Improved offset-free model predictive control utilizing learned model-plant mismatch map

Sang Hwan Son * **Jong Woo Kim** ** **Tae Hoon Oh** ***
GiBaek Lee **** **Jong Min Lee** ***

** School of Chemical and Biomolecular Engineering, Pusan National
University, Busan 46241, South Korea (e-mail: shson@ pusan.ac.kr)
** Technische Universität Berlin, KIWI Biolab, Ackerstrasse 76, 13355
Berlin, Germany (e-mail: jong.w.kim@tu-berlin.de)
*** School of Chemical and Biological Engineering, Seoul National
University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea (e-mail:
rozenk@snu.ac.kr, jongmin@snu.ac.kr)
**** Department of Applied Chemical and Energy Engineering, Korea
National University of Transportation, 50, Daehak-ro, Chungju-si,
Chungcheongbuk-do, Korea (e-mail: glee@ut.ac.kr)*

Abstract: The requirement for a framework that effectively overcomes the limitation of model-based and data-driven control strategies by combining both methods continues to grow. In this study, we propose an approach that learns the model-plant mismatch map and utilizes it based on the offset-free model predictive control (MPC). Specifically, the mismatch map is learned via general regression neural network (GRNN) that has been applied in broad range of fields based on the data from the process, and then the learned mismatch information is provided to the MPC system. In addition, since the approximated mismatch map via GRNN cannot be perfect, an additional supplementary disturbance estimator is utilize to ensure the zero-offset tracking property. Finally, the learned and supplementary disturbance signals are applied to the target problem and the optimal control problem based on the offset-free MPC framework. The effectiveness of the proposed combined model-based and data driven framework is demonstrated by closed-loop simulation. The result shows that the proposed framework can improve the closed-loop tracking performance by utilizing both the learned mismatch information from GRNN and the stabilizing property of the supplementary disturbance estimator.

Keywords: Model predictive control, general regression neural network, model-plant mismatch, offset-free tracking.

1. INTRODUCTION

Model predictive control (MPC) predicts the propagation of physical system in a horizon with known dynamics of the system and derive an effective and reliable solution based on the model Kim et al. (2019); Jeong and Lee (2020); Kim et al. (2020c). Thus, the closed-loop performance of MPC is directly related to the accuracy of model Kim et al. (2018); Jeong (2020). However, since model-plant mismatch and unmeasured disturbances always exist in real systems, MPC usually cannot achieve optimal performance. Machine learning (ML) performs a specific task such as classification, regression, clustering and policy learning from the behavior of the real system Oh et al. (2021); Kim et al. (2020a). Thus, ML does not require any given model or prior assumption about the system, and can implicitly manage uncertainties. However, this model-free pure learning approach without any prior information of the system is often limited because it requires a large amount of data and exploratory behavior can intrinsically damage the system Kim and Lee (2020); Kim et al. (2020b, 2021).

Since the advantages and limitations of model-based approach and data-driven approach are complementary to each other, the combination of MPC and ML is an emerging area of research. Zhang et al. (2016), Chen et al. (2018), and Hertneck et al. (2018) derive a reliable policy by supervised learning from a nominal MPC policy to avoid the system failure from exploratory policy. Williams et al. (2017), Gillespie et al. (2018), Kaiser et al. (2018), and Thuruthel et al. (2018) improve the optimality of MPC by continuously updating the dynamic model for MPC with sampled data. Koryakovskiy et al. (2018) improves the closed-loop performance of MPC by deriving a direct compensatory control action using reinforcement learning with the same performance measure of MPC. Terzi et al. and Manzano et al. Terzi et al. (2019); Manzano et al. (2020), propose learning-based robust predictive control

frameworks that identify prediction model with bounded uncertainty utilizing data from the process and design robust MPC system. Vaupel et al. Vaupel et al. (2020) improves nonlinear MPC computation by approximating the optimal control policy via a neural network model and applying the policy model to initialize the optimal control system.

This study proposes a framework that performs predictive control based on a model-based on prior knowledge of the process, learns model-plant mismatch map from process data, and applies it for closed-loop reference tracking. The proposed framework is developed based on the offset-free MPC that is one of the most popular methods for compensating for the model-plant mismatch in the mode-based control system design Muske and Badgwell (2002); Pannocchia and Rawlings (2003); Pannocchia and Bemporad (2007). The disturbance estimator approach, which tracks the reference trajectory in the presence of plant-model mismatch or unmeasured nonzero mean disturbances by augmenting the disturbance model to the plant model and deriving compensatory disturbance from the estimator Maeder et al. (2009); Maeder and Morari (2010); Son et al. (2022), is utilized in the proposed framework. Specifically, in the proposed framework, the general regression neural network (GRNN) proposed by Specht (1991) is utilized to approximate a map for model-plant mismatch in the controlled variable space using the plant data. Subsequently, the learned disturbance value from the approximated mismatch map is applied to the target problem and the optimal control problem of the offset-free MPC system. However, since the trained mismatch map cannot provide a perfect model-plant mismatch information, the supplementary disturbance estimator is also constructed to ensure the zero-offset tracking property by introducing additional supplementary disturbance signal to handle the transient state and the change in the characteristics of the process during the operation. By this, the proposed method can efficiently improve the closed-loop reference tracking performance of the model-based control system by utilizing both the learned model-plant mismatch information from the GRNN map and the stabilizing property of the disturbance estimator.

2. PRELIMINARIES

2.1 Linear offset-free model predictive control

We consider the following discrete time-invariant plant

$$\begin{cases} x_p(k+1) = f_p(x_p(k), u(k)) \\ y_p(k) = g_p(x_p(k)) \\ z_p(k) = H y_p(k) \end{cases} \quad (1)$$

where $x_p \in \mathbb{R}^{n_{x_p}}$, $u \in \mathbb{R}^{n_u}$, $y_p \in \mathbb{R}^{n_y}$, and $z_p \in \mathbb{R}^{n_z}$ denote the state, input, output, and controlled variables of the plant, respectively. Generally, we assume that the matrix H has full row rank.

The following process constraints are also considered.

$$u \in \mathcal{U}, \ x_p \in \mathcal{X} \quad (2)$$

$\mathcal{U}$ and $\mathcal{X}$ are compact polyhedral constraint sets for input and plant state, respectively.

In linear offset-free MPC, to drive the plant controlled variables z_p to the reference trajectory, a control system

is designed based on the linear time-invariant state-space model in (3).

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) \end{cases} \quad (3)$$

where $x \in \mathbb{R}^{n_x}$ and $y \in \mathbb{R}^{n_y}$ denote the state and output of the linear system, respectively. The reference signal $r(k)$ is assumed to converge to a constant r_∞ as $k \to \infty$. Additionally, the linear system in (3) is also assumed to be controllable and observable.

One of the most commonly used way to accomplish the zero-offset tracking of the reference trajectory by compensating for the mismatch between the the model in (3) and plant in (1) is to introduce an additional integrating state, which is usually called disturbance as follows (Pannocchia and Rawlings, 2003; Maeder et al., 2009):

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + B_d d(k) \\ d(k+1) = d(k) \\ y(k) = Cx(k) + C_d d(k) \end{cases} \quad (4)$$

where $d \in \mathbb{R}^{n_d}$ denotes the disturbance variable, and $B_d \in \mathbb{R}^{n_x \times n_d}$ and $C_d \in \mathbb{R}^{n_y \times n_d}$ are the matrices describing the disturbance dynamics.

The estimator for state and disturbance variable can be constructed as follows:

$$\begin{bmatrix} \hat{x}(k+1) \\ \hat{d}(k+1) \end{bmatrix} = \begin{bmatrix} A & B_d \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{x}(k) \\ \hat{d}(k) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(k)$$
$$+ \begin{bmatrix} L_x \\ L_d \end{bmatrix} \left(-y_p(k) + [C \ C_d] \begin{bmatrix} \hat{x}(k) \\ \hat{d}(k) \end{bmatrix} \right) \quad (5)$$

where $L_x \in \mathbb{R}^{n_x \times n_y}$ and $L_d \in \mathbb{R}^{n_d \times n_y}$ denote the gains for state and disturbance variable, respectively. Then, proper gain values are selected to make the estimator dynamics stable.

Subsequently, the target state $\bar{x}$ and target input $\bar{u}$ values are derived by solving the following target problem in (6) based on the estimated state $\hat{x}$ and disturbance $\hat{d}$ values obtained from (5). In the target problem, a desired input and output pair $(\bar{u}_s, \bar{y}_s)$ and process constraints are also considered.

$$\min_{\bar{x}, \bar{u}} \quad ||\bar{u} - \bar{u}_s||_{Q_{\bar{u}}}^2 + ||\bar{y} - \bar{y}_s||_{Q_{\bar{y}}}^2 \quad (6a)$$

$$\text{s.t.} \quad \begin{bmatrix} A - I & B \\ HC & 0 \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{u} \end{bmatrix} = \begin{bmatrix} -B_d \hat{d} \\ \bar{z} - HC_d \hat{d} \end{bmatrix} \quad (6b)$$

$$\bar{y} = C\bar{x} + C_d \hat{d} \quad (6c)$$

$$\bar{u} \in \mathcal{U}, \ \bar{x} \in \mathcal{X} \quad (6d)$$

where $Q_{\bar{u}} \in \mathbb{R}^{n_u \times n_u}$ and $Q_{\bar{y}} \in \mathbb{R}^{n_y \times n_y}$ denote the weighting matrices.

Then, a finite-horizon optimal control problem in (7) is solved to obtain the optimal input to drive the controlled variables to the reference values.

$$\min_{u_i} \quad \phi_t(x_N, \bar{x}) + \sum_{i=0}^{N-1} \phi(x_i, u_i, \bar{x}, \bar{u}) \quad (7a)$$

$$\text{s.t.} \quad x_0 = \hat{x}, \ d = \hat{d} \quad (7b)$$

$$x_{i+1} = Ax_i + Bu_i + B_d d \quad (7c)$$

$$u_i \in \mathcal{U}, \ x_{i+1} \in \mathcal{X}, \ x_N \in \mathcal{X}_N \quad (7d)$$

$$i = 0, \ldots, N - 1.$$

where $\mathcal{X}_N$ denotes the terminal constraint set, $\phi(\cdot)$ is the single stage cost, and $\phi_t(\cdot)$ is the terminal stage cost, respectively, as in (8a) and (8b).

$$\phi(x_i, u_i, \bar{x}, \bar{u}) := |x_i - \bar{x}|^2_{Q_x} + |u_i - \bar{u}|^2_{Q_u} \qquad (8a)$$

$$\phi_t(x_N, \bar{x}) := |x_N - \bar{x}|^2_{Q_x^N} \qquad (8b)$$

where $|v|^2_Q := v^\top Q v$, $Q_x \in \mathbb{R}^{n_x \times n_x}$, and $Q_u \in \mathbb{R}^{n_u \times n_u}$ denote the weighting matrices.

3. LEARNING AND UTILIZATION OF MISMATCH MAP IN OFFSET-FREE MPC

The linear offset-free MPC introduced in Section 2 can accomplish zero-steady state offset by compensating for the model-plant mismatch in proper condition. However, since this method should gradually estimate the disturbance variable from the measured prediction error of the output, considerable delay in estimating of a proper disturbance value can be occurred. Moreover, since the model-plant mismatch values are not saved, even in the case where the system reaches a certain state point that it has reached already at past, the proper disturbance value at that point should be gradually estimated again. These limitations make it difficult for the nominal linear offset-free MPC scheme to efficiently compensate for the model-plant mismatch. Therefore, reference tracking performance of the control system can be considerably degraded during the operation. To address these limitations, in this section, an improved offset-free MPC framework that learns the intrinsic model-plant mismatch map from the process data and utilizes the information from the learned mismatch map is developed.

3.1 Learning of model-plant mismatch map

In this study, learning and utilization of a reduced model-plant mismatch map only on the steady-state manifold at the controlled variable space is proposed. Since, the steady-state model-plant mismatch map is a very tiny part of the entire map, this reduced mismatch map can be efficiently obtained only using a small amounts of data and computation.

When a set-point $\bar{r}$ can be reached at the process, subsequently, (9) and (10) are satisfied at that steady-state point based on (5).

$$\hat{x}_\infty = (A + L_x C)\hat{x}_\infty + (B_d + L_x C_d)\hat{d}_\infty$$
$$+ Bu_\infty - L_x y_{p,\infty} \qquad (9)$$

$$\hat{d}_\infty = L_d C \hat{x}_\infty + (I + L_d C_d)\hat{d}_\infty - L_d y_{p,\infty} \qquad (10)$$

where $y_{p,\infty}$ is the plant output at steady-state point that satisfies $\bar{r} = H y_{p,\infty}$, u_∞ is the input, $\hat{x}_\infty$ denotes the system state, and $\hat{d}_\infty$ is the disturbance value at steady-state.

Then, a $(\hat{x}_\infty, \hat{d}_\infty)$ pair in (11) can be derived from the $(y_{p,\infty}, u_\infty)$ pair by rearranging (9) and (10).

$$\begin{bmatrix} \hat{x}_\infty \\ \hat{d}_\infty \end{bmatrix} = \begin{bmatrix} A - I + L_x C & B_d + L_x C_d \\ L_d C & L_d C_d \end{bmatrix}^{-1} \begin{bmatrix} L_x & -B \\ L_d & 0 \end{bmatrix} \begin{bmatrix} y_{p,\infty} \\ u_\infty \end{bmatrix}. \qquad (11)$$

Now, we define the intrinsic relation between $\bar{r}$ and $\hat{d}_\infty$ as follows:

$$\hat{d}_\infty = f_d(\bar{r}). \qquad (12)$$

The function $f_d : \mathbb{R}^{n_z} \to \mathbb{R}^{n_d}$ represents the intrinsic model-plant mismatch. Then, the model-plant mismatch $f_d(\cdot)$ in (12) is approximated via GRNN using the disturbance data from the process.

GRNN is a variation of neural network models based on radial basis function (RBF) for non-parametric regression proposed in Specht (1991). We can also interpret GRNN as a normalized RBF network with hidden units centered at every training sample. In GRNN, the predicted output $o(i)$ from the input i can be described as a weighted average value of outputs for the training set as follows:

$$o(i) = \frac{\sum_{s=1}^{N_s} o_s \omega(i, i_s)}{\sum_{s=1}^{N_s} \omega(i, i_s)} \qquad (13)$$

where N_s denotes the number of training samples, and $\omega(i, i_s)$ is the weight. Each weight value is an RBF output that is the exponential of the negatively scaled distance value between the new pattern and each given training pattern as in (14)

$$\omega(i, i_s) = e^{-(i-i_s)^\top(i-i_s)/2\sigma^2} \qquad (14)$$

where σ denotes the smoothing factor that represents the width of RBF.

Back-propagation neural networks (BPNNs) commonly require forward and backward pass training, but in the case of GRNN, it is a single-pass learning network that does not requires back-propagation. The only adjusted parameter is the smoothing factor ρ in GRNN (Sun et al., 2019). Therefore, GRNN takes significantly less time to train the network compared to BPNNs. Due to this significant benefit of rapid training, GRNN is known to be suitable for on-line training systems that require minimal computation (Rooki, 2016). However, since GRNN has a structure that has the same number of nuerons in the hidden layer identical to the number of training samples, it has a shortcoming that the size of network can be huge.

3.2 Utilization of model-plant mismatch map

In this section, the framework to utilize the following learned model-plant mismatch map via GRNN.

$$\hat{d}^\ell = \hat{f}_d(\bar{r}) \qquad (15)$$

where $\hat{f}_d(\cdot)$ represents an approximated function of $f_d(\cdot)$ in (12) with GRNN.

Since the mismatch map in (15) is not the entire map but a reduced map as described in the previous section, this map does not provide the perfect mismatch information for the entire system state. Therefore, an additional supplementary disturbance estimator is introduced to exploit the stabilizing property of the disturbance estimator. At this revised form of estimator, the supplementary disturbance variable $\hat{d}^s$ is updated considering the disturbance value $\hat{d}^\ell$ from the mismatch map as in (16).

$$\begin{bmatrix} \hat{x}(k+1) \\ \hat{d}^s(k+1) \end{bmatrix} = \begin{bmatrix} A & B_d \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{x}(k) \\ \hat{d}^s(k) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} B_d \\ 0 \end{bmatrix} \hat{d}^\ell(k)$$
$$+ \begin{bmatrix} L_x \\ L_d \end{bmatrix} (-y_p(k) + C\hat{x}(k) + C_d(\hat{d}^\ell(k) + \hat{d}^s(k))) \qquad (16)$$

where $\hat{x}^{\ell,s}$ is the state value estimated considering the disturbance values from the map $\hat{d}^\ell$ and supplementary estimator $\hat{d}^s$.

Then, we can exploit the stabilizing property of the supplementary disturbance estimator in (16). By this, even in the cases where the learned mismatch map $\hat{f}_d$ in (15) is not perfect or the intrinsic relation between model and plant has been changed by system transformation or additional unknown disturbance, the zero steady-state offset can be accomplished.

Then, the target values for the state $\bar{x}$ and input $\bar{u}$ values are derived from the revised form of target problem considering both the disturbance value from the mismatch map $\hat{d}^\ell$ and the supplementary estimator $\hat{d}^s$ as in (17).

$$\min_{\bar{x},\bar{u}} \quad ||\bar{u} - \bar{u}_s||^2_{Q_{\bar{u}}} + ||\bar{y} - \bar{y}_s||^2_{Q_{\bar{y}}} \tag{17a}$$

$$\text{s.t.} \quad \begin{bmatrix} A - I & B \\ HC & 0 \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{u} \end{bmatrix} = \begin{bmatrix} -B_d(\hat{d}^\ell + \hat{d}^s) \\ \bar{z} - HC_d(\hat{d}^\ell + \hat{d}^s) \end{bmatrix} \tag{17b}$$

$$\bar{y} = C\bar{x} + C_d(\hat{d}^\ell + \hat{d}^s) \tag{17c}$$

$$\bar{u} \in \mathcal{U}, \ \bar{x} \in \mathcal{X}. \tag{17d}$$

Now, a finite-horizon optimal control problem is designed as a state regulation problem based on the derived target values $(\bar{x}, \bar{u})$ from the target problem (17) to obtain the optimal input to drive the controlled variable of the system to the desired reference value considering the disturbance value from the mismatch map $\hat{d}^\ell$ and the supplementary estimator $\hat{d}^s$:

$$\min_{u_i} \quad \phi_t(x_N, \bar{x}) + \sum_{i=0}^{N-1} \phi(x_i, u_i, \bar{x}, \bar{u}) \tag{18a}$$

$$\text{s.t.} \quad x_0 = \hat{x}, \ d^\ell = \hat{d}^\ell, \ d^s = \hat{d}^s \tag{18b}$$

$$x_{i+1} = Ax_i + Bu_i + B_d(d^\ell + d^s) \tag{18c}$$

$$u_i \in \mathcal{U}, \ x_{i+1} \in \mathcal{X}, \ x_N \in \mathcal{X}_t \tag{18d}$$

$$i = 0, \ldots, N-1.$$

where $\phi(\cdot)$ and $\phi_t(\cdot)$ denote the single stage cost in (8a) and terminal stage cost in (8b), respectively.

By this, the proposed offset-free MPC framework can efficiently improve the closed-loop reference tracking performance of the control system compared to the standard offset-free MPC introduced in *Section 2* by enhancing the accuracy of prediction of future state using both the information mismatch from the learned mismatch map and the stabilizing property of the supplementary estimator.

4. NUMERICAL EXAMPLE

In this section, a closed-loop numerical simulation result is presented to demonstrate the efficacy of the proposed offset-free MPC framework compared to the standard method at a control system for reference tracking of multiple controlled variables where the set-point values change periodically.

A continuous stirred-tank reactor (CSTR) presented in Rawlings et al. (2017) where a first-order reaction, $A \to B$ takes place in the liquid phase, and the temperature of the reactor is controlled by manipulating of the external cooling jacket is utilized as a virtual plant.

The control objective of the system is to control the outlet concentration of the reactant, c, to track the reference values while regulating the reactor temperature, T, at a fixed value by manipulating the jacket temperature, T_c and the outlet flow rate, F. The dynamics of the CSTR is described below.

$$\begin{cases} \dfrac{dc}{dt} = \dfrac{F_0(c_0 - c)}{\pi r^2 h} - k_0 \exp(-\dfrac{E}{RT})c \\ \dfrac{dT}{dt} = \dfrac{F_0(T_0 - T)}{\pi r^2 h} - \dfrac{\Delta H}{\rho C_p}k_0 \exp(-\dfrac{E}{RT})c \\ \qquad + \dfrac{2U}{r\rho C_p}(T_c - T) \\ \dfrac{dh}{dt} = \dfrac{F_0 - F}{\pi r^2} \end{cases} \tag{19}$$

Then, a state-space model is derived by linearization at the steady-state:

$$c^s = 0.878 \text{ kmol/m}^3, \quad T^s = 324.5 \text{ K}, \quad h^s = 0.659 \text{ m}$$
$$T_c^s = 300 \text{ K}, \quad F^s = 0.1 \text{ m}^3/\text{min}$$

Subsequently, a discretized control-oriented model in (20) is derived with sampling instant 1 min.

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) \\ z(k) = Hy(k) \end{cases} \tag{20}$$

Process constraints are also considered in the design of control system:

$$0.83 \le c \le 0.92, \ 320 \le T \le 330, \ 0.4 \le h \le 1.2,$$
$$295 \le T_c \le 310, \ 0.07 \le F \le 0.13$$

The terminal state constraint set $\mathcal{X}_N$ is obtained as the maximal controlled invariant set of the process.

The set-point values for c and T are set to change periodically for every 15 min along the ranges $[0.84, 0.91]$ and $[321, 329]$, respectively.

We also implement a process disturbance (i.e., the outlet concentration from the virtual plant is set as 3% higher than the average concentration in CSTR, $F \to 1.03\,F$) to increase the model-plant mismatch by introducing an additional error source.

Fig. 1 illustrates the approximation result of the model-plant mismatch map f_d using GRNN from the 100 data points of the process. Since the disturbance values at steady-state points (d_1^{ss}, d_2^{ss}) are dependent on the set-point values for the controlled variables (c_{ref}, T_{ref}), the learned mismatch maps for d_1^{ss} and d_2^{ss} values are illustrated separately for the set-point pair of controlled variables. The mismatch maps for d_1^{ss} and d_2^{ss} learned via GRNN show proper prediction performance for the model-plant mismatch.

The graphs on the left in Fig. 2 represent the closed-loop trajectories of the controlled variables from the developed and standard offset-free MPC frameworks. Both frameworks show the zero steady-state offset tracking, but the developed framework show considerably superior reference tracking performance compared to the standard method by utilizing the learned model-plant mismatch information from the constructed GRNN model illustrated in Fig. 1. Furthermore, nonetheless the approximated model-plant
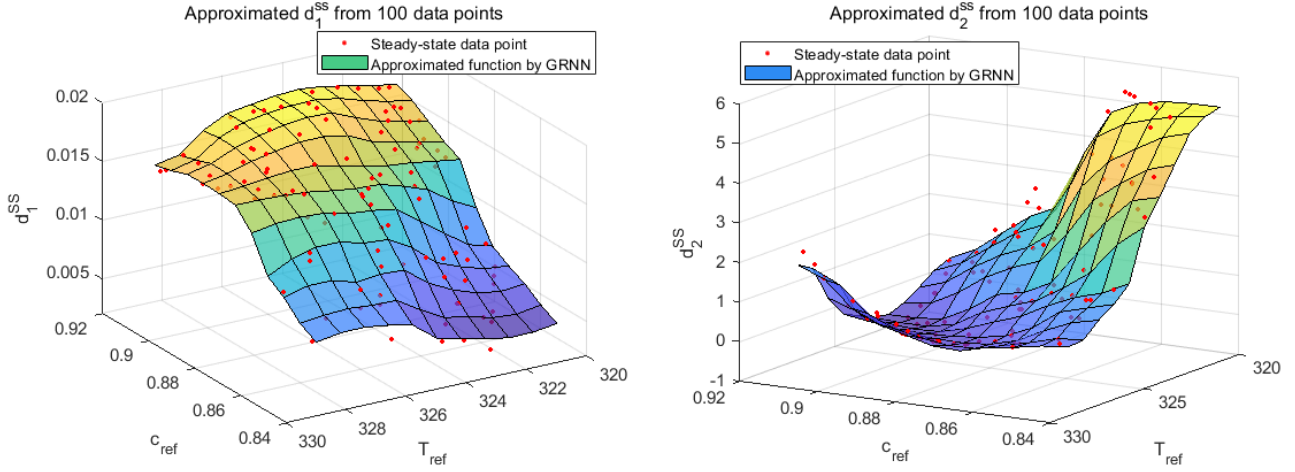
Fig. 1. Approximated model-plant mismatch map through GRNN modeling using 100 data points from the process.
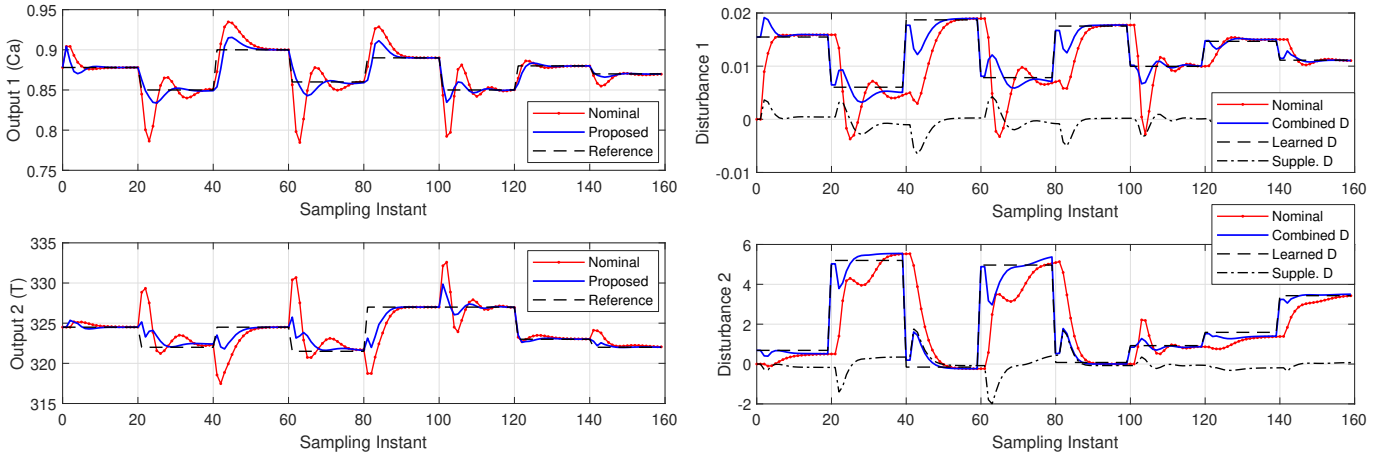

Fig. 2. Controlled variable and estimated disturbance trajectories of the developed and standard offset-free MPC under set-point change.

mismatch map via GRNN model with 100 data points does not provide the perfect prediction as shown in Fig. 1, the closed-loop trajectories of controlled variables from the developed controller show zero-offset tracking by utilizing the stabilizing property of the supplementary disturbance estimator.

The graphs on the right in Fig. 2 show the estimated value of the disturbance variables from the developed and standard offset-free MPC schemes. The blue-colored line represents the combined disturbance value from the developed framework that is the summation of the black-colored learned disturbance value and supplementary disturbance value. The learned disturbance value generally matches the steady-state disturbance values at each set-points. However, we can see a small amount of deviations of learned disturbance value from the proper disturbance value that is shown by red and blue line (e.g., near 40 and 80 mins). This implies that the information from the mismatch map in Fig. 1 is not perfect at the related set-point. In this case, the supplementary disturbance supports obtaining the proper combined disturbance values, thus, we can see that supplementary value at steady-state has non-zero value near 40 and 80 mins.

5. CONCLUSION

We propose an improved offset-free MPC framework that learns the intrinsic model-plant mismatch via GRNN from the estimated steady-state disturbance data and exploits them in the controller. As shown in the closed-loop simulation result, we can effectively improve the closed-loop performance of the model-based controller.

Moreover, since we combined machine learning and model-based control based on offset-free MPC scheme, we can exploit its own model-plant mismatch compensating property through the supplementary estimator design. Therefore, developed method can effectively improve the reference tracking performance without using enormous data, unlike existing schemes for improving performance of the model-based controller that update the entire process dynamics or directly learning the input signal to compensate for the mismatch.

ACKNOWLEDGEMENTS

REFERENCES

Chen, S., Saulnier, K., Atanasov, N., Lee, D.D., Kumar, V., Pappas, G.J., and Morari, M. (2018). Approximating explicit model predictive control using constrained neural networks. In *2018 American Control Conference (ACC)*, 1520–1527. IEEE.

Gillespie, M.T., Best, C.M., Townsend, E.C., Wingate, D., and Killpack, M.D. (2018). Learning nonlinear dynamic models of soft robots for model predictive control with neural networks. In *2018 IEEE International Conference on Soft Robotics (RoboSoft)*, 39–45. IEEE.

Hertneck, M., Köhler, J., Trimpe, S., and Allgöwer, F. (2018). Learning an approximate model predictive controller with guarantees. *IEEE Contr. Syst. Lett.*, 2(3), 543–548.

Jeong, D.H. (2020). Weighted similarity based just-in-time model predictive control for batch trajectory tracking. *Chemical Engineering Research and Design*, 162, 137–148.

Jeong, D.H. and Lee, J.M. (2020). Ensemble learning based latent variable model predictive control for batch trajectory tracking under concept drift. *Computers & Chemical Engineering*, 139, 106875.

Kaiser, E., Kutz, J.N., and Brunton, S.L. (2018). Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proc. Math. Phys. Eng. Sci.*, 474(2219), 20180335.

Kim, J.W., Oh, T.H., Son, S.H., Jeong, D.H., and Lee, J.M. (2020a). Convergence analysis of the deep neural networks based globalized dual heuristic programming. *Automatica*, 122, 109222.

Kim, J.W., Park, B.J., Oh, T.H., and Lee, J.M. (2021). Model-based reinforcement learning and predictive control for two-stage optimal control of fed-batch bioreactor. *Computers & Chemical Engineering*, 154, 107465.

Kim, J.W., Park, B.J., Yoo, H., Oh, T.H., Lee, J.H., and Lee, J.M. (2020b). A model-based deep reinforcement learning method applied to finite-horizon optimal control of nonlinear control-affine system. *Journal of Process Control*, 87, 166–178.

Kim, Y. and Lee, J.M. (2020). Model-based reinforcement learning for nonlinear optimal control with practical asymptotic stability guarantees. *AIChE Journal*, 66(10), e16544.

Kim, Y., Oh, T.H., Park, T., and Lee, J.M. (2019). Backstepping control integrated with Lyapunov-based model predictive control. *Journal of Process Control*, 73, 137–146.

Kim, Y., Park, T., Jung, C., Kim, C.H., Kim, Y.W., and Lee, J.M. (2018). Hybrid nonlinear model predictive control of lnt and urealess SCR aftertreatment system. *IEEE Transactions on Control Systems Technology*, 27(5), 2305–2313.

Kim, Y., Thierry, D.M., and Biegler, L.T. (2020c). Serial advanced-multi-step nonlinear model predictive control using an extended sensitivity method. *Journal of Process Control*, 96, 82–93.

Koryakovskiy, I., Kudruss, M., Vallery, H., Babuška, R., and Caarls, W. (2018). Model-plant mismatch compensation using reinforcement learning. *IEEE Robot. Autom. Lett.*, 3(3), 2471–2477.

Maeder, U., Borrelli, F., and Morari, M. (2009). Linear offset-free model predictive control. *Automatica*, 45(10), 2214–2222.

Maeder, U. and Morari, M. (2010). Offset-free reference tracking with model predictive control. *Automatica*, 46(9), 1469–1476.

Manzano, J.M., Limon, D., de la Peña, D.M., and Calliess, J.P. (2020). Robust learning-based MPC for nonlinear constrained systems. *Automatica*, 117, 108948.

Muske, K.R. and Badgwell, T.A. (2002). Disturbance modeling for offset-free linear model predictive control. *J. Process Control*, 12(5), 617–632.

Oh, T.H., Kim, J.W., Son, S.H., Kim, H., Lee, K., and Lee, J.M. (2021). Automatic control of simulated moving bed process with deep Q-network. *Journal of Chromatography A*, 1647, 462073.

Pannocchia, G. and Bemporad, A. (2007). Combined design of disturbance model and observer for offset-free model predictive control. *IEEE Trans. Autom. Control*, 52(6), 1048–1053.

Pannocchia, G. and Rawlings, J.B. (2003). Disturbance models for offset-free model-predictive control. *AIChE J.*, 49(2), 426–437.

Rawlings, J.B., Mayne, D.Q., and Diehl, M. (2017). *Model predictive control: theory, computation, and design*, volume 2. Nob Hill Publishing Madison, WI.

Rooki, R. (2016). Application of general regression neural network (GRNN) for indirect measuring pressure loss of Herschel–Bulkley drilling fluids in oil drilling. *Measurement*, 85, 184–191.

Son, S.H., Oh, S.K., Park, B.J., Song, M.J., and Lee, J.M. (2022). Idle speed control with low-complexity offset-free explicit model predictive control in presence of system delay. *Control Engineering Practice*, 119, 104990.

Specht, D.F. (1991). A general regression neural network. *IEEE Trans. Neural Netw.*, 2(6), 568–576.

Sun, X., Liu, J., Zhu, K., Hu, J., Jiang, X., and Liu, Y. (2019). Generalized regression neural network association with terahertz spectroscopy for quantitative analysis of benzoic acid additive in wheat flour. *Royal Society open science*, 6(7), 190485.

Terzi, E., Fagiano, L., Farina, M., and Scattolini, R. (2019). Learning-based predictive control for linear systems: A unitary approach. *Automatica*, 108, 108473.

Thuruthel, T.G., Falotico, E., Renda, F., and Laschi, C. (2018). Model-based reinforcement learning for closed-loop dynamic control of soft robotic manipulators. *IEEE Trans. Robot.*, 35(1), 124–134.

Vaupel, Y., Hamacher, N.C., Caspari, A., Mhamdi, A., Kevrekidis, I.G., and Mitsos, A. (2020). Accelerating nonlinear model predictive control through machine learning. *J. Process Control*, 92, 261–270.

Williams, G., Wagener, N., Goldfain, B., Drews, P., Rehg, J.M., Boots, B., and Theodorou, E.A. (2017). Information theoretic MPC for model-based reinforcement learning. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 1714–1721. IEEE.

Zhang, T., Kahn, G., Levine, S., and Abbeel, P. (2016). Learning deep control policies for autonomous aerial vehicles with MPC-guided policy search. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 528–535. IEEE.