Applied Mathematical Modelling 38 (2014) 738–746

Contents lists available at SciVerse ScienceDirect

Applied Mathematical Modelling

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

State-space modeling for seismic signal analysis

Hui Peng [a,f,*], Genshiro Kitagawa [b], Tetsuo Takanami [c,d], Norio Matsumoto [e]

[a] School of Information Science & Engineering, Central South University, Changsha, Hunan 410083, China
[b] Research Organization of Information and Systems, Kamiyacho Central Place 2F, 4-3-13 Toranomon, Minato-ku, Tokyo 105-0001, Japan
[c] Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington, DC 20015-1305, USA
[d] The Institute of Statistical Mathematics, 10-3 Midoricho, Tachikawa, Tokyo 190-8562, Japan
[e] Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8567, Japan
[f] Hunan Engineering Laboratory for Advanced Control and Intelligent Automation, Changsha, Hunan 410083, China

ARTICLE INFO

Article history:
Received 26 September 2012
Received in revised form 11 June 2013
Accepted 2 July 2013
Available online 17 July 2013

Keywords:
Time series modeling
Geodetic signal analysis
Signal decomposition
Parameter estimation
Strain data
Groundwater level data

ABSTRACT

The time series utilized for geodetic signal analysis, such as strain and groundwater level data, usually is largely affected by barometric pressure, earth tide and precipitation, and also suffer from missing observations due to instrument maintenance or breakdown. To detect informative geodetic signal from heavily noise-affected data, one must build a time series model for decomposition of the data taking into account the characteristics of effects from these covariates. This paper proposes a new modeling method for detecting geodetic signal from earthquake-related time series data by introducing pole-restricted precipitation model, jump component and pre-processing with AR model for interpolating missing observations. Using the proposed method, a geodetic sample data can be decomposed stably into several components including geodetic trend signal, barometric pressure response, earth tidal response, precipitation response and data level shift due to mechanical maintenance or breakdown. The decomposition of the time series and the interpolation of the missing observations are performed very efficiently by using the state-space representation and the Kalman filter/smoother. Finally, case studies of real geodetic sample data demonstrate the effectiveness of the proposed modeling method that lead to some important findings in seismology.

1. Introduction

Continuously observed strain and groundwater level time series can be used to analyze earthquake phenomena, to understand the changes of geodetic data before and after earthquakes occur, and furthermore, to help people predict earthquakes [1,2]. However, looking at this kind of raw data, one can see that it is very difficult to find clear effects of earthquakes in the data, because the raw data are usually seriously affected by various aseismic factors, such as barometric pressure, earth tide, precipitation, and also suffer from missing observations due to instrument maintenance or breakdown [1,3–6]. Therefore, in order to detect coseismic effects for possible use in earthquake prediction, deletion of the aseismic response of strain or groundwater level data is very meaningful and important work.

[7–10] estimated barometric pressure and earth tide response using a Bayesian model, but they did not include the effect of precipitation. [11,12] considered the effect of precipitation using a non-Bayesian regression model. [13–15] proposed a state-space model approach for decomposition of groundwater level data to delete the effects of barometric pressure, earth

* Corresponding author at: School of Information Science & Engineering, Central South University, Changsha, Hunan 410083, China. Tel./fax: +86 731 88830642.

E-mail address: huipeng@mail.csu.edu.cn (H. Peng).

0307-904X/$ - see front matter
http://dx.doi.org/10.1016/j.apm.2013.07.008

tide, and precipitation simultaneously, in which the missing and outlying observations in raw groundwater level data were treated with the first order non-Gaussian trend model.

In the method mentioned above, a problem remained. Namely, in that method, the precipitation effect is expressed by an ARMAX type model, and one has to carefully select its AR parameters for guaranteeing the stability of the model, so as to obtain reasonable response function from precipitation to the water level. Although the fitting of decomposition model for the trend, pressure effect and earth tide effect is rather simple, inclusion of the precipitation effect component makes it somewhat tricky, since the ARMAX model may have complex characteristic roots. This, in general, brings oscillatory geodetic response to the precipitation and further makes it difficult to estimate the coefficients of the ARMAX model by numerical optimization procedure.

On the other hand, the missing data treatment approach used in the previous approach can only cope with general missing observation without large jump that may occur in case of instrument breakdown for strain data. In the filtering for the estimation of unknown state, missing observations of the geodetic time series can be reasonably treated very naturally [13,16]. However, if there are missing observations in the covariates such as barometric pressure, it may cause severe effects since the observation at one point affects at least m time points because of the regression type model shown later in model (3).

This paper proposes a new decomposition method for the analysis of geodetic strain or groundwater level data. In this method, the data related to earthquakes are decomposed into several components including geodetic trend signal, barometric pressure response, earth tide response, precipitation response and the effect of missing data or jump type disturbance. A model structure and the optimization approach to guaranteeing the stability of the precipitation component model is presented in this paper. The estimation of the parameters of the proposed model is implemented by the Kalman filter, and the maximum likelihood method. The main contributions of the paper are the use of pole-restricted ARMAX model for precipitation response, inclusion of the jump component and the interpolation method for missing observation of regression covariates. In case studies, two examples are shown to illustrate the effectiveness of the method proposed in this paper.

2. Modeling method

2.1. Model structure

We propose the following model to represent the behavior of geodetic data such as borehole strain time series or groundwater level time series

$$\begin{cases} y_n = t_n + P_n + E_n + R_n + S_n + w_n, \\ w_n \sim N(0, \sigma^2), \quad n = 1, \ldots, N, \end{cases} \tag{1}$$

where N is the number of observations, y_n is the raw observation of strain or groundwater level and, t_n, P_n, E_n, R_n, S_n and w_n are the trend, barometric pressure, earth tide, precipitation effect components, jump effect and the observation noise, respectively. Note that this model is an extension of the model proposed in [13].

The trend component t_n is modeled by the following first-order trend model

$$t_n = t_{n-1} + u_n, \quad u_n \sim N(0, \tau^2). \tag{2}$$

It is also possible to use the second order trend model $t_n = 2t_{n-1} - t_{n-2} + u_n$, which is known to yield smoother trend. On the other hand, the first order trend model has an ability to adapt to sudden changes of the trend. Since the objective of the modeling in this paper is to detect abrupt changes of the trend component by removing the effects of other covariates, hereafter we use the first order trend model.

The barometric pressure response P_n is assumed to be represented by a regression model

$$P_n = \sum_{i=0}^{m} a_i p_{n-i}, \tag{3}$$

where p_n is the observed barometric pressure, and the coefficients $a_i (i = 0, 1, \ldots, m)$ are unknown parameters to be estimated. The earth tidal response E_n is modeled by a regression model

$$E_n = \sum_{i=0}^{\ell} b_i e_{n-i}, \tag{4}$$

where e_n is the given theoretical earth tide data, and $b_i (i = 0, 1, \ldots, l)$ are the parameters to be estimated. If one cannot obtain the theoretical earth tide time series, the following model including several major constituents of tides can be used to approximate the effect of earth tide in the geodetic data

$$e_n = \sum_{i=1}^{r} h_i \sin(\omega_i n + \omega_{0,i}), \tag{5}$$

where h_i, ω_i and $\omega_{0,i}$ are the amplitude (cm), frequency (degree/hour) and initial phase of the i-th major constituent, respectively. The frequency of the constituent ω_i is known, but h_i and $\omega_{0,i}$ are estimated by maximizing the likelihood of model (1).

The response to precipitation R_n is assumed to be represented by an ARMAX-type model

$$R_n = \sum_{i=1}^{k} c_i R_{n-i} + \sum_{i=0}^{k-1} d_i r_{n-i}^{\rho}, \tag{6}$$

where r_n is the observed precipitation, and the autoregressive coefficients c_i are assumed to satisfy

$$1 - c_1 q^{-1} - c_2 q^{-2} - \cdots - c_k q^{-k} = (1 - \alpha_1 q^{-1})(1 - \alpha_2 q^{-1}) \cdots (1 - \alpha_k q^{-1}). \tag{7}$$

The coefficients $c_i,\ i = 1, \ldots, k$ are obtained by

$$\begin{cases} c_1 = (-1)^2 \sum_{i=1}^{k} \alpha_i, \\ c_2 = (-1)^3 \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \alpha_i \alpha_j, \\ \cdots \\ c_k = (-1)^{k+1} \alpha_1 \alpha_2 \cdots \alpha_k. \end{cases} \tag{8}$$

The poles of the characteristic equation of the AR operator α_i and the coefficient d_i in (6) are the parameters to be estimated. ρ is the nonlinear factor. Using a different value from (1), one can build model that expresses nonlinear response from precipitation to the geodetic time series. $\alpha_i (i = 1, \ldots, k)$ are estimated by maximizing the likelihood of model (1) subject to the constrains $0 < \alpha_i < 1,\ i = 1, 2, \ldots, k$, to guarantee the stability of model (6) and to avoid unfavorable oscillatory behavior in the response function.

In model (1), S_n is the jump of the level due to instrument maintenance or breakdown, which is represented by

$$S_n = \sum_{i=1}^{n_s} \eta_i s_{i,n}, \tag{9}$$

where η_i is an unknown jump amplitude, and $s_{i,n}$ is the step function defined by, $s_{i,n} = 0$ for $i \leqslant n_j$ and $s_{i,n} = 1$ for $i > n_j$. Here, n_j is the jump occurring time of the j-th jump and is assumed to be known.

2.2. State space representation

On the basis of model (1) and the component models (2–9), one can build the following state space model to represent the behavior of the geodetic time series

$$\begin{cases} \mathbf{X}_n = \mathbf{A}(\theta)\mathbf{X}_{n-1} + \mathbf{B}(\theta) r_n^{\rho} + \mathbf{G} u_n, \\ y_n = \mathbf{C}_n \mathbf{X}_n + S_n + w_n, \end{cases} \tag{10}$$

where $\mathbf{X}_n = [t_n, a_0 \cdots a_m, b_0 \cdots b_\ell, R_{n,1} \cdots R_{n,k}]^{\mathrm{T}},\ u_n \sim N(0, \tau^2)$ and $w_n \sim N(0, \mathbf{\Phi}_n)$, and the components $R_{n,j};\ j = 1, \ldots, k$ are defined by

$$R_{n,1} = R_n,$$
$$R_{n,j} = \sum_{i=j}^{k} c_i R_{n+j-i-1} + \sum_{i=j-1}^{k-1} d_i r_{n+j-i-1}^{\rho}, \quad \text{for } j = 2, \ldots, k.$$

The matrices and the vectors used in (10) are defined by

$$\mathbf{A}(\theta) = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c_1 & 1 & & \\ & & & \vdots & \vdots & \ddots & \\ & & & c_{k-1} & 0 & \cdots & 1 \\ & & & c_k & 0 & \cdots & 0 \end{bmatrix},$$

$$\mathbf{B}(\theta) = [0\ \ 0\ \ \cdots\ \ 0\ \ d_0\ \ \cdots\ \ d_{k-2}\ \ d_{k-1}]^{\mathrm{T}},$$
$$\mathbf{C}_n = [1\ \ p_n\ \ \cdots\ \ p_{n-m}\ \ e_n\ \ \cdots\ \ e_{n-1}\ \ 1\ \ 0\ \ \cdots\ \ 0],$$
$$\mathbf{G} = [1\ \ 0\ \ \cdots\ \ 0]^{\mathrm{T}},$$
$$\Phi_n = \sigma_n^2,$$

where the parameter is defined by $\theta = (d_0, \ldots, d_{k-1}, \alpha_1, \ldots, \alpha_k, h_1, \ldots, h_r, \omega_{0,1}, \ldots, \omega_{0,r}, \sigma^2, \tau^2, \rho, \eta_1, \ldots, \eta_{n_s})$ with the constraints $0 < \alpha_i < 1,\ i = 1, 2, \ldots, k$. Note that σ_n^2 is usually a constant σ^2. But if there are missing observations, and are estimated by interpolation algorithm, it should be set to a larger value as will be discussed in Section 2.5.

2.3. State prediction and filtering

The following Kalman filter is used to estimate the state vector $\mathbf{X}_n$ from observed time data [17]. Since the state vector contains the trend component t_n, response coefficients a_i, b_i and the precipitation response $R_{n,1} = R_n$, by estimating the state vector, we can decompose the observed time series into various components. Further, the likelihood of the model and AIC can be computed by using the one-step ahead predictor [16].

Prediction: Let $\mathbf{X}_{n|n-1} = E\{\mathbf{X}_n|\mathbf{Z}_{n-1}\}$ and $\mathbf{V}_{n|n-1} = E\left\{(\mathbf{X}_n - \mathbf{X}_{n|n-1})(\mathbf{X}_n - \mathbf{X}_{n|n-1})^{\mathrm{T}}\right\}$ denote the conditional mean and conditional covariance of $\mathbf{X}_n$ given $\mathbf{Z}_{n-1} = \{y_1, y_2, \ldots, y_{n-1}\}$, then the one-step-ahead predictor is obtained as follows

$$\begin{cases} \mathbf{X}_{n|n-1} = \mathbf{A}(\theta)\mathbf{X}_{n-1|n-1} + \mathbf{B}(\theta)r_n^{\rho}, \\ \mathbf{V}_{n|n-1} = \mathbf{A}(\theta)\mathbf{V}_{n-1|n-1}\mathbf{A}(\theta)^{\mathrm{T}} + \mathbf{G}\tau^2\mathbf{G}^{\mathrm{T}}. \end{cases} \tag{11}$$

Filtering: Let $\mathbf{X}_{n|n} = E\{\mathbf{X}_n|\mathbf{Z}_n\}$ and $\mathbf{V}_{n|n} = E\left\{(\mathbf{X}_n - \mathbf{X}_{n|n})(\mathbf{X}_n - \mathbf{X}_{n|n-1})^{\mathrm{T}}\right\}$ denote the conditional mean and conditional covariance of $\mathbf{X}_n$, given $\mathbf{Z}_n = \{y_1, y_2, \ldots, y_n\}$, then the filtered state may be given as

$$\begin{cases} \gamma_n = y_n - \mathbf{C}_n\mathbf{X}_{n|n-1} - S_n, \\ \Psi_n = E\{\gamma_n^2\} = \mathbf{C}_n\mathbf{V}_{n|n-1}\mathbf{C}_n^{\mathrm{T}} + \Phi_n, \\ \mathbf{K}_n = \mathbf{V}_{n|n-1}\mathbf{C}_n^{\mathrm{T}}/\Psi_n, \\ \mathbf{X}_{n|n} = \mathbf{X}_{n|n-1} + \mathbf{K}_n\gamma_n, \\ \mathbf{V}_{n|n} = (\mathbf{I} - \mathbf{K}_n\mathbf{C}_n)\mathbf{V}_{n|n-1}. \end{cases} \tag{12}$$

2.4. Parameters estimation and model order selection

After obtaining the optimal state estimates $\mathbf{X}_n^* = \{\mathbf{X}_1^*, \mathbf{X}_2^*, \ldots, \mathbf{X}_n^*\}, n = 1, 2, \ldots, N$, the (-2) log-likelihood of the model can be obtained as follows using the one-step-ahead prediction error γ_n and its variance Ψ_n

$$\begin{aligned} \ell(\theta) &= (-2)\log p(y_N, \ldots, y_1|\theta, \mathbf{X}_N^*) \\ &= \sum_{n=1}^{N}(-2)\log p(y_n|y_{n-1}, \ldots, y_1, \theta, \mathbf{X}_n^*) \\ &= \sum_{n=1}^{N}(-2)\log p(\gamma_n|y_{n-1}, \ldots, y_1, \theta, \mathbf{X}_n^*) \\ &= \sum_{n=1}^{N}\left\{\log \Psi_n^* + (\gamma_n^*)^2/\Psi_n^*\right\} + N\log 2\pi. \end{aligned} \tag{13}$$

Therefore, the optimal estimate of the parameters of model (10) are obtained as the solution of the following minimization problem

$$\theta^*, \mathbf{X}_{0|0}, \mathbf{V}_0 = \arg\min \ell(\theta) \tag{14}$$

subject to the constraints $0 < \alpha_i < 1$ for $i = 1, \ldots, k$.

The model orders, m, ℓ, k in (3), (4), (6) may be selected by minimizing the AIC of model (10) [18] defined by

$$\mathrm{AIC} = -2\max \ell(\theta) + 2k(\theta), \tag{15}$$

where $k(\theta)$ is the dimension of the parameter vector θ, and is given by $2k + 2r + n_s + 3$.

2.5. Missing data treatment

Before the parameter estimation of model (10) can be carried out, the missing data in strain or groundwater level and the covariates such as the barometric pressure time series have to be processed. In this paper, we use AR model to rebuild the missing data of each time series where the data before and after the missing data period are used to estimate the AR parameters, and the missing observations are interpolated by the estimated AR model and the fixed interval smoothing algorithm.

Assume that the uni-variate time series y_n, say strain, is expressed by an autoregressive model

$$y_n = \sum_{j=1}^{p}\alpha_j y_{n-j} + w_n, \quad w_n \sim N(0, \sigma^2). \tag{16}$$

Then, it can be expressed by a state space model

$$\begin{cases} \mathcal{X}_n = \mathbf{F}\mathcal{X}_{n-1} + \mathbf{G}v_n, \\ y_n = \mathbf{H}\mathcal{X}_n + w_n, \end{cases} \tag{17}$$

where $v_n \sim N(0, \mu^2)$ and

$$\mathbf{F} = \begin{bmatrix} \alpha_1 & 1 & & \\ \vdots & & \ddots & \\ \alpha_{p-1} & & & 1 \\ \alpha_p & & & \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{H} = [1 \ \ 0 \ \ \cdots \ \ 0], \quad \mathcal{X}_n = [y_n \ \ y_{n-1} \ \ \cdots \ \ y_{n-p+1}]^{\mathrm{T}}.$$

In the Kalman filtering for this state space model, if the observation y_n is missing, we have the relation that $\mathcal{X}_{n|n} = \mathcal{X}_{n|n-1}$, indicating that we just need to skip the filtering step in the Kalman filter. Then by the smoothing algorithm [16],

$$\begin{cases} \mathbf{A}_n = \bar{\mathbf{V}}_{n|n}\mathbf{F}^T\bar{\mathbf{V}}_{n+1|n}^{-1}, \\ \mathcal{X}_{n|N} = \mathcal{X}_{n|n} + \mathbf{A}_n\left(\mathcal{X}_{n+1|N} - \mathcal{X}_{n+1|n}\right), \\ \overline{\mathbf{V}}_{n|N} = \overline{\mathbf{V}}_{n|n} + \mathbf{A}_n\left(\overline{\mathbf{V}}_{n+1|N} - \overline{\mathbf{V}}_{n+1|N}\right)\mathbf{A}_n^{\mathrm{T}} \end{cases} \tag{18}$$

for $n = N-1, \ldots, 1$, we can obtain the smoothed estimate of the state vector even when there are missing observations. Note that the initial value $\mathcal{X}_{N|N}$ is obtained as the end point of the Kalman filter.

Then, we can estimate the missing observation, say y_n, by $y_{n|N} = \mathbf{H}\mathcal{X}_{n|N}$ and its variance is given by

$$E\left(y_n - y_{n|N}\right)^2 = \mathbf{H}\overline{\mathbf{V}}_{n|N}\mathbf{H}^{\mathrm{T}} + \sigma^2.$$

This means that in fitting the decomposition model considered in the previous subsections, the missing observation is replaced by $y_{n|N}$ and its variance by $\mathbf{H}\overline{\mathbf{V}}_{n|N}\mathbf{H}^{\mathrm{T}} + \sigma^2$. Namely, in the state-space model (10), we set $\Phi_n = \mathbf{H}\overline{\mathbf{V}}_{n|N}\mathbf{H}^{\mathrm{T}} + \sigma^2$ if y_n is a missing observation.

3. Case studies

3.1. The strain data, including the 2003 Tokachi-oki earthquake

The study area is located in the southern corner of Kuril arc and is a seismically active region where large inter-plate earthquakes have occurred repeatedly due to subduction of the Pacific plate. On September 25, 2003 at 19:50 (UTC), a great thrust-type Tokachi-oki earthquake (M8.0) occurred off Tokachi of Hokkaido at the junction of the Kuril and the Japan trenches, which is shown in Fig.1.

The Sacks–Evertson strainmeter (dilatometer) in the borehole of depth 110 m [19] located at the Urakawa Seismological Observatory of Hokkaido University, Kami-kineusu (station code KMU) in the neighborhood of the ruptured zone of the 2003 Tokachi-oki earthquake has been working since 1982 [20]. The data are sent to the Institute of Seismology and Volcanology, Hokkaido University, Sapporo by telemetry system.

The observation of strain, barometric pressure, and precipitation after interpolating missing data is shown in Fig. 2, where the one hour sampled data are used from June 01, 2003 to November 30, 2003. The earth tide component over this period is obtained by using model (5) with 4 major constituents of tides, M_2 ($\omega = 28.984$), K_1 ($\omega = 15.041$), S_2 ($\omega = 30$) and O_1 ($\omega = 13.943$).

In order to detect changes of strain due to geodetic movement, the raw data was decomposed into trend, earth tide, barometric pressure, precipitation responses and level jump by using the modeling method proposed in this paper, and the main results are shown in Fig. 3. According to the AIC values of model (10) given in Table 1, the orders of estimated model for the strain data are chosen as $m = 20, \ \ell = 2, \ k = 3$. The decomposition obtained by model (10) and the Kalman filter are shown in Fig. 3. From the figure, one can see a very smooth geodetic trend is extracted except at around $n = 2800$, from the complicated raw data.

The large abrupt change in the geodetic trend is a very interesting phenomenon and was explained by the two slow slip events (M $\sim$ 7) that occurred following the 2003 M8.0 Tokachi-oki earthquake off Hokkaido [23,24].

It can be seen that the most of the variations around the trend can be explained as the effects of either barometric pressure or earth tide. The effect of precipitation is also very large for this strain data. The difference between the trend and the original data can be explained with this component.

This example shows that although the observed geodetic data is strongly affected by the effects of various covariates, the use of model (10) could provide us with an interesting geodetic movement, and may finally lead to important discovery such as the slow slips that occurred immediately after the 2003 Tokachi-oki earthquake.

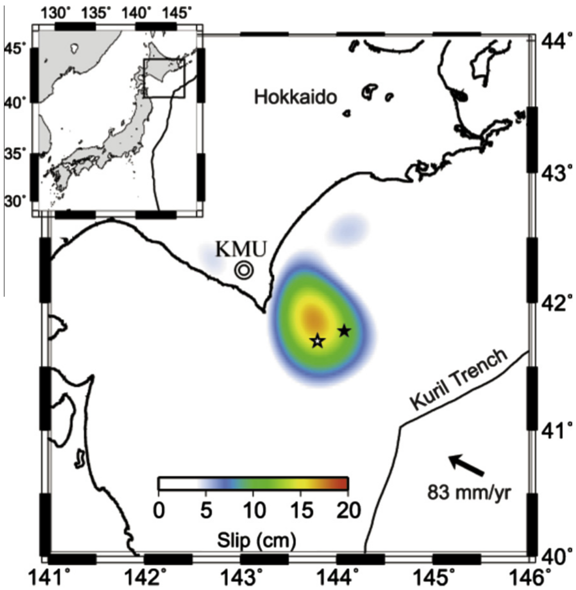

Fig. 1. Location map of the 2003 Tokachi-oki earthquake (M8.0). Solid and open stars denote epicenters of the main shock and the largest aftershock, respectively. A double circle labeled KMU shows the location of the borehole strainmeter. Colors represent slip distribution model for 5 h following the main shock by Fukuda et al. [21]. Thick arrow shows the relative motion of the Pacific plate with respect to the North American plate by DeMets et al. [22].

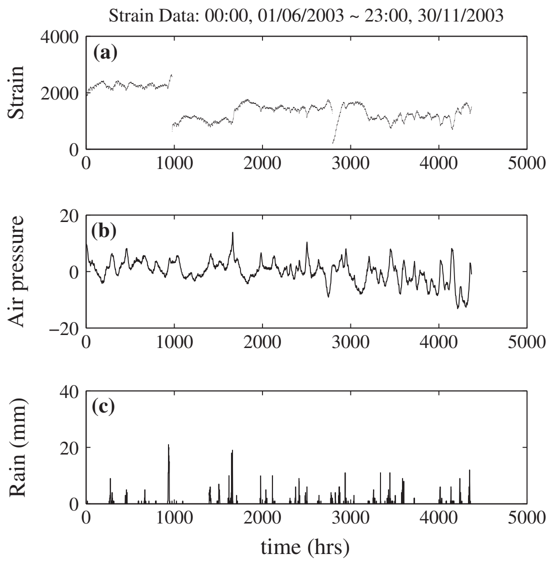

Fig. 2. 2003 KMU strain data: (a) strain, (b) air pressure, and (c) precipitation; amplitudes of strain and air pressure are digital counts from A/D converter.

3.2. The 1984–1985 Haibara groundwater level data analysis

The 1984–1985 Haibara groundwater level data, barometric pressure, theoretical earth tide and precipitation time series from 03/19/1984 to 02/15/1985 at the Haibara well, Shizuoka Prefecture, Central Japan after interpolating the missing observations are shown in Fig. 4. This area was specified as one of two important observation areas for earthquake prediction, and the Geological Survey of Japan, AIST, has been observed groundwater level and other time series for over 30 years. The actual sampling interval is 2 minuites, but we used the data sampled at every one hour. The same data was previously analyzed by [13–15] and found clear coseismic effects of groundwater to the earthquakes. The problem left in these analyses was the difficulty in fitting precipitation model, because in these previous papers, ordinary ARMAX model was used. In this example, we used the precipitation model proposed in this paper. The decomposition of the data using the proposed state-space model (10) and the proposed optimization method are shown in Fig. 5. From the figure, one can see that very smooth ground-

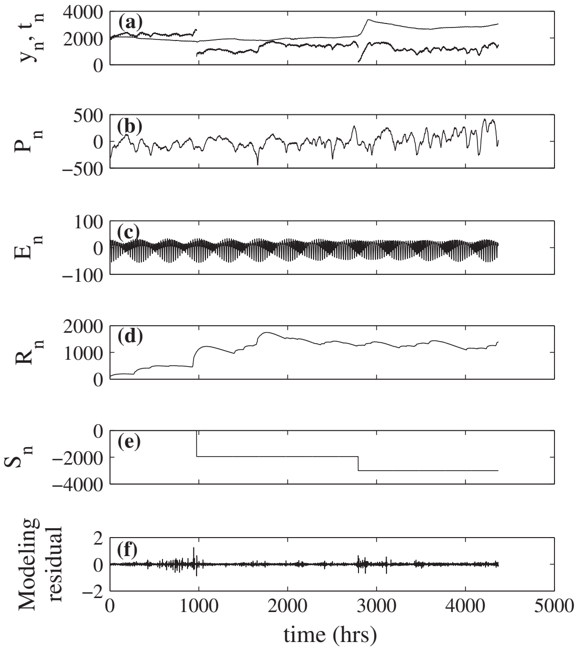

Fig. 3. Decomposition by the estimated model (10): (a) the raw data with jumps recorded at KMU in 2003 and the final geodetic trend (smooth curve); (b) air pressure effect; (c) earth tide effect; (d) precipitation response; (e) the level jump component; and (f) the modeling residuals.

Table 1
AIC of the estimated model. The bold value is the smallest one in all AIC values in the table.

Order of R_n	r^1	$r^{1.2}$	$r^{1.4}$	$r^{1.6}$	$r^{1.8}$
2	32827.77	32850.34	32888.59	32946.51	32987.88
3	**32793.56**	32813.52	32840.94	32841.14	32888.50
4	32798.56	32819.20	32841.35	32850.54	32906.82
5	32804.58	32824.38	32838.07	32860.24	32901.52

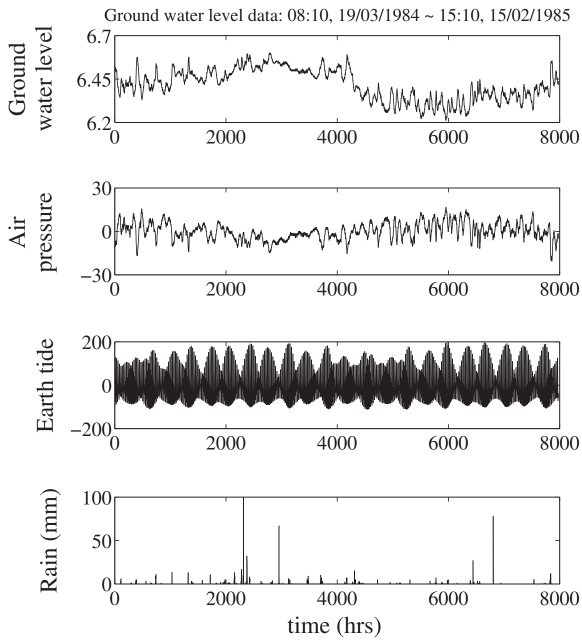

Fig. 4. Observation of 1984–1985 Haibara groundwater level data.

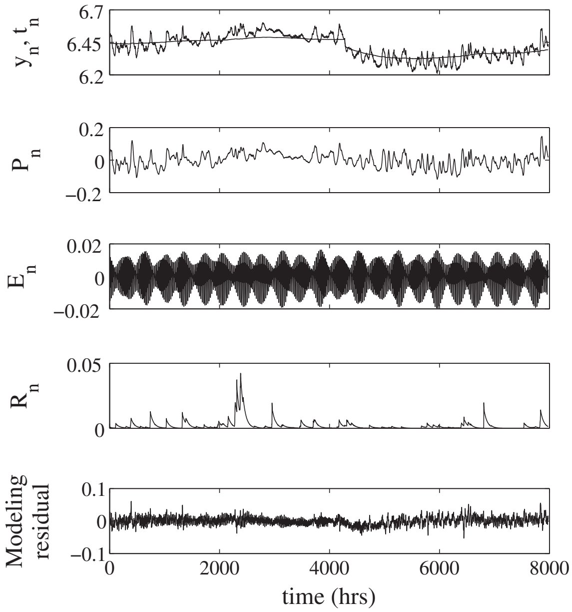

Fig. 5. Components decomposed by model (10) for 1984–1985 Haibara groundwater level data.

water level trend is extracted from the heavily contaminated data except at around $n = 4,200$ (September 14, 1984) when Nagano-ken Seibu Earthquake (magnitude 6.8, distance 126 km) occurred.

It can be seen that the range of the barometric pressure effect is about 0.2 m and most of the variations in the observed groundwater level data ascribe to this effect. Both the earth tide effect and the precipitation effect are minuscule with ranges of about 0.04 m, but are clearly extracted and contributed to make the estimated trend component very smooth. It is noteworthy that the effect of the earthquake is very long lasting for over 100 days.

The impulse response from the precipitation to the groundwater level computed by the estimated precipitation model obtained by assuming that all poles are positive real numbers, is shown in the top plot of Fig. 6. On the other hand, the middle plot of Fig. 6 shows the impulse response of the precipitation model obtained by allowing complex number poles [13], which shows unfavorable oscillatory response function.

The restricted real roots model proposed in this paper has significant merits that it always has a natural response function and makes the optimization for model parameters estimation quite easy.

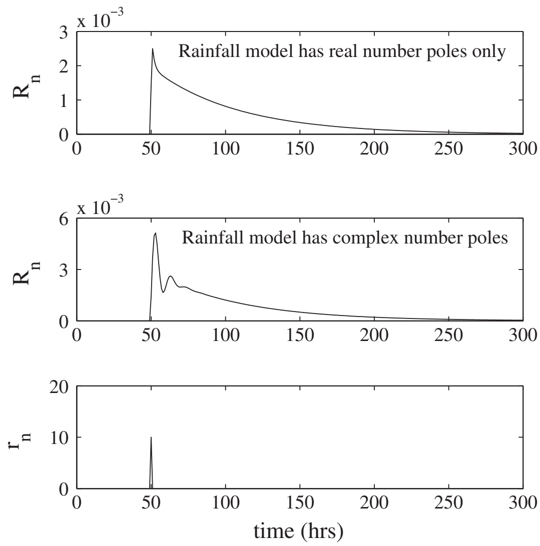

Fig. 6. Impulse response of rainfall model (6) with different type poles.

4. Conclusions

On the basis of previous work, we improved the model structure for decomposition of earthquake-related geodetic time series, such as strain data, and groundwater level data. The proposed model includes several components, namely, the trend component, that we want for the time series decomposition, the barometric pressure effect, the earth tide effect, the precipitation effect, and the jump effect caused by measurement device breakdown. The AR parameters of the precipitation effect component model are estimated by maximizing likelihood of the model under the restriction that all poles are real numbers between 0 and 1.

This restricted model has a significant merit that the stability of the precipitation response is guaranteed in the model optimization process and makes the parameter estimation quite easy. In the response function type or ARMAX type models, the missing observations in the covariates may cause severe effects because one missing observation may affect many time points. This problem was mitigated by estimating the missing observation by interpolation with state space representation of the AR model. Inclusion of the jump process makes it possible to adjust automatically to abrupt changes of trend component due to malfunction or breakdown of the measurement system whatever the reason.

Future work is to use non-Gaussian noise model and Monte Carlo filtering method for the decomposition of earthquake-related time series, which may further improve modeling abilities.

Acknowledgments

The authors thank the editor and the anonymous referees for their valuable comments and suggestions, which improved the original manuscript. This work was supported by the Institute of Statistical Mathematics, Tokyo, Japan, and by the National Natural Science Foundation of China under Grant 71271215 and 71221061, and the International Science & Technology Cooperation Program of China under Grant 2011DFA10440.

References

[1] E. Roeloffs, Hydrologic precursors to earthquakes: a review, Pure Appl. Geophys. 126 (1988) 177–209.
[2] E. Roeloffs, Evidence for aseismic deformation rate changes prior to earthquakes, Annu. Rev. Earth Planet. Sci. 34 (2006) 591–627.
[3] R. Hart, M. Gladwin, R. Gwyther, D. Agnew, F. Wyatt, Tidal calibration of borehole strain meters: removing the effects of small-scale inhomogeneity, J. Geophys. Res. 101 (1996). 25–25.
[4] T. Narasimhan, B. Kanehiro, P. Witherspoon, Interpretation of earth tide response of three deep, confined aquifers, J. Geophys. Res. 89 (1984) 1913–1924.
[5] E. Roeloffs, Tidal calibration of plate boundary observatory borehole strainmeters: roles of vertical and shear coupling, J. Geophys. Res. 115 (2010) B06405.
[6] G. Van der Kamp, J. Gale, Theory of earth tide and barometric effects in porous formations with compressible grains, Water Resour. Res. 19 (1983) 538–544.
[7] M. Ishiguro, H. Akaike, M. Ooe, S. Nakai, A Baysian approach to the analysis of earth tides, in: Proceedings of Ninth International Symposium on Earth Tides, pp. 283–292.
[8] M. Ishiguro, T. Sato, Y. Tamura, M. Ooe, Tidal data analysis, an introduction to BAYTAP, Proc. Inst. Stat. Math. 32 (1984) 71–85.
[9] G. Igarashi, H. Wakita, Tidal responses and earthquake-related changes in the water level of deep wells, J. Geophys. Res. 96 (1991) 4269–4278.
[10] Y. Tamura, T. Sato, M. Ooe, M. Ishiguro, A procedure for tidal analysis with a Bayesian information criterion, Geophys. J. Int. 104 (1991) 507–516.
[11] N. Matsumoto, M. Takahashi, G. Kitagawa, Development of a linear multivariate regression model to detect coseismic changes of ground water level, Bull. Geol. Surv. Jpn. 40 (1989) 613–623. In Japanese with English abstract.
[12] N. Matsumoto, Regression analysis for anomalous changes of ground water level due to earthquakes, Geophys. Res. Lett. 19 (1992) 1193–1196.
[13] G. Kitagawa, N. Matsumoto, Detection of coseismic changes of underground water level, J. Am. Stat. Assoc. 91 (1996) 521–528.
[14] N. Matsumoto, G. Kitagawa, Extraction of hydrological anomalies related to earthquakes, in: T. Takanami, G. Kitagawa (Eds.), Methods and Applications of Signal Processing in Seismic Network Operations, Springer-Verlag, Berlin, 2003, pp. 235–250.
[15] N. Matsumoto, G. Kitagawa, E. Roeloffs, Hydrological response to earthquakes in the Haibara well, central Japan – I. Groundwater level changes revealed using state space decomposition of atmospheric pressure, rainfall and tidal responses, Geophys. J. Int. 155 (2003) 885–898.
[16] G. Kitagawa, Introduction to Time Series Modeling, CRC Press, New York, 2010.
[17] B.D. Anderson, J.B. Moore, Optimal Filtering, Prentice Hall, New Jersey, 1979.
[18] S. Konishi, G. Kitagawa, Information Criteria and Statistical Modeling, Springer, New York, 2008.
[19] I.S. Sacks, S. Suyehiro, D.W. Evertson, Sacks–Evertson strainmeter, its installation in Japan and some preliminary results concerning strain steps, in: Proceedings of the Japan Academy, vol. 47, pp. 707–712.
[20] T. Takanami, T. Ogawa, I. Sacks, T. Linde, I. Nakanishi, Long-period volume-strain seismogram of the eighth August 1993 Esashi-oki earthquake, off southwest of Hokkaido, Japan and its source mechanism, Journal of the Faculty of Science, Hokkaido University. Series 7, Geophysics 11 (1998) 523–543.
[21] J. Fukuda, K. Johnson, K. Larson, S. Miyazaki, Fault friction parameters inferred from the early stages of afterslip following the 2003 Tokachi-oki earthquake, J. Geophys. Res. 114 (2009).
[22] C. DeMets, R. Gordon, D. Angus, S. Stein, Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions, Geophys. Res. Lett. 21 (1994) 2191–2194.
[23] T. Takanami, I.S. Sacks, T.A. Linde, A Strain Event Related to Aftershock Activity Following the 2003 Tokachi-oki (M8.0) Earthquake, AGU Fall, Meeting Abstracts, 2009, p. G33A-0619.
[24] T. Takanami, I. Sacks, T.A. Linde, H. Peng, G. Kitagawa, A strain event related to aftershock activity following the 2003 Tokachi-oki earthquake (M8.0), Abstract of Japan Geoscience Union Meeting 2010, 2010, p. SSS014–04.