Letter

A Model Predictive Control Algorithm Based on Biological Regulatory Mechanism and Operational Research

Jinying Yang, Yongjun Zhang, Tanju Yildirim, and Jiawei Zhang

Dear Editor,

This letter presents an intelligent model predictive control algorithm inspired by biological regulatory mechanism and operational research. In terms of overall architecture, based on biological regulatory system and operational research theory, priority factor module and central coordination module are innovatively added on the basis structure of heuristic dynamic programming to carry out overall regulation of the system. In internal structure, the neural network is integrated with the biofeedback mechanism, and a new multi-level feedback neural network that can obtain more feedback information is proposed. The network is applied to the model network, action network and critic network of the algorithm. The convergence speed is greatly improved and the predictive control speed for nonlinear time-varying systems is improved on the premise of ensuring the control accuracy. The effectiveness and superiority of the proposed algorithm in prediction and control are verified by experiment.

Related work: With the increasing nonlinearity and timely variation of the process control target, in order to meet the control requirements of complex controlled objects, some control algorithms inspired by computer technology have been developed. In 1978, Testud *et al.* [1] proposed the model predictive control algorithm (MPC) and established the concept of rolling time domain optimization. MPC have been widely used in various fields due to its good applicability and robustness, scholars from all over the world have produced a variety of MPC. According to the structural model, MPC can be roughly divided into non-parametric model prediction [2], [3], predictive control based on adaptive control theory [4]−[6], and predictive control based on structural design [7]−[10]. With the development of science, the increasingly complex controlled objects make it difficult for traditional MPC to accurately control nonlinear and time-varying uncertain system. Zhang *et al.* [11], [12] combined MPC with other structures, such as neural network and load observer to improve the response speed and robustness of the algorithm. However, these algorithms are still in the theoretical simulation stage and need to be validated in practice. Yang *et al.* [13] used the distributed MPC to regulate the Yellow River Basin, and the control effect is stable. But, the model needs to set parameters such as river area in advance, which are difficult to measure. Islam applied the MPC to the flood control of Ukai dam in India, and the algorithm could predict the flood and implement the corresponding control strategy [14]. Nevertheless, the system model was fixed and was not suitable for time-varying systems. The relevant research on the MPC in the biological regulatory mechanism and operational research has not gained adequate attention, which is one of the current research motivations. The algorithm proposed in this letter does not need to obtain the system's physical parameters, it only needs the data of the con-

Corresponding author: Yongjun Zhang.

J. Y. Yang and Y. J. Zhang are with the Institute of Engineering Technology, University of Science and Technology Beijing, Beijing 100083, China (e-mail: celia_y7@163.com; zhangyj@ustb.edu.cn).

T. Yildirim is with National Institute for Materials Science, Ibaraki 305-0047, Japan (e-mail: tanjuyildirim1992@hotmail.com).

J. W. Zhang is with the School of Electrical and Information Engineering, University of Sydney, Sydney 2000, Australia (e-mail: jzha4631@uni.sydney.edu.au).

trolled variables and related variables, and over time, the algorithm can update the model to avoid model mismatch.

The main contributions can be summarized as follows. 1) A new intelligent MPC, called a biological-inspired intelligent heuristic dynamic programming (Bio-int-HDP), is proposed by integrating biological regulatory mechanisms and operational decision. 2) The ultra-short feedback mechanism of the endocrine system and operational decision is innovatively incorporated into a neural network and algorithm's new modules, its convergence speed and predictive control speed for nonlinear time-varying systems is increased by improving the structure of such neural network and algorithm's modules. 3) The Bio-int-HDP is applied to addressing a real-world challenge-regulating a multi-tributary system. Simulation demonstrates that the algorithm can successfully achieve high-precision, real-time flow controls in the upper tributaries of the Murray River, which provides an effective and accurate flood warning and control method for the flood prone multi-tributary area.

Proposed model prediction control:

Overall structure of algorithm: Considering that the system is controlled by multiple variables, based on the heuristic dynamic programming (HDP), the priority factor module is introduced to optimize the systems control process based on the operational research decision theory, and the priority factor is adjusted based on a biological system's regulation mechanism. At the same time, inspired by the biological central nervous system, the central coordination module is introduced to make real-time correction to each module according to the deviation between the output value and the set value. The system structure diagram of the Bio-int-HDP is shown in Fig. 1. The dotted box part in Fig. 1 is the overall structure of the Bio-int-HDP. The model network module acts as the prediction model part of the controller. The action network module and the priority factor module act as the rolling optimization part of the controller. The critic network module acts as the feedback correction part of the controller. At the same time, the central coordination module coordinates the whole system according to the ultra-short feedback mechanism, making the predictive control more rapid, stable and accurate.
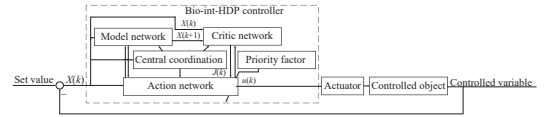

Fig. 1. Structure diagram of the Bio-int-HDP predictive control system.

Structural of multi-level feedback neural network: The traditional neural network only adjusts the parameters through the deviation between the output value and the real value. The feedback data is less, the convergence speed and accuracy of the network is not satisfactory. The ultra-short feedback mechanism of hormone regulation refers to the rapid regulation of gland hormone concentration on gland secretion, which can regulate the secretion of hormone before the regular feedback action.

Based on the ultra-short feedback regulation mechanism, the structure of a BP neural network is improved. The structure of the improved network is shown in Fig. 2. The improved neural network adds feedback information within and between layers, which enables the network to quickly perceive the internal information and improves the convergence speed and accuracy. In the Bio-int-HDP, the model network module, the action network module and the critic network module all choose the improved biological neural network.

Calculation of the proposed algorithm:

Model network module: The forward calculation of the module is divided into five steps

$$M_m = [u(k), x(k)] \tag{1}$$

$$
\begin{aligned}
m_{h1} = W_{m1} \times (M_m - \alpha_{m1} \times (1 - e^{-W_{m1} \times M_m}/1 + e^{-W_{m1} \times M_m}) \\
- \alpha_{m2} \times W_{m2} \times (1 - e^{-W_{m1} \times M_m}/1 + e^{-W_{m1} \times M_m}))
\end{aligned}
\tag{2}
$$

$$\hat{x}(k+1) = W_{m2} \times [(1 - e^{-W_{m1} \times M_m})/(1 + e^{-W_{m1} \times M_m})] \tag{3}$$

$$m_{h2} = (1 - e^{-m_{h1}})/(1 + e^{-m_{h1}}) \tag{4}$$

$$x(k+1) = W_{m2} \times m_{h2} + \vartheta_m \times \hat{x}(k+1) \tag{5}$$

where M_m is the input variable, W_{m1} is the weight matrix from input layer to hidden layer, W_{m2} is the weight matrix from hidden layer to output layer, m_{h1} is the inactive value of hidden layer, m_{h2} is the acti-
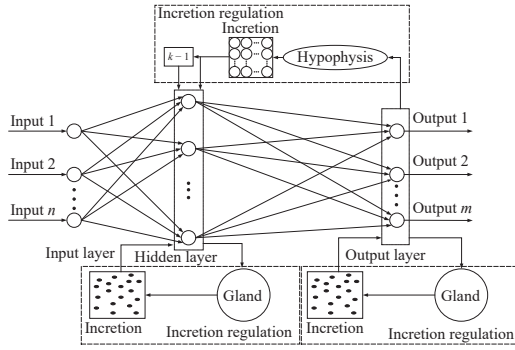
Fig. 2. Structure of a three-layer self-feedback BP neural network.

vated value of hidden layer, and $\hat{x}(k+1)$ is the pre-output, $x(k+1)$ is the final output, α_{m1} and α_{m2} are the priority factors regulated by the priority factor module, and ϑ_m is the optimization coefficient.

Gradient descent method is used to adjust the weight matrix of the module. The error of the model network module is defined as

$$e_{mp} = x_p(k+1) - x(k+1) \tag{6}$$

$$E_{mp} = (1/2)e_{mp}^2 = (1/2)[x_p(k+1) - x(k+1)]^2 \tag{7}$$

where e_{mp} is the error between the real value and the predicted value, and $x_p(k+1)$ is the real value of the system at the next moment. After forward calculation, the value of the weight matrix is adjusted according to following formulas:

$$\Delta W_{m1} = l_{mp} \times -(\partial E_{mp}/\partial W_{m1}) \tag{8}$$

$$W_{m1} = W_{m1} + \Delta W_{m1} \tag{9}$$

$$\Delta W_{m2} = l_{mp}[-\partial E_{mp}/\partial W_{m2}] \tag{10}$$

$$W_{m2} = W_{m2} + \Delta W_{m2} \tag{11}$$

where $l_{mp} \in (0, 1)$ is the learning rate of the module. In the process of system operation, l_{mp} will be regulated in real time by the central coordination module according to different states of the system.

Action network module: The forward calculation of the module is divided into four steps

$$a_{h1} = W_{a1} \times (x(k) - \alpha_{a1} \times (1 - e^{-W_{a1} \times x(k)}/1 + e^{-W_{a1} \times x(k)})$$
$$- \alpha_{a2} \times W_{a2} \times (1 - e^{-W_{a1} \times x(k)}/1 + e^{-W_{a1} \times x(k)})) \tag{12}$$

$$a_{h2} = (1 - e^{-a_{h1}})/(1 + e^{-a_{h1}}) \tag{13}$$

$$\hat{u} = W_{a2} \times [(1 - e^{-W_{a1} \times x(k)})/(1 + e^{-W_{a1} \times x(k)})] \tag{14}$$

$$u = W_{a2} \times a_{h2} + \vartheta_a \times \hat{u} \tag{15}$$

where $x(k)$ is the input of the module, W_{a1} is the weight matrix from input layer to hidden layer, W_{a2} is the weight matrix from hidden layer to output layer, a_{h1} is the inactive value of the hidden layer, a_{h2} is the activated value of the hidden layer, $\hat{u}$ is the pre-output, u is the final output, α_{a1} and α_{a2} are the priority factors regulated by the priority factor module, and ϑ_a is the optimization coefficient.

The goal of the module is to minimize the system performance index J, and the gradient descent method is used to adjust the weights.

$$\Delta W_{a2} = l_{ap} \times [-(\partial \hat{J}/\partial W_{a2})] = l_{ap} \times [-(\partial \hat{J}/\partial u) \times (\partial u/\partial W_{a2})] \tag{16}$$

$$W_{a2} = W_{a2} + \Delta W_{a2} \tag{17}$$

$$\Delta W_{a1} = l_{ap} \times [-(\partial \hat{J}/\partial W_{a1})] = l_{ap} \times [-(\partial J/\partial u) \times (\partial u/\partial W_{a1})] \tag{18}$$

$$W_{a1} = W_{a1} + \Delta W_{a1} \tag{19}$$

where $l_{ap} \in (0,1)$ is the learning rate of the module. In the process of system operation, l_{ap} will be regulated in real time by the central coordination module according to different states of the system.

Critic network module: The forward calculation of the module is divided into three steps

$$c_{h1} = W_{c1} \times (x - \alpha_c \times W_{c1}) \tag{20}$$

$$c_{h2} = (1 - e^{-c_{h1}})/(1 + e^{-c_{h1}}) \tag{21}$$

$$J = W_{c2} \times c_{h2} \tag{22}$$

where x is the input, W_{c1} is the weight matrix from the input layer to the hidden layer, W_{c2} is the weight matrix from the hidden layer to the output layer, c_{h1} is the inactivated value of the hidden layer, c_{h2} is the activated value of the hidden layer, J is the performance index

and α_c is the feedback coefficient of the hidden layer.

The objective function is the error function as follow:

$$e_c(k) = \hat{J}(k) - U(k) - \xi \times \hat{J}(k+1). \tag{23}$$

The goal of the module is to minimize the following formula:

$$E_c(k) = (1/2)e_c(k)^2. \tag{24}$$

The gradient descent method is used to adjust the weights:

$$\Delta W_{c1} = l_c \times -(\partial E_c/\partial W_{c1}) =$$
$$- l_c \times e_c \times [W_{c2}' \times (1 - m_{h2} \times m_{h2})] \times x \tag{25}$$

$$W_{c1} = W_{c1} + \Delta W_{c1} \tag{26}$$

$$\Delta W_{c2} = l_c \times [-(\partial E_c/\partial W_{c2})] = -l_c \times e_c \times c_{h2} \tag{27}$$

$$W_{c2} = W_{c2} + \Delta W_{c2} \tag{28}$$

where $l_c \in (0, 1)$ is the learning rate of the module. In the process of system operation, l_c will be real-time regulated by the central collaborative module according to different system states.

Priority factor module: Priority factor module adjusts priority factors according to the importance of each variable, so as to improve the overall control effect of the control system. Priority factor module firstly obtains the average value of each input variable of the module as follow:

$$EX = \sum_{i=1}^{n} x(i)/n. \tag{29}$$

Calculate the variance of each variable

$$DX = \sum_{i=1}^{n} (x(i) - EX)^2/n. \tag{30}$$

Calculate the covariance of variable and controlled variable

$$Cov(X, Y) = E(X - EX)(Y - EY) = EXY - EX \times EY. \tag{31}$$

Finally, obtain the correlation coefficient between each variable and the controlled variable

$$\rho_{XY} = Cov(X, Y)/(\sqrt{DX} \times \sqrt{DY}). \tag{32}$$

According to the ultra-short feedback regulation rule of biological regulation mechanism, priority factors can be adjusted in real time.

$$(\rho_{X_iY}) = \begin{cases} f_{up}(\rho_{X_iY}), & 0 \leq \rho_{X_iY} \leq 1 \\ f_{down}(\rho_{X_iY}), & -1 \leq \rho_{X_iY} < 0 \end{cases} \tag{33}$$

that,

$$f_{up}(\rho_{X_iY}) = \text{sign}(\rho_{X_iY}) \times [|\rho_{X_iY}|/(\upsilon_1 + |\rho_{X_iY}|)] \tag{34}$$

$$f_{down}(\rho_{X_iY}) = \text{sign}(\rho_{X_iY}) \times [\upsilon_2/(\upsilon_2 + |\rho_{X_iY}|)] \tag{35}$$

υ_1 is constant when $\rho_{X,Y} \geq 0$, and υ_2 is constant when $\rho_{X,Y} < 0$.

The priority factors α_{m1}, α_{m2}, α_{a1}, α_{a2} are real-time regulated along with system control according to (29)–(35).

Central coordination module: In the traditional HDP, the learning rates are fixed values selected by experience. However, learning rate has the corresponding optimal value according to the different state. Based on the regulation mechanism of biological hormone secretion, the central coordination module adjusts the learning rate according to the control deviation of the control system, so as to improve the overall control effect.

The learning rate of model network module l_m can be written as

$$l_m = a_m \times f(\alpha) + \beta_m \tag{36}$$

where α is the difference between the output value of action network module and the set value, β_m is the basic learning rate of model network module and a_m is the constant coefficient of model network module learning rate. $f(\alpha)$ can be written as

$$f(\alpha) = \begin{cases} f_{up}(\alpha), & \alpha \geq 0 \\ f_{down}(\alpha), & \alpha < 0 \end{cases} \tag{37}$$

so that,

$$f_{up}(\alpha) = \text{sign}(\alpha) \times [|\alpha|/(\delta_1 + |\alpha|)] \tag{38}$$

$$f_{down}(\alpha) = \text{sign}(\alpha) \times [1/(\delta_2 + |\alpha|)] \tag{39}$$

where δ_1 is constant when $\alpha \geq 0$, and δ_2 is constant when $\alpha < 0$.

The learning rate of critic network module l_c can be written as

$$l_c = a_c \times f(\alpha) + \beta_c \tag{40}$$

where $f(\alpha)$ is shown in (37), β_c is the basic learning rate of the critic network module, α_c is the constant coefficient of the critic network module learning rate.

The learning rate of action network module l_a can be written as

$$l_a = a_a \times f(\alpha) + \beta_a \tag{41}$$

where $f(\alpha)$ is shown in (37), β_a is the basic learning rate of the action network module, a_a is the constant coefficient of the learning rate for the action network module.

An illustrative example: In order to verify the predictive control effect of the Bio-int-HDP in a nonlinear time-varying system, the upper tributaries of the Murray River as shown in Fig. 3 is selected as the research objects.
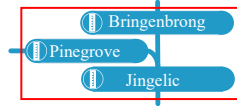

Fig. 3. Simplified diagram of relationship between rivers.

After pretreatment of relevant watershed data, six groups of data were selected as input based on correlation coefficients: the flow $Bc(k-1)$ and level $Bcl(k-1)$ of Bringenbrong, the flow $Pc(k-1)$ and level $Pcl(k-1)$ of Pinegrove, the flow $Jc(k-1)$ and level $Jcl(k-2)$ of Jingellic. The flow $Jc(k)$ was selected as output. A total of 3000 sets of data were selected as training data, and 300 sets of data were selected as test data. The flow prediction curve is shown in Fig. 4, and the comparison of flow prediction is shown in Table 1. It can be seen that compared with the Bio-int-HDP, both self-feedback BP and BP have a large deviation in predicting the flow. The above results show that the prediction model in the algorithm can predict the future flow well and meet the requirement of predictive control.
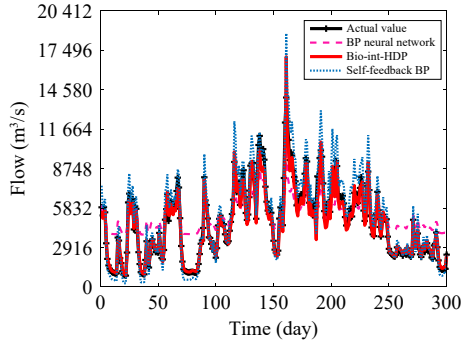

Fig. 4. Comparison of prediction effect.

Table 1. Comparison of Flow Prediction

Algorithm	MAE (m^3/s)	MRE (%)	RMSE (m^3/s)
BP	1080.21	0.3727	1303.30
Self-feedback BP	604.95	0.1348	817.93
Bio-int-HDP	250.91	0.0491	396.77

In this letter, Bio-int-HDP is compared with internal model control (IMC) and HDP. The results are shown in Fig. 5. The end of the algorithm optimization process should satisfy one of the following termination criteria. 1) The iteration number has reached the maximum generation number. 2) The fitness value of global best solution is smaller than the set value, which is called iteration convergence. Based on the above termination criteria, the training time of the three algorithms is statistically analyzed under the same termination criteria. In order to eliminate the data contingency, the single point training time of three algorithms was carried out for 50 simulation experiments, and the averaged results are shown in Table 2.

It can be seen that the Bio-int-HDP responds quickly to the predictive control requirements of complex time-varying systems and is regulated in a timely manner. Therefore, Bio-int-HDP is a new intelligent algorithm superior to traditional MPC.

Conclusions: This letter proposed a model predictive control algorithm based on biological regulatory mechanism and operational research decision, called Bio-int-HDP. The structure of the neural network and HDP is improved by introducing a biological regulation theory and operational research decision theory. Experimental results show that the Bio-int-HDP can respond quickly to the predictive control requirements of complex time-varying systems and regulate in a timely manner. The algorithm has potential in flood warning and control since Bio-int HDP has higher prediction and control accuracy in multi-tributary watershed flow prediction and control. Due to
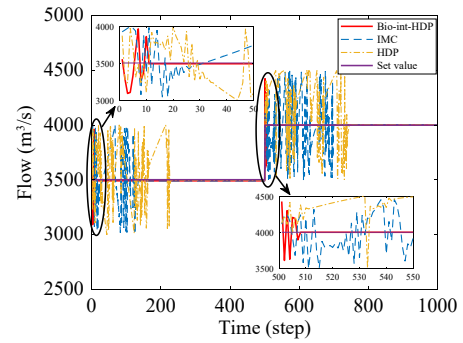

Fig. 5. Comparison of control effect.

Table 2. Comparison of Forecast Controller's Training Time

Evaluation index	IMC	HDP	Bio-int-HDP
Training time (s)	15.84	21.11	0.80

the complexity of the actual system and difficulty in obtaining data, future research will focus on methods to reduce the amount of data and the complexity of controlled objects while ensuring the high accuracy and applicability of the algorithm.

References

[1] J. Testud, J. Richalet, A. Rault, and J. Papon, "Model predictive heuristic control: Applications to industial processes," *Automatica*, vol. 14, no. 5, pp. 413–428, 1978.

[2] H. Yu *et al.*, "Nonparametric model prediction for intelligent regulation of human cardiorespiratory system to prescribed exercise medicine," *IEEE Access*, vol. 8, pp. 224621–224630, 2020.

[3] J. Liu, Y. Gu, Y. Cheng, and X. Wang, "Prediction of breast cancer pathogenic genes based on multi-agent reinforcement learning," *Acta Autom. Sinica*, vol. 48, pp. 1–13, 2022.

[4] R. R. Nair and L. Behera, "Robust adaptive gain higher order sliding mode observer based control-constrained nonlinear model predictive control for spacecraft formation flying," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 1, pp. 367–381, Jan. 2018.

[5] G. Miloud, S. Hicham, and B. Youcef, "Sensorless speed control of synchronous reluctance motors using model predictive control associated with model reference adaptive system," in *Proc. IEEE 1st Int. Maghreb Meeting Conf. Sciences and Techniques of Autom. Control and Computer Engineering MI-STA*, 2021, pp. 230–233.

[6] M. Wang, Y. Wang, T. Chai, and X. Zhang, "External generalized predictive cascade control for main steam temperature based on weight factor self-regulating," *Acta Autom. Sinica*, vol. 48, no. 2, pp. 418–433, 2022.

[7] Y. Cheng, L. Huang, and X. Wang, "Authentic boundary proximal policy optimization," *IEEE Trans. Cyber.*, 2021. DOI: 10.1109/TCYB.2021.3051456.

[8] J. Salt, J. Alcaina, Á. Cuenca, and A. Baños, "Multirate control strategies for avoiding sample losses. Application to UGV path tracking," *ISA Trans.*, vol. 101, pp. 130–146, 2020.

[9] G. Xia, Y. Hua, X. Tang, L. Zhao, and W. Chen, "Internal-model control of vehicle chassis based on wavelet-network dynamic inversion method," *Int. J. Vehicle Autonomous Syst.*, vol. 14, no. 2, pp. 170–195, 2018.

[10] Q. Yang, J. Sun, and J. Chen, "Output consensus for heterogeneous linear multiagent systems with a predictive event-triggered mechanism," *IEEE Trans. Cyber.*, vol. 51, no. 4, pp. 1993–2005, 2021.

[11] X. Zhang, Z. Yuan, H. Cui, and P. Luo, "Application of neural network predictive control in cement combined grinding," in *Proc. Chinese Control And Decision Conf.*, 2020, pp. 2414–2418.

[12] J. Thomas and A. Hansson, "Enumerative nonlinear model predictive control for linear induction motor using load observer," in *Proc. UKACC Int. Conf. Control*, 2014, pp. 373–377.

[13] F. Yang, Y. Shi, Y. Zhang, L. Liu, and W. Guan, "Research on the control of the yellow river diversion project based on the decentralized model predictive control," in *Proc. 32nd Youth Academic Annual Conf. Chinese Association of Automation*, 2017, pp. 681–686.

[14] A. Islam, "Model predictive control for flood regulation," in *Proc, Int. Conf. Intelligent Computing, Instrumentation and Control Technologies*, 2017, pp. 275–279.