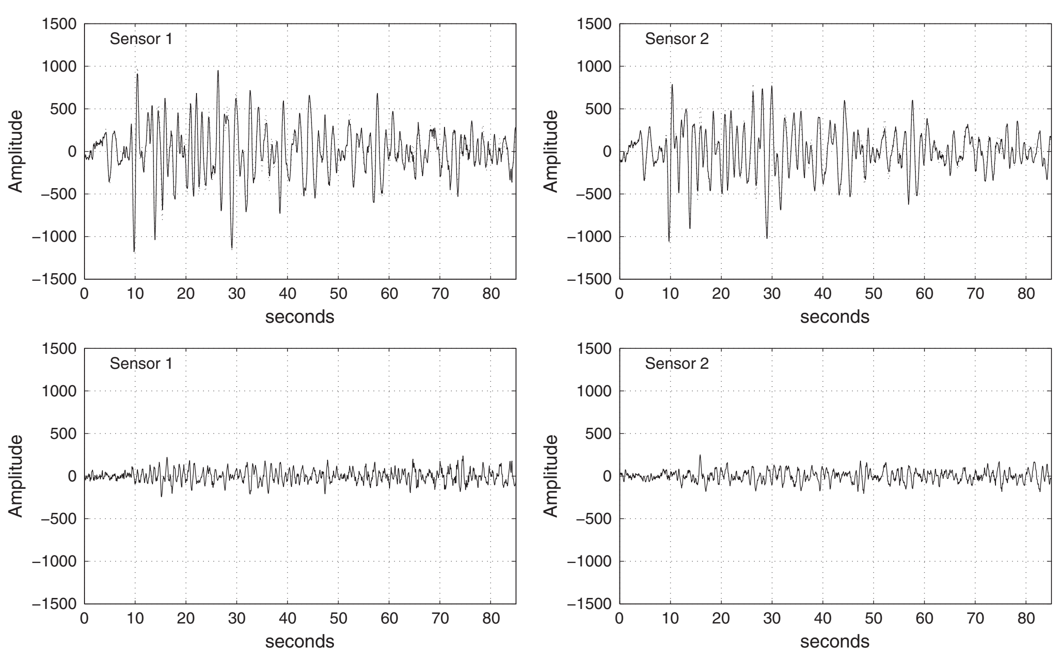

Figure 9. For the Chinese earthquake data, the observed (solid line) and recovered (dotted line) series of the first two sensors in the top panels and the residuals in the bottom panels

  J. Time. Ser. Anal. **37**: 837–850 (2016)
DOI: 10.1111/jtsa.12197

channel. Suppose $\boldsymbol{u}^{(k)} = \left(u_1^{(l)}, \cdots, u_N^{(l)}\right)$ is the vector of the first arrival times at the channels in seconds. We may write $u_i = \boldsymbol{r}_i'\boldsymbol{\theta} + e_i$, where $\boldsymbol{r}_i$ is the coordinate of the i-th sensor, and $\boldsymbol{\theta}$ is slowness in seconds per cycle. Minimizing $\sum_{i=1}^{N}\left(u_i^{(l)} - \boldsymbol{r}_i'\boldsymbol{\theta}\right)^2$ leads to an estimate of $\boldsymbol{\theta}$. The mean of the estimated slowness parameters $\left\{\widehat{\boldsymbol{\theta}}^{(l)}, l = 1, \cdots, L\right\}$ is $(-0.0876, 0.1706)'$. Converting the slowness parameters to azimuths and velocities, it gives that the azimuth of the signal is 152.8° with the 95% credible interval 151.7–154.5°, and the velocity is 5.22 km/second with the 95% credible interval 4.62–5.36 km/second. Figure 8 plots a contour map of the first arrival times. The first wave of the signal arrives at the points on each straight line at the same time. The correlation and beam power estimator gave an azimuth of about 148° and a velocity of 6.3 km/second respectively. The true azimuth from the event to the array is 150.5°. Both estimated values from the Bayesian method and the beam power estimator are slightly off from the true angle.

The observed and recovered series of the sensors are shown in Figure 9, and we note the excellent fit between the observed and recovered values, shown in the top panels for the first two sensors. The dotted and solid lines are very close but not identical as can be noted by the residual plots, shown in the bottom panels of Figure 6. The maximum amplitudes on all of the original channels are approximately 1000, and we can note that the root mean square error values range from 56 on the fourth sensor to 103 on the seventh sensor, implying that the signal-to-noise ratio is almost exclusively 10 or more. The amplitudes on the residual noise scale are mostly between −200 and 200, and the noise is relatively white.

6. CONCLUSIONS

We have considered here a new approach to deconvolution. The time series on an array is modelled as filtered versions of a signal that can be represented as an exponentially modulated AR process. The aim is to let the filter functions sort out the salient parts of the signal using a mixture of a Bernoulli process and a truncated normal distribution as a model. The modulated AR(p) model is often used for earthquakes in the seismic damage field.

Although the estimated signal depends on an MCMC analysis, hence unsuitable for real-time processing, the results seem to offer promise for producing waveforms that are improvements over those that one obtains from routine processing. In the case of the seismic event, the estimated waveform shows more clearly the component phases and focuses attention on a given frequency. We also note that improvements in the location capabilities of the seismic array may be possible using the time delays between the arrivals predicted by the amplitude multiplier locations.

We note that proposed models (3) and (4) can be extended to model other types of multivariate or panel time series. In fact, our model is a one-factor dynamic factor model, but with delay and echo, resulting in a specific and often sparse loading matrix. It can be easily extended to a multifactor model that is commonly considered in the economics literature (Stock and Watson, 2012; Bai and Ng, 2008; Stock and Watson, 2006). The additional features of delay and echo in our model are generally not considered in dynamic factor model research. However, such effects certainly exist in some applications. For example, the same economic factor or financial event may have impact on different countries with different delay. Echoes may occur in seasonal time series.

ACKNOWLEDGEMENTS

Chen's research is partially supported by NSF grants DMS-1209085 and DMS-1513409. Lin's research is partially supported by National Science Foundation of China (NSFC) grants 11101341 and 71131008. We are grateful to two anonymous referees, an Associate Editor and the Editor for their many insightful comments and suggestions that led to significant improvement of the article.

REFERENCES

Andrieu C, Doucet A, Godsill SJ. 1998. Bayesian blind marginal separation of convolutively mixed discrete sources. *The 8th IEEE Workshop on Neural Networks for Signal Processing,* Cambridge, UK; 43–52.

DOI: 10.1111/jtsa.12197

Bai J, Ng S. 2008. Large dimensional factor analysis. *Foundations and Trends in Econometrics* **3**: 89–163.

Benesty J, Gansler T, Morgan DR, Sondhi MM, Gay SL. 2001. *Advances in Network and Acoustic Echo Cancellation*. New York: Springer-Verlag.

Carter C, Kohn R. 1994. On Gibbs sampling for state space models. *Biometrika* **81**: 541–553.

Casella G, George EI. 1992. Explaining the Gibbs sampler. *The American Statistician* **46**: 167–174.

Chen R, Liu JS. 2000. Mixture Kalman filters. *Journal of Royal Statistical Society, Series B* **62**: 493–509.

Chen R, Liu JS, Wang X. 2002. Convergence analyses and comparisons of Markov chain Monte Carlo algorithms in digital communications. *IEEE Transactions on Signal Processing* **50**: 255–270.

Cheng Q, Chen R, Li T-H. 1996. Simultaneous wavelet estimation and deconvolution of reflection Seismic signals via Gibbs sampler. *IEEE Transactions on Geoscience and Remote Sensing* **34**: 377–384.

Chib S, Ergashev B. 2009. Analysis of multifactor affine yield curve models. *Journal of the American Statistical Association* **104**: 1324–1337.

Chib S, Nardari F, Shephard N. 2006. Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics* **134**: 341–371.

Chib S, Ramamurthy S. 2010. Tailored randomized block MCMC methods with application to DSGE models. *Journal of Econometrics* **155**: 19–38.

Früwirth-Schnatter S. 1994. Data augmentation and dynamic linear models. *Journal of Time Series Analysis* **15**: 183–202.

Gelfand A, Smith A. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**: 398–409.

Gibbons S, Schweitzer J, Ringdal F, Kvaerna T, Mykkelveit S, Paulson B. 2011. Improvements to seismic monitoring of the European Arctic using three-component array processing at SPITS. *Bulletin of the Seismological Society of America* **101**: 2737–2754.

Gilks W, Richardson S, Spiegelhalter D. 1995. *Markov Chain Monte Marlo in Practice*. Boca Raton: Chapman & Hall/CRC.

Godsill SJ, Andrieu C. 1999. Bayesian separation and recovery of convolutively mixed autoregressive sources. *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 3, Phoenix, AZ; 1733–1736.

Goodman J, Sokal A. 1989. Multigrid Monte Carlo method: conceptual foundations. *Physical Review D* **40**: 2035–2071.

Hastings W. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**: 97–109.

Lawrence C, Altschul S, Boguski M, Liu A, Neuwald JS, Wootton J. 1993. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* **262**: 208–214.

Lee SJ, Rangarajan A, Gindi GR. 1995. Bayesian image reconstruction in SPECT using higher order mechanical models as priors. *IEEE Transactions on Medical Imaging* **14**: 669–680.

Liu JS. 1994. The collapsed Gibbs sampler with applications to a gene regulation problem. *Journal of the American Statistical Association* **89**: 958–966.

Liu JS. 2001. *Monte Carlo Strategies in Scientific Computing*. New York: Springer.

Liu JS, Sabatti C. 2000. Generalized Gibbs sampler and multigrid Monte Carlo for Bayesian computation. *Biometrika* **87**: 353–369.

Marinari E, Parisi G. 1992. Simulated tempering: a new Monte Carlo scheme. *Europhysics Letters* **19**: 451–458.

Murano K, Unagami S, Amano F. 1990. Echo cancellation and applications. *IEEE Communications Magazine* **28**: 49–55.

Robert R, Casella G. 1999. *Monte Carlo Statistical Methods*. New York: Springer,.

Shumway R. 1971. On detecting a signal in N stationarily correlated noise series. *Technometrics* **13**: 499–519.

Shumway R, Dean W. 1968. Best linear unbiased estimation for multivariate stationary processes. *Technometrics* **10**: 523–534.

Shumway R, Der Z. 1985. Deconvolution of multiple time series. *Technometrics* **27**: 385–393.

Shumway R, Smart E, Clauter D. 2008. Mixed signal processing for regional and teleseismic arrays. *Bulletin of the Seismological Society of America* **98**: 36–51.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* **64**: 583–639.

Stock JH, Watson MW. 2006. Forecasting with many predictors. In *Handbook of Economic Forecasting*, Elliott G, Granger C, Timmermann A (eds.) North Holland: Amsterdam, pp. 515–554.

Stock JH, Watson MW. 2012. Dynamic factor models. In *The Oxford Handbook of Economic Forecasting*, Clements MP, Hendry DF (eds.) Oxford University Press: Oxford, pp. 35–60.

Suess E, Shumway R, Chen R. 1998. Bayesian deconvolution of seismic array data using the Gibbs sampler. *Proceedings of the American Statistical Association*: Dallas, Texas; 74–79.

Troughton PT, Godsill SJ. 1997. Bayesian model selection for time series using Markov chain Monte Carlo. *Proceedings of the 1997 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 5: Munich, Germany; 3733–3736.

Vaseghi SV. 2006. *Advanced Digital Signal Processing and Noise Reduction*, 2nd ed. West Sussex, England: Wiley.

Verhofen M. 2005. Markov chain Monte Carlo methods in financial econometrics. *Financial Markets and Portfolio Management* **19**: 397–406.

Winkler G. 1995. *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Berlin: Springer.

  J. Time. Ser. Anal. **37**: 837–850 (2016)
DOI: 10.1111/jtsa.12197