Journal of System Simulation

Volume 29 | Issue 3 Article 17

6-2-2020

Filtering Based Maximum Likelihood Stochastic Gradient Prediction on Wind Power Curtailment

Ziyun Wang
Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education),College of the Internet of Things, Jiangnan University, Wuxi 214122, China;

Zhicheng Ji
Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education),College of the Internet of Things, Jiangnan University, Wuxi 214122, China;

Filtering Based Maximum Likelihood Stochastic Gradient Prediction on Wind Power Curtailment

Abstract

Abstract: The parameter estimation problem of Hammerstein finite impulse response models and its application on the wind curtailment prediction field were considered. *By adopting the maximum likelihood principle, the maximum likelihood estimate was obtained by minimizing the likelihood function. To reduce the impact of the unknown noise term, the maximum likelihood idea and the filtering theory were combined by changing the coupled nonlinear model into a parameter-independent model and to derive a filtering based maximum likelihood stochastic gradient algorithm for the Hammerstein system modeling on wind power curtailment prediction.* The given simulation validates that the proposed algorithm can identify the wind power characteristic curve accurately and contributes to calculate the wind power curtailment prediction that shows its good practicability.

Keywords

system identification, stochastic gradient, filtering theory, maximum likelihood, wind power curtailment prediction

Recommended Citation

Wang Ziyun, *Ji Zhicheng*. Filtering Based Maximum Likelihood Stochastic Gradient Prediction on Wind Power Curtailment[J]. Journal of System Simulation, 2017, 29(3): 589-594.

第 29 卷第 3 期 系统仿真学报© Vol. 29 No. 3
2017 年 3 月 Journal of System Simulation Mar., 2017

Filtering Based Maximum Likelihood Stochastic Gradient Prediction on Wind Power Curtailment

Wang Ziyun, Ji Zhicheng

(Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education),
College of the Internet of Things, Jiangnan University, Wuxi 214122, China)

Abstract: The parameter estimation problem of Hammerstein finite impulse response models and its application on the wind curtailment prediction field were considered. *By adopting the maximum likelihood principle, the maximum likelihood estimate was obtained by minimizing the likelihood function. To reduce the impact of the unknown noise term, the maximum likelihood idea and the filtering theory were combined by changing the coupled nonlinear model into a parameter-independent model and to derive a filtering based maximum likelihood stochastic gradient algorithm for the Hammerstein system modeling on wind power curtailment prediction.* The given simulation validates that the proposed algorithm can identify the wind power characteristic curve accurately and contributes to calculate the wind power curtailment prediction that shows its good practicability.

Keywords: system identification; stochastic gradient; filtering theory; maximum likelihood; wind power curtailment prediction

基于滤波极大似然随机梯度的弃风电量预测

王子赟，纪志成

（江南大学，物联网工程学院，轻工过程先进控制教育部重点实验室，无锡 214122）

摘要：研究了一类 Hammerstein 有限脉冲响应模型的建模方法，并用于风电场弃风电量预测领域。*采用极大似然估计律对似然方程进行最小化，同时为了减少有色噪声对建模过程的干扰，结合极大似然估计方法和滤波过程，将原本耦合的非线性模型转变为独立参数的辨识模型，进而推导了一类基于滤波的极大似然随机梯度辨识算法，并将该方法用于风电场弃风电量的预测领域。*仿真结果表明提出的算法可以精确的辨识实际风电场的风电功率特性曲线，并能很好的预测风电场的弃风电量情况，具有很强的实用性。

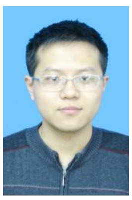
关键词：系统建模；随机梯度；滤波算法；极大似然估计；风电弃风电量预测

中图分类号：TP13 文献标识码：A 文章编号：1004-731X (2017) 03-0589-06

DOI: 10.16182/j.issn1004731x.joss.201703017

Introduction

Parameter estimation methods have been widely studied for designing identification algorithms[1-2], modeling complex systems[3-4] and solving matrix equations[5-6]. For example, Liu et al. derived a decomposition based recursive least squares identification algorithm using the hierarchical identification principle for the lifted input-output representation of general dual-rate sampled-data



Received: 2016-08-11 **Revised:** 2017-01-03;
Foundation: China Postdoctoral Science Foundation Funded Project (2016M591765), Jiangsu Postdoctoral Science Foundation Funded Project (1601252C), Fundamental Research Funds for the Central Universities (JUSRP115A30);
Biography: Ziyun Wang (1989-) Fuzhou, Jiangxi Province, China. PhD, research interest include nonlinear system modeling theory, parameter estimation and the prediction and evaluation methods for wind power

systems[7]; Ding considered the state filtering and parameter estimation problems for state space systems with scarce output availability and proposed a least squares based algorithm and an observer based parameter estimation algorithm to estimate the system parameter matrices and states when the scarce states are available[8].

Recently, some existing works on deriving the identification methods have significant influence, e.g., Li et al. proposed a maximum likelihood least squares identification method for input nonlinear finite impulse response moving average systems by directly estimating the parameters of the linear and nonlinear parts of the Hammerstein system without using the over-parameterization technique[9]; Wang et al. derived a recursive maximum likelihood least squares identification algorithm for systems with autoregressive moving average noises based on the maximum likelihood principle and proved that the maximum of the likelihood function was equivalent to minimizing the least squares cost function[10].

Differing from the works in [9-10], this paper considers the parameter estimation problem of nonlinear Hammerstein systems and derives a maximum likelihood stochastic gradient algorithm for Hammerstein finite impulse response model. In order to diminish the impact of unknown noise term, a two-stage filtering based maximum likelihood stochastic gradient algorithm is proposed.

The purpose of this paper is presenting an identification algorithm to directly estimate the parameters of nonlinear models, and shows its effectiveness in the wind power curtailment prediction simulation via true sampled data.

Briefly, the rest of this paper is organized as follows. Section 1 presents the identification model of the Hammerstein finite impulse response systems. Section 2 gives the modeling formula of calculate wind power curtailment. Section 3 proposes a filtering based maximum likelihood stochastic gradient algorithm by transferring a finite impulse response moving average model to a controlled autoregressive model. Section 4 provides a wind power curtailment prediction case to show the effectiveness of the proposed algorithm in system modeling. Finally, some concluding remarks are offered in Section 5.

1 Problem formulation

Consider the following nonlinear system:

$$y(t)=B(z)\sum_{i=1}^{n_c}c_i f_i(u(t))+D(z)v(t) \tag{1}$$

where $u(t)$ and $y(t)$ are the input and output sequences of the Hammerstein finite impulse response moving average (FIR-MA) system, respectively, $v(t)$ is the stochastic white noise with zero mean and variance σ^2, $B(z)$ and $D(z)$ are polynomials in the unit backward shift operator z^{-1} [i.e. $z^{-1}y(t)=y(t-1)$]:

$$B(z):=1+b_1z^{-1}+b_2z^{-2}+\cdots+b_{n_b}z^{-n_b},$$

$$D(z):=1+d_1z^{-1}+d_2z^{-2}+\cdots+d_{n_d}z^{-n_d}.$$

Assume that the orders n_b, n_c and n_d are known, $y(t)=0, u(t)=0$ for $t\le 0$. The aim of identifying Hammerstein system is to consistently estimate the parameter vectors

$$b:=[b_1,b_2,\cdots,b_{n_b}]^{\mathrm{T}}\in\mathbf{R}^{n_b},$$

$$c:=[c_1,c_2,\cdots,c_{n_c}]^{\mathrm{T}}\in\mathbf{R}^{n_c},$$

$$d:=[d_1,d_2,\cdots,d_{n_d}]^{\mathrm{T}}\in\mathbf{R}^{n_d}$$

from the measured data $\{u(t),y(t)\}_{t=1}^{N}$. From Equation (1), the identification model can be rewritten as

https://dc-china-simulation.researchcommons.org/journal/vol29/iss3/17
DOI: 10.16182/j.issn1004731x.joss.201703017

$$y(t) = B(z)\bar{u}(t) + D(z)v(t) = \sum_{i=1}^{n_b} b_i\bar{u}(t-i) + \sum_{i=1}^{n_c} c_i f_i(u(t)) + \sum_{i=1}^{n_d} d_i v(t-i) + v(t) = \phi^{\mathrm{T}}(t)\theta + v(t) \quad (2)$$

2 The wind power curtailment modeling

The calculation of wind power curtailment is the basis of well-planing on the power usage. The mainly used way to get the power generation is by getting the system model on the wind power firstly before calculating its curtailment. By modeling the wind power generation curve on the sampled wind power data and wind speed data, the wind power curtailment in the near future can be obtained.

In the modeling of wind power generation system, the generated wind power can be calculated by a formula that yields from (2), that is

$$p(t) = \sum_{i=0}^{m} \theta_i v^i(t) + \sum_{j=1}^{n} \vartheta_j e^j(t)$$

In the formula, $P(t)$ is the sampled wind power in a short time, i.e., 48 hours, $v(t)$ stands for the wind speed and $e(t)$ is modeled as white noise that shows the disturbance of sampling process. The main object is to get the unknown vectors θ_i and ϑ_j.

3 The filtering based maximum likelihood stochastic gradient algorithm

For a given set of measurements $\{u_N, y_N\} := \{u(i), y(i)\}_{i=1}^{N}$, let the likelihood function $L(y_N | u_{N-1}, \theta)$ equals the probability density function $p(y_N | u_{N-1}, \theta)$. By maximizing the likelihood function, the maximum likelihood estimate $\hat{\theta}_{ML}$ can be written as[9]

$$\hat{\theta}_{ML} = \arg\max_{\theta} L(y_N | u_{N-1}, \theta) \quad (3)$$

Where $L(y_N | u_{N-1}, \theta) = \prod_{t=1}^{N} p(\phi^{\mathrm{T}}(t)\theta + v(t) | y_{t-1}, u_{t-1}, \theta)$. (4)

If the input $\bar{u}(t)$ and the output $y(t)$ are filtered by a rational function $D^{-1}(z)$, the FIR-MA model in (1) will be filtered into a controlled auto-regressive model. Multiplying both sides of (1) by $D^{-1}(z)$ yields

$$\frac{1}{D(z)} y(t) = B(z)\frac{1}{D(z)}\bar{u}(t) + v(t)$$

Or $y_f(t) = B(z)\bar{u}_f(t) + v(t)$

Where $\bar{u}_f(t) := \frac{1}{D(z)}\sum_{i=1}^{n_c} c_i f_i(u(t)) = \sum_{i=1}^{n_c} c_i U_i(t)$

$$y_f(t) := \frac{1}{D(z)} y(t)$$

$$U_j(t) := \frac{1}{D(z)} f_j(u(t))$$

Defining and minimizing the following cost function

$$J(\theta_n) := \sum_{i=1}^{L} [w(i) - \varphi_n^{\mathrm{T}}(i)\theta_n]^2 | \theta_n \quad (5)$$

The filtering based maximum likelihood stochastic gradient (F-ML-SG) algorithm for Hammerstein models can be summarized as follows:

$$\hat{\theta}_s(t) = \hat{\theta}_s(t-1) - \frac{\mathrm{grad}[v(t)]|_{\hat{\theta}_s(t-1)}}{r_2(t)} [\hat{y}_f(t) - \varphi_f^{\mathrm{T}}(t)\hat{\theta}_s(t-1)] \quad (6)$$

$$r_2(t) = r_2(t-1) + \| \mathrm{grad}[v(t)]|_{\hat{\theta}_s(t-1)} \|^2 \quad (7)$$

$$\mathrm{gard}[v(t)]|_{\hat{\theta}_s(t-1)} = -\hat{\varphi}_f(t) \quad (8)$$

$$\hat{\theta}_n(t) = \hat{\theta}_n(t-1) - \frac{\mathrm{grad}[v(t)]|_{\hat{\theta}_n(t-1)}}{r_1(t)} [\hat{w}(t) - \hat{\varphi}_n^{\mathrm{T}}(t)\hat{\theta}_n(t-1)] \quad (9)$$

$$r_1(t) = r_1(t-1) + \left\| \mathrm{grad}[v(t)]|_{\hat{\theta}_n(t-1)} \right\|^2 \quad (10)$$

$$\mathrm{gard}[v(t)]|_{\hat{\theta}_n(t-1)} = -\hat{\varphi}_n(t) \quad (11)$$

$$\hat{\varphi}_f(t) = [\hat{\bar{u}}_f(t-1), \hat{\bar{u}}_f(t-2), \cdots, \hat{\bar{u}}_f(t-n_b), \hat{U}_1(t), \hat{U}_2(t), \cdots, \hat{U}_{n_c}(t)]^{\mathrm{T}} \quad (12)$$

$$\hat{\bar{u}}_f(t) = \hat{c}_1(t)\hat{U}_1(t) + \hat{c}_2(t)\hat{U}_2(t) + \cdots + \hat{c}_{n_c}(t)\hat{U}_{n_c}(t) \quad (13)$$

$$\hat{U}_j(t) = -\hat{d}_1(t)\hat{U}_j(t-1) - \hat{d}_2(t)\hat{U}_j(t-2) - \cdots - \hat{d}_{n_d}(t)\hat{U}_j(t-n_d) + f_j(u(t)) \quad (14)$$

第 29 卷第 3 期
2017 年 3 月
系统仿真学报
Journal of System Simulation
Vol. 29 No. 3
Mar., 2017

$$\hat{\varphi}_s(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b), f_1(u(t)), f_2(u(t)),\cdots, f_{n_c}(u(t))]^{\mathrm{T}} \quad (15)$$

$$\hat{\varphi}_n(t)=[\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}} \quad (16)$$

$$\hat{\bar{u}}(t)=\hat{c}_1(t)f_1(u(t))+\hat{c}_2(t)f_2(u(t))+\cdots+\hat{c}_{n_c}(t)f_{n_c}(u(t)) \quad (17)$$

$$\hat{y}_f(t)=-\hat{d}_1(t)\hat{y}_f(t-1)-\hat{d}_2(t)\hat{y}_f(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{y}_f(t-n_d)+y(t) \quad (18)$$

$$\hat{w}(t)=y(t)-\hat{\varphi}_s^{\mathrm{T}}(t)\hat{\theta}_s(t-1) \quad (19)$$

$$\hat{v}(t)=\hat{y}_f(t)-\hat{\varphi}_f^{\mathrm{T}}(t)\hat{\theta}_s(t) \quad (20)$$

The steps involved in the F-ML-SG algorithm for Hammerstein systems are listed in the following:

1. To initialize, let $t=1$, set the initial values of the parameter estimation variables and co-variance matrices as follows:

$\hat{\theta}_s(i)=\mathbf{1}_{n_b+n_c}/p_0$, $\hat{\theta}_n(i)=\mathbf{1}_{n_d}/p_0$, $\hat{y}_f(i)=1/p_0$, $\hat{w}(i)=1/p_0$, $\hat{\bar{u}}_f(i)=1/p_0$, $\hat{\bar{u}}(i)=1/p_0$, $\hat{v}(i)=1/p_0$ for $i\le 0$, $\hat{U}j(i)=1/p_0$ for $i\le 0$ and $j=1,2,\cdots,n_c$, $p_0=10^6$, and give the basis functions $f_j(\cdot)$.

2. Collect the input-output data $u(t)$ and $y(t)$, construct the information vectors $\hat{\varphi}_s(t)$ by (15), $\hat{\varphi}_n(t)$ by (16), respectively. Compute $\hat{w}(t)$ by (19).

3. Compute $\mathrm{grad}[v(t)]|_{\hat{\theta}_n(t-1)}$ and $r_1(t)$ by (11) and (10). Update parameter vector $\hat{\theta}_n(t)$ by (9).

4. Compute $\hat{U}_j(t)$ and $\hat{y}_f(t)$ by (14) and (18), and form $\hat{\varphi}_f(t)$ by (12).

5. Compute $\mathrm{grad}[v(t)]|_{\hat{\theta}_s(t-1)}$ and $r_2(t)$ by (8) and (7). Update parameter vector $\hat{\theta}_s(t)$ by (6).

6. Compute $\hat{\bar{u}}_f(t)$, $\hat{\bar{u}}(t)$ and $\hat{v}(t)$ by (13), (17) and (20).

7. Increase t by 1 and go to Step 2.

In addition, some published works have proved that when the rational function $\left[D^{-1}(z)-\frac{1}{2}\right]$ is filtered by the input-output data, the colored noise can be filtered.

4 Application and example

Consider a wind power curtailment prediction problem. The sampled data is from a wind farm in Jiangsu Province, China. The installed capacity of this wind farm is 22.50 MW that contains 18 wind turbines. This simulation is applied on one of the wind turbines and the cut-in wind speed for this wind farm is settled 3 m/s and the rated wind speed is settled 10 m/s.

A series of wind speed data is shown in Fig. 1 and the given sampled wind speed is collected in continuous 72 hours from March 1, 0:00:00, 2015 to March 3, 23:50:00, 2015 with 10 minutes for one set. The sampled array includes 432 sets of wind speed and wind power data.

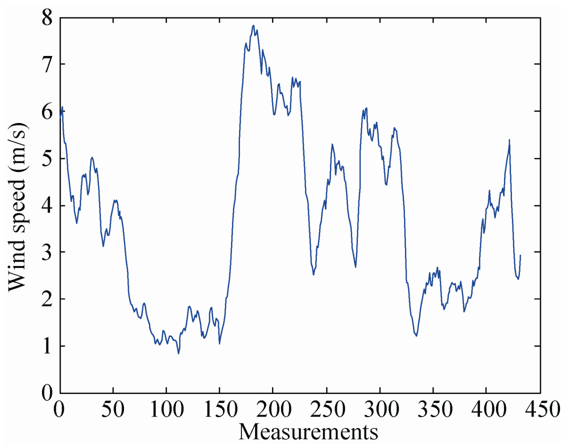


Fig. 1 Sampled wind speed curve in 72 hours

From Fig. 1, some sampled wind speed data are under the cut-in wind speed set point. It is obvious that the wind power characteristic curve is modeled by the wind speed between the cut-in wind speed and the rated wind speed, thus it is essential to deal with the collected data and the wind speed-power data set should be removed from the sampled data array. Thus, the processed wind speed and wind power measurements satisfy the wind power curtailment forecasting that are shown in Fig. 2.

DOI: 10.16182/j.issn1004731x.joss.201703017

Fig. 2 shows 249 wind speed and wind power data sets, the wind power curve reflects the following phenomenon with the wind speed. One common method to model the wind power characteristic curve uses the nonlinear system with colored noise as the disturbance. The nonlinear model is set as in Section 2:

$$p(t)=\sum_{i=0}^{m}\theta_i v^i(t)+\sum_{j=1}^{n}\vartheta_j e^j(t)$$

In this simulation, we adopt the fifth order nonlinear function and the F-ML-SG algorithm in (6) to (20) for identifying the unknown parameters of the wind power characteristic curve.

$$P(t)=0.32312v^5(t)-7.64046v^4(t)+68.57059v^3(t)-289.29765v^2(t)+590.90435v(t)-471.59402+0.96403e(t)+0.80749e^2(t)$$

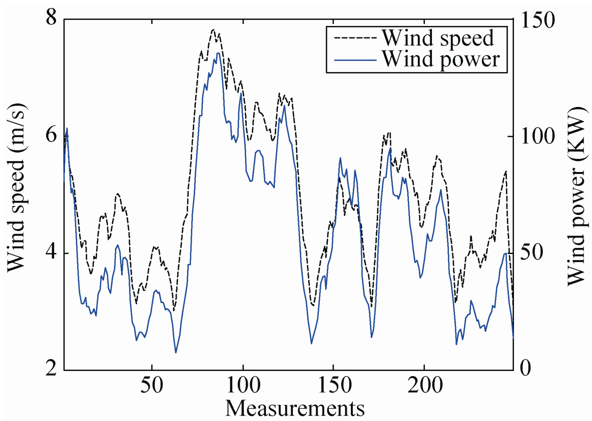


Fig. 2 Processed wind speed and wind power curve

The wind power prediction can be calculated at each time t and the comparison between the collected wind power and the wind power forecast is shown in Fig. 3, that shows the predicted wind power has highly consistency with the wind power sampled data. By comparison, the maximum likelihood stochastic gradient algorithm is also applied to forecast the wind power generation and its simulation on the same example is shown in Fig. 3.

From Fig. 3, it can be seen that the proposed F-ML-SG algorithm has a better performance on modeling the wind power generation curve, especially on the big fluctuation sample data. The main reason is that the presented F-ML-SG algorithm can filter the colored unknown noise at the modeling step, thus, the influence of the noise interference can be reduced to the minimum.

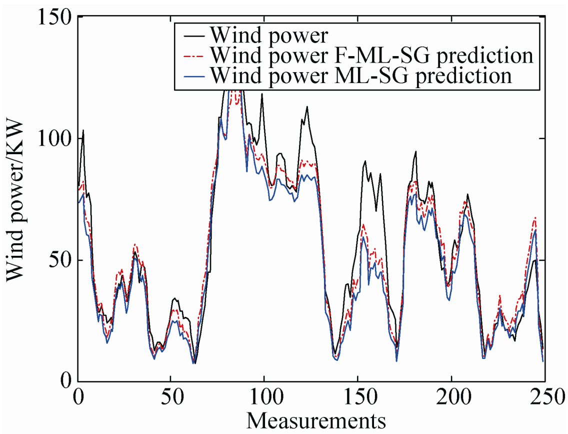


Fig. 3 Comparison between sampled wind power and prediction

The prediction of wind curtailment depends on the different value between the wind power prediction and the wind power sampled value. In Figure 3, the prediction of wind curtailment is the sum of absolute values of the inter-space between two curves. The wind curtailment prediction

$$P_c=\int_1^t\left|P-P_f\right|=\sum_{t=1}^{t}\left|P-P_f\right|$$

In this case, through forcasting the wind power during next 72 hours, the wind power curtailment is 1905 KW that only has 7.653% accumulated error compared with the true value.

5 Conclusions

This paper derived a maximum likelihood stochastic gradient algorithm for the Hammerstein nonlinear systems based on the filtering theory and applied it on the wind power curtailment prediction.

Journal of System Simulation, Vol. 29 [2017], Iss. 3, Art. 17

The original FIR-MA nonlinear system is transformed into a controlled auto-regressive moving average model when the input-output data is filtered by the rational function. The proposed algorithms are not only applied on the wind power forecast field, but can be also used for designing adaptive strategy for input nonlinear systems and dealing with the estimation of rotor power coefficient.

References:

[1] Huang J, Shi Y, Huang H, et al. Filtering for multirate nonlinear sampled-data systems using T-S fuzzy models. Digital Signal Processing (S1095-4333), 2013, 23(1): 418-426.

[2] Liu Y J, Ding F. Convergence properties of the least squares estimation algorithm for multivariable system. Applied Mathematical Modelling (S0307-904X), 2013, 37(1/2): 476-483.

[3] Zhang Y, Cui G M. Bias compensation methods for stochastic systems with colored noise. Applied Mathematical Modelling (S0307-904X), 2011, 35(4): 1709-1716.

[4] Liao H S, Sampath P V, Curtis Z K, et al. Hierarchical modeling of a groundwater remediation capture system [J]. Journal of Hydrology (S0022-1694), 2015, 527: 196-211.

[5] Dehghan M, Hajarian M. Efficient iterative method for solving the second-order sylvester matrix equation EVF-AVF-CV=BW [J]. IET Control Theory and Applications (S1751-8644), 2009, 3(10): 1401-1408.

[6] Dehghan M, Hajarian M. SSHI methods for solving general linear matrix equations [J]. Engineering Computations (S0264-4401), 2011, 28(8): 1028-1043.

[7] Liu Y J, Ding F, Shi Y. An efficient hierarchical identification method for general dual-rate sampled-data systems [J]. Automatica (S0005-1098), 2014, 50 (3): 962-970.

[8] Ding F. State filtering and parameter estimation for state space systems with scarce measurements [J]. Signal Processing (S0165-1684), 2014, 104(6): 369-380.

[9] Li J H, Ding F, Yang G W. Maximum likelihood least squares identification method for input nonlinear finite impulse response moving average systems [J]. Mathematical and Computer Modelling (S0895-7177), 2012, 55(3/4): 442-450.

[10] Wang W, Ding F, Dai J Y. Maximum likelihood least squares identification for systems with autoregressive moving average noise [J]. Applied Mathematical Modelling (S0307-904X), 2012, 36(5): 1842-1853.

(上接第 588 页)

[12] 许立, 王长林. 基于遗传算法列车自动运行速度曲线的优化 [J]. 铁路计算机应用, 2013, 22(10): 46-49. (Xu Li, Wang Changlin. Optimization of automatic train operation speed curve based on Genetic Algorithm [J]. Railway Computer Application (S1005-8451), 2013, 22(10): 46-49.)

[13] 李欢. 地铁列车运行仿真及策略优化 [D]. 成都: 西南交通大学, 2012. (Li Huan. Simulation of Subway Train Operation and Optimization Strategy [D]. Chengdu, China: Southwest Jiaotong University, 2012.)

[14] 张茜. 城市轨道交通列车自动运行(ATO)多目标控制策略研究 [D]. 北京交通大学, 2014. (Zhang Xi. Study on Multi-Objective Control Strategies of Automatic Train Operation in Urban Rail Transit [D]. Beijing, China: Jiaotong University, 2014.)

[15] 程军. 自适应模糊 PID 控制器在可调式射流泵供水系统中的应用 [J]. 武汉大学学报(工学版), 2011, 44(5): 668-670. (Cheng Jun. Adaptive fuzzy PID controller's application to technical water supply system based on adjustable jet pump [J]. Engineering Journal of Wuhan University (S1671-8836), 2011, 44(5): 668-670.)

[16] 石辛民, 郝整清. 模糊控制及其 MATLAB 仿真 [M]. 北京: 清华大学出版社/北京交通大学出版社, 2008: 94-101. (Shi Xinmin, Hao Zhengqing. Fuzzy control and MATLAB simulation [M]. Beijing, China: Tsinghua university press/Beijing Jiaotong University press, 2008: 94-101.)

DOI: 10.16182/j.issn1004731x.joss.201703017