Genomic Data Assimilation for Estimating Hybrid Functional Petri Net from Time-Course Gene Expression Data

Masao Nagasaki[*1]
masao@ims.u-tokyo.ac.jp

Rui Yamaguchi[*1]
ruiy@ims.u-tokyo.ac.jp

Ryo Yoshida[1]
yoshidar@ims.u-tokyo.ac.jp

Seiya Imoto[1]
imoto@ims.u-tokyo.ac.jp

Atsushi Doi[1]
doi@ims.u-tokyo.ac.jp

Yoshinori Tamada[2,3]
tamada@ism.ac.jp

Hiroshi Matsuno[4]
matsuno@sci.yamaguchi-u.ac.jp

Satoru Miyano[1]
miyano@ims.u-tokyo.ac.jp

Tomoyuki Higuchi[2,3]
higuchi@ism.ac.jp

[1] Human Genome Center, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo, 108-8639, Japan
[2] Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo, 106-8569, Japan
[3] Japan Science Technology Agency, Tokyo, Japan
[4] Faculty of Science, Yamaguchi University, 1677-1 Yoshida, Yamaguchi, 753-8512, Japan

Abstract

We propose an automatic construction method of the hybrid functional Petri net as a simulation model of biological pathways. The problems we consider are how we choose the values of parameters and how we set the network structure. Usually, we tune these unknown factors empirically so that the simulation results are consistent with biological knowledge. Obviously, this approach has the limitation in the size of network of interest. To extend the capability of the simulation model, we propose the use of data assimilation approach that was originally established in the field of geophysical simulation science. We provide genomic data assimilation framework that establishes a link between our simulation model and observed data like microarray gene expression data by using a nonlinear state space model. A key idea of our genomic data assimilation is that the unknown parameters in simulation model are converted as the parameter of the state space model and the estimates are obtained as the maximum a posteriori estimators. In the parameter estimation process, the simulation model is used to generate the system model in the state space model. Such a formulation enables us to handle both the model construction and the parameter tuning within a framework of the Bayesian statistical inferences. In particular, the Bayesian approach provides us a way of controlling overfitting during the parameter estimations that is essential for constructing a reliable biological pathway. We demonstrate the effectiveness of our approach using synthetic data. As a result, parameter estimation using genomic data assimilation works very well and the network structure is suitably selected.

Keywords: data assimilation, hybrid functional Petri net, state space model, particle filter, circadian rhythm

1 Introduction

There are two main orientations in studies of biological pathways. One is to estimate the network structures of the pathways by employing statistical methods with observed data, e.g. microarray

*These authors equally contributed to this study.

gene expression data. The other is to build models of pathways for simulations with well-established biological knowledge and to promote better understanding of the networks through the simulations. While our intent is close to the latter one, in this study we propose an automatic construction method of the simulation models with the observation data. In this sense, this study makes a connection between the two approaches for studies of biological pathways.

Simulation studies on biological pathways give us a scope to understanding the regulatory mechanism in cells. In computational biology, a variety of biological pathway models has been proposed, e.g. Hybrid Functional Petri Net (HFPN) [1, 11, 12], and Hybrid Functional Petri Net with extension (HFPNe) [13] and differential equations [15]. They were used for modeling wide variety of biological pathways and succeeded in reproducing consistent time-developing profiles of biological elements such as amounts of mRNAs and proteins. These simulation models are usually governed by several parameters, i.e. initial values, reaction speeds, and threshold of activities and so on. Usually, they are tuned empirically by experts to fit the simulated elements with observed results [2]. However, with the increase of elements in biological pathway, it will be too difficult to tune these parameters by hand. Thus, computational methods that automatically determine these parameters with observed data will be necessary. There is another problem regarding evaluation of goodness of models. Suppose we have several models with different pathways. In that case, even if a model gives better fit with the observation data than the other models, we cannot simply determine the model as the best one because complex models will generally overfit to the data. Thus, another method that takes into account of complexities of models will be necessary.

Data assimilation (DA) [19] is one of promising ways to solve the above problems in a comprehensive manner. DA is a concept firstly emerged in geophysics that combines observations with numerical simulation models [19]. From the statistical point of view, DA can be realized by solving an inverse problem to estimate unknown parameters of a simulation model with the observed data. We can formulate parameter estimation problems with the *nonlinear state space model* (SSM). Once one incorporate a simulation model into an SSM, one can estimate unknown parameters by a proper estimation procedure and also compare models by a statistical criterion. Recently application areas of DA have been diverging, e.g. oceanology [17] and space physics [14]. There are are some studies of estimating parameters in a biological simulation with observed data, e.g. [6, 16], in contrast, the DA concept that we propose is more general mainly from the following two points.

Firstly, DA can deal with not only an estimation problem of parameters, but also an estimation problem of output variables from the simulation model. For a simulation model of a biological pathway, the output variables may represent evolving biological elements, e.g. concentrations of mRNAs and proteins. In an SSM, such variables are called *state variables*. In other methods, parameters like initial parameters are once given, the simulation model simply runs and outputs time-course data of biological elements without referring the data. On the contrary, in DA, once the parameters are given, the simulation also runs. But during the running, at every time when observed value is obtained, the information in the data is utilized to refine the state variables, that is the origin of the name **d**ata **a**ssimilation. This case is expected to give more plausible result that represents real phenomena. We note that other methods, which do not update state variables with observed data, is a special case of DA.

Secondly, DA deals with a problem of model selection in a statistical sense. This problem was not intensively dealt in other studies. The second property of DA allows us to objectively compare a number of hypothesized models and to extract the best model among them. Usually, molecular biologists and medical scientists have new hypotheses that will contribute to update the current model of targeting biological pathways. This property of DA allows us to directly compare these hypothesized models and suggests the plausible model among them.

The purpose of this study is to newly apply DA to a biological simulation and evaluate the applicability for parameter estimations and model selections. As a simulation model, we use a HFPN model for circadian rhythm in *mouse* [10]. To investigate the applicability of the method, we analyze a

synthetic data set generated from the simulation model. For a better approximation of real microarray experimental data, the synthetic data includes sufficient noise.

In Section 2, we briefly explain HFPN and HFPNe as simulation platforms for biological systems. We note that HFPN is a subclass of HFPNe. We then introduce a model for circadian rhythm represented by HFPN, i.e. HFPNe. In Section 3, we show a formulation of nonlinear state space models for DA and explain estimation procedure of state variables and parameters in a context of Bayesian statistics. In Section 3.2, a computational method for estimating state variables and parameters based on sequential Monte Carlo simulation, which is called particle filter, is explained. A model selection criteria using results from the particle filter is shown in Section 3.3. In Section 4, we carry out several numerical experiments using synthetic data sets. In Section 4.1, we deal with parameter estimation problems. We estimate different types of parameters in the simulation model, e.g. parameters for initial values, reaction speeds, and thresholds, which are explained in Section 2. We then examine whether difficulties of the estimation would change depending the types of parameters. In Section 4.2, we deal with a model selection problem. Four different models for circadian rhythm which have different networks in their pathways are used. We introduce a score to compare the goodness of the models and select the best model. Finally, we provide concluding remarks in Section 5.

2 Hybrid Functional Petri Net and Its Extension

Petri net [20] is a kind of graphical programming language invented in the 1960's and it has been extensively studied for modeling concurrent control systems and applied to industrial systems. The model consists of *place*, *transition*, *arc*, and *token*. A place can hold tokens as its content. A transition has arcs coming from places and arcs going out from the transition to some places. A transition connected with these arcs defines a *firing rule* in terms of the contents of the place to/from which the arcs are attached. In modeling with Petri net, a place represents the amount/density of some biological molecule/object and a transition defines the speed/condition/mechanism of interaction/reaction/transfer among the places connected by the arcs. However, the elemental models can be applied to discrete features. Due to this limitation of Petri net, the notions of *Hybrid Functional Petri net* [1, 11] and *Hybrid Functional Petri net with extension* [13] with the notion of object corresponding to the Java class and method by extending the notion of hybrid Petri net were proposed. In addition, for intuitive notations, the place, transition, and arc is renamed as *entity*, *process*, and *connector*, respectively. The HFPN is the complete subset of HFPNe. For the circadian clock model in this paper, we just use the HFPN components. Although the DA concept is also applicable to HFPNe models.

HFPN has two kinds of entities, namely, discrete and continuous and two kinds of processes, discrete and continuous. The concepts of *discrete entity* and *discrete process* are the same as those in the traditional discrete Petri net. A *continuous entity* can hold a real number as its content. A *continuous process* fires continuously at the speed of the parameter assigned to the continuous process. Three types of connectors are used in HFPN. A specific value is assigned to each connector as a *threshold*. When a *process connector* with threshold w is attached to a discrete/continuous process, a certain number of tokens are transferred through the normal connector only if the content of the entity at the source of the process connector exceeds the threshold w. The firing rule of a *association connector* is the same as that of process connector in terms of the threshold, but the content of the entity at the source of the association connector is not consumed by firing. An *inhibitory connector* with threshold w enables the process to fire only if the content of the entity at the source of the connector is less than or equal to w. The formal definition of HFPN and HFPNe is in Nagasaki *et al.* [13]

Figure 1 shows a model of circadian rhythm with HFPN [10] in which each of the connectors is labeled by c_i $(i = 1, \ldots, 45)$. The entities/reaction speeds of processes/thresholds of connectors are summarized in Tables 1, 2, and 3, respectively.

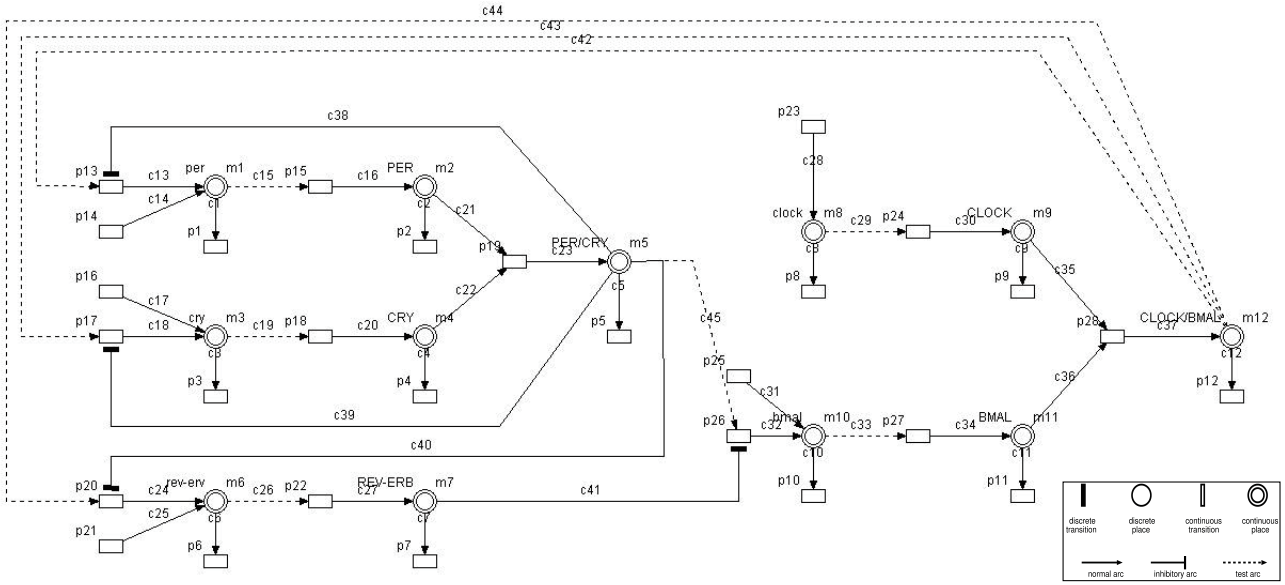

Figure 1: Circadian clock model with HFPN. The legend for the elements in HFPN is shown in the bottom right rectangle. Note that discrete places and discrete transitions are not used in this model. The places, transitions, arcs are also called entities, processes, and connectors, respectively.

Table 1: Biological entities in the HFPN of Figure 1. Variable $m_i(t)$ $(i = 1, \ldots, 12)$ indicates a concentration of the ith entity at time t. $m_i(0)$ $(i = 1, \ldots, 12)$ is the initial value of $m_i(t)$.

Entity Name	Variable	Initial Value	Biological Type
per	$m_1(t)$	$m_1(0)$	mRNA
PER	$m_2(t)$	$m_2(0)$	protein
cry	$m_3(t)$	$m_3(0)$	mRNA
CRY	$m_4(t)$	$m_4(0)$	protein
PER/CRY	$m_5(t)$	$m_5(0)$	complex protein
rev-erv	$m_6(t)$	$m_6(0)$	mRNA
REV-ERV	$m_7(t)$	$m_7(0)$	protein
clock	$m_8(t)$	$m_8(0)$	mRNA
CLOCK	$m_9(t)$	$m_9(0)$	protein
bmal	$m_{10}(t)$	$m_{10}(0)$	mRNA
BMAL	$m_{11}(t)$	$m_{11}(0)$	protein
CLOCK/BMAL	$m_{12}(t)$	$m_{12}(0)$	complex protein

Table 2: Processes and their speeds in the HFPN. #1: Type of Biological Process. (DGR: Degradation, TSC: Transcription, TSL: Translation, and BND: Binding). Each k_i $(i = 1, \ldots, 5)$ in process speeds $v_j(t)$ at time t $(j = 1, \ldots, 28)$ is a common parameter to control speeds of similar biological processes: k_1 for mRNA degradation, k_2 for mRNA base transcription, k_3 for protein base transcription, k_4 for protein degradation, and k_5 for mRNA transcription. Some process speeds include a numerical factor (e.g. 1.42 for v_2 and v_9, and 0.67 for v_5 and v_{12}). These factors represent relative stabilities of the corresponding proteins and mRNAs (cf. [10]).

Process Name	#1	Speed of corresponding processes	Process Name	#1	Speed of corresponding processes
p1	DGR	$v_1(t) = m_1(t) \times k_1$	p15	TSL	$v_{15}(t) = m_1(t) \times k_3$
p2	DGR	$v_2(t) = m_2(t) \times k_4 \times 1.42$	p16	TSC	$v_{16}(t) = k_2$
p3	DGR	$v_3(t) = m_3(t) \times k_1$	p17	TSC	$v_{17}(t) = k_5$
p4	DGR	$v_4(t) = m_4(t) \times k_4$	p18	TSL	$v_{18}(t) = m_3(t) \times k_3$
p5	DGR	$v_5(t) = m_5(t) \times k_4 \times 0.67$	p19	BND	$v_{19}(t) = m_2(t) \times m_4(t) \times 0.1$
p6	DGR	$v_6(t) = m_6(t) \times k_1$	p20	TSC	$v_{20}(t) = k_5$
p7	DGR	$v_7(t) = m_7(t) \times k_4$	p21	TSC	$v_{21}(t) = k_2$
p8	DGR	$v_8(t) = m_8(t) \times k_1$	p22	TSL	$v_{22}(t) = m_6(t) \times k_3$
p9	DGR	$v_9(t) = m_9(t) \times k_4 \times 1.42$	p23	TSC	$v_{23}(t) = k_2 \times 10$
p10	DGR	$v_{10}(t) = m_{10}(t) \times k_1$	p24	TSL	$v_{24}(t) = m_8(t) \times k_3 \times 0.2$
p11	DGR	$v_{11}(t) = m_{11}(t) \times k_4$	p25	TSC	$v_{25}(t) = k_2$
p12	DGR	$v_{12}(t) = m_{12}(t) \times k_4 \times 0.67$	p26	TSC	$v_{26}(t) = k_5 \times 1.1$
p13	TSC	$v_{13}(t) = k_5$	p27	TSL	$v_{27}(t) = m_{10}(t) \times k_3 \times 0.4$
p14	TSC	$v_{14}(t) = k_2$	p28	BND	$v_{28}(t) = m_9(t) \times m_{11}(t) \times 0.1$

Table 3: Threshold parameters s_i $(i = 1, \ldots, 5)$ and the corresponding connectors in the HFPN. Note that the other connectors do not have the threshold parameters, which means the corresponding regulations are always effective.

Threshold parameters	Name of regulation	Corresponding connector
s_1	bmal active regulation	c45
s_2	bmal inhibitory regulation	c41
s_3	rev-erv inhibitory regulation	c40
s_4	cry inhibitory regulation	c39
s_5	per inhibitory regulation	c38

3 Data Assimilation with Nonlinear State Space Models

3.1 Probabilistic Formulation of HFPN

A biological pathway model represented by the HFPN describes a dynamics in evolving time course of the p biological entities $\boldsymbol{m}(t) = (m_1(t), \ldots, m_p(t))$ at the evenly spaced discrete time points. Series of the biological entities, e.g. concentrations of proteins, expression levels of mRNA transcripts, follows a set of first-order difference equations, $m_i(t+n/N) = g_i(\boldsymbol{m}(t+(n-1)/N), \boldsymbol{\theta})$ $i = 1, \ldots, p$ for $t = 1, \ldots, T$ and $n = 1, \ldots, N$ where the parameter vector $\boldsymbol{\theta}$ contains, for example, protein degradation rates k, some threshold values s which regulate structural changes in the pathway. For ease of explanation, time course data that we use for estimation of the model parameters are assumed to be measured at only integer time points $t = 1, \ldots, T$. As an illustration, consider a simple HFPN model $\mathcal{S}_1$ of five biological entities as follows:

$$
\begin{aligned}
m_1\Big(t+\frac{n}{N}\Big) &= m_1\Big(t+\frac{n-1}{N}\Big) + \frac{k_1}{N} m_1\Big(t+\frac{n-1}{N}\Big),\\
m_2\Big(t+\frac{n}{N}\Big) &= m_2\Big(t+\frac{n-1}{N}\Big) + \frac{k_2}{N} m_2\Big(t+\frac{n-1}{N}\Big),\\
m_3\Big(t+\frac{n}{N}\Big) &= m_3\Big(t+\frac{n-1}{N}\Big) + \frac{k_3}{N} m_1\Big(t+\frac{n-1}{N}\Big) - \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big)\chi,\\
m_4\Big(t+\frac{n}{N}\Big) &= m_4\Big(t+\frac{n-1}{N}\Big) + \frac{k_4}{N} m_2\Big(t+\frac{n-1}{N}\Big) - \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big)\chi,\\
m_5\Big(t+\frac{n}{N}\Big) &= m_5\Big(t+\frac{n-1}{N}\Big) + \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big)\chi,
\end{aligned}
\tag{1}
$$

where

$$
\chi = I\left[m_3\Big(t+\frac{n-1}{N}\Big) > s_1\right] I\left[m_4\Big(t+\frac{n-1}{N}\Big) > s_2\right], \tag{2}
$$

and the $I[\,]$ denotes an indicator function which takes value one if the argument is true, otherwise zero. In this model, the full parameters to be estimated are composed of rates of biological reactions k_i, $i = 1, \ldots, 5$ threshold values s_1, s_2 and initial states $m_i(0)$, $i = 1, \ldots, 5$. We let $m_1(t)$ and $m_2(t)$ be concentrations of mRNAs and the rest entity corresponds to a particular protein. Especially $m_5(0)$ is a concentration of a protein complex made of proteins corresponding to $m_3(t)$ and $m_4(t)$. Suppose that a series of mRNA transcription levels, $\boldsymbol{y}_t$, $t = 1, \ldots, T$ is profiled at the successive time points by the time course microarray experiments. Then, the task to be addressed is to tune the parameter values so as to mimic the real observations.

Furthermore, let us consider that we alternatively hypothesize $\mathcal{S}_2$ as

$$
\begin{aligned}
m_1\Big(t+\frac{n}{N}\Big) &= m_1\Big(t+\frac{n-1}{N}\Big) + \frac{k_1}{N} m_1\Big(t+\frac{n-1}{N}\Big),\\
m_2\Big(t+\frac{n}{N}\Big) &= m_2\Big(t+\frac{n-1}{N}\Big) + \frac{k_2}{N} m_2\Big(t+\frac{n-1}{N}\Big) I\left[m_5\Big(t+\frac{n-1}{N}\Big) \le s_3\right],\\
m_3\Big(t+\frac{n}{N}\Big) &= m_3\Big(t+\frac{n-1}{N}\Big) + \frac{k_3}{N} m_1\Big(t+\frac{n-1}{N}\Big) - \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big)\chi,\\
m_4\Big(t+\frac{n}{N}\Big) &= m_4\Big(t+\frac{n-1}{N}\Big) + \frac{k_4}{N} m_2\Big(t+\frac{n-1}{N}\Big) - \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big)\chi,\\
m_5\Big(t+\frac{n}{N}\Big) &= m_5\Big(t+\frac{n-1}{N}\Big) + \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big)\chi,
\end{aligned}
\tag{3}
$$

by the expert knowledge, where χ is the same as Equation (2) and s_3 is a threshold value. Then, we address the problem of testing the hypothesis $\mathcal{S}_2$ with respect to $\mathcal{S}_1$ by applying a technique of statistical model comparison.

Whereas the problem amounts to statistical point estimation and model selection, some difficulties remain to be resolved. For instance, in most cases, concentrations of the p entities $m_i(t)$ are directly unobservable, therefore, should be inferred from a given data set. Particularly, we only monitor up to quantities of the mRNAs by gene expression profiles, e.g. microarray data. Moreover, it is likely that true concentration values of mRNAs are also unmeasurable and to be estimated from data because gene expression profiles contain a variety of experimental noises to be filtered out from the analysis. This sometimes leads to occurrence of overfitting during the parameter estimation process, mainly, due to the fact that total number of the estimands, i.e. $\dim(\boldsymbol{\theta}) + pT$ is much greater than the number of observations $T\dim(\boldsymbol{y}_t)$. From this point of view, controlling overfitting is a key for success in the construction of a biologically meaningful pathway model.

In sequel, we present a statistical estimation technique based on a nonlinear state space model, e.g. [8], which is organized according to the following equations:

$$\boldsymbol{m}_t = \boldsymbol{f}(\boldsymbol{m}_{t-1}, \boldsymbol{w}_t, \boldsymbol{\theta}), \tag{4}$$

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{m}_t + \boldsymbol{\epsilon}_t. \tag{5}$$

Here, we use an abbreviation $\boldsymbol{m}(t) \equiv \boldsymbol{m}_t$. Equation (4) denotes a simulation devise for generating a series of concentrations of p entities. Note that the vector-valued function $\boldsymbol{f} = (f_1, \ldots, f_p)$, in its argument, contains a quasi-system noise $\boldsymbol{w}_t$ which follows a white noise process. Hence dynamics of the p entities turns to a stochastic process rather than deterministic one as the original HFPN. This aspect is a key for controlling degree of fitness in the parameter estimations as will be discussed in later. For example, the HFPN model of Equation (1) is transformed into Equation (4) by devising the multiplicative system noise as

$$\begin{aligned}
m_1\Big(t+\frac{n}{N}\Big) &= m_1\Big(t+\frac{n-1}{N}\Big) + \frac{k_1}{N} m_1\Big(t+\frac{n-1}{N}\Big) w_1\Big(t+\frac{n-1}{N}\Big),\\
m_2\Big(t+\frac{n}{N}\Big) &= m_2\Big(t+\frac{n-1}{N}\Big) + \frac{k_2}{N} m_2\Big(t+\frac{n-1}{N}\Big) w_2\Big(t+\frac{n-1}{N}\Big),\\
m_3\Big(t+\frac{n}{N}\Big) &= m_3\Big(t+\frac{n-1}{N}\Big) + \frac{k_3}{N} m_1\Big(t+\frac{n-1}{N}\Big) w_3\Big(t+\frac{n-1}{N}\Big)\\
&\quad - \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big) \chi w_5\Big(t+\frac{n-1}{N}\Big),\\
m_4\Big(t+\frac{n}{N}\Big) &= m_4\Big(t+\frac{n-1}{N}\Big) + \frac{k_4}{N} m_2\Big(t+\frac{n-1}{N}\Big) w_4\Big(t+\frac{n-1}{N}\Big)\\
&\quad - \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big) \chi w_5\Big(t+\frac{n-1}{N}\Big),\\
m_5\Big(t+\frac{n}{N}\Big) &= m_5\Big(t+\frac{n-1}{N}\Big) + \frac{k_5}{N} m_3\Big(t+\frac{n-1}{N}\Big) m_4\Big(t+\frac{n-1}{N}\Big) \chi w_5\Big(t+\frac{n-1}{N}\Big).
\end{aligned} \tag{6}$$

The system noises $w_i(t)$, $i = 1, \ldots, 5$ are here independently and identically distributed (i.i.d.) according to the log-normal distribution and are multiplicatively mixed into the simulation process. The system model Equation (4) represents time evolution of a probabilistic HFPN as like Equation (6) for $n = 1, \ldots, N$. The second equation (Equation (5)) describes a system of generating observed data $\boldsymbol{y}_t \in \mathcal{R}^d$ from concentrations of the p biological entities $\boldsymbol{m}_t \in \mathcal{R}^p$ with $\boldsymbol{H} \in \mathcal{R}^d \times \mathcal{R}^p$ and the observational noise $\boldsymbol{\epsilon}_t$ i.i.d. from the Gaussian distribution $N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$. As was mentioned in the above, it is a typical that $d < p$ because a part of expression values of the p entities is unobservable. Following this fact, the observation matrix H is determined in the following way: $(H)_{ij}$ takes value one if the expression value of the jth entity is observed by the ith element of the data vector $\boldsymbol{y}_t$, otherwise zero.

The probabilistic modeling of the HFPN enables us to handle the parameter estimation and the model selection within a unified framework of Bayesian statistical inference. Suppose that we construct a pathway model $\mathcal{S}$ with the probabilistic HFPN. According to the first-order Markov property of the

state space model, the joint distribution of concentrations of the p entities $\mathcal{M}_T = \{\boldsymbol{m}_1, \ldots, \boldsymbol{m}_p\}$ and the observations $\mathcal{Y}_T = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T\}$ are represented by

$$p(\mathcal{Y}_T, \mathcal{M}_T|\boldsymbol{\theta}) = p(\boldsymbol{m}_0) \prod_{t=1}^{T} p(\boldsymbol{y}_t|\boldsymbol{m}_t, \sigma^2) p(\boldsymbol{m}_t|\boldsymbol{m}_{t-1}, \boldsymbol{\theta}). \tag{7}$$

Here, the $p(\boldsymbol{m}_0)$ denotes the prior distribution of initial state of the p biological entities and, by definition, the $p(\boldsymbol{y}_t|\boldsymbol{m}_t, \sigma^2)$ corresponds to the observational model of the Gaussian distributions $N(\boldsymbol{H}\boldsymbol{m}_t, \sigma^2\boldsymbol{I})$. The system model with the parameters $\boldsymbol{\theta}$ defines the form of the density function $p(\boldsymbol{m}_t|\boldsymbol{m}_{t-1}, \boldsymbol{\theta})$.

In the Bayesian approach, before proceeding to the estimation of $\boldsymbol{\theta}$, we are required to set a prior distribution of the parameters $p(\boldsymbol{\theta})$. Given the prior, all statistical inferences are addressed based on the joint posterior distribution which is represented by

$$p(\mathcal{M}_T, \boldsymbol{\theta}|\mathcal{Y}_T) \propto p(\mathcal{Y}_T, \mathcal{M}_T|\boldsymbol{\theta}) p(\boldsymbol{\theta}) \tag{8}$$

up to the normalizing constant. For example, the model parameters and concentrations of the biological entities are conventionally estimated with the posterior expectations, e.g. $\mathrm{E}[\boldsymbol{\theta}|\mathcal{Y}_T]$ and $\mathrm{E}[\boldsymbol{m}_t|\mathcal{Y}_T]$, or the maximum a posteriori estimator $\mathrm{argmax}_{\theta,m} p(\mathcal{M}_T, \boldsymbol{\theta}|\mathcal{Y}_T)$. Unfortunately, exact computations of these Bayes estimators are essentially impossible except in very special scenarios, such as the linear-Gaussian dynamic system [4, 18] that are amenable to the Kalman smoother [5, 7] Since the HFPN yields the system model having a high nonlinearity, use of the computational intensive methods is indeed necessary to evaluate the Bayes estimators. In this paper, we present a sequential Monte Carlo method, known as the particle filtering algorithm [3, 9] which recursively generate a simulated sample from the joint posterior distribution Equation (8) to approximately compute the Bayes estimators of interest.

3.2 Particle Filtering

The particle filtering is a generic and simple computational approach to approximate the joint posterior distribution of the nonlinear state space models Equation (8) by the empirical distribution of the Monte Carlo sample which is referred to as the particles. The method proceed by follows:

- Set the observational variance parameter σ^2 at a certain value and draw a set of M particles, $\{\boldsymbol{\theta}^{(j)}\}_{j=1}^{M}$ and $\{\boldsymbol{m}_0^{(j)}\}_{j=1}^{M}$ from the prior distributions $p(\boldsymbol{\theta})$ and $p(\boldsymbol{m}_0)$.
- Repeat the following steps for $t = 1, \ldots, T$.
  - Generate the M system noises $\{\boldsymbol{w}_t^{(j)}\}_{j=1}^{M}$ i.i.d. from the log-normal distribution $q(\boldsymbol{w}_t)$.
  - Repeat simulations of the probabilistic HFPN at the N successive time points according to the system model

    $$\boldsymbol{m}_t^{(j)} = \boldsymbol{f}(\boldsymbol{m}_{t-1}^{(j)}, \boldsymbol{\theta}^{(j)}, \boldsymbol{w}_t^{(j)})$$

    for each particle $j = 1, \ldots, M$
  - Compute the importance weight of each particle by

    $$\alpha^{(j)} \propto p(\boldsymbol{y}_t|\boldsymbol{m}_t^{(j)}, \sigma^2),$$

    where these M weights are normalized so as to sum to one $\sum_{j=1}^{m} \alpha^{(j)} = 1$.
  - Update the current particles by resampling of $\{\boldsymbol{\theta}^{(j)}\}_{j=1}^{M}$ and $\{\boldsymbol{m}_t^{(j)}, \ldots, \boldsymbol{m}_0^{(j)}\}_{j=1}^{M}$ with probabilities $\alpha^{(1)}, \ldots, \alpha^{(M)}$.

The joint posterior distribution of the model parameters of HFPN is consistently approximated by the empirical distribution of the M particles

$$\hat{p}(\boldsymbol{\theta}|\mathcal{Y}_T) = \frac{1}{M}\sum_{j=1}^{M}\delta(\boldsymbol{\theta}-\boldsymbol{\theta}^{(j)}), \quad \forall\boldsymbol{\theta}\in\boldsymbol{\Theta}.$$

Under some mild conditions, it holds that the $\hat{p}(\boldsymbol{\theta}|\mathcal{Y}_T)$ is a consistent estimator of the posterior distribution $p(\boldsymbol{\theta}|\mathcal{Y}_T)$. Furthermore, one consistently evaluates the posterior expectation of the model parameters and the maximum a posteriori estimator by taking average and mode of the M particles, respectively, where in this study we estimate the model parameters with the approximated maximum a posteriori estimators $\hat{\boldsymbol{\theta}} \approx \boldsymbol{\theta}^* = \mathrm{argmax}_\theta p(\boldsymbol{\theta}|\mathcal{Y}_T)$.

3.3 Model Selection and Controlling Smoothness

One of the most important issues in the genomic data assimilation is to improve the current pathway model by exploiting the temporal structure of the real observations. Suppose that the expert knowledge hypothesizes a set of K pathway models $\mathcal{S}_1,\ldots,\mathcal{S}_K$. In the Bayesian statistical inference, a natural way it is that these hypotheses are tested based on the statistical evidences with the marginal likelihoods

$$\log p(\mathcal{Y}_T|\mathcal{S}_k,\sigma^2) = \log\int p(\mathcal{Y}_T,\mathcal{M}_T|\boldsymbol{\theta},\mathcal{S}_k,\sigma^2)p(\boldsymbol{\theta}|\mathcal{S}_k)\; d\boldsymbol{\theta}\; d\boldsymbol{m}_1\ldots d\boldsymbol{m}_T, \tag{9}$$

for $k=1,\ldots,K$. The evidences of the K pathway models are ordered according to magnitude of the scores $\log p(\mathcal{Y}_T|\mathcal{S}_k,\sigma^2)$, $k=1,\ldots,K$. Unfortunately, for the nonlinear state space model, analytic forms of the marginal likelihoods are usually unknown due to the intractable integrals in Equation (9). Therefore we have to exploit some numerical integral techniques. In this study, we apply an approximation method based on the particle filtering as

$$\log\hat{p}(\mathcal{Y}_T|\mathcal{S}_k,\sigma^2) = \sum_{t=1}^{T}\log\frac{1}{M}\sum_{j=1}^{M}p(\boldsymbol{y}_t|\boldsymbol{m}_t^{(j)},\sigma^2), \tag{10}$$

where the M particles are generated during the particle filtering of a pathway model $\mathcal{S}_k$. Under some regularity conditions, the approximated marginal likelihoods of the form Equation (10) tends to true one as T and M go to infinity. We assign Equation (10) to each hypothesis for assessing the statistical evidence.

In our approach, one more task to be addressed is the determination of the observational variance parameter σ^2 in Equation (5) which plays a key role in avoiding overfitting in the parameter estimation. It should be stressed here that in practice of the genomic data assimilation, amount of data available is often fairly small, for example, length of time course gene expression profiles is extremely short, typically, less than 20. Therefore, problem controlling overfitting is a key for success in constructing a realistic biological pathway model. To clarify the mechanism of the proposed parameter estimation, consider again the joint posterior distribution

$$p(\mathcal{M}_T,\boldsymbol{\theta}|\mathcal{Y}_T) \propto \prod_{t=1}^{T}p(\boldsymbol{y}_t|\boldsymbol{m}_t,\sigma^2)\prod_{t=1}^{T}p(\boldsymbol{m}_t|\boldsymbol{m}_{t-1},\boldsymbol{\theta})p(\boldsymbol{\theta}).$$

In this equation the term $p(\mathcal{M}_T,\boldsymbol{\theta}) = \prod_{t=1}^{T}p(\boldsymbol{m}_t|\boldsymbol{m}_{t-1},\boldsymbol{\theta})p(\boldsymbol{\theta})$ corresponds to the prior distribution whereas the $p(\mathcal{Y}_T|\boldsymbol{\theta},\mathcal{M}_T,\sigma^2) = \prod_{t=1}^{T}p(\boldsymbol{y}_t|\boldsymbol{m}_t,\sigma^2)$ represents the likelihood of the estimands $\boldsymbol{m}_1,\ldots,\boldsymbol{m}_T$ and $\boldsymbol{\theta}$ with respect the data. In usual, as the prior distribution becomes diffuse compared to the likelihood, the estimated pathway model tends to fit well to the observations, but the estimations are sometimes biased if amount of the data is small, and vice versa. The smoothness of

the estimated pathway model is controlled by tuning the observational noise variance σ^2 which determines the trade-off between goodness of fit and penalty for the parameter estimation. Particularly, as σ^2 becomes a small value, the estimated pathway model tends to be smooth. In this study, to control the smoothness, we apply an empirical Bayes method which chooses a variance parameter as to maximize the approximated marginal likelihood $l_{mg} = \log \hat{p}(\mathcal{Y}_T|\mathcal{S}_k, \sigma^2)$ in Equation (10) over σ^2, i.e. $\hat{\sigma}^2 = \arg\max_{\sigma^2} l_{mg}$.

4 Application and Result

We demonstrate the applicability of the proposed method with applications to the concentrations of the biological pathway model for the circadian rhythm with the HFPN in Figure 1. We generated a synthetic data by the following procedure: (i) generate a time-course of 12 entities $\boldsymbol{m}_{org,t}$ according to the HFPN with a fixed parameter vector $\boldsymbol{\theta}^{org}$ (see Table 4 and the solid lines in Figure 5). The system noises were set to the log-normal deviate $w_{it} = \exp(u_{it})$ where $u_{it} \sim \mathcal{N}(0, \tau^2)$. (ii) sample data points from the simulated time-course $\boldsymbol{m}_{org,t}$ with a fixed sampling interval Δ, and (iii) add a Gaussian white noise ($\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \boldsymbol{R}_{org})$), to the sampled data, where $\boldsymbol{R}_{org} = \sigma^2_{org}\boldsymbol{I}_{12}$. There is a restriction for observation data, that is, they must be positive because they represents concentrations of mRNAs or proteins. Thus, we generated the observation noise $\boldsymbol{\epsilon}_t$ again, if the resultant $\boldsymbol{y}_{syn,t}$ has negative elements. As a result we obtained the synthetic data set $\mathcal{Y}_{syn} = \{\boldsymbol{y}_{syn,t0,}, \boldsymbol{y}_{syn,t0+\Delta}, \boldsymbol{y}_{syn,t0+2\Delta}, \ldots, \boldsymbol{y}_{syn,t0+12\Delta}\}$ having 13 time points (see $\times$s in Figure 5). Figure 2 shows a schematic view of the relationship among the entities represented in the HFPN, the state space model, the simulated time-course and the synthetic time-course.

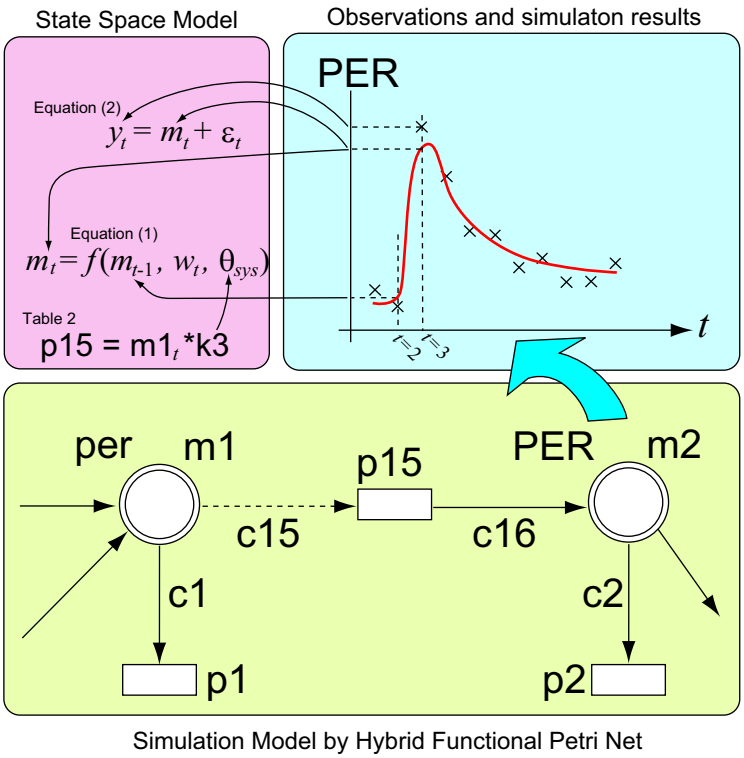

Figure 2: A schematic view of the relationship among the entities represented in the HFPN and the state space model, and the simulated and the observed (the synthetic) time-courses. See also Figure 5.

Table 4: Parameter values for $\boldsymbol{\theta}^{org}$ and σ_{org}.

Values in $\boldsymbol{\theta}^{org}$ and σ_{org}			
$m_1(0)$	0.2655	k_1	0.2
$m_2(0)$	0.5997	k_2	0.05
$m_3(0)$	0.2680	k_3	0.5
$m_4(0)$	0.9579	k_4	0.1
$m_5(0)$	1.9292	k_5	1.0
$m_6(0)$	0.2511	s_1	2.2
$m_7(0)$	1.2507	s_2	1.1
$m_8(0)$	2.4944	s_3	1.4
$m_9(0)$	1.5474	s_4	1.5
$m_{10}(0)$	0.2507	s_5	1.5
$m_{11}(0)$	0.1967	τ	0.1
$m_{12}(0)$	0.4545	σ_{org}	0.2

4.1 Parameter Estimation

The problem we consider here is to estimate unknown parameters in the simulation model and σ for the observation noise by using the method in Section 3. There are 24 parameters in total as shown in Table 4. In our preliminary examinations, we had found that it had been still difficult to estimate the all of them simultaneously. Thus, we tried to estimate a part of them with considering the other parameters as known using values in $\boldsymbol{\theta}^{org}$. There are three types of parameters in $\boldsymbol{\theta}^{org}$, i.e. the initial values for concentrations of mRNAs and proteins $(m_1(0), \ldots, m_{12}(0))$, the values to control process speeds $(k_1, \ldots, k_5)$, and the thresholds for transcriptional regulations $(s_1, \ldots, s_5)$. Since their roles in the model are different, it is meaningful to examine whether there are differences for difficulties of estimations depending on the types of parameters. We considered following four cases, in each case three parameters were blinded, i.e. (a) only initial parameters: $\boldsymbol{\theta}^a = [m_5(0), m_8(0), m_9(0)]'$, (b) only speed parameters: $\boldsymbol{\theta}^b = [k_2, k_3, k_5]'$, (c) only thresholds: $\boldsymbol{\theta}^c = [s_1, s_3, s_5]'$, and (d) mixed case (initial parameters and thresholds): $\boldsymbol{\theta}^d = [m_6(0), m_7(0), s_2]'$. In each case, we also assumed σ was unknown. Using the synthetic data $\boldsymbol{y}_{syn}$, we estimated parameters in the four cases. As a result we obtained fairly good estimations for all cases. Figure 3 shows profiles of the approximated marginal likelihood l_{mg} (Equation (10)) as a function of the standard deviation σ with boxplots. In order to evaluate the accuracy of the estimations, we repeated the same experiments nine times for each σ. For all cases the l_{mg}s took maximum value at $\hat{\sigma} = 0.2$, this is the same as σ_{org}. Figure 4 shows profiles of the posterior distributions of the parameters, $P(\boldsymbol{\theta}^{\dagger}|\mathcal{Y}_{syn}, \hat{\sigma})$, where $\dagger \in \{a, b, c, d\}$. We can observe that each of the profile has the mode near the corresponding value in $\boldsymbol{\theta}^{org}$. Note that each of the plots shows a continuous density function, however, they are only for clear presentations, which was made by fitting a continuous function to the actual results represented by discrete particles having the same weight. Running the simulation model with the estimated parameters, we made predicted time-series, i.e. a long-term prediction without filtering. Figure 5 shows the results. We see that all cases of the experiments show the good prediction.

4.2 Model Selection

In order to examine the applicability of the proposed model selection procedure, we compared the maximum marginal likelihoods obtained from different pathway models for the circadian rhythm. We hypothesized four slightly different models, i.e. Model1 ($\mathcal{S}_1$), Model2 ($\mathcal{S}_2$), Model3 ($\mathcal{S}_3$), and Model4 ($\mathcal{S}_4$). Model2 is the original model. The HFPN is the same in Figure 1. The unknown parameters is

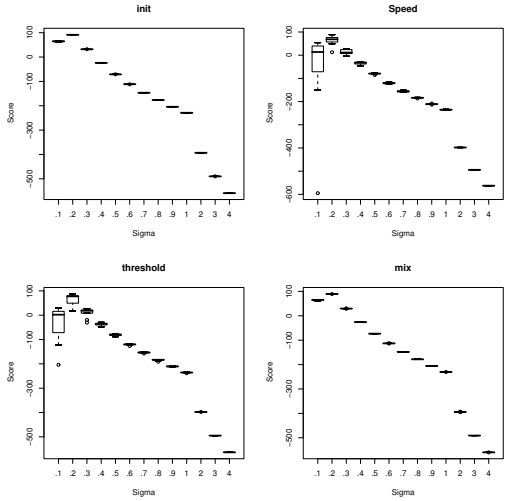

Figure 3: The approximated likelihood l_{mg} as a function of standard deviation parameter σ for the cases to estimate (a) initial parameters $\boldsymbol{\theta}^a = [m_5(0), m_8(0), m_9(0)]'$ (top left); (b) speed parameters $\boldsymbol{\theta}^b = [k_2, k_3, k_5]'$ (top right); (c) thresholds $\boldsymbol{\theta}^c = [s_1, s_3, s_5]'$ (bottom left); and (d) mixed case $\boldsymbol{\theta}^d = [m_6(0), m_7(0), s_2]'$ (bottom right).

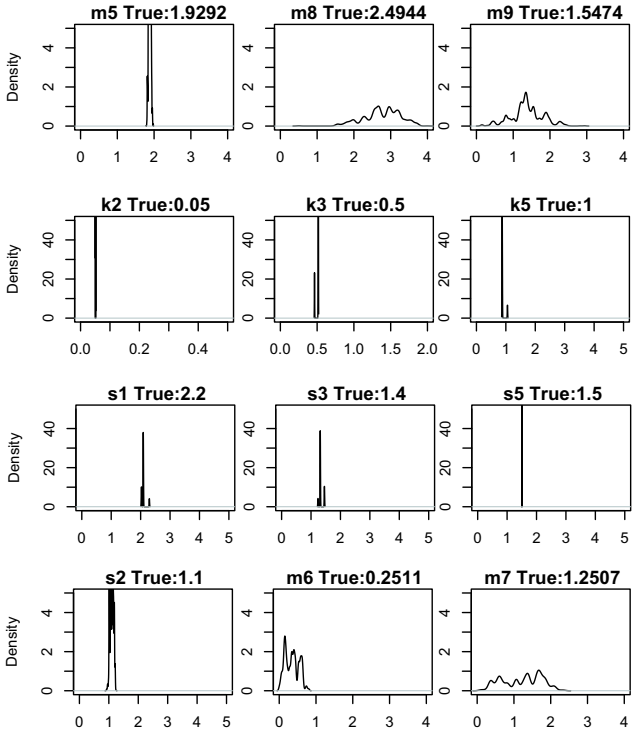

Figure 4: Posterior distributions of parameters ($P(\boldsymbol{\theta}^\dagger|\mathcal{Y}_{syn}, \hat{\sigma})$), where $\dagger \in \{a, b, c, d\}$. See note in the text. The panels on the first, the second, the third, and the fourth rows are for $\boldsymbol{\theta}^a$, $\boldsymbol{\theta}^b$, $\boldsymbol{\theta}^c$, and $\boldsymbol{\theta}^d$, respectively.

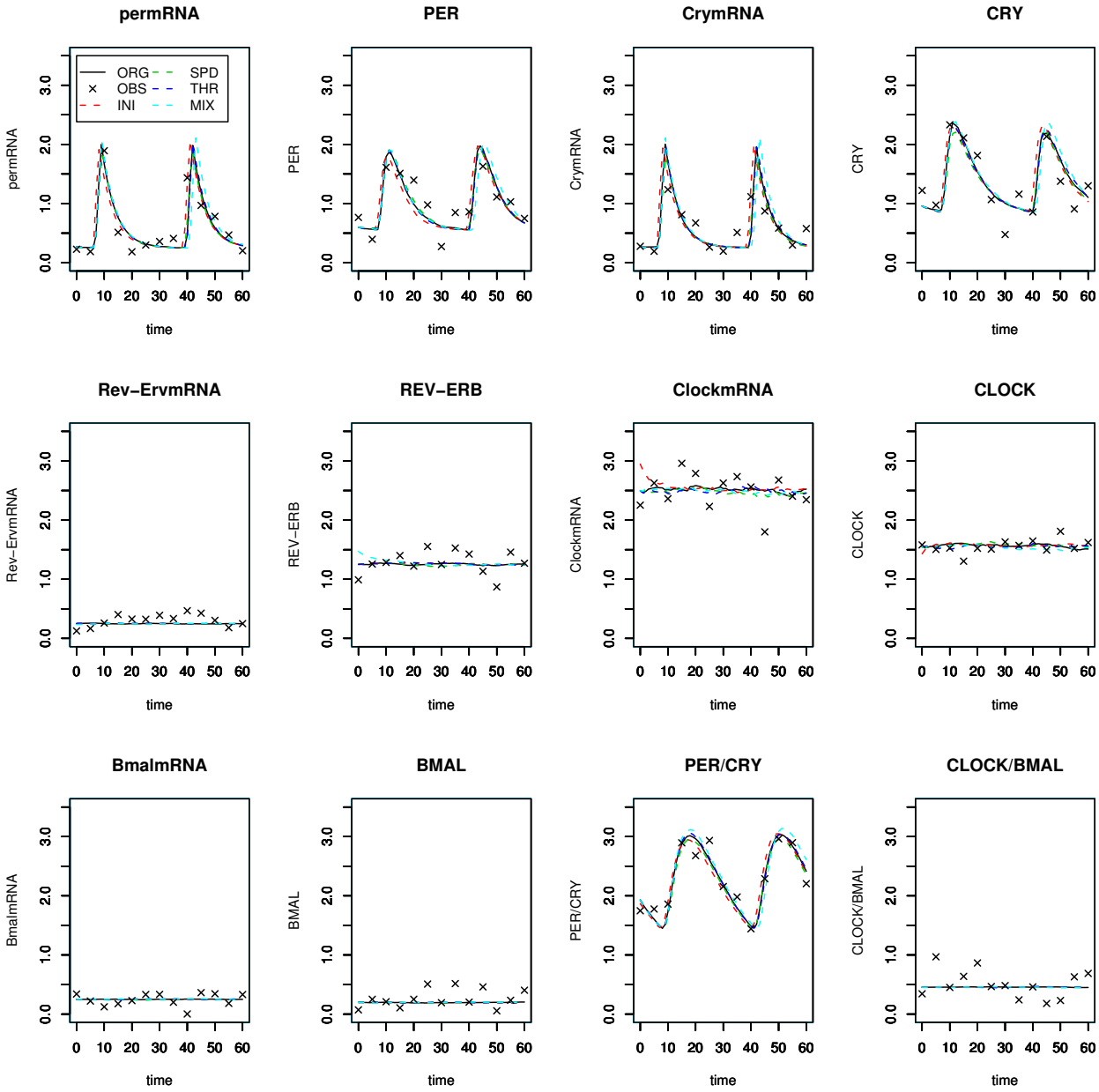

Figure 5: The predicted time-courses of the 12 entities ($\boldsymbol{m}_t$) in the HFPN with the estimated parameters for the four cases: (a) with the initial parameters $\boldsymbol{\theta}^a$ (red dotted lines), (b) with the speed parameters $\boldsymbol{\theta}^b$ (green dotted lines), (c) with the threshold parameters $\boldsymbol{\theta}^c$ (blue dotted lines), and (d) with the two initial parameters and one threshold parameter $\boldsymbol{\theta}^d$ (cyan dotted lines). The simulated time-course $\boldsymbol{m}_{org,t}$ with the parameters $\boldsymbol{\theta}^{org}$ (solid lines) and the synthetic data, $\mathcal{Y}_{syn}$ (×s), are also shown as references.

assumed as $\boldsymbol{\theta}^{Model2} = [m_5(0), s_3]'$ and σ. Model1 is a subtractive case, which is deleted a connector (c40 in Figure 1) representing the rev-erv inhibitory regulation with the threshold s_3 from the original HFPN model. For this model, we assumed that the unknown parameters are only $\boldsymbol{\theta}^{Model1} = [m_5(0)]$ and σ. Model3 is an additive case. For this model we added to the original model a connector from m_7 to $p17$ to represent a cry active regulation with a new threshold parameter s_6. Then the unknown parameters are assumed as $\boldsymbol{\theta}^{Model3} = [m_5(0), s_3, s_6]'$ and σ. Finally, Model4 is also additive case and the same places of the Model3 are connected but by an inhibitory connector, which represents a cry inhibitory regulation with a new threshold parameter s_7. We assumed the unknown parameters for this model as $\boldsymbol{\theta}^{Model4} = [m_5(0), s_3, s_7]'$ and σ.

We applied these models to the same data set $\mathcal{Y}_{syn}$ and estimated the parameters and obtained the maximum marginal likelihood $l_{mg}^{\max} = \log \hat{p}(\mathcal{Y}_{syn}|\mathcal{S}_k, \hat{\sigma}^2)$ for each model $(k = 1, \ldots, 4)$. The profiles of l_{mg} as a function of σ for Model2, Model3, and Model4 were correctly took the maximum at $\sigma = 0.2$, like those in Figure 3 (not shown). On the other hand, for Model1 the subtractive model, the l_{mg} do not show the maximum but shows monotonically increasing as σ increases (not shown). Then we compared the $l_{mg}^{\max}$s obtained from each model. The best result we should expect is the l_{mg}^{max} obtained from Model2 becomes the largest due to the synthetic data was originally generated from the HFPN of Model2 with θ^{org}. As shown in Figure 6, the result is consistent with the expectation.

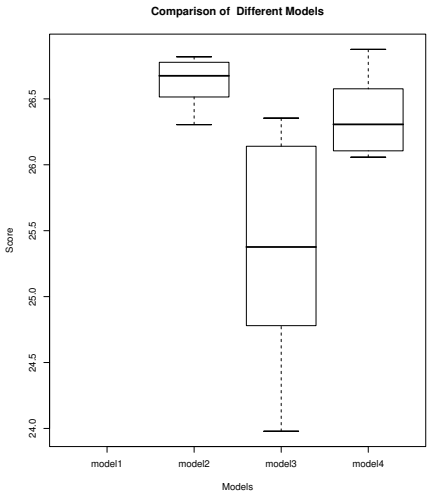

Figure 6: The maximum approximated marginal likelihood $l_{mg}^{\max}$ for Model1 (subtractive model), Model2 (original one), Model3 (additive one with an active regulator), and Model4 (additive one with an inhibitory regulator). The $l_{mg}^{\max}$ for Model1 ran off the lower edge.

5 Concluding Remarks

In this study, we applied DA approach for the first time, to a biological simulation model to solve parameter estimation and model selection problems. DA is a concept to combine a simulation model and data which can be formulated as a state estimation problem by nonlinear state space models (SSMs). Thus, we can employ effective statistical methods to solve these problems. We explained an estimation method called particle filter, based on sequential Monte Carlo simulations and method to select better models by comparing the approximated maximum marginal likelihood l_{mg}^{max} in Equation (10).

In order to investigate the applicability of the method to biological simulation models, we carried out several numerical experiments using the circadian rhythm model with HFPN. For parameter estimation problems, this method can successfully estimate the parameters regardless of the types of

parameters as shown in Section 4.1. Although the number of parameters are currently somewhat small, this property should be helpful. Because some parameter estimation problems in biological pathways, most parts of the parameters can be tuned manually with biological background knowledge except for a small number of the parameters. For the model selection problem, we considered four models with different HFPNs and applied the proposed method. The models were hypothesized by considering possible situations for model buildings, in which slightly different models were need to be compared. With the comparison of the l_{mg}^{max} scores that are obtained from these models, we could choose the true model as the best one. It suggests that we can evaluate the goodness of models more objectively. It may enhance interactions between experimental biologists and computational biologists.

Taking into account these results, we can foresee the next problems to be challenged. The first improvement is to increase the number of parameters to be estimated simultaneously. It might be achieved by improving the algorithm of particle filter to more accurate estimation. The second improvement is to apply these methods for real experimental data *in vivo* and *vitro*. Although we did not focused in this paper, in the real data case, estimating state vectors and extract information from them will be more important. We note that the methods explained in this paper can be applied not only to HFPN (HFPNe) but also to the other simulation platforms, e.g. differential equations.

DA has developed in the area of geophysics. In that area, the problems are much different in many regards comparing with biology, e.g. the scale of simulation, the amount of data. Especially for the biological problem, the length of time courses are much shorter than that of other fields of DA. Overcoming these problems, we can expect to device a powerful tool such as to discover insightful biological knowledge and to forecast reactions of the biological system to the drags and chemical compounds in more plausible way based on the combined information from the both simulation models and observation data.

References

[1] Doi, A., Fujita, S., Matsuno, H., Nagasaki, M., and Miyano., S., Constructing biological pathway models with hybrid functional Petri nets, *In Silico Biol.*, 4:271–291, 2004.

[2] Doi, A., Nagasaki, M., Matsuno, H., and Miyano, S., Simulation-based validation of the p53 transcriptional activity with hybrid functional Petri net, *In Silico Biol.*, 6:0001, 2006.

[3] Gordon, N. J., Salmond, D. J., and Smith., A. F. M., Nobel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings-F*, 140(2):107–113, 1993.

[4] Harvey, A. C., *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, New York, 1989.

[5] Kalman, R. E., A new approach to linear filtering and prediction problems, *Trans. Amer. Soc. Mech. Eng., J. Basic Engineering*, 82:35–45, 1960.

[6] Kikuchi, S., Tominaga, D., Arita, M., Takahashi, K., and Tomita, M., Dynamic modeling of genetic networks using genetic algorithm and S-system, *Bioinformatics*, 19(6):643–650, 2003.

[7] Kitagawa, G. and Gersch, W., A smoothness priors-state space modeling of time series with trend and seasonality. *J. Amer. Statist. Assoc.*, 79(386):378–389, 1984.

[8] Kitagawa, G., Non-Gaussian state-space modeling of nonstationary time series, *J. Amer. Statist. Assoc.*, 82(400):1032–1063, 1987.

[9] Kitagawa, G., Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *J. Comput. Graph. Statist.*, 5:1–25, 1996.

[10] Matsuno, H., Inouye, S.-I. T., Okitsu, Y., and Fujii, Y., A new regulatory interactions suggested by simulations for circadian genetic control mechanism in mammals, *J. Bioinform. Comp. Biol.*, 4(1):139–153, 2006.

[11] Matsuno, H., Tanaka, Y., Aoshima, H., Doi, A., Matsui, M., and Miyano, S., Biopathways representation and simulation on hybrid functional Petri net, *In Silico Biol.*, 3:389–404, 2003.

[12] Nagasaki, M., Doi, A., Matsuno, H., and Miyano, S., Integrating biopathway databases for large-scale modeling and simulation, In: The Second Asia-Pacific Bioinformatics Conference, *Conferences in Research and Practice in Information Technology*, Australian Computer Society, 29:43–52, 2004.

[13] Nagasaki, M., Doi, A., Matsuno, H., and Miyano. S., A versatile Petri net based architecture for modeling and simulation of complex biological processes, *Genome Inform.*, 15(1):180–197, 2004.

[14] Nakano, S., Ueno, G., Ebihara, Y., Fok, M.-C., Ohtani, S., Brandt, P. C., and Higuchi, T., Data assimilation of global ENA data to estimate the ring current structure and the inner-magnetospheric electric field, (submitted).

[15] Savageau, M. A., Biochemical systems analysis II: The steady state solution for an n-pool system using a power law approximation, *J. Theor. Biol.*, 25:370–379, 1969.

[16] Tominaga, D., Okamoto, M., Maki, Y., Watanabe, S., and Eguchi, Y., Nonlinear numerical optimization technique based on a genetic algorithm for inverse problems: Towards the inference of genetic networks, *Proc. German Conf. Bioinformatics (GCB '99)*, 127–140, 1999.

[17] Ueno, G., Higuchi, T., Kagimoto, T., and Hirose, N., Application of the ensemble Kalman filter and smoother to a coupled atmosphere-ocean model, (submitted).

[18] West, M. and Harrison, J., *Bayesian Forecasting and Dynamic Models*, 2nd ed., Springer-Verlag, New York, 1997.

[19] Wunsch, C., *The Ocean Circulation Inverse Problem*, Cambridge University Press, Cambridge, 1996.

[20] `http://www.daimi.au.dk/PetriNets/tools/`