# Dynamic Grouped Mixture Models for Intermittent Multivariate Sensor Data

Naoya Takeishi<sup>1</sup>, Takehisa Yairi<sup>1</sup>, Naoki Nishimura<sup>2</sup>, Yuta Nakajima<sup>2</sup> and Noboru Takata<sup>2</sup>

<sup>1</sup> The University of Tokyo, Tokyo, Japan, takeishi@ailab.t.u-tokyo.ac.jp,
<sup>2</sup> Japan Aerospace Exploration Agency, Tsukuba, Japan

Abstract. For secure and efficient operation of engineering systems, it is of great importance to watch daily logs generated by them, which mainly consist of multivariate time-series obtained with many sensors. This work focuses on challenges in practical analyses of those sensor data: temporal unevenness and sparseness. To handle the unevenly and sparsely spaced multivariate time-series, this work presents a novel method, which roughly models temporal information that still remains in the data. The proposed model is a mixture model with dynamic hierarchical structure that considers dependency between temporally close batches of observations, instead of every single observation. We conducted experiments with synthetic and real dataset, and confirmed validity of the proposed model quantitatively and qualitatively.

**Keywords:** multivariate time-series, unevenly spaced time-series, mixture models, latent factor models, sensor data

# 1 Introduction

For secure and efficient operation of engineering systems, such as industrial plants, vehicles and artificial satellites, it is of great importance to watch daily logs generated by them. Those logs are useful for various tasks including fault detection and isolation, maintenance prediction, and operation optimization. They mainly consist of multivariate time-series obtained with many sensors equipped to the system, and their amount keeps on increasing as the system goes on running. To deal with such sensor data with a machine learning approach, one can utilize a statistical model that represents their characteristic behavior. For example, state space models (SSM) have been widely used for multivariate time-series like the sensor data obtained from engineering systems (see e.g. [7, 23]).

This work focuses on challenges in practical analyses of the sensor data: temporal unevenness and sparseness. Practically, observation grid of the sensor data is often unevenly and sparsely spaced as illustrated in Fig. 1(a), and it makes the traditional models such as the SSM inappropriate. One of the challenges, the temporal unevenness, is often the case not only with the sensor data of the engineering systems but also with many kinds of scientific and industrial



**Fig. 1.** Concepts of the intermittent multivariate time-series (IMT); (a) multivariate time-series with a large part of observations missing (represented with dotted lines), and (b) results of periodical surveys where observations in a single survey (enclosed by the dotted square) are not ordered. The different colors denote different variables.

time-series. Many researchers have been studying on analyses of the unevenly spaced time-series (see e.g. [12, 24, 20]), though the range of application is limited. With regard to another challenge, the temporal sparsity, the situation of practical analyses can be tragic; the sampling rate of each sensor may vary from a few seconds to several days, the measurements may be unsynchronized, and the sensors may be stopped for a long time and restarted by event-driven measurements. Moreover, data that contain faulty behaviors may not be utilized, which leads to further loss of data and their temporal sequentiality. Consequently it is quite difficult to model the whole sensor data as successive time-series. In the following, we will refer to such unevenly and sparsely spaced sequential data as *intermittent multivariate time-series* (IMT).

One of our motivating examples of the IMT is telemetry data obtained from small artificial satellites. Generally, artificial satellites retrieve readings of hundreds of sensors such as voltmeters, thermometers, accelerometers, gyroscopes, and star sensors, as well as indicators of satellite's status including on/off of equipment, error flags, and operating modes. Hence the telemetry data from the artificial satellites would be generated as multivariate time-series. With regard to some satellites, however, the retrieved sensor readings are not always recorded nor transmitted to the ground due to limitation of memory size and transmission capability. Therefore the telemetry data that we finally obtain are often intermittent, that is, data series in interest are obtained at very low and uneven sampling rates. This is especially true for small satellites and microsatellites, and nanosatellites, which are expected to play a key role in space development. An instance of an observation grid of the telemetry data obtained from a small artificial satellite is shown in Fig. 2, whose vertical axis corresponds to different sensor types and horizontal axis is along time. One can see that the white cells, where measurement is recorded, are unevenly and sparsely placed.

An example of the IMT other than the sensor data is results of longitudinal studies or periodical surveys. These data generally have temporal sequentiality in some large scales such as years or decades, but observations in a single survey do not have any temporal ordering. This property suits the idea of the IMT,



Fig. 2. An observation grid of telemetry data obtained from a small satellite. The white cells denote timestamps where a measurement was recorded, and the black cells are ones where a measurement was *not* recorded. The rows represent different sensors, and the columns are along time. One can see the data are unevenly and sparsely spaced.

in the sense that there occurs certain amount of observations intermittently, as illustrated in Fig. 1(b).

This work presents a novel method to model the intermittent multivariate time-series, which roughly captures temporal information of the data. The proposed method is a mixture model with dynamic hierarchical structure that considers dependency between temporally close batches of observations, instead of every single observation. Our approach can be a useful option in practical sensor data analyses, because it models correlation along time as well as correlation among multiple variables (sensors), without evenly spaced measurements nor interpolation of the data.

In the rest of this paper, related work is briefly reviewed in Sect. 2, the proposed model is explained in Sect. 3, and experimental results are presented in Sect. 4. This paper is concluded in Sect. 5.

### 2 Related Work

#### 2.1 Discrete-time Latent Variable Models

With regard to *evenly* spaced multivariate time-series, one of the most widely accepted models would be discrete-time state space models (SSM), which are also referred to as linear dynamical system (LDS) or dynamic factor models (see e.g. [7,23]). State space models assume (often low-dimensional) continuous latent factors with temporal correlation behind observations, and every observation is conditionally independent given the latent factors. Another model often used is hidden Markov models (HMM) and their variants such as factorial hidden Markov models [9] and factor analyzed hidden Markov models [18], which assume discrete latent variables with temporal dependency.

These models, in discrete time setting, are capable of dealing with the *unevenly* spaced time-series by skipping some measurement updates within filtering procedures [12]. However, skipping too many measurements can cause a significant bias on model estimation, especially when the sampling rate is very low or a large part of observations is missing.

#### 2.2 Processing of Unevenly Spaced Time-series

Many researches have explicitly focused on the unevenly spaced time-series. For example, Zumbach and Müller presented some basic operators such as the moving average [24], and Erdogan applied the autoregressive model [5]. Some other studies handled the unevenly spaced time-series with continuous-time models (see e.g. [12, 3]). In astronomy, periodicity analysis of the unevenly spaced timeseries has been intensely discussed (see e.g. [20, 6]). Note that most of those studies are on basic operators and models for univariate and stationary timeseries, and do not aim to handle the sensor data of engineering systems, which are multivariate and possibly nonstationary as well as intermittent.

Another major way to deal with the unevenly spaced time-series is interpolation.<sup>3</sup> However, it has been pointed out that interpolation of time-series generates a bias on statistic estimation (see e.g. [11, 17]). Moreover, interpolation does not make much sense if an observation interval is too large.

#### 2.3 Practical Alternative: Ignoring Temporal Dependencies

Practically there is a powerful alternative to model the IMT: just ignore the temporal dependency and regard them as a set of i.i.d. observations! If we decided to adopt this i.i.d. assumption, possible modeling approaches would be to use latent variable models such as mixtures of Gaussian, the principal component analysis (PCA), almost equivalently the factor analysis (FA), their mixture versions [22, 8], the independent component analysis (ICA), the canonical correlation analysis (CCA), Gaussian process latent variable models (GPLVM) [14], restricted Boltzmann machines (RBM), and an autoencoder and its variants.

One justification of this compromise is that important information of the sensor data lies in inter-variable (inter-sensor) relationships than in temporal relationships. For example, if usually correlated sensors lose the correlation, a failure of the system can be easily suspected. Moreover, learning the models that assume i.i.d. observations is usually faster and less likely to be trapped in local minima. Of course, such models will completely miss temporal information of the data, but this drawback can be mitigated to some extend by methods like moving average. Actually, the models with the i.i.d. assumption have been widely used for sensor data analyses such as fault detection. See literatures [4, 13] for example, though there would be numerous similar cases in practice.

# 3 Dynamic Grouped Mixtures of Factor Analyzers

In the previous section, we briefly introduced some methods that can be used for the unevenly spaced time-series. These would be useful in many applications, but they mainly treat univariate time-series with a moderate range of observation intervals, which is not the case with the IMT sensor data. We also mentioned the practical compromise, that is, i.i.d. assumption. Although this is acceptable in many cases, it wastes temporal information that still remains in the IMT.

In this section, we present a novel method that especially focuses on two important characteristics of the IMT sensor data of engineering systems. First,

<sup>&</sup>lt;sup>3</sup> Detailed description on interpolation can be found in surveys such as [1, 17].



**Fig. 3.** Graphical models of (a) mixtures of factor analyzers (MFA), and (b) dynamic grouped mixtures of factor analyzers (DGMFA) for three time slices. Quantities other than probabilistic variables are omitted to simplify the figure.

the IMT data maintain temporal information at *some* scale; while the timedependency at the finest level cannot be traced due to the sparseness, it can be observed in a larger scale. For example, physical condition of satellites in orbits around the earth gradually changes according to the revolution of the earth, which can be observed with sensor data even if only a single measurement occurs a day. Second, the sensor data of engineering systems often behave nonlinearly because of physical phenomena, control laws and changes of operational modes. The nonlinearity can be handled with nonlinear models or mixtures of linear models, and we adopt the latter in this work.

The proposed model is a dynamical extension of mixture models. Our idea presented below works with any types of mixture models, though we focus on mixtures of factor analyzers (MFA) [8] for their simplicity and good performance in this work. The MFA is a latent variable model with two types of latent variables z and x as shown in Fig. 3(a), where z denotes a cluster to which each data point y belongs, and x is a low-dimensional representation behind each data point y. For the details, see the literature [8].

#### 3.1 Data Grouping

An essential procedure of the proposed method is grouping of the time-series. We propose to partition the IMT data into multiple batches of data points, on a scale where we would like to model the temporal correlation of the IMT data. This time-series grouping can be formulated as follows: an observation  $\boldsymbol{y}_m$   $(m = 1, \ldots, M)$  is labeled with t  $(t = 1, \ldots, T)$  when the timestamp of  $\boldsymbol{y}_m$  is within the range from  $\tau_t^{\text{begin}}$  to  $\tau_t^{\text{end}}$ , where  $\tau_t^{\text{begin}} = \tau_{t-1}^{\text{end}}$ . In this work, the grouping interval is set evenly  $\tau_t^{\text{end}} = \tau_t^{\text{begin}} + \Delta$  for a fixed value of  $\Delta$ . Thus, the amount of observations within the t-th group,  $N_t$   $(\sum_{t=1}^T N_t = M)$ , differs for each t generally. A suitable granularity of the partition depends on nature of the data and should be tuned empirically or with validation.

#### 3.2 Generative Model

A data generation procedure of the proposed model for the *t*-th data-group can be described as follows:

- 1. Sample a prior of cluster assignment  $\eta_t | \eta_{t-1} \sim \mathcal{N}(\eta_{t-1}, \Lambda)$ .
- 2. For  $n = 1, ..., N_t$ :
  - (a) Sample a cluster assignment  $\boldsymbol{z}_{t,n} | \boldsymbol{\eta}_t \sim \text{Categorical} \left( \mathcal{S} \left( \boldsymbol{\eta}_t \right) \right)$

  - (b) Sample a latent factor  $\boldsymbol{x}_{t,n,k} \sim \mathcal{N}(\boldsymbol{0}, I)$ , where  $k = z_{t,n}$ . (c) Sample a data point  $\boldsymbol{y}_{t,n} | \boldsymbol{z}_{t,n}, \boldsymbol{x}_{t,n,k} \sim \mathcal{N}(L_k \boldsymbol{x}_{t,n,k} + \boldsymbol{b}_k, \boldsymbol{\Psi}_k)$ .

The subscript n denotes the number of a data point within a group (n = $1, \ldots, N_t$ , and k is the number of clusters or mixture components  $(k = 1, \ldots, K)$ . The parameter  $\Lambda$  controls temporal transition of the cluster assignment priors  $\eta$ , while the parameters  $\{L_k, b_k, \Psi_k\}$  are a loading matrix, bias and residual variance of the k-th factor analyzer, respectively. Note that  $\mathcal{S}(\cdot)$  is a softmax function whose k-th element is denoted by  $\mathcal{S}_k(\cdot)$ . A graphical model for probabilistic variables of the proposed model is shown in Fig. 3(b). In the following, we term this model dynamic grouped mixtures of factor analyzers (DGMFA).

#### EM Algorithm with Variational Approximation 3.3

An objective function to learn the DGMFA is a incomplete-data likelihood:

$$p\left(\boldsymbol{y}_{1:T,1:N}\right) = \int \mathrm{d}\boldsymbol{\eta}_{1:T} \left[\prod_{t=1}^{T} \mathcal{N}\left(\boldsymbol{\eta}_{t}; \boldsymbol{\eta}_{t-1}, \boldsymbol{\Lambda}\right) \prod_{n=1}^{N_{t}} \sum_{k=1}^{K} \mathcal{S}_{k}\left(\boldsymbol{\eta}_{t}\right) \int \mathrm{d}\boldsymbol{x}_{t,n,k} \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{I}\right) \mathcal{N}\left(\boldsymbol{L}_{k}\boldsymbol{x}_{t,n,k} + \boldsymbol{b}_{k}, \boldsymbol{\Psi}_{k}\right)\right].$$
(1)

As analytical maximization of this function is intractable, we adopt EM algorithm with variational approximation closely related to one presented by [2], which computes variational posteriors of  $\eta$  by Kalman filtering and smoothing with variational pseudo observations. The most part of the algorithm is the same with one presented in the literatures [2, 8]. Readers should be careful only about an update procedure of variational parameters for z; they are updated with Gaussian likelihoods of observation given an expectation of other latent variables.

#### 4 Experiments

We conducted three experiments with synthetic and real dataset: (1) Denoising of synthetic time-series and (2) visualization and (3) anomaly detection of the IMT sensor data of an artificial satellite.

#### 4.1Simulation: Denoising of Multivariate Time-series

To confirm validity of the proposed model quantitatively, we conducted denoising experiments with synthetically-generated time-series. We prepared noise free time-series  $\{\bar{\boldsymbol{y}}_1, \dots, \bar{\boldsymbol{y}}_M\}$  and noised time-series  $\{\boldsymbol{y}_1, \dots, \boldsymbol{y}_M\}$ , where

$$\boldsymbol{y}_{m} = \bar{\boldsymbol{y}}_{m} + \boldsymbol{e}_{m} \quad \text{where} \quad \boldsymbol{e}_{m} \sim \mathcal{N}\left(\boldsymbol{0}, \nu^{2} I\right),$$
(2)



Fig. 4. One variable of the noised time-series generated by a linear dynamical system and a Gaussian noise. Note that data actually used in the experiment is 8-dimensional. (a) is the original time-series with no subsampling, and (b) is subsampled by rate of 0.2, i.e., 20% of observations was disposed.

for m = 1, ..., M. Now the task is to recover the noise free time-series  $\{\bar{y}_1, ..., \bar{y}_M\}$  from the noised time-series  $\{y_1, ..., y_M\}$ . We generated the noise-free time-series by a linear dynamical system as follows:

$$\boldsymbol{x}_m = \begin{bmatrix} -0.6 & 0.2 \\ -0.1 & 0.5 \end{bmatrix} \boldsymbol{x}_{i-1} + \boldsymbol{w}_m \quad \text{for} \quad m = 2, \dots, M,$$
$$\bar{\boldsymbol{y}}_m = A\boldsymbol{x}_m,$$

where  $\boldsymbol{w}_i$  is a noise that follows  $\mathcal{N}(\mathbf{0}, 0.1^2 I)$ , and A is an  $8 \times 2$  emission matrix whose elements were sampled randomly from  $\mathcal{N}(0, I)$  independently. Therefore the time-series were in 8 dimension, and the noised version was made following (2) with  $\nu = 0.1$ . We generated the data for 5000 timestamps and used 3000 timestamps for training, 1000 for validation and another 1000 for testing. The intermittent situation was simulated by randomly subsampling the original timeseries, and its rate was varied from 0 (no subsampling) to 0.8 (80% of observations was disposed). We showed a part of the generated time-series at Figs. 4(a) and 4(b), with different subsampling rates.

The parameters of the proposed model and other baselines were learned with training set of data, model settings (the number of mixtures K and the number of observation groups T) were selected by a grid-search with the validation set. Note that we do not need to select the latent dimension  $d_x$  in this case, because we know  $d_x = 2$  originally.

Denoising performances were evaluated using the test set of data. The performances in terms of root mean squared errors are shown in Table 1 for different models and different subsampling rates. As baseline, we tried some methods that are widely used in practice: the hidden Markov model (HMM), the linear dynamical system (LDS), the mixture of factor analyzers (MFA). The HMM and LDS are learned skipping missing observations. One can confirm that the proposed model, DGMFA, performed well even for high subsampling rates, while the LDS, with which the original data were generated, failed with the high subsampling rates. In Fig. 5, a part of denoising results and the ground truth are plotted for

Table 1. Denoising performances in RMS errors.

Subsampling rate	HMM	LDS	MFA	DGMFA
0 (no subsampling)	$2.80 \times 10^{-1}$	$1.29 imes10^{-1}$	$1.77 \times 10^{-1}$	$1.24 imes10^{-1}$
0.2 (20%  disposed)	$3.04 \times 10^{-1}$	$1.30 imes10^{-1}$	$1.69 \times 10^{-1}$	$1.26 imes10^{-1}$
0.4 (40%  disposed)	$3.15 \times 10^{-1}$	$1.39 \times 10^{-1}$	$1.70 \times 10^{-1}$	$1.25 imes10^{-1}$
0.6 ( $60\%$ disposed)	$3.19 \times 10^{-1}$	$1.46 \times 10^{-1}$	$1.98 \times 10^{-1}$	$1.24 imes10^{-1}$
$0.8 \ (80\% \text{ disposed})$	$3.45 \times 10^{-1}$	$2.13 \times 10^{-1}$	$1.92 \times 10^{-1}$	$1.25 imes10^{-1}$



Fig. 5. A part of denoising results for test data with subsampling rate (s.r.) 0 and 0.2.

the LDS, MFA, and DGMFA. The simple MFA produced large errors at some timestamps, while the DGMFA achieved the compatible result with the LDS.

### 4.2 Application: Visualization of Sensor Data

The task addressed in this section is to visualize sensor data of an artificial satellite. It is indispensable to watch behaviors of an artificial satellite at all times for secure operation, though it is a tremendous task for human operators since the amount of the data is getting huge. Thus visualizing those data will be a great help to see the picture of satellite's behaviors. We adopted two unsupervised learning techniques, clustering and dimensionality reduction, and tried to use the MFA and the proposed model DGMFA.

The telemetry sensor data visualized here are obtained from SDS-4 [15], which is an working small satellite operated by Japan Aerospace Exploration Agency. The data contain sensor readings such as equipment's voltage, current, temperature. Their sampling rate varies approximately from one second to five



Fig. 6. Visualization of sensor data with cluster assignment rates of each day learned by MFA (left) and DGMFA (right). Each point corresponds to each day. The coloring scheme is by satellite's operational modes (upper) and chronological order (lower).

minutes and differs by sensors. Moreover, the observations are often missing for a long period, approximately from 3 to 12 hours. The asynchrony was slightly compensated by a zero-order hold that lasted up to only 10 seconds. We used 92 types of sensor readings that took continuous values in this experiment, i.e., the data were 92-dimensional intermittent time-series. Remember the fact that satellite's operation consists of a nominal mode and five types of experimental modes, since it is emphasized in the following.

Model settings were empirically selected without any validation because the aim of this experiment is just to show visualization capability of the proposed model qualitatively. The data were partitioned into groups by days and the period of the data is from 1 January 2015 to 30 June 2015, hence T = 183. Consequently, the number of data points within a day was from 300 to 1300 approximately. The intrinsic dimensionality was set  $d_x = 6$  with an intrinsic dimensionality estimator [10], and the number of the mixture components was chosen empirically K = 10 just for visualization clarity.

Results of the visualization are presented in Fig. 6, where the left column is by the simple MFA and the right is by its extension, DGMFA. Figure 6 was drawn as follows. First, we calculated a cluster assignment rate of each group (day) of the data:  $\sum_{\tau_t^{\text{begin}} \leq \text{timestamp}(\boldsymbol{y}_m) < \tau_t^{\text{end}} \mathbb{E}[\boldsymbol{z}_m]$  for MFA and  $\mathbb{E}[\boldsymbol{\eta}_t]$  for DGMFA, from t = 1 to t = 183. It is a 10-dimensional quantity since the number of clusters K was set to 10. Second, we further reduced the dimensionality of these



Fig. 7. ROC curves of anomaly detection. The proposed model, DGMFA (yellow), is compatible with the one-class SVM (blue) that is cheating in parameter tuning.

quantities with Sammon mapping [19] into two dimensions. Finally, we plotted the results of the Sammon mapping with two types of coloring scheme. The plots in the upper row of Fig. 6 are colored according to the six types of operational modes of the satellite, and ones in the lower row are colored according to the chronological order. Note that each point in these figures corresponds to each day of satellite's operation, because the data were partitioned by days.

In Fig. 6(d), the points along the horizontal axis are arranged by the color gradation, i.e. approximately ordered by date. On the other hand, the points in Fig. 6(c) are not ordered geometrically by the color (date) at all. This means that the proposed model was more successful at learning the dynamical behavior of the satellite than the classical MFA was. Since the horizontal axis in Figs. 6(b) and 6(d) would denote chronological order (left to right) and the vertical axis would correspond to operational modes of the satellite, it can be anticipated that one of the experimental operation (denoted by orange points, referred to as "Exp.1") was not conducted after the end of March.

#### 4.3 Application: Anomaly Detection

Another experiment was conducted with the telemetry sensor data obtained from the satellite to quantitatively confirm availability of the proposed model. We focused on semi-supervised anomaly detection, where models were trained using data without anomalous behaviors, and another range of data that may contain anomalies were tested by the trained models. We used the telemetry sensor data during three months: a month for training, another month for validation, and still another month for testing. It has been known that two anomalous (novel) events occurred within the test range: one is an unexpected change of satellite's attitude, and another is a first-ever type of operation of the satellite. We expected to detect these events with as few false alarms as possible.

The proposed model were compared with two practical baselines: one-class SVM [21] and the MFA. The parameters of MFA and DGMFA were tuned using the validation data that contain no anomalies, while this is suboptimal for the detection task. On the other hand, the parameter of the one-class SVM was tuned

with cheating, that is, one with the best detection performance for the test data was adopted. We investigated negative log likelihoods by the one-class SVM and an absolute value of reconstruction errors by the MFA and DGMFA, and a part of the data whose scores exceed a threshold was reported as an anomaly.

The detection performances in ROC curves are shown in Fig. 7, and AUC of each curve is 0.8998 for the one-class SVM, 0.8307 for the MFA, and 0.9105 for the DGMFA. Note that the proposed model was compatible with the one-class SVM that is cheating in parameter tuning.

## 5 Conclusion

In this paper, we introduced a novel way to handle the multivariate intermittent time-series (IMT), where observations were unevenly and sparsely spaced. The proposed method consists of partitioning the time-series into multiple batches of observations, and a model that roughly captures the temporal information of the batches of observations. The partition is done on a scale in which the temporal dependency can be observed even if finer-level information is lost. The proposed model performed better than simple time-series models such as HMM and LDS on a denoising task with the IMT data. Also, it successfully visualized the sensor data obtained from an small artificial satellite, which are difficult to model appropriately with standard time-series models or simple i.i.d. models.

One of the most demanding extensions to the method proposed in this work is automatic partition of the data. The constant interval was adopted in this work, though an optimal granularity of the partition is usually unknown and may change in different parts of time-series.

It is noteworthy that the idea to set the dynamic priors on the cluster assignment of batches of observations is widely known because it was introduced in the work on dynamic topic models (DTM) [2]. In the DTM, each word in a document follows a mixture of multinomials and each document has a prior over mixture component assignment, with the priors having temporal sequentiality.<sup>4</sup> It shall be interesting to combine the proposed model with the DTM, which would enable us to analyze the IMT sensor data and text simultaneously.

### References

- Adorf, H.M.: Interpolation of irregularly sampled data series a survey. In: Astronomical Data Analysis Software and Systems IV (1995)
- Blei, D.M., Lafferty, J.D.: Dynamic topic models. In: Proceedings of the 23rd International Conference on Machine Learning. pp. 113–120 (2006)
- 3. Brockwell, P.J.: Lévy-driven continuous-time ARMA processes. In: Handbook of Financial Time Series. Springer Berlin Heidelberg (2009)
- Dunia, R., Qin, S.J., Edgar, T.F., McAvoy, T.J.: Identification of faulty sensors using principal component analysis. AIChE Journal 42(10), 2797–2812 (1996)

<sup>&</sup>lt;sup>4</sup> From the document-modeling point of view, it is also interesting to compare the MFA with the mixtures of unigrams [16].

- 5. Erdogan, E.: Statistical models for unequally spaced time series. In: Proceedings of the 5th SIAM International Conference on Data Mining. pp. 626–630 (2005)
- Foster, G.: Wavelets for period analysis of unevenly sampled time series. The Astronomical Journal 112, 1709–1729 (1996)
- 7. Geweke, J.: The dynamic factor analysis of economic time-series models. In: Latent Variables in Socio-Economic Models. North-Holland (1977)
- 8. Ghahramani, Z., Hinton, G.E.: The EM algorithm for mixtures of factor analyzers. Tech. rep., University of Toronto (1996)
- Ghahramani, Z., Jordan, M.I.: Factorial hidden markov models. Machine Learning 29(2-3), 245–273 (1997)
- Gupta, M.D., Huang, T.S.: Regularized maximum likelihood for intrinsic dimension estimation. In: Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence (2010)
- Hayashi, T., Yoshida, N.: On covariance estimation of non-synchronously observed diffusion processes. Bernoulli 11, 359–379 (2005)
- 12. Jones, R.: Time series analysis with unequally spaced data. In: Handbook of Statistics, vol. 5. Elsevier Science (1985)
- Kermit, M., Tomic, O.: Independent component analysis applied on gas sensor array measurement data. IEEE Sensors Journal 3(2), 218–228 (2003)
- Lawrence, N.: Probabilistic non-linear principal component analysis with gaussian process latent variable models. Journal of Machine Learning Research 6, 1783–1816 (2005)
- Nakamura, Y., Nishijo, K., Murakami, N., Kawashima, K., Horikawa, Y., Yamamoto, K., Ohtani, T., Takhashi, Y., Inoue, K.: Small demonstration satellite-4 (SDS-4): Development, flight results, and lessons learned in JAXAs microsatellite project. In: Proceedings of the 27th Annual AIAA/USU Conference on Small Satellites (2013)
- Nigam, K., McCallum, A.K., Thrun, S., Mitchell, T.: Text classification from labeled and unlabeled documents using EM. Machine Learning 39(2/3), 103–134 (2000)
- Rehfeld, K., Marwan, N., Heitzig, J., Kurths, J.: Comparison of correlation analysis techniques for irregularly sampled time series. Nonlinear Processes in Geophysics 18, 389–404 (2011)
- Rosti, A.V.I., Gales, M.J.F.: Factor analysed hidden Markov models for speech recognition. Computer Speech and Language 18(2), 181–200 (2004)
- Sammon, J.W.: A nonlinear mapping for data structure analysis. IEEE Transactions on Computers 18, 401–409 (1969)
- Scargle, J.: Studies in astronomical time series analysis. II Statistical aspects of spectral analysis of unevenly spaced data. The Astrophysical Journal 343, 874–887 (1982)
- Schölkopf, B., Platt, J.C., Shawe-Taylor, J.C., Smola, A.J., Williamson, R.C.: Estimating the support of a high-dimensional distribution. Neural Computation 13(7), 1443–1471 (2001)
- 22. Tipping, M., Bishop, C.: Mixtures of probabilistic principal component analysers. Neural Computation 11(2), 443–482 (1999)
- Watson, M.W., Engle, R.F.: Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. Journal of Econometrics 23, 385–400 (1983)
- Zumbach, G., Müller, U.: Operators on inhomogeneous time series. International Journal of Theoretical and Applied Finance 4, 147–177 (2001)