Model Predictive Control of Electric Power Systems Based on Gaussian Process Predictors

Tomohiro Hachino, Hitoshi Takata, Seiji Fukushima, and Yasutaka Igarashi
Kagoshima University, Kagoshima, Japan
Email: {hachino, takata, fukushima, igarashi}@eee.kagoshima-u.ac.jp

***Abstract*—This paper presents a model predictive control of electric power systems based on the multiple Gaussian process predictors. The Gaussian process model is a nonparametric model and the output of the model has Gaussian distribution with mean and variance. The multistep ahead predictors for the phase angle in transient state of the electric power system are formed by training multiple Gaussian process models in accordance with the direct approach. Based on these predictors, model predictive control is accomplished, where the input signal is optimized so that the error between the predicted future output and the reference signal becomes small taking the uncertainty of the predicted future output into account. Simulation results for a simplified electric power system are shown to illustrate the effectiveness of the proposed model predictive control.**

***Index Terms*—model predictive control, electric power system, multistep ahead prediction, Gaussian process model, direct method**

I. Introduction

In recent years, model predictive control has been widely applied to both process systems and servo systems [1]-[5]. In general, the performance of model predictive control greatly depends on the accuracy of the model used for prediction. The Gaussian process (GP) model is one of the attractive models for multistep ahead prediction. The GP model is a nonparametric model and fits naturally into Bayesian framework [6]-[8].This model has recently attracted much attention for system identification [9], [10], time series forecasting [11]-[13], and predictive control [3], [14], [15]. Since the GP model gives us not only the mean value but also the variance of the conditionally expected value of the output, it is useful for model predictive control considering the uncertainty of model. Moreover, the GP model has far fewer parameters to describe the nonlinearity than the parametric models such as radial basis function model, neural network model, and fuzzy model.

This paper proposes a novel model predictive control of electric power systems based on the multiple GP predictors. The concept of the multiple GP predictors for electric power systems was presented in [16]. In this approach, multistep ahead prediction for the phase angle in transient state of the electric power system is directly performed by using the multiple trained GP models as every step ahead predictor. Since this direct method uses not only one-step ahead predictor but also all-step ahead predictors, the prediction errors are not accumulated so much as the prediction horizon increases. Moreover, these multiple predictors give the predictive values of the phase angle and uncertainty of the predictive values as well. In the stage of control, the input signal which is the increment of excitation voltage is optimized so that the error between the predicted future output (phase angle) and the reference signal becomes small taking the uncertainty of the predicted future output into account. The information about uncertainty of the predicted future output is used as a constraint. This means that the input signal which causes large uncertainty of the predicted future output is excluded in the optimization. In this paper, the genetic algorithm (GA) [17] is applies for this constrained optimization problem.

This paper is organized as follows. In section II, the objective electric power system is described. In section III, the multiple GP prior models are derived for every step ahead predictors, and the direct multistep ahead prediction is given using the GP posterior distribution. In section VI, model predictive control based on the multiple GP predictors is presented using the GA. In section V, simulation results are shown to illustrate the effectiveness of the proposed model predictive control. Finally, conclusions are given in section VI.

II. Electric Power Systems

Consider a single machine power system described by

$$\begin{cases} \widetilde{M}\ddot{\delta}(t) + \widetilde{D}\dot{\delta}(t) + P_e = P_{in} \\ P_e = \dfrac{e_t E_{fd}}{X_e}\left(1 + \Delta E_{fd}(t)\right)\sin\delta(t) \\ y(t) = \delta(t) + v(t) \end{cases} \tag{1}$$

where $\delta(t)$: phase angle, $y(t)$: phase angle corrupted by the measurement noise $v(t)$, $\Delta E_{fd}(t)$: increment of excitation voltage, $\widetilde{M}$: inertia coefficient, $\widetilde{D}$: damping coefficient, P_e: generator output power, P_{in}: turbine output power, E_{fd}: excitation voltage, e_t: infinite bus voltage, and X_e: system impedance. The measurement noise $v(t)$ is zero mean white Gaussian noise with variance σ^2.Itis assumed that the input $u(k) = \Delta E_{fd}(kT_s)$ and the noisy measurement of the output $y(k) =$

Manuscript received July 1, 2014; revised November 21, 2014.

©2015 Engineering and Technology Publishing

doi: 10.12720/joace.3.5.418-424

$y(kT_s)$at $t = kT_s$are available when the multistep ahead predictors are trained, where T_s is the samplingperiod.

III. Multistep ahead Prediction

There are two approaches to multistep ahead prediction. One is the direct method that makes multistep ahead prediction directly by using a specific step ahead predictor. The other is the iterated method that repeats one-step ahead prediction up to the desired step. The iterated multistep ahead predictions with propagation of the prediction uncertainty based on the GP model were presented in [11], [12].Although the computational burden of this approach is not so heavy during the training phase, unacceptable prediction errors are gradually accumulated as the prediction horizon increases especially in the presence of measurement noise. Therefore, in this paper, the direct approach is considered [16].

The problem of multistep ahead prediction is usually to estimate the future outputs given the past input and output data. The optimal predictor can be written as

$$\hat{y}(k+j) = \mathrm{E}[y(k+j)|\boldsymbol{x}(k)] \tag{2}$$

where $\mathrm{E}[\cdot]$ is the expectation operator, and

$$\begin{aligned}\boldsymbol{x}(k) = \big[&y(k), y(k-1), \cdots, y\big(k-L_y+1\big),\\ &u(k), u(k-1), \cdots, u(k-L_u+1)\big]^{\mathrm{T}}\end{aligned} \tag{3}$$

is the state vector consisting of the past outputs and inputs up to the prespecified lags$L_y$and L_u. Actually, with the GP framework, not only estimates $\hat{y}(k+j)$but also its uncertainty,i.e.,thevariance $\hat{\sigma}^2(k+j)$ are estimated.Therefore, the problem here is to construct the following probability distributions for the multistep ahead prediction

$$\begin{aligned}y(k+j)|\boldsymbol{x}(k) \sim N\big(\hat{y}(k+j), \hat{\sigma}^2(k+j)\big)\\ (j = 1,2,\cdots,M)\end{aligned} \tag{4}$$

And to carry out multistep ahead prediction up to *M* step based on these distributions, by using the GP framework.

A. Derivation of GP Prior Models

Consider a *j*-step ahead predictor as

$$\begin{aligned}y(k+j) = f_j\big(\boldsymbol{x}(k)\big) + \varepsilon_j(k)\\ (j = 1,2,\cdots,M)\end{aligned} \tag{5}$$

where $f_j(\cdot)$is a function which is assumed to be stationary and smooth. $\varepsilon_j(k)$ is zero mean Gaussian noise with unknown variance σ_j^2.

Putting $k = k_s,\ k_s+1, \cdots,\ k_s+N-1$ on (5) yields

$$\boldsymbol{y}_j = \boldsymbol{f}_j + \boldsymbol{\varepsilon}_j \tag{6}$$

where

$$\begin{aligned}\boldsymbol{y}_j &= [y(k_s+j), y(k_s+j+1),\ \cdots, y(k_s+j+N-1)]^{\mathrm{T}}\\ \boldsymbol{f}_j &= \big[f_j(\boldsymbol{x}_1), f_j(\boldsymbol{x}_2), \cdots, f_j(\boldsymbol{x}_N)\big]^{\mathrm{T}}\\ \boldsymbol{\varepsilon}_j &= \big[\varepsilon_j(k_s), \varepsilon_j(k_s+1), \cdots, \varepsilon_j(k_s+N-1)\big]^{\mathrm{T}}\end{aligned} \tag{7}$$

$$\begin{aligned}\boldsymbol{X} &= [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N]^{\mathrm{T}}\\ &= [\boldsymbol{x}(k_s), \boldsymbol{x}(k_s+1), \cdots, \boldsymbol{x}(k_s+N-1)]^{\mathrm{T}}\end{aligned}$$

$\boldsymbol{y}_j$ and $\boldsymbol{f}_j$ are the vector of model outputs and the vector of function values for the *j*-step ahead predictor, respectively. $\boldsymbol{X}$ is the model input matrix and is common for every step ahead predictors. $\{\boldsymbol{X}, \boldsymbol{y}_j\}$ is the training input and output data for the*j*-step ahead predictor.

A GP is a Gaussian random function and is completely described by its mean function and covariance function. We can regard it as a collection of random variables which has joint multivariable Gaussian distribution. Therefore, the vector of function values $\boldsymbol{f}_j$ can be represented by the GP as

$$\boldsymbol{f}_j \sim N\big(\boldsymbol{m}_j(\boldsymbol{X}), \boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{X})\big) \tag{8}$$

where $\boldsymbol{m}_j(\boldsymbol{X})$ is the*N*-dimensional mean function vector and $\boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{X})$ is the*N*-dimensional covariance matrix evaluated at all pairs of the training input data. Equation (8) means that $\boldsymbol{f}_j$ has a Gaussian distribution with the mean function vector $\boldsymbol{m}_j(\boldsymbol{X})$ and the covariance matrix $\boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{X})$.

The mean function is often represented by a polynomial regression [8]. In this paper, the mean function vector $\boldsymbol{m}_j(\boldsymbol{X})$ is expressed by the first order polynomial, i.e., a linear combination of the model input:

$$\boldsymbol{m}_j(\boldsymbol{X}) = \big[m_j(\boldsymbol{x}_1), m_j(\boldsymbol{x}_2), \cdots, m_j(\boldsymbol{x}_N)\big]^{\mathrm{T}} = \tilde{\boldsymbol{X}}\boldsymbol{\theta}_j \tag{9}$$

where $\tilde{\boldsymbol{X}} = [\boldsymbol{X}, \boldsymbol{e}]$,and $\boldsymbol{e} = [1,1,\cdots,1]^{\mathrm{T}}$is the *N*-dimensional vector consisting of ones, and $\boldsymbol{\theta}_j = \big[\theta_{jo}, \theta_{j1}, \cdots, \theta_{j(L_y+L_u)}\big]^{\mathrm{T}}$ is the unknown weighting parameter vector of the mean function to be trained.

The covariance matrix $\boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{X})$is constructed as

$$\boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{X}) = \begin{bmatrix}\Sigma_{j(1,1)} & \cdots & \Sigma_{j(1,N)}\\ \vdots & \ddots & \vdots\\ \Sigma_{j(N,1)} & \cdots & \Sigma_{j(N,N)}\end{bmatrix} \tag{10}$$

where the element $\Sigma_{j(p,q)} = \mathrm{cov}\big(f_j(\boldsymbol{x}_p), f_j(\boldsymbol{x}_q)\big) = s_j(\boldsymbol{x}_p, \boldsymbol{x}_q)$ is a function of $\boldsymbol{x}_p$ and $\boldsymbol{x}_q$. Under the assumption that the process is stationary and smooth, the following Gaussian kernel is utilized for $\Sigma_{j(p,q)}$:

$$\begin{aligned}\Sigma_{j(p,q)} &= s_j(\boldsymbol{x}_p, \boldsymbol{x}_q)\\ &= \rho_j^2 \exp\left(-\frac{\|\boldsymbol{x}_p - \boldsymbol{x}_q\|^2}{2\ell_j^2}\right)\end{aligned} \tag{11}$$

where ρ_j^2 is the signal variance, ℓ_j is the length scale, and $\|\cdot\|$ denotes the Euclidean norm.The free parameters ρ_j and ℓ_j of (11) and the noise standard deviation σ_j are called *hyperparameters* and construct the hyperparameter vector $\boldsymbol{h}_j = \big[\rho_j, \ell_j, \sigma_j\big]^{\mathrm{T}}$. ρ_j can control the overall variance of the random function $f_j(\cdot)$ and determines the magnitude of the function $f_j(\cdot)$. ℓ_j can change the characteristic length scale so that the axis about the model inputchanges.

Since $\boldsymbol{y}_j$ is noisy observation, we have the following GP model for *j*-step ahead prediction from (6) and (8) as

©2015 Engineering and Technology Publishing

$$\boldsymbol{y}_j \sim N\big(\boldsymbol{m}_j(\boldsymbol{X}), \boldsymbol{K}_j(\boldsymbol{X},\boldsymbol{X})\big) \tag{12}$$

where

$$\begin{aligned} &\boldsymbol{K}_j(\boldsymbol{X},\boldsymbol{X}) = \boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{X}) + \sigma_j^2 \boldsymbol{I}_N \\ &\boldsymbol{I}_N\text{: } N \times N \text{ identity matrix} \end{aligned} \tag{13}$$

In the following, $\boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{X})$ and $\boldsymbol{K}_j(\boldsymbol{X},\boldsymbol{X})$ are written as $\boldsymbol{\Sigma}_j$ and $\boldsymbol{K}_j$, respectively.

B. Training of GP Prior Models

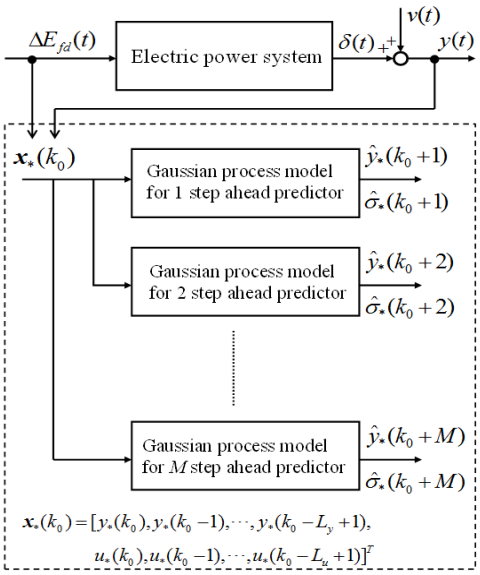


Figure 1. The multistep ahead prediction scheme

To perform multistep ahead prediction, the proposed direct approach needs1to M step ahead prediction modelsas shown in Fig. 1. The accuracy of prediction greatly depends on the unknown parameter vector $\boldsymbol{\vartheta}_j = [\boldsymbol{\theta}_j^{\mathrm{T}}, \boldsymbol{h}_j^{\mathrm{T}}]^{\mathrm{T}}$ and therefore $\boldsymbol{\vartheta}_j$ has to be optimized.This training is carried out by minimizing the negative log marginal likelihood of the training data:

$$\begin{aligned} J(\boldsymbol{\vartheta}_j) &= -\log p(\boldsymbol{y}_j|\boldsymbol{X},\boldsymbol{\vartheta}_j) \\ &= \frac{1}{2}\log|\boldsymbol{K}_j| + \frac{1}{2}\big(\boldsymbol{y}_j - \boldsymbol{m}_j(\boldsymbol{X})\big)^{\mathrm{T}} \boldsymbol{K}_j^{-1} \\ &\quad \times \big(\boldsymbol{y}_j - \boldsymbol{m}_j(\boldsymbol{X})\big) + \frac{N}{2}\log(2\pi) \\ &= \frac{1}{2}\log|\boldsymbol{K}_j| + \frac{1}{2}\big(\boldsymbol{y}_j - \tilde{\boldsymbol{X}}\boldsymbol{\theta}_j\big)^{\mathrm{T}} \boldsymbol{K}_j^{-1}\big(\boldsymbol{y}_j - \tilde{\boldsymbol{X}}\boldsymbol{\theta}_j\big) \\ &\quad + \frac{N}{2}\log(2\pi) \end{aligned} \tag{14}$$

Since the cost function $J(\boldsymbol{\vartheta}_j)$ generally has multiple local minima, this training problem becomes a nonlinear optimization one. However, we can separate the linear optimization part and the nonlinear optimization part for this optimization problem. The partial derivative of (14) with respect to the weighting parameter vector $\boldsymbol{\theta}_j$ of the mean function is as follows:

$$\frac{\partial J(\boldsymbol{\vartheta}_j)}{\partial \boldsymbol{\theta}_j} = -\tilde{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{K}_j^{-1}\boldsymbol{y}_j + \tilde{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{K}_j^{-1}\tilde{\boldsymbol{X}}\boldsymbol{\theta}_j \tag{15}$$

Note that if the hyperparameter vector $\boldsymbol{h}_j$ of the covariance function is given, then the weighting parameter $\boldsymbol{\theta}_j$ can be estimated by the linear least-squares method putting $\partial J(\boldsymbol{\vartheta}_j)/\partial \boldsymbol{\theta}_j = \boldsymbol{0}$:

$$\boldsymbol{\theta}_j = \big(\tilde{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{K}_j^{-1}\tilde{\boldsymbol{X}}\big)^{-1}\tilde{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{K}_j^{-1}\boldsymbol{y}_j \tag{16}$$

However, even if the weighting parameter vector $\boldsymbol{\theta}_j$ is known, the optimization with respect to hyperparametervector $\boldsymbol{h}_j$ is a complicated nonlinear problem and might suffer from the local minima problem. Therefore, the unknown parameter vector $\boldsymbol{\vartheta}_j$ is determined by the separable least-squares (LS) approachcombining the linear LS method and the GA [17], as

$$\begin{aligned} \boldsymbol{\vartheta}_{j[best]} &= \big[\boldsymbol{\theta}_{j[best]}^{\mathrm{T}}, \boldsymbol{h}_{j[best]}^{\mathrm{T}}\big]^{\mathrm{T}} \\ &= \big[\boldsymbol{\theta}_{j[best]}^{\mathrm{T}}, \rho_{j[best]}, \ell_{j[best]}, \sigma_{j[best]}\big]^{\mathrm{T}}. \end{aligned}$$

C. Multistep ahead Prediction by GP Posterior

We have already obtained the GP prior models for j (j = 1, 2,…M) step ahead predictors. In the direct approach, multistep ahead prediction up to M step is carried out directly using every GP prior models as shown in Fig. 1.

For a new given test input

$$\boldsymbol{x}_* = \boldsymbol{x}_*\ k_0\ = \ y_*\ k_0\ , y_*\ k_0 - 1\ , \cdots, y_*\ k_0 - L_y + 1\big), u_*(k_0), u_*(k_0-1), \cdots, u_*(k_0 - L_u + 1)\big]^{\mathrm{T}}$$

And corresponding test output $y_*(k_0 + j)$ $(j = 1,2,\cdots,M)$, we have the followingjointGaussian distribution:

$$\begin{bmatrix} \boldsymbol{y}_j \\ y_*(k_0+j) \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{m}_j(\boldsymbol{X}) \\ m_j(\boldsymbol{x}_*) \end{bmatrix}, \begin{bmatrix} \boldsymbol{K}_j & \boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{x}_*) \\ \boldsymbol{\Sigma}_j(\boldsymbol{x}_*,\boldsymbol{X}) & s_j(\boldsymbol{x}_*,\boldsymbol{x}_*) + \sigma_{j[best]}^2 \end{bmatrix}\right) \quad (j = 1,2,\cdots,M) \tag{17}$$

where k_0 is the starting step for prediction, and $\boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{x}_*) = \boldsymbol{\Sigma}_j^{\mathrm{T}}(\boldsymbol{x}_*,\boldsymbol{X})$ is the N-dimensional covariance vector evaluated at all pairs of the training and test data. $s_j(\boldsymbol{x}_*,\boldsymbol{x}_*)$ is the variance of the test data. $\boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{x}_*)$ and $s_j(\boldsymbol{x}_*,\boldsymbol{x}_*)$ are calculated by the trained covariance function.

From the formula for conditioning a joint Gaussian distribution [18], the posterior distribution for a specific test data is

$$\begin{aligned} &y_*(k_0+j)|\boldsymbol{X},\boldsymbol{y}_j,\boldsymbol{x}_* \sim N(\hat{y}_*(k_0+j), \hat{\sigma}_*^2(k_0+j)) \\ &(j = 1,2,\cdots,M) \end{aligned} \tag{18}$$

where

$$\begin{aligned} \hat{y}_*(k_0+j) &= m_j(\boldsymbol{x}_*) + \boldsymbol{\Sigma}_j(\boldsymbol{x}_*,\boldsymbol{X})\boldsymbol{K}_j^{-1}(\boldsymbol{y}_j - \boldsymbol{m}_j(\boldsymbol{X})) \\ \hat{\sigma}_*^2(k_0+j) &= s_j(\boldsymbol{x}_*,\boldsymbol{x}_*) - \boldsymbol{\Sigma}_j(\boldsymbol{x}_*,\boldsymbol{X})\boldsymbol{K}_j^{-1}\boldsymbol{\Sigma}_j(\boldsymbol{X},\boldsymbol{x}_*) \\ &\quad + \sigma_{j[best]}^2 \end{aligned} \tag{19}$$

©2015 Engineering and Technology Publishing

Are the predictive mean and the predictive variance at the j-step ahead, respectively.

IV. Model Predictive Control

In this section, model predictive control is carried out using the multiple GP predictors described in the previous section. Model predictive control is generally based on the receding horizon strategy. At the current time step k_0, the optimal input signal on the control horizon $[k_0, k_0 + N_u - 1]$ is determined so that the error between the predicted future output and the reference signal on the prediction horizon $[k_0 + 1, k_0 + N_y]$ becomes small. Next, only the first input signal at $k = k_0$ is applied to the objective system and the next output at $k = k_0 + 1$ is measured. Then, the above procedure is repeated, putting the time step forward on one step. In this paper, the problem of model predictive control:

$$\min_{u_*(k_0)} \sum_{j=1}^{M} \left(y_{ref}(k_0 + j) - \hat{y}_*(k_0 + j) \right)^2 \tag{20}$$

Subject to

$$|u_*(k_0)| \le U_{max} \tag{21}$$

$$\sum_{j=1}^{M} \hat{\sigma}_*^2(k_0 + j) \le S_{max} \tag{22}$$

Which is considered, where $y_{ref}(k)$ is the reference signal, U_{max} is the constraint for the input signal, and S_{max} is the constraint for the uncertainty about the predicted future output. The constraint condition (22) means that the input signal which causes large uncertainty of the predicted future output is excluded in the optimization. Although the constraint condition (22) may not influence the performance of control when the prediction model is quite accurate, it would work when the uncertainty of the prediction model is large. The prediction horizon is taken to be $[k_0 + 1, k_0 + M]$, and the control horizon is set to be $[k_0, k_0]$, namely, only one input signal is determined in the proposed method.

The detailed algorithm of the proposed model predictive control using the GA is as follows:

1) Step 1: Initialization for time step

Set the current time stepas $k = k_0$.

2) Step 2: Initialization for GA

Generate an initial population of$Q$binary strings with L bits for the input signal $u_*(k_0)$ randomly.

Set the initial generation g to 0.

3) Step3: Decoding

Decode Q strings into real values $u_{*[i]}(k_0)(i = 1,2,\cdots,Q)$ as follows:

$$u_{*[i]}(k_0) = \frac{2U_{max}}{2^L - 1} U_{[i]} - U_{max} \tag{23}$$

where $U_{[i]}$ is the decimal value converted from the corresponding binary representation. This coding way limits the search range of $u_{*[i]}(k_0)$ to $[-U_{max}, U_{max}]$ so that the constraint condition (21) is satisfied.

4) Step4: Construction of state vector

Construct the candidates of the state vector:

$$\boldsymbol{x}_{*[i]} = \boldsymbol{x}_{*[i]}(k_0) = \left[y_*(k_0), y_*(k_0 - 1), \cdots, y_*(k_0 - L_y + 1), u_{*[i]}(k_0), u_*(k_0 - 1), \cdots, u_*(k_0 - L_u + 1) \right]^{\mathrm{T}}$$

$(i = 1,2,\cdots,Q)$.

5) Step5: Prediction of future output

Predict the future output on the prediction horizon $[k_0 + 1, k_0 + M]$ using (19) as follows:

$$\begin{aligned} \hat{y}_{*[i]}(k_0 + j) &= m_j(\boldsymbol{x}_{*[i]}) \\ &\quad + \boldsymbol{\Sigma}_j(\boldsymbol{x}_{*[i]}, \boldsymbol{X}) \boldsymbol{K}_j^{-1} (\boldsymbol{y}_j - \boldsymbol{m}_j(\boldsymbol{X})) \\ \hat{\sigma}_{*[i]}^2(k_0 + j) &= s_j(\boldsymbol{x}_{*[i]}, \boldsymbol{x}_{*[i]}) \\ &\quad - \boldsymbol{\Sigma}_j(\boldsymbol{x}_{*[i]}, \boldsymbol{X}) \boldsymbol{K}_j^{-1} \boldsymbol{\Sigma}_j(\boldsymbol{X}, \boldsymbol{x}_{*[i]}) + \sigma_{j[best]}^2 \\ &\quad (j = 1,2,\cdots,M) \end{aligned} \tag{24}$$

6) Step6: Fitness value calculation

Calculate the error between the candidates of predicted future output and the reference signal from (25):

$$J_{[i]} = \sum_{j=1}^{M} \left(y_{ref}(k_0 + j) - \hat{y}_{*[i]}(k_0 + j) \right)^2 \tag{25}$$

And the fitness values $F_{[i]} = D - J_{[i]}$, where D is a positive constant value.

If $\sum_{j=1}^{M} \hat{\sigma}_*^2(k_0 + j) > S_{max}$, then let $F_{[i]} = 0$. Namely, the candidates of input signal that do not satisfy the constraint condition (22) become lethal genes.

7) Step7: Reproduction

Reproduce each of individual strings with probability of $F_{[i]} / \sum_{l=1}^{Q} F_{[l]}$. Practically, the linear fitness scaling [17] is utilized to avoid undesirable premature convergence.

8) Step8: Crossover

Select two strings randomly and decide whether or not to cross them over according to the crossover probability P_c. Exchange strings at a crossing position if the crossover is required. The crossing position is chosen randomly.

9) Step9: Mutation

Alter a bit (0 or 1) of string according to the mutation probability P_m.

10) Step10: Repetition for GA

If g is less than the prespecified g_{max}, then $g = g + 1$ and go to *step 3*.

11) Step11: Determination of optimal input signal

Determine the optimal input signal $u_{*[best]}(k_0)$ by thestringwith the best fitness value over all the past generation.

12) Step12: Measurement of next output

Measure the next output $y_*(k_0 + 1)$ by applying the input $u_{*[best]}(k_0)$to the objective system (1).

13) Step13: Putting the time step forward

Let $k_0 = k_0 + 1$ and go to step 2.

V. Simulations

Consider a simplified electric power system [19] described by

©2015 Engineering and Technology Publishing

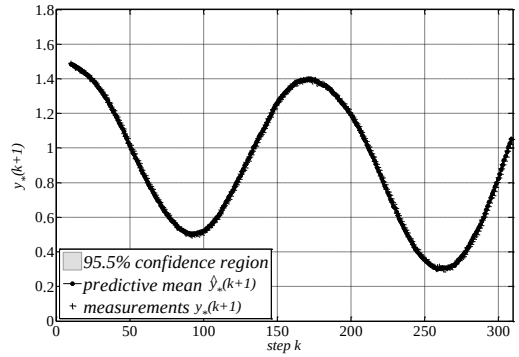


Figure 2. Prediction result for 1 step ahead prediction ($\sigma = 0.004$).

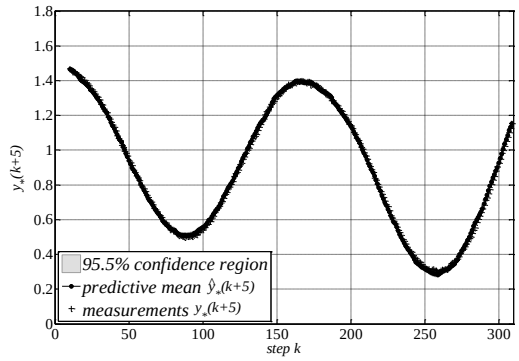


Figure 3. Prediction result for 5 step ahead prediction ($\sigma = 0.004$).

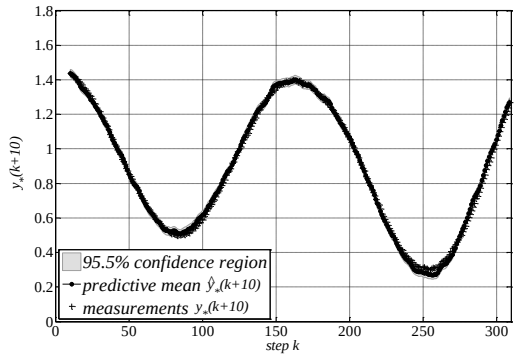


Figure 4. Prediction result for 10 step ahead prediction ($\sigma = 0.004$).

$$\begin{cases} \tilde{M}\ddot{\delta}(t) + \tilde{D}\dot{\delta}(t) + P_e = P_{in} \\ P_e = \dfrac{e_t E_{fd}}{X_e}\left(1 + \Delta E_{fd}(t)\right)\sin\delta(t) \\ y(t) = \delta(t) + v(t) \end{cases} \tag{26}$$

where $\tilde{M} = 0.06, \tilde{D} = 0.06, E_{fd} = 1.0, e_t = 1.0, X_e = 1.0,$ and $P_{in} = 0.8$.These are all per unit values.The training data are sampled with sampling period $T_s = 0.01$ as $u(k) = \Delta E_{fd}(kT_s)$ and $y(k) = y(kT_s)$at $t = kT_s$. The measurement noise$v(t)$ is set to be 0 (a noise free case), and zero mean Gaussian noise with standard deviation $\sigma = 0.004$ (a case of noise to signal ratio: 1%).The lags for the state vector (3) are chosen as $L_y = 2$ and $L_u = 1$ in the noise free case, and $L_y = 10$ and $L_u = 1$ in the case of $\sigma = 0.004$, respectively. The number of the training input and output data is taken to be $N = 300$ for training each j ($j = 1,2,\cdots,10$) step ahead predictor.

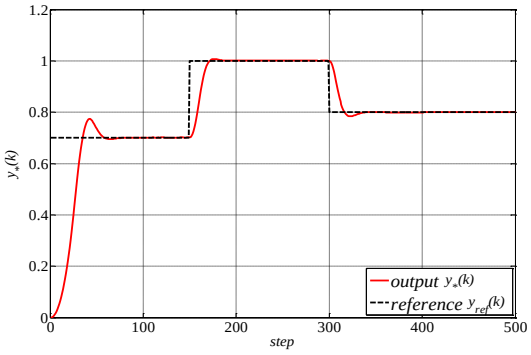
For the case of $\sigma = 0.004$, to validate the results of training the predictors, the prediction results for 1, 5 and 10 step ahead predictors are shown in Fig. 2-Fig. 4. In these figures, the circles with lines show the predictive mean $\hat{y}_*(k+j)$, the crosses show the measurements (test output) $y_*(k+j)$, and the shaded areas give the double standard deviation confidence interval (95.5% confidence region). From these figures, we can confirm that the error between the test data and the predictive mean is quite small for every step ahead predictors and it does not become so large as the prediction horizon increases.

Next, model predictive control is carried out using the trained multiple GP predictors. The constraints for the cost function are taken to be $U_{max} = 2.5$ and $S_{max} = 0.001$, respectively. The design parameters for the GA are as follows:

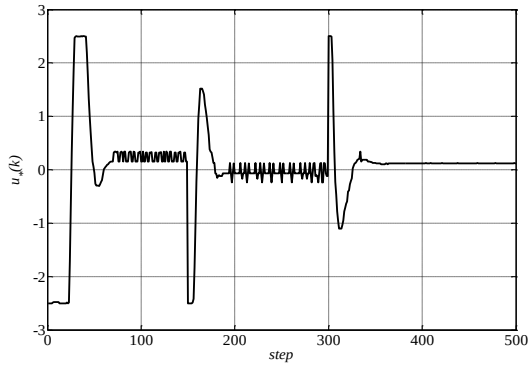
Population size: $Q = 30$
String length: $L = 10$
Crossover probability: $P_c = 0.8$
Mutation probability: $P_m = 0.03$
Maximum generation number: $g_{max} = 10$

The reference signal $y_{ref}(k)$ is changed stepwise as follows:

$$y_{ref}(k) = \begin{cases} 0.7 & (0 \le k < 150) \\ 1.0 & (150 \le k < 300) \\ 0.8 & (300 \le k \le 500) \end{cases} \tag{27}$$



(a) Output $\boldsymbol{y_*(k)}$



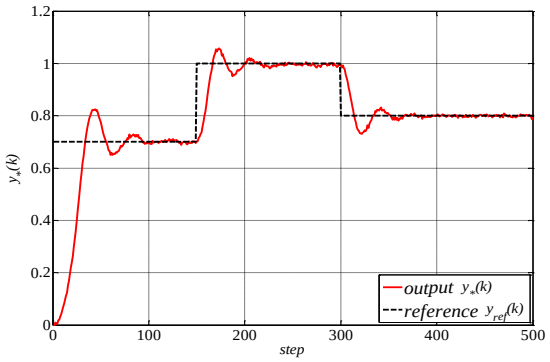
(b) Input $\boldsymbol{u_*(k)}$

Figure 5. Time responses in the case of noise free.

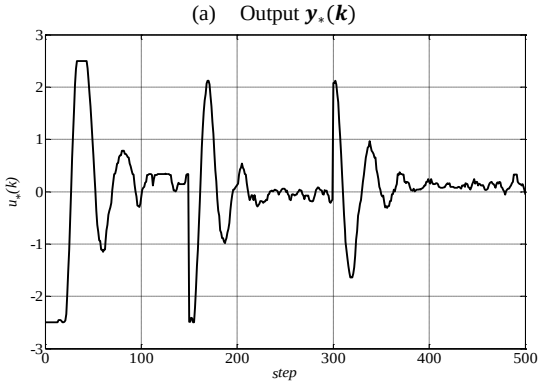
Fig. 5 shows the time responses of output $y_*(k)$and the input signal $u_*(k)$ in the case of noise free. Fig. 6 shows the time responses of output $y_*(k)$ and the inputsignal $u_*(k)$ in the case of the noise standard deviation $\sigma = 0.004$. It can be seen that the output $y_*(k)$ can track to the reference signal $y_{ref}(k)$ well in both noise free and noisy cases. These results indicate that the proposed

©2015 Engineering and Technology Publishing

model predictive control could be successfully applied to the electric power systems.



(a) Output $\boldsymbol{y}_*(\boldsymbol{k})$



(b) Input $\boldsymbol{u}_*(\boldsymbol{k})$

Figure 6. Time responses in the case of $\sigma = 0.004$.

VI. Conclusions

In this paper, a novel model predictive control of electric power systems based on the multiple Gaussian process predictors has been presented. Using the trained multiple GP predictors for multistep ahead prediction, the optimal input signal is determined so that the constrained cost function is minimized in the framework of the receding horizon strategy. Through the numerical simulations for the simplified electric power system, it has been experimentally demonstrated that the good control performance is obtained by the proposed model predictive control scheme. The analysis of stability condition for this control scheme is one of the future works.

References

[1] H. Seki, M. Ogawa, S. Ooyama, K. Akamatsu, M. Ohshima, and W. Yang, "Industrial application of a nonlinear model predictive control to polymerization reactors," *Control Engineering Practice*, vol. 9, no. 8, pp. 819-828, 2001.

[2] Z. K. Nagy, B. Mahn, R. Franke, and F. Allgöwer, "Evaluation study of an efficient output feedback nonlinear model predictive control for temperature tracking in an industrial batch reactor," *Control Engineering Practice*, vol. 15, no. 7, pp. 839-850, 2007.

[3] B. Likar and J. Kocijan, "Predictive control of a gas-liquid separation plant based on a Gaussian process model," *Computers and Chemical Engineering*, vol. 31, no. 3, pp. 142-152, 2007.

[4] T. Geyer, "Generalized model predictive direct torque control: Long prediction horizons and minimization of switching losses," in *Proc. IEEE Conference on Decision and Control*, 2009, pp. 6799-6804.

[5] J. Scoltock, T. Geyer, and U. K. Madawala, "A Comparison of model predictive control schemes for MV induction motor drives," *IEEE Trans. Industrial Informatics*, pp. 909-919, 2013.

[6] A. O'Hagan, "Curve fitting and optimal design for prediction (with discussion)," *J. Royal Statistical Society B*, vol. 40, pp. 1-42, 1978.

[7] C. K. I. Williams and C. E. Rasmussen, "Gaussian processes for regression," *Advances in Neural Information Processing Systems*, vol. 8, pp. 514-520, MIT Press, 1996.

[8] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006.

[9] J. Kocijan, A. Girard, B. Banko, and R. Murray-Smith, "Dynamic systems identification with Gaussian processes," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 11, no. 4, pp. 411-424, 2005.

[10] T. Hachino and H. Takata, "Identification of continuous-time nonlinear systems by using a Gaussian process model," *IEEJ Trans. Electrical and Electronic Engineering*, vol. 3, no. 6, pp. 620-628, 2008.

[11] A. Girard, C. E. Rasmussen, J. Q. Candela, and R. Murray-Smith, "Gaussian process priors with uncertain inputs -application to multiple-step ahead time series forecasting," *Advances in Neural Information Processing Systems*, vol. 15, pp. 542-552, MIT Press, 2003.

[12] J. Q. Candela, A. Girard, J. Larsen, and C. E. Rasmussen, "Propagation of uncertainty in Bayesian kernel models-application to multiple-step ahead forecasting," in *Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing*, 2003, pp. 701-704.

[13] T. Hachino and V. Kadirkamanathan, "Multiple Gaussian process models for direct time series forecasting," *IEEJ Trans. Electrical and Electronic Engineering*, vol. 6, no. 3, pp. 245-252, 2011.

[14] J. Kocijan, R. Murray-Smith, C. E. Rasmussen, and B. Likar, "Predictive control with Gaussian process models," in *Proc. IEEE Eurocon 2003: Int. Conference on Computer as a Tool*, 2003, pp. 352-356.

[15] R. Murray-Smith, D. Sbarbaro, C. E. Rasmussen, and A. Girard, "Adaptive, cautions, predictive control with Gaussian process priors," in *Proc. 13th IFAC Symposium on System Identification*, 2003, pp. 1195-1200.

[16] T. Hachino, K. Naritomi, H. Takata, S. Fukushima, and Y. Igarashi, "Multistep ahead prediction of electric power systems using multiple Gaussian process models," *Journal of Automation and Control Engineering*, to appear.

[17] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.

[18] R. Vonmises, *Mathematical Theory of Probability and Statistics*, Academic Press, 1964.

[19] H. Takata, "An automatic choosing control for nonlinear systems," in *Proc. IEEE Conference on Decision and Control*, 1996, pp. 3453-3458.

Tomohiro Hachino received the B.S., M.S. and Dr. Eng. degrees in electrical engineering from Kyushu Institute of Technology in 1991, 1993, and 1996, respectively. He is currently an Associate Professor at the Department of Electrical and Electronics Engineering, Kagoshima University. His research interests include nonlinear control and identification, signal processing, and evolutionary computation. Dr. Hachino is a member of IEEJ, SICE, and ISCIE.

Hitoshi Takata received the B.S. degree in electrical engineering from Kyushu Institute of Technology in 1968 and the M.S. and Dr. Eng. degrees in electrical engineering from Kyushu University in 1970 and 1974, respectively. He is currently a Professor Emeritus and a part-time lecturer at Kagoshima University. His research interests include the control, linearization, and identification of nonlinear systems. Dr. Takata is a member of IEEJ and RISP.

©2015 Engineering and Technology Publishing

Seiji Fukushima received the B.S., M.S., and Ph.D. degrees in electrical engineering from Kyushu University in 1984, 1986, and 1993, respectively. He is currently a Professor at the Department of Electrical and Electronics Engineering, Kagoshima University. His research interests include photonics/radio hybrid communication systems and their related devices. Dr. Fukushima is a member of IEICE, IEEE/Photonics Society, Japan Society of Applied Physics, Japanese Liquid Crystal Society, and Optical Society of America.

Yasutaka Igarashi received the B.E., M.E., and Ph.D. degrees in information and computer sciences from Saitama University in 2000, 2002, and 2005, respectively. He is currently an Assistant Professor at the Department of Electrical and Electronics Engineering, Kagoshima University. His research interests include optical CDMA communication systems and the cryptanalysis of symmetric-key cryptography. Dr. Igarashi is a member of IEICE and RISP.

©2015 Engineering and Technology Publishing