

Data-Driven Control of Nonlinear Process Systems Using a Three-Degree-of-Freedom Model-on-Demand Model Predictive Control Framework

Sarasij Banerjee,* Owais Khan,* Mohamed El Mistiri,* Naresh N. Nandola,* Eric Hekler,* and Daniel E. Rivera*



Cite This: *Ind. Eng. Chem. Res.* 2025, 64, 8847–8864



Read Online

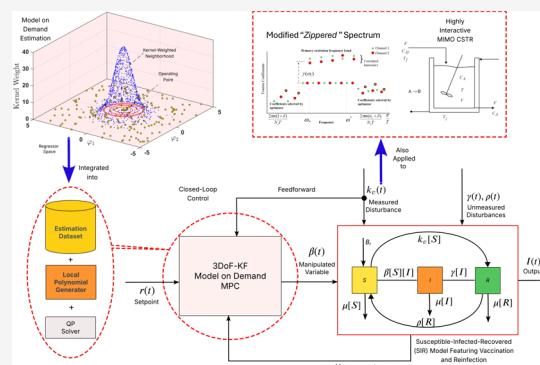
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: This paper presents a Model-on-Demand (MoD) approach to system identification and its integration with a three-degree-of-freedom Kalman filter-based Model Predictive Control (3DoF-KF MPC) framework. MoD estimation represents a hybrid of local and global modeling techniques, judiciously formulated to take advantage of both while not being computationally demanding. The 3DoF-KF MPC algorithm enables responses to set point changes and measured and unmeasured disturbances to be tuned intuitively and independently, thereby providing superior performance and ease of use over tuning with move suppression and error weights as done with conventional MPC algorithms. The algorithm proposed in this paper involves estimating MoD-based predictive models that are seamlessly integrated into 3DoF-KF MPC to generate control actions that vary with operating conditions. This results in notable performance enhancements in the context of both SISO and MIMO control compared to conventional ARX models. Performance and robustness of the 3DoF-KF MoD MPC framework are demonstrated in this paper through two case studies involving (i) epidemic control of a variant of the widely used SISO Susceptible-Infected-Removed (SIR) model and (ii) a nonlinear, highly interactive MIMO Continuous Stirred Tank Reactor (CSTR) model. The second case study further provides guidelines for designing informative databases for effective MoD-based MIMO identification and implementing 3DoF-KF MPC-based control for a demanding class of systems. Overall, this paper demonstrates technological and practical improvements in system identification and control of nonlinear SISO and MIMO systems through the synergistic integration of MoD estimation and 3DoF-KF MPC, providing an effective approach for operating complex nonlinear process systems.



1. INTRODUCTION

Data-driven approaches have gained increasing popularity in the modeling and control of dynamic systems due to their ability to address highly nonlinear processes without relying heavily on *a priori* information.^{1,2} These data-centric methods have contributed particularly to improving model-based control schemes such as model predictive control (MPC).^{3,4} Machine learning (ML)-based approaches have also seen significant application in the context of MPC.^{5,6} Typically, these control schemes rely on global nonlinear models such as recurrent neural networks (RNNs) to capture nonlinearities in dynamic systems; recent efforts propose online updating techniques to improve RNN models in real time.⁷ Although these can deliver desired performance, they come with added challenges of significant engineering and computational effort.⁸ Model-on-Demand (MoD) estimation provides an attractive alternative because it can address these practical challenges with reduced effort, while displaying improved performance over linear MPC. The concept of MoD arises from “just-in-time” modeling, first proposed by Cybenko⁹ and extensively studied by Stenman et

al.^{10,11} and Braun et al.¹² MoD relies on a user-selected regressor structure to generate time-varying local polynomial models over adaptive neighborhoods of the current operating point. Additionally, MoD reduces the number of arbitrary decisions related to data-weighing otherwise demanded by strictly local estimation algorithms. Consequently, MoD enables the user to place emphasis on generating informative databases through judicious experiment design while devoting less concern with estimator complexity. MoD-MPC has been previously studied in the context of conventional MPC algorithms which use move-suppression-based tuning.^{10,13–15} This formulation adopts a Generalized Predictive Control (GPC)-like approach to

Received: December 2, 2024

Revised: April 9, 2025

Accepted: April 10, 2025

Published: April 21, 2025



optimize the MPC objective through the use of CARIMA models. However, a superior and more robust, intuitive formulation can be achieved by integrating the three-degree-of-freedom Kalman filter-based model predictive control (3DoF-KF MPC) framework¹⁶ with MoD estimation, which allows users to independently adjust the speed-of-response related to set point tracking, measured disturbance rejection, and unmeasured disturbance rejection. This provides greater flexibility than conventional move-suppression-based tuning and demonstrates multiple benefits in terms of usability and performance while leveraging the benefits of locally linear MoD models estimated in real time, as demonstrated in the paper.

This work combines the 3DoF-KF MPC framework with Multi-Input-Multi-Output (MIMO) MoD¹⁷ estimation generating local state-space models that are updating over time, effectively accomplishing nonlinear control while displaying modest engineering complexity, comparable to traditional linear MPC. A comprehensive guide is provided to the user on designing an informative database suitable for MIMO identification, accomplishing MoD-based model estimation and validation, culminating in control of nonlinear systems using the 3DoF-KF MoD MPC; initial efforts in this regard have been demonstrated in Banerjee et al.¹⁸ The efficacy of the 3DoF-KF MoD MPC framework is illustrated in the paper through two case studies: (i) control of a modified version of the widely accepted SIR-based epidemiological model in a SISO setting,¹⁹ and (ii) MIMO control of the temperature and concentration of a CSTR model. This second case study, in particular, demonstrates the ability of the MoD-based estimator and 3DoF-KF MPC-based control to handle nonlinearities in a demanding, highly interacting multivariable setting.

This paper is organized as follows: Section 2 provides a brief overview of the MoD estimator framework and some general guidelines for generating predictive models using MoD. Section 3 presents the controller problem formulation in terms of a process model description, filter design, and the structure of 3DoF-KF MPC; it also defines the optimization problem for generating the candidate control sequence for MoD MPC. Sections 4.1 and 4.2 present the case studies involving the SISO SIR model and the MIMO CSTR models, respectively, and elaborate on the overall process involving input signal design, model estimation, and validation, culminating in the 3DoF-KF MPC-based control. Finally, Section 5 provides a synopsis of the presented work and some statements regarding future work.

2. NONLINEAR SYSTEM IDENTIFICATION USING A MODEL-ON-DEMAND APPROACH

The Model-on-Demand estimation algorithm systematically blends global (e.g., prediction error modeling) and local (e.g., nearest neighbors) modeling for estimating dynamic models of complex nonlinear systems; this is achieved through local polynomial estimation of data selected using an adaptive bandwidth parameter and weighted through a kernel function. These operations are performed online at each sampling instant, allowing MoD to retain the simplicity of local modeling techniques while matching the predictive ability of global regression models. The ensuing subsections provide a detailed description of the MoD formulation and its integration into the 3DoF-KF MPC framework.

2.1. Model-on-Demand Estimation. 2.1.1. *The Concept.* Given a database of N observations $(X_k, Y_k)_{k=1}^N$ and a current operating point c , the basic premise behind MoD estimation is to generate a locally linear or quadratic ARX model with a regressor

structure $[n_a \ n_b \ n_k]$ over a suitable neighborhood of data around c , as depicted in Figure 1.^{10,14} At each time instant, model

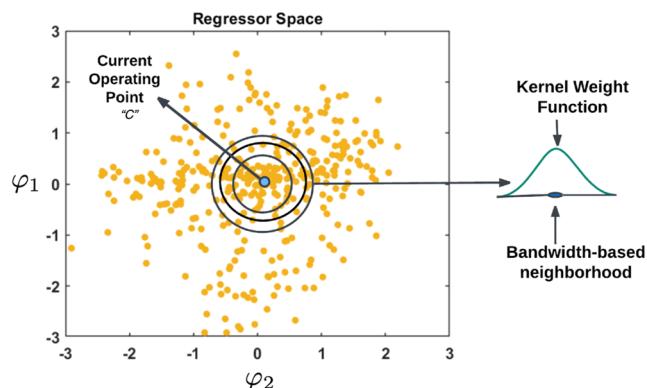


Figure 1. Schematic illustrating data selection for local polynomial estimation in MoD.

parameters are obtained from the least-squares solution to the optimization problem:

$$\hat{\theta} = \arg \min_{\theta} \sum_{k=1}^N l(y_k - \hat{m}(\tilde{\varphi}_k, \theta)) K_h \left(\frac{\|\varphi_k - \varphi_c\|_M}{h} \right) \quad (1)$$

where $\varphi_c = [y_{c-1} \cdots y_{c-n_a} \bar{u}_{c-n_b} \cdots \bar{u}_{c-n_b-n_k+1}]^T$ denotes the regressor vector at the operating condition c and φ_k ($k = 1, 2, \dots, N$) represents individual regressor vectors obtained from the estimation data set at each sampling instant k . $\tilde{\varphi} = \varphi_k - \varphi_c$ is a deviation term that defines the distance between the regressor at the current operating point and the k^{th} regressor from the estimation data set y_c and \bar{u}_c refer to values of the output and input variables, respectively, at the current operating point. $\hat{m}(\varphi_k, \theta)$ is a local polynomial function of the form:

$$\hat{m}(\tilde{\varphi}_k, \theta) = \theta_0 + \theta_1^T (\varphi_k - \varphi_c) + (\varphi_k - \varphi_c)^T \theta_2 (\varphi_k - \varphi_c) + \dots \quad (2)$$

where $\theta = [\theta_0 \ \theta_1 \ \theta_2 \ \dots]^T$ represents the vector of coefficients. Equation 2 can assume a simple (e.g., linear or quadratic) form depending on the user preference for the local series expansion. $l(\cdot)$ denotes a scalar-valued norm function, typically a 2-norm. $\|\cdot\|_M$ is a scaled distance function. h describes the bandwidth parameter dictating the local neighborhood size and is determined via the local Goodness-of-Fit (GOF) measure. $K_h(\cdot)$ is the kernel function for assigning weights and can be chosen to be tricube, Gaussian, uniform, or Epanechnikov, among others. The tricube kernel given by

$$K(u) = \begin{cases} (1 - |u|^3)^3, & \text{if } |u| < 1 \\ 0, & \text{if } |u| \geq 1 \end{cases}$$

is used for this particular study. The continuously differentiable nature of the tricube function, added with the cubic dependence on distance from the current operating point, makes it good for smooth curve fitting in the presence of outliers. $\hat{\theta} = [\hat{\theta}_0 \ \hat{\theta}_1]$ denotes the set of parameters of the extended design matrix $[1 \ \tilde{\varphi}^T \ \dots \ (\tilde{\varphi}^{p_{\text{reg}}})^T]$ arising from the least-squares solution of polynomial regression. p_{reg} refers to the polynomial order and $\hat{\theta}_1 \in \mathbb{R}^{p_{\text{reg}}}$ is the vector of the polynomial coefficients. p_{reg} is determined by the length of the regressor vector (d_{reg}). For the linear polynomial $p_{\text{reg}} = d_{\text{reg}}$ and for the quadratic polynomial $p_{\text{reg}} = d_{\text{reg}} + d_{\text{reg}}(d_{\text{reg}} + 1)/2$.

2.1.2. Bias-Variance Trade-off in MoD. Deciding on MoD parameters includes choosing reasonable regressor structures and local polynomial orders (e.g., 0, 1, or 2) to avoid under- or overparameterization. Furthermore, the decisions to obtain a suitable bandwidth involve the choice of a modified goodness-of-fit (GoF) criterion (e.g., AIC, FPE, GCV) to balance model fit and the number of parameters involved. The modification is obtained by varying the size of the estimation data set instead of the number of model parameters while evaluating the GoF criteria, as done in traditional estimation. This can be illustrated more clearly with an example: consider Akaike's Information Criterion (AIC):

$$\min_{d,\theta} \left(1 + \frac{2d}{N} \right) \sum_{i=1}^N e^2(t_i, \theta) \quad (3)$$

where N denotes the length of estimation data, θ denotes the vector of regressors, d denotes the dimension of θ and $e^2(t, \theta)$ symbolizes the one-step-ahead prediction error for a given θ . While the typical implementation of such an information criterion involves the variation of the number of parameters d for a fixed size of the data set N , the bias-variance trade-off for MoD is obtained through the adjustment of the data length, or in other words, the neighborhood size. This is specified through a user-defined search range for the number of data points in the neighborhood k_{\min} and k_{\max} .

Increasing neighborhood size decreases variance but simultaneously increases bias arising from fitting a restricted complexity (e.g., simple linear or quadratic) model to a large neighborhood of data. The optimal bandwidth is, therefore, obtained when the corresponding neighborhood size achieves a reasonable bias-variance trade-off. This demands prudent choices of $[k_{\min}, k_{\max}]$ on the user's end. k_{\min} should be large enough for the estimation problem to be well-conditioned, the upper limit of k_{\max} can be the size of the data set, in which case the entire data set is weighted and used for estimation.

Choosing a sensible regressor structure is crucial to avoiding under- or overparameterization. Local polynomial orders (e.g., 0, 1, or 2) and goodness-of-fit criteria (e.g., AIC, FPE, GCV) can be chosen as per the requirements of the problem.^{10,14} These choices allow the formulation of an adaptive bandwidth parameter as follows:

- At each sampling instant, the variable bandwidth parameter h is initialized with a small value h_0 corresponding to a well-posed problem.
- The number of iterations N_{iter} are then obtained from $[k_{\min}, k_{\max}]$ over which h exponentially increases as $h_c = \left(1 + \frac{0.3}{d_{\text{reg}}} \right) h_{c-1}$ and calculates the value of the GOF measure.
- An optimum pair $(k_{\text{opt}}, h_{\text{opt}})$ of neighborhood size and bandwidth, respectively, are finally chosen to correspond to the lowest GOF cost for the estimation.

Typically, the neighborhood size ($k_{\text{opt}}/h_{\text{opt}}$) at each operating point will depend on the extent of noise in the estimation data set around the current operating point. A higher value of $(k_{\text{opt}}, h_{\text{opt}})$ is suitable for neighborhoods with greater noise to allow a reduction of the error due to variance, and a lower value for the pair is suitable for reducing the error due to bias by fitting over a smaller neighborhood in the presence of less noise. In practice, the choice of these parameter values is simple

and intuitive, which facilitates the implementation of this algorithm in real-world settings.

2.2. Model on Demand Estimation Based Predictive Modeling. The generalized MIMO MoD identification is a manifold process that involves the estimation of individual outputs through Multi-Input-Single-Output (MISO) models that consider system inputs and other outputs, followed by the concatenation of these models to form a MIMO structure that takes into account interaction terms among various outputs.²⁰ Assuming $\hat{m}_i(\varphi_k, \hat{\theta}) = \hat{\theta}_{i,0} + \hat{\theta}_{i,1}^T (\varphi_k - \varphi_c)$ to be a locally linear model for the i^{th} output y_i and $\hat{\theta}_{i,0}$ and $\hat{\theta}_{i,1}$ to be the optimal regressor coefficients, a time-series-based one-step ahead prediction of y_i can be evaluated according to

$$\hat{y}_{i,k} = \alpha_i + \hat{\theta}_{i,1}^T \varphi_k \quad (4)$$

where $\alpha_i = \hat{\theta}_{i,0} - \hat{\theta}_{i,1}^T \varphi_c$. These equations can be organized into a MIMO ARX-like model with a structure of the form $F(q^{-1})y_k = G(q^{-1})u_k + \alpha$ where the polynomial matrices $F(q^{-1})$ and $G(q^{-1})$ are defined as

$$F(q^{-1}) = I + F_1q^{-1} + F_2q^{-2} + \dots + F_{n_a}q^{-n_a} \quad (5)$$

$$G(q^{-1}) = G_1q^{-n_k} + G_2q^{-n_k-2} + \dots + G_{n_b}q^{-n_k-n_b+1} \quad (6)$$

where y_k is the output vector at time instant k (dimension $n_y \times 1$), u_k is the input vector at time instant k (dimension $n_u \times 1$), q^{-1} is the backward shift operator. I is the $n_y \times n_y$ identity matrix, F_i are $n_y \times n_y$ coefficient matrices, G_j are $n_y \times n_u$ coefficient matrices, n_a is the order of the autoregressive part, n_b is the order of the exogenous input part, n_k denotes the lag orders, and $\alpha_k = [\alpha_i]_{i=1}^{n_y}$ denotes the bias term. This process of generating an ARX-based time-series model from the MoD-based estimation can be used to generate state-space matrices needed for defining a MIMO piecewise affine predictive structure of the form:

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k \bar{u}_k + \alpha_k \quad (7)$$

$$\hat{y}_k = C_k \hat{x}_k + \nu_k \quad (8)$$

where $\bar{u}_k \in \mathbb{R}^{\bar{n}_u}$ with $\bar{n}_u = n_u + n_{\text{dist}}$. Here n_u and n_{dist} correspond to the number of manipulated and disturbance inputs, respectively. $\hat{x}_k \in \mathbb{R}^{n_x}$, and $\hat{y}_k \in \mathbb{R}^{n_y}$ are the states and outputs of the system, respectively. Furthermore, the identification procedure aims to generate predictive models for control purposes. The control input, \bar{u}_k can be split into n_u number of manipulated variables ($u_k \in \mathbb{R}^{n_u}$) and n_{dist} measured disturbance variables ($d_k \in \mathbb{R}^{n_{\text{dist}}}$). At each sampling instant k , the state-space matrices $\{A_k, B_k, C_k\}$ are evaluated from θ_1 . Additionally, the matrix B_k can be split into two components, $B_{u,k}$ and $B_{d,k}$ conforming to u_k and d_k respectively. For notational convenience, the subscript k will be dropped from $\{A_k, B_k, C_k\}$ and α_k in the derivations following this. $\nu_k \in \mathbb{R}^{n_y}$ is a stochastic signal that represents the lumped effects of noise and the unmeasured disturbances present in the outputs. Assuming that the system under consideration is open-loop stable, ν_k can be expressed in a state-space form as follows:

$$\zeta_{k+1} = A_w \zeta_k + B_w w_k \quad (9)$$

$$\nu_k = C_w \zeta_k \quad (10)$$

where w_k represent an integrating white noise. It is further assumed that ν_k has uncorrelated components, i.e., $B_w = C_w = I$ and $A_w = \text{diag}\{\Lambda_1, \dots, \Lambda_{n_y}\}$, where I stands for

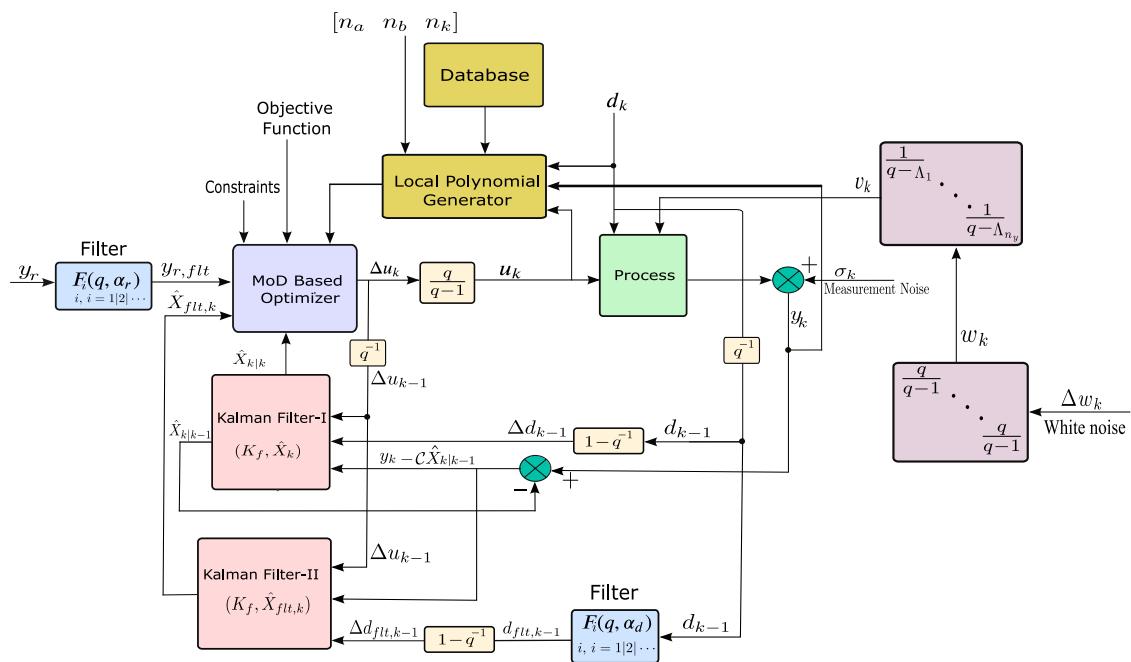


Figure 2. MoD-based MPC for set point tracking of a process subject to process constraints, measured and unmeasured disturbances. The controller receives filtered signals of measured disturbances and reference changes as inputs and is configured for 3DoF. Kalman Filter-I refers to (21)–(22), while Kalman Filter-II refers to (23)–(24).

identity matrix of appropriate dimension. The diagonal values of matrix A_w are 0 for single-integrating and 1 for double-integrating disturbances. Such formulation creates an extended state-space model that enables zero offset to integrating noise and unmeasured disturbances. It is given by

$$\tilde{X}_{k+1} = \mathcal{A}\tilde{X}_k + \mathcal{B}_u\Delta u_k + \mathcal{B}_d\Delta d_k + \mathcal{B}_w\Delta w_k \quad (11)$$

$$\hat{y}_k = C\tilde{X}_k \quad (12)$$

where

$$\tilde{X}_k = \begin{bmatrix} \Delta x_k \\ \Delta \zeta_k \\ \hat{y}_k \end{bmatrix}, \quad \mathcal{A} = \begin{bmatrix} A & 0 & 0 \\ 0 & A_w & 0 \\ CA & C_w A_w & I \end{bmatrix}, \quad \mathcal{B}_i = \begin{bmatrix} B_i \\ 0 \\ CB_i \end{bmatrix}, \quad i = u, d$$

$$\mathcal{B}_w = \begin{bmatrix} 0 \\ I \\ I \end{bmatrix}, \quad C = [0 \quad 0 \quad I]$$

The model per eqs 11–12 is utilized in the MPC-based controller formulation presented in Section 3.

3. MODEL PREDICTIVE CONTROL FORMULATION

The key components of the 3DoF-KF MoD MPC architecture include a database that stores the open-loop estimation data set, a local polynomial generator, and an MoD-based optimizer, as illustrated in Figure 2. As operating conditions change over time, the MoD estimator locates a suitable neighborhood of data points in the regressor space of the estimation data set that best represents a local neighborhood of the current operating point. This is followed by the estimation of local polynomial-based predictive models, which are subsequently utilized by the optimizer along with other system constraints to generate physically realizable control actions.

3.1. Optimization Problem. The 3DoF-KF MPC formulated in this paper relies on minimizing the following quadratic cost function to generate optimal control sequences:

$$\min_{u_{k+1}^{i_0,1,\dots,m-1}} J_k \triangleq \sum_{j=1}^p \left\| (\hat{y}_{k+j} - y_{r,k+j}) \right\|_{W_y}^2 + \sum_{i=0}^{m-1} \left\| (u_{k+i} - u_{r,k+i}) \right\|_{W_u}^2 + \sum_{i=0}^{m-1} \left\| (\Delta u_{k+i}) \right\|_{W_{\Delta u}}^2 \quad (13)$$

subject to equality constraints represented by eqs 11–12 and the following inequality constraints:

$$y_{\min} \leq \hat{y}_{k+j} \leq y_{\max}, \quad \forall j = 1, 2, \dots, p \quad (14)$$

$$u_{\min} \leq u_{k+i} \leq u_{\max}, \quad \forall i = 0, 1, \dots, m-1 \quad (15)$$

$$\Delta u_{\min} \leq \Delta u_{k+i} \leq \Delta u_{\max}, \quad \forall i = 0, 1, \dots, m-1 \quad (16)$$

where p is the prediction horizon, m is the control horizon, $\|\cdot\|_{W_*} \triangleq \sqrt{(\cdot)^T W_*(\cdot)}$ is the weighted 2-norm with the corresponding weight matrix given by $W_* > 0$. $(\cdot)_r$ denotes the vector of reference trajectories for the system outputs. An extension of this formulation that includes discrete-valued inputs leads to developing a hybrid MPC framework.¹⁶ However, this feature is not present in the case studies of Section 4.

The optimization problem stated in eqs 13–16, while it summarizes the MPC problem, involves additional considerations that rely on the 3DoF-KF framework to independently tune responses to different control modes, as discussed in the ensuing section.

3.2. The Three-Degree-of-Freedom Kalman Filter (3DoF-KF) Formulation. The 3DoF-KF MPC relies on an intuitive time-constant-based tuning to independently adjust the speed of response corresponding to set point tracking and

measured and unmeasured disturbance rejection. In accomplishing multiple degree-of-freedom tuning, the 3DoF-KF MPC control system mimics Internal Model Control, as described in Morari and Zafriou.²¹ This is described in the ensuing subsections.

3.2.1. Reference Trajectory Tracking and Measured Disturbance Rejection. The output reference trajectory $y_{r,k+i}$ in eq 13 is filtered before passing it to the controller to influence the speed of response to set point changes as demonstrated in Figure 2. The filtered trajectory $y_{r,\text{flt}}$ is given by

$$\frac{y_{r,\text{flt},k+i}}{y_{r,k+i}} = F(q, \alpha_r^j), j = 1, \dots, n_y, i = 1, \dots, p \quad (17)$$

$F(q, \alpha_r^j)$ is chosen to be a discrete-time Type-I filter corresponding to the j -th output reference as given in Morari et al.²¹

$$F(q, \alpha_r^j) = \frac{(1 - \alpha_r^j)q}{q - \alpha_r^j}, 0 \leq \alpha_r^j < 1, j = 1, \dots, n_y \quad (18)$$

α_r^j is the tuning parameter to adjust the speed of response of each output j to set point change, where a low value for α_r^j yields a fast output set point tracking response. α_r^j can further be specified in terms of the desired closed-loop time constant τ_r^j as follows $\alpha_r^j = e^{-T_s/\tau_r^j}$. This adjustment is more intuitive than adjusting move suppression weights ($W_{\Delta u}$), which directly affect the manipulated variables, making it more difficult to predict the effect on a specific controlled variable response. When a forecast of the measured disturbance is available over the prediction horizon p , a separate filter $F(q, \alpha_d^j)$, as illustrated in Figure 2, is implemented to independently adjust the speed of response to changes in the measured disturbances. It is given by

$$\frac{d_{\text{flt},k+i}}{d_{k+i}} = F(q, \alpha_d^j), j = 1, \dots, n_{\text{dist}}, i = 0, 1, \dots, p - 1 \quad (19)$$

$$F(q, \alpha_d^j) = \frac{(1 - \alpha_d^j)q}{q - \alpha_d^j}, 0 \leq \alpha_d^j < 1 \quad (20)$$

The speed of response in rejecting each measured disturbance signal j can be tuned by adjusting $\alpha_d^j \in [0, 1]$ in $F(q, \alpha_d^j)$. Faster responses are obtained when the value of α_d^j approaches zero, and the speed of response goes down as α_d^j approaches 1. The time constant representation of α_d^j is given by $\alpha_d^j = e^{-T_s/\tau_d^j}$, where τ_d^j denotes the desired closed-loop time constant for measured disturbance rejection. The filter transfer function $F(q, \alpha_d^j)$ depends on whether the disturbance is an asymptotic step or ramp (i.e., Type-I or Type-II signal, as defined in Morari and Zafriou²¹). In this paper, a Type-I filter is used to filter step disturbances.

3.2.2. Kalman Filter: Unmeasured Disturbance Rejection. To predict future outputs \hat{y}_{k+i} ($i = 1, 2, \dots, p$) of the model in eqs 11–12 over the predictions horizon p , for the eq 13, an initial value of \tilde{X}_k is required at each sampling instant. As not all states in \tilde{X}_k are measurable, a state observer needs to be designed for the state estimation of the unmeasurable states. This entails the use of multiple Kalman filters involving a two-step process for separately capturing the effects of the measured and the unmeasured disturbances on the output; this is illustrated by Kalman Filter-I in Figure 2:

$$\hat{X}_{k|k-1} = \mathcal{A}\hat{X}_{k-1|k-1} + \mathcal{B}_u \Delta u_{k-1} + \mathcal{B}_d \Delta d_{k-1} \quad (21)$$

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + \mathcal{K}_f(y_k - C\hat{X}_{k|k-1}) \quad (22)$$

where \mathcal{K}_f is the Kalman gain matrix, and \hat{X}_k is the estimated value of \tilde{X}_k arising from the state prediction equation of the Kalman filter (eq 21). The first step of state estimation, as outlined in (21) and (22), utilizes the unfiltered measured disturbance d_k . The correction term in (22) specifically accounts for the effects of unmeasured disturbances only.

The second step utilizes the filtered measured disturbance signal d_{flt} and the error correction term from Kalman Filter-I, as illustrated by Kalman Filter-II in Figure 2.

$$\hat{X}_{\text{flt},k|k-1} = \mathcal{A}\hat{X}_{\text{flt},k-1|k-1} + \mathcal{B}_u \Delta u_{k-1} + \mathcal{B}_d \Delta d_{\text{flt},k-1} \quad (23)$$

$$\hat{X}_{\text{flt},k|k} = \hat{X}_{\text{flt},k|k-1} + \mathcal{K}_f(y_k - C\hat{X}_{\text{flt},k|k-1}) \quad (24)$$

Equations 23–24 allow the estimate $\hat{X}_{\text{flt},k|k}$ to account for both measured (filtered) and unmeasured disturbances (through the first and the second term of eq 24 respectively) in a manner that measured disturbance rejection tuning parameter α_d^j does not affect the unmeasured disturbance rejection, while the unmeasured disturbance filter matrix \mathcal{K}_f does not affect measured disturbance rejection. The gain matrix \mathcal{K}_f can be tuned to regulate the speed of unmeasured disturbance rejection. Optimal values of \mathcal{K}_f can be obtained through the solution of an algebraic Riccati equation, which, however, requires covariance matrices for unmeasured disturbances and measurement noise that may not be precisely known. Instead, the parametrization described in Lee et al.²² is utilized, allowing independent tuning of the speed of unmeasured disturbance rejection for each output channel in the presence of set point tracking and measured disturbance rejection. This Kalman filter gain is given by

$$\mathcal{K}_f = [0 \quad F_w^T \quad F_y^T]^T \quad (25)$$

where

$$F_w = \text{diag}\{f_w^1, \dots, f_w^{n_y}\}, \quad F_y = \text{diag}\{f_y^1, \dots, f_y^{n_y}\}$$

$$f_w^j = \frac{(f_y^j)^2}{1 + \Lambda_j - \Lambda_j f_y^j}, \quad j = 1, \dots, n_y$$

$f_y^j \in [0, 1]$ is a tuning parameter directly proportional to the degree of unmeasured disturbance rejection for each output j . If f_y^j is set to a small value close to zero, the resultant control action assigns a very small weight to the current measurement and primarily relies on the deterministic model expressed in (21) and the feedforward anticipation. Conversely, as f_y^j approaches 1, the state estimator attempts to compensate for prediction error arising from the unmeasured disturbance, resulting in aggressive control action. This enables the users to directly dictate the speed of unmeasured disturbance rejection for each output response on its own, ensuring control system robustness in a way that is more intuitive than the typical move-suppression-based MPC tuning.¹⁶ A time constant equivalence (τ_u) can be defined similarly to τ_r^j and τ_d^j using the relationship $f_y^j = 1 - e^{-T_s/\tau_u^j}$ for simpler tuning. Increasing (τ_u) decreases the controller's sensitivity to output changes due to unmeasured disturbances.

3.3. Formulation of the Overall Optimization Problem.

Minimizing the constrained MPC objective function mentioned in (13) over the prediction horizon p requires propagating the predictions described in (11)–(12) p steps into the future. This entails two sets of equations arising from filtered (for objective function) as well as unfiltered measured disturbances (for constraints). These are given as

$$\hat{y}_{k+1} = \Phi \hat{X}_k + \mathcal{H}_u \mathcal{U}_k + \mathcal{H}_d \mathcal{D}_k - \mathcal{H}_{u1} u_{k-1} - \mathcal{H}_{d1} d_{k-1} \quad (26)$$

$$\hat{y}_{\text{fl},k+1} = \Phi \hat{X}_{\text{fl},k} + \mathcal{H}_u \mathcal{U}_k + \mathcal{H}_d \mathcal{D}_{\text{fl},k} - \mathcal{H}_{u1} u_{k-1} - \mathcal{H}_{d1} d_{\text{fl},k-1} \quad (27)$$

where

$$\hat{\mathcal{Y}}_{k+1} = [\hat{y}_{k+1}^T \quad \hat{y}_{k+2}^T \quad \cdots \quad \hat{y}_{k+p}^T]^T \quad (28)$$

$$\mathcal{Y}_{\text{fl},k+1} = [y_{\text{fl},k+1}^T \quad y_{\text{fl},k+2}^T \quad \cdots \quad y_{\text{fl},k+p}^T]^T \quad (29)$$

$$\mathcal{U}_k = [u_k^T \quad u_{k+1}^T \quad \cdots \quad u_{k+m-1}^T]^T \quad (30)$$

$$\mathcal{D}_k = [d_k^T \quad d_{k+1}^T \quad \cdots \quad d_{k+p-1}^T]^T \quad (31)$$

$$\mathcal{D}_{\text{fl},k} = [d_{\text{fl},k}^T \quad d_{\text{fl},k+1}^T \quad \cdots \quad d_{\text{fl},k+p-1}^T]^T \quad (32)$$

and \hat{X}_k and $\hat{X}_{\text{fl},k}$ denote $\hat{X}_{k|k}$ and $\hat{X}_{\text{fl},k|k}$ for notational convenience. The motivation behind using two prediction equations can be understood as follows: the MPC objective function relies on the filtered output prediction given by (27), where the filtering allows the user to dictate how aggressively the controller responds to changes in the set point and measured disturbances. Meanwhile, the unfiltered true prediction, as given in (26), is used to formulate constraints, ensuring that the estimated value of the true output remains within the system's physical limitations. The matrices $\Phi, \mathcal{H}_u, \mathcal{H}_d, H_{ii} \forall i = u, d$ in eqs 26 and 27 are given in Section 1 of the Supporting Information. The optimization problem can be reformulated in vector form as

$$\begin{aligned} \min_{\mathcal{U}_k} J_k \triangleq & \|(\hat{\mathcal{Y}}_{\text{fl},k+1} - \mathcal{Y}_{r,\text{fl},k+1})\|_{\hat{W}_y}^2 + \|(\mathcal{U}_k - \mathcal{U}_{r,k})\|_{\hat{W}_u}^2 \\ & + \|(R_u \mathcal{U}_k - R_{u0} u_{k-1})\|_{\hat{W}_{\Delta u}}^2 \end{aligned} \quad (33)$$

subject to eqs 26–27 and the following inequality constraints:

$$\mathcal{Y}_{\min} \leq \hat{y}_{k+1} \leq \mathcal{Y}_{\max} \quad (34)$$

$$\mathcal{U}_{\min} \leq \mathcal{U}_k \leq \mathcal{U}_{\max} \quad (35)$$

$$\Delta \mathcal{U}_{\min} \leq \Delta \mathcal{U}_k \leq \Delta \mathcal{U}_{\max} \quad (36)$$

where $\hat{W}_* = \text{diag}(W_*)$, and

$$\mathcal{Y}_{r,\text{fl},k+1} = \begin{bmatrix} y_{r,\text{fl},k+1} \\ y_{r,\text{fl},k+2} \\ \vdots \\ y_{r,\text{fl},k+p} \end{bmatrix}, R_u = \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ -I & I & \cdots & 0 & 0 \\ 0 & -I & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -I & I \end{bmatrix}, R_{u0} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\mathcal{U}_{r,k} = [u_{r,k}^T \quad u_{r,k+1}^T \quad \cdots \quad u_{r,k+m-1}^T]^T$$

Substituting \hat{y}_{k+1} and $\hat{y}_{\text{fl},k+1}$ from (26) and (27) respectively into (33) and (34), eliminating the constant terms and regrouping the results in the modified cost function gives

$$\min_{\mathcal{U}_k} J_k \triangleq \frac{1}{2} \mathcal{U}_k^T \mathcal{H} \mathcal{U}_k + \mathcal{G}^T \mathcal{U}_k \quad (37)$$

$$\Gamma \mathcal{U}_k \leq \mathcal{M} \quad (38)$$

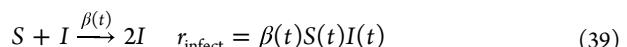
where \mathcal{U}_k is the decision variable of the optimization problem. \mathcal{H} and \mathcal{G} correspond to the coefficient matrix corresponding to the quadratic and the linear terms, respectively. The coefficient matrices Γ and \mathcal{M} are used to express constraints. Equations 37–38 constitute a linearly constrained quadratic programming problem that can be solved using standard solvers such as quadprog from MATLAB. See the Supporting Information document for the matrix expressions for $\mathcal{H}, \mathcal{G}, \Gamma$, and \mathcal{M} . A complete controller description is now available, allowing us to examine some case studies in Section 4.

4. CASE STUDIES

This section is devoted to describing the implementation of the 3DoF-KF MoD MPC algorithm to two nonlinear process systems involving (a) a SISO SIR model for infection transmission and (b) a MIMO CSTR model involving highly interacting dynamics. Both case studies involve systematic treatment of the process systems through judicious input signal design, MoD-based data-centric modeling, and finally, closed-loop control. Both problems make use of the locally linear MoD models to obtain predictions for the 3DoF-KF MPC framework. While the MoD estimation algorithm allows estimating locally quadratic models with higher estimation accuracies, the linear form is of greater appeal in these case studies owing to its ability to be integrated into a linear MPC framework without incurring additional structural complexity and computational cost. Furthermore, the 3DoF-KF control framework provides a robust and efficient solution that effectively mitigates plant-model mismatch due to system nonlinearities, even with locally linear models. This ensures a balance between performance and computational efficiency in these case studies.

While the first problem addresses the bulk of the issues associated with the 3DoF-KF MoD MPC, the second problem illustrates the ability of the proposed identification and control framework in a highly demanding situation involving multi-variable control with highly interactive dynamics. These two problems illustrate how MoD identification and control can manage to estimate complex system dynamics through judicious use of data and generate the desired closed-loop performance with modest engineering effort.

4.1. The SIR Model: A SISO Case Study. The classic Susceptible-Infected-Removed (SIR) epidemic model, analogous to a process-systems-based Continuous Stirred Tank Reactor, provides a basis for understanding the propagation of infectious diseases through a set of differential equations. The nonlinear nature of the problem makes it a suitable starting point for validating the 3DoF MoD MPC approach. The standard SIR model compartmentalizes the overall population into three classes: the *Susceptible*, the *Infected*, and the *Recovered*. A susceptible person, when exposed to an infected person, becomes infected themselves. However, no transfer takes place in the reverse direction. This process draws an analogy from an irreversible autocatalytic reaction, as shown below:

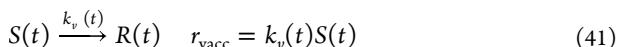


The removal from the infected population can happen through either recovery or death. A comparison of the recovery process

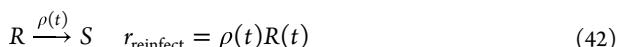
can be made with the irreversible degradation of the catalyst in a reaction with a rate constant $\gamma(t)$ as follows:



The parameter $\gamma(t)$ considers the fraction of the infected population recovered per day through treatment, while μ represents the fraction of the overall population removed from the system through death. Death, characterized by μ , can occur in any of the three classes of the population. Intuitively, an increase in $\gamma(t)$ and μ causes $I(t)$ to decrease. The value of μ is held constant for the study. The susceptible population $S(t)$ can be vaccinated at a daily rate of $k_v(t)$, preventing a person from being infected and transported to the recovered compartment as per:



where $k_v(t)$ denotes the daily vaccination rate. A variation of this problem is considered in this study, where the recovered population may suffer from loss of immunity, subsequently returning to the pool of susceptible individuals. The reinfection or loss of immunity of the recovered population possesses similar dynamics as the vaccination, with a rate constant $\rho(t)$, and can be described as follows:



The rates of reinfection $\rho(t)$ and of vaccination $k_v(t)$ oppose each other, and they can be considered analogous to the rate constants for the forward reaction from *Susceptible* to *Recovered* and the reverse reaction from *Recovered* to *Susceptible* respectively. The overall process schematic is depicted in Figure 3. These conditions give rise to the following set of nonlinear

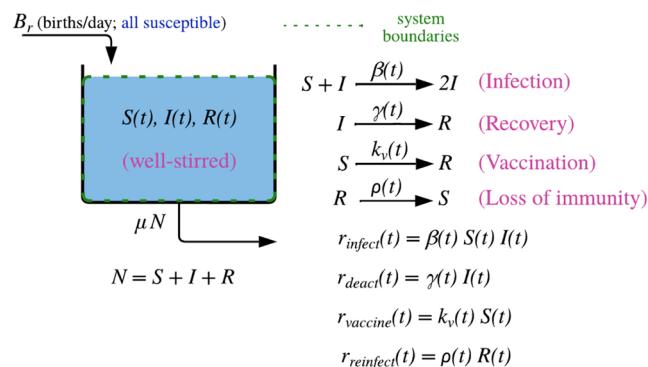


Figure 3. SIR model with births, deaths, vaccination, and loss of immunity, represented as a CSTR. Infection is conceptualized as an autocatalytic reaction, while recovery reflects a deactivation reaction. Vaccination and Reinflection are characterized by first-order kinetics. Endemic steady-state conditions (where $I \neq 0$ at $t = 0$) are considered.

lumped parameter models through species accounting that describe the infection dynamics in a parsimonious manner:

$$\begin{aligned} \frac{dS(t)}{dt} = & \frac{B_r}{\text{Birth rate}} - \frac{\beta(t)S(t)I(t)}{\text{Infection rate}} - \frac{\mu S(t)}{\text{Mortality rate}} \\ & - \frac{k_v(t)S(t)}{\text{Vaccination}} + \frac{\rho(t)R(t)}{\text{Loss of Immunity/Reinfection}} \end{aligned} \quad (43)$$

$$\frac{dI(t)}{dt} = \frac{\beta(t)S(t)I(t)}{\text{Infection rate}} - \frac{\gamma(t)I(t)}{\text{Deactivation}} - \frac{\mu I(t)}{\text{Mortality rate}} \quad (44)$$

$$R(t) = N - S(t) - I(t), \quad N = \frac{B_r}{\mu} \quad (45)$$

B_r denotes the birth rate (assumed constant), $\beta(t)$ is the time-varying transmission rate of the infection, $\gamma(t)$ refers to the recovery rate, μ is the mortality rate, and $k_v(t)$ denotes the vaccination rate. As B_r and μ are held constant throughout the study, the total population $N = B_r/\mu = S(t) + I(t) + R(t)$ will remain constant as well.

The SIR model is further characterized by an endemic steady-state condition with a nonzero infected population at the start of the study. The nontrivial endemic steady-states \bar{S}, \bar{I} are given by

$$\bar{S} = \frac{(\mu + \bar{\gamma})}{\bar{\beta}} \quad (46)$$

$$\bar{I} = -\frac{(\bar{\gamma}\mu^2 + \bar{k}_v\mu^2 + \bar{\rho}\mu^2 + \mu^3 - B_r\bar{\beta}\mu - B_r\bar{\beta}\bar{\rho} + \bar{\gamma}\bar{k}_v\mu + \bar{\gamma}\mu\bar{\rho})}{\bar{\beta}(\bar{\gamma}\mu + \bar{\rho}\mu + \mu^2)} \quad (47)$$

where $\bar{*}$ denote the initial values. The complete set of plant parameters for the SIR model can be obtained from Table 1. The

Table 1. Base Parameters for Plant Dynamics of the SIR Model

parameters	units	values	parameters	units	values
B_r	Persons/Day	500	$\bar{\beta}$	Person ⁻¹ Day ⁻¹	0.001
μ	Day ⁻¹	0.08	$\bar{\gamma}$	Day ⁻¹	0.35
k_v	Day ⁻¹	0	$\bar{\rho}$	Day ⁻¹	0.05

control problem is as follows: the infected population $I(t)$ is the controlled variable, $\beta(t)$ acts as the manipulated variable, $k_v(t)$ is a measured disturbance, while $\gamma(t)$ and $\rho(t)$ are regarded as unmeasured disturbances to the system.

4.1.1. Input Signal Design. The dynamics of the nonlinear SIR model vary greatly with operating conditions, making it necessary to generate a database that spans all regions of the infected population. This is achieved by generating input signals that excite the plant sufficiently to capture these dynamics over the required operation ranges; data generation is followed by parameter estimation and model validation. The input signals must be “plant-friendly”;²³ for this purpose, multisine signals generated with minimum crest-factor approach developed in Guillaume et al.²⁴ are utilized, which are deterministic and are simple to implement. The multisine signals are defined as follows:

$$u_k = \lambda \sum_{i=1}^{n_s} \sqrt{2\alpha_i} \cos(\omega_i k T_s + \phi_i) \quad (48)$$

where N_s is the signal length, T_s denotes the sampling time, $\omega_i = 2\pi i/N_s T_s$, and $n_s \leq N_s/2$. The primary frequency band of interest spanned by these signals lies in the range $[\omega_*, \omega^*]$ given by

$$\omega_* = \frac{1}{\beta_s \tau_{\text{dom}}^H} \leq \omega \leq \frac{\alpha_s}{\tau_{\text{dom}}^L} = \omega^* \quad (49)$$

where $\tau_{\text{dom}}^H, \tau_{\text{dom}}^L$ denote the high and low estimates of the dominant time constants of the system, respectively, and α_s and β_s are parameters related to the high and low-frequency ranges of

interest.²⁵ The number of sinusoids (n_s) is chosen to be sufficiently high to achieve persistent excitation. The parameter values used in the input signal design are provided in Table 2,

Table 2. Signal Parameters for the Design of the Multisine Signal for the SIR Problem

parameter	value	parameter	value
α_s	2	β_s	3
τ_{dom}^L	5	τ_{dom}^L	10

with the resultant power spectrum of the signal illustrated in Figure 4. While the input corresponding to a manipulated

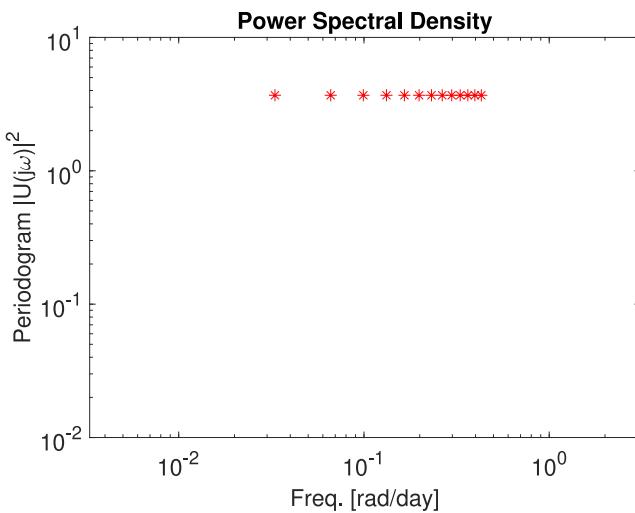


Figure 4. Power spectrum of the designed $\beta(t)$ input signal highlighting the frequency band of interest for the SIR model; this is obtained from the *a priori* plant information, per the specifications in Table 2.

variable $\beta(t)$ is designed with multisine signals, the measured disturbance data is generated using a filtered white noise signal. For this purpose, the signal k_v is designed as $k_{v,k} = 0.9k_{v,k-1} + a_{k-1}$, where a_k denotes a band-limited white noise signal with a noise power of $\sigma_a^2 = 0.1$. The estimation and validation data for

creating the database are generated from different realizations of these input signals having the same frequency content and spanning the same range of operations, as depicted in Figure 5. While the plot illustrates single cycles of length 152 days for the input signal for the purpose of clarity, three cycles of each signal (for a total of 456 days) were used for the data set generation.

4.1.2. Open-Loop Model-on-Demand Estimation. The input-output data is used to achieve MoD-based estimation that forms the basis for the predictive control framework. Performance is first tested in the open-loop setting by evaluating various validation criteria. The MoD estimates are contrasted with ARX-based global polynomial models on the same data set. The linear ARX-based global model can be integrated with the linear 3DoF-KF MPC framework in a manner similar to the MoD models, making it a suitable benchmark for comparison with the locally linear MoD models both in open-loop and closed-loop settings. The results presented in this paper use the localized Akaike Information Criterion (AIC) to evaluate the goodness-of-fit for MoD-based estimation. The range $[k_{\min}, k_{\max}]$ for neighborhood size is chosen to be $[30, 50]$ based on the length of the database, along with a variance penalty of 3 and a tricube kernel function is chosen to impose weights on these neighboring points. The normalized root-mean-square error (NRMSE) is calculated over the validation data set, which is given by

$$\text{NRSME fit (\%)} = 100 \times \left(1 - \frac{\|y_k - \hat{y}_k\|_2}{\|y_k - \bar{y}\|_2} \right) \quad (50)$$

where y_k denotes the measured output, \hat{y}_k refers to the estimated value of the output, \bar{y} is the mean value of the measured y_k signal, and $\|\cdot\|_2$ signifies an l_2 norm. Based on a regressor structure of $[n_a \ n_{b_1} \ n_{b_2} \ n_{k_1} \ n_{k_2}]$ of $[2 \ 2 \ 2 \ 1 \ 1]$, and a locally linear polynomial model of the form

$$\begin{aligned} I_k = & \theta_{0,k} + \theta_{1,k}I_{k-1} + \theta_{2,k}I_{k-2} + \theta_{3,k}\beta_{k-1} + \theta_{4,k}\beta_{k-2} \\ & + \theta_{5,k}k_{v,k-1} + \theta_{6,k}k_{v,k-2} \end{aligned} \quad (51)$$

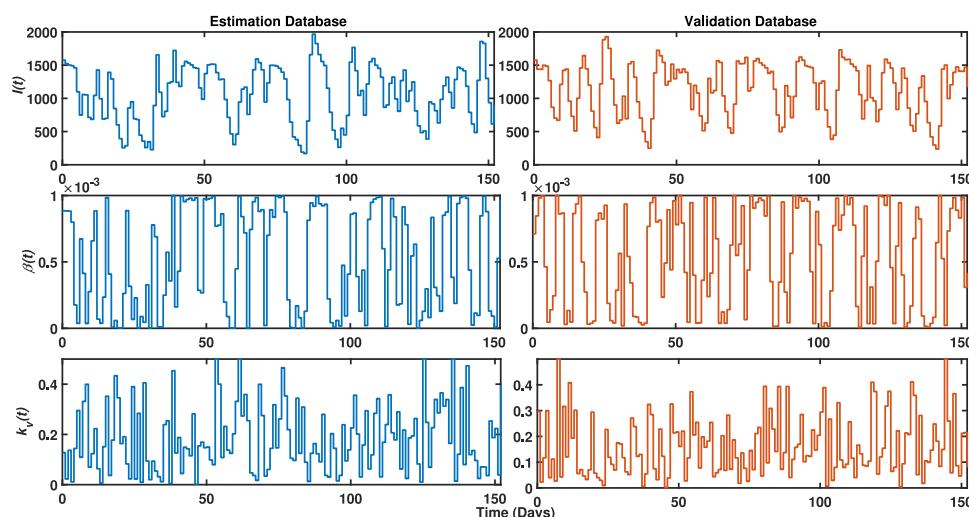


Figure 5. Plot illustrating a single cycle of each input signal for MoD database generation for the SIR case study. Different realizations of multisine signals with the same frequency content are used for the manipulated variable $\beta(t)$. The measured disturbance $k_v(t)$ is characterized by a band-limited white noise with autoregressive filtering.

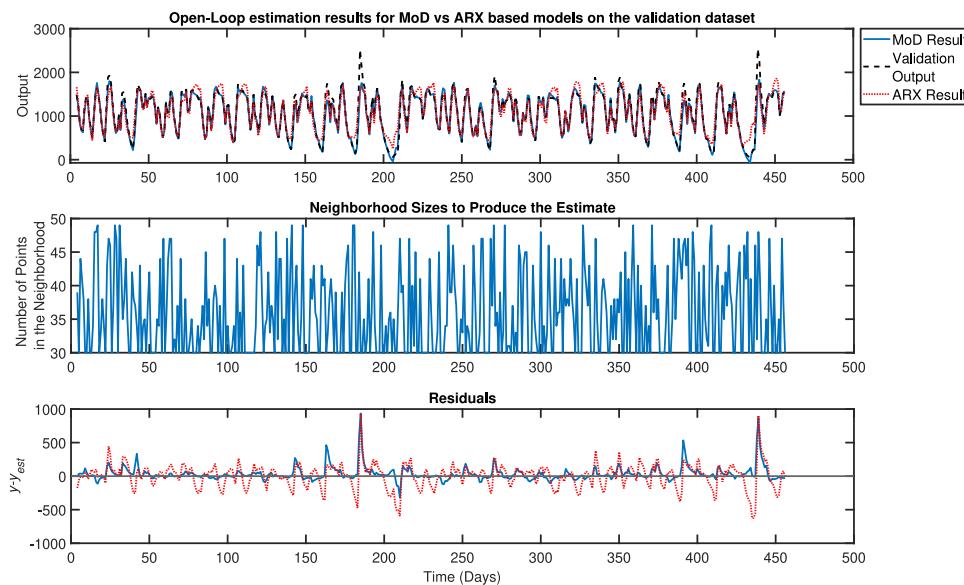


Figure 6. A comparison of the open-loop performance of the MoD-based estimator and the ARX-based estimator for the controlled variable ($I(t)$, top) and the residual (bottom) for the SIR case study, with structure [2 2 1 1]. The middle plot describes neighborhood sizes for the MoD estimate.

MoD displays an open-loop NRMSE fit percentage of 75.95% over the validation data, while the ARX model provides a fit percentage of 47.14%, as illustrated in Figure 6. Furthermore, the MoD-based estimation predicts the validation data with comparatively lower root-mean-square (RMS) and maximum errors, as summarized in Table 3. While nearly 24% of the

Table 3. Open-Loop Estimation Results for MoD vs ARX for the SIR Problem

method	NRMSE fit (%)	RMS error ($\sqrt{\sum e^2}$)	max error ($ e_{\max} $)
MoD	75.95	113	887
ARX	47.14	248	1314

variability in the data remains unmodelled in MoD, MoD estimation of this magnitude will be “good enough” (as seen from the results in the ensuing section).

4.1.3. Model-on-Demand with Three-Degree-of-Freedom MPC. The Model-on-Demand algorithm is now compared to ARX-based estimation in a closed-loop setting by implementing the 3DoF-KF MPC architecture described in Section 3. A set point change of 90% reduction in the infected population is given to both the closed-loop systems at $t = 5$ days, with $[k_{\min}, k_{\max}]$ for the MoD chosen to be [320, 450]. The measured disturbance k_v experiences a +0.3 step increase at $t = 30$ days, and the unmeasured disturbances $\gamma(t)$ and $\rho(t)$ are given step increases of +0.3 and +0.2 from their steady-states at $t = 120$ and 140 days, respectively. The controllers have lower bounds on $\beta(t)$ and $I(t)$ to be $u_{\min} = 0$ and $y_{\min} = 0$, respectively. All controller parameters are included in Table 4.

Figure 7 illustrates the advantage of 3DoF-KF MoD MPC-based closed-loop control of the nonlinear SIR plant over the 3DoF-KF ARX MPC. 3DoF-KF MoD MPC efficiently achieves set point tracking with a fast settling time of 15 days ($t = 20$ days), with zero offset and minimal undershoot. This demonstrates the controller’s ability to implement necessary policies precisely, avoiding unnecessary restrictions while effectively preventing unwanted infection spread. Such performance of the 3DoF-KF MoD MPC can be ascribed to the adaptive

Table 4. Control Design Parameters for the SIR Problem

parameter	value	parameter	value
p	50	y_{\min}	0
m	10	y_{\max}	∞
τ_r	4 Days	Δu_{\min}	$-0.15\bar{\beta}$
τ_d	1 Days	Δu_{\max}	$0.15\bar{\beta}$
τ_u	1 Day	W_y	1
u_{\min}	0 Day ⁻¹	$W_{\Delta u}$	0.01
u_{\max}	$[\infty \infty]$	T_{sampling}	1 Day

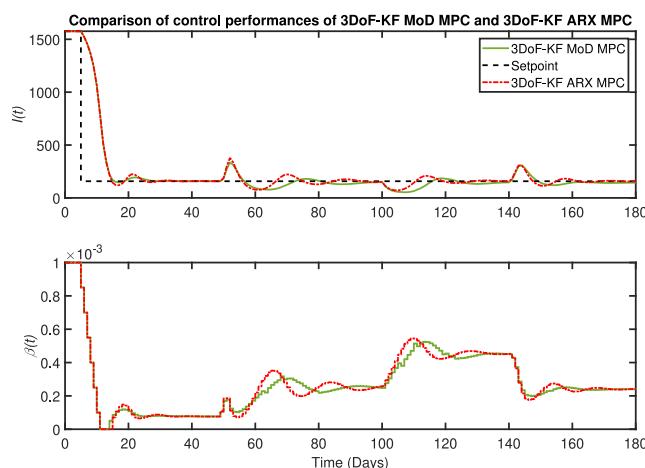


Figure 7. Comparison of the closed-loop performance between the 3DoF-KF MoD MPC vs the 3DoF ARX MPC for the SIR problem, featuring smoother set point tracking and effective disturbance rejection achieved by the 3DoF-KF MoD MPC through the manipulation of $\beta(t)$ with noticeably lesser variance. Measured disturbance $\Delta k_v(t) = +0.3$ is introduced at $t = 50$ days, and unmeasured disturbances $\Delta \gamma(t) = +0.3$ and $\Delta \rho(t) = +0.2$ are introduced at $t = 120$ days and $t = 140$ days, respectively.

nature of the coefficients of the MoD models in the face of changing closed-loop conditions, as shown in Figure 8. While there are varying degrees of fluctuations among the coefficients associated with various time-lagged terms, their combined effect

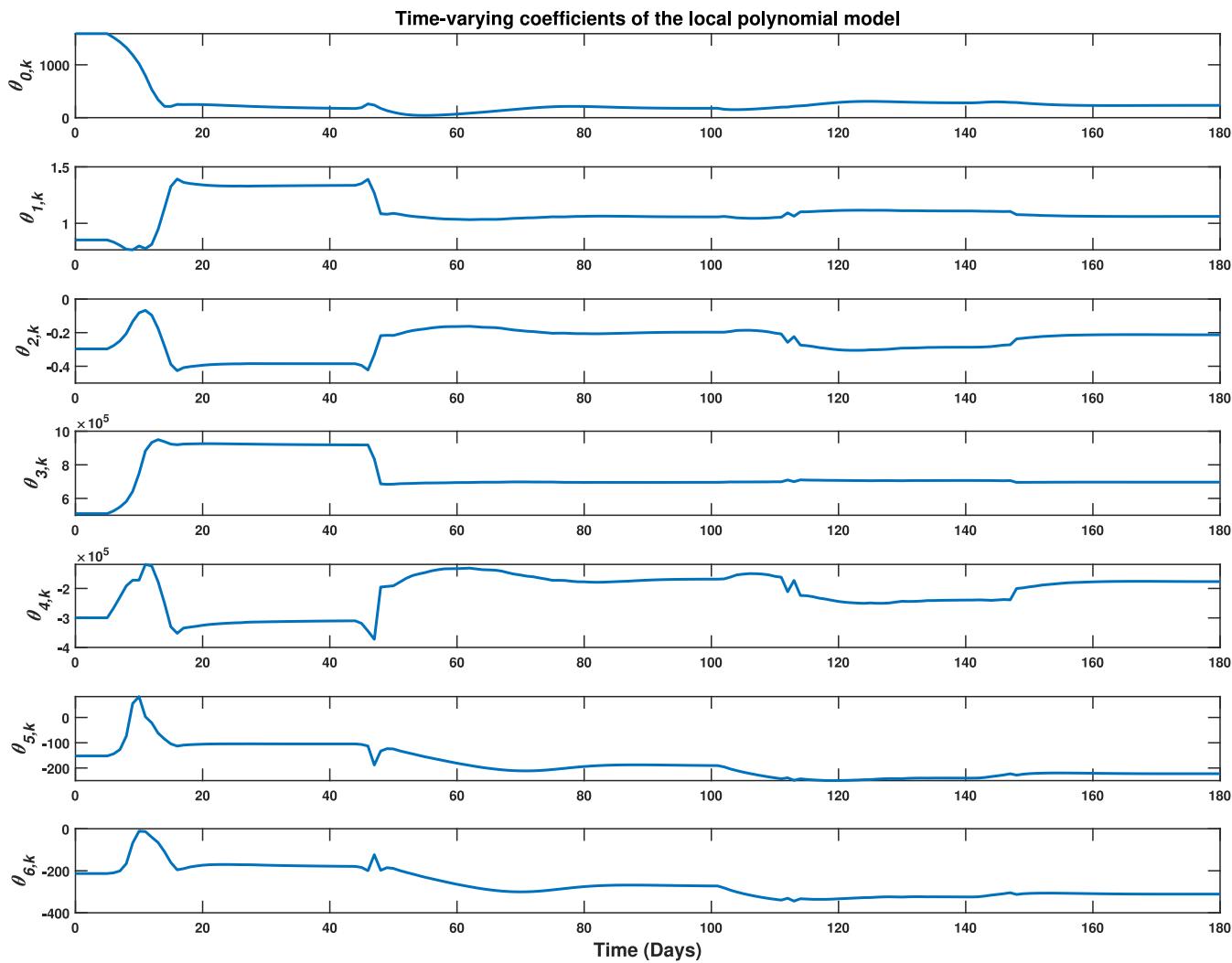


Figure 8. Time-varying coefficients of the locally linear MoD model described in eq 51 generated during closed-loop operation of the 3DoF-KF MoD MPC for infection control of the nonlinear SIR plant.

on the output predictions results in the superior control action of the 3DoF-KF MoD MPC algorithm over the ARX-based controller framework. In contrast, the 3DoF-KF ARX MPC fails to provide a good fit over the data associated with the nonlinearities in the lower regions of the infected dynamics and hence forces the controller to have an oscillatory behavior for the same set of constraints imposed on the MoD-MPC. Additionally, the 3DoF-KF MoD MPC displays faster mitigations in the face of disturbances than the ARX-based controller, signifying its superior ability to dynamically adapt and stabilize the infected dynamics in the presence of known or unforeseen events. It also demonstrates notably more feasible manipulated variable response with lesser oscillations while increasing $\beta(t)$ as a response to step changes in vaccination and recovery. The lower variance in the $\beta(t)$ signal translates to smoother adjustments in distancing policies (e.g., lockdown, masking), thereby avoiding abrupt and burdensome changes that may eventually lead to compliance fatigue across society. This results in more relaxed policies that allow restoration to normal activities. The better control is an outcome of the fact that the MoD-based predictive model can make better use of the information contained in the estimation database, as demonstrated in Figure 8.

A second example is demonstrated to particularly highlight the robustness of the 3DoF framework in the presence of plant-model mismatch and disturbances. For this, a MoD-based predictive model is utilized, and a comparative case study is presented by contrasting the performance of the 3DoF-KF MoD MPC framework with the conventional move-suppression-based single-DoF framework (SDoF-KF MoD MPC) under two sets of move suppression values: (a) $W_{\Delta u} = 0.01$ and (b) $W_{\Delta u} = 10$, the other controller and plant parameters remain the same as the previous case. As evident from Figure 9, while the 3DoF formulation attains fast set point tracking with minimal undershoot along with smooth rejection of measured and unmeasured disturbances, the SDoF cases demonstrate inferior performance in set point tracking as well as disturbance rejection. The SDoF controller with $W_{\Delta u} = 0.01$ further demonstrates an offset in response to the disturbances and higher variance in set point tracking. To reduce this, a higher move suppression is imposed, which significantly affects the overall profile of the manipulated variable and, consequently, the infection trajectory. While it reduces high-frequency variance, $W_{\Delta u} = 10$ restricts the movement of $\beta(t)$, resulting in noticeably slower responses and an increase in low-frequency oscillations that limit the controller's performance. These results clearly indicate the benefits of separately tuning the set point and the

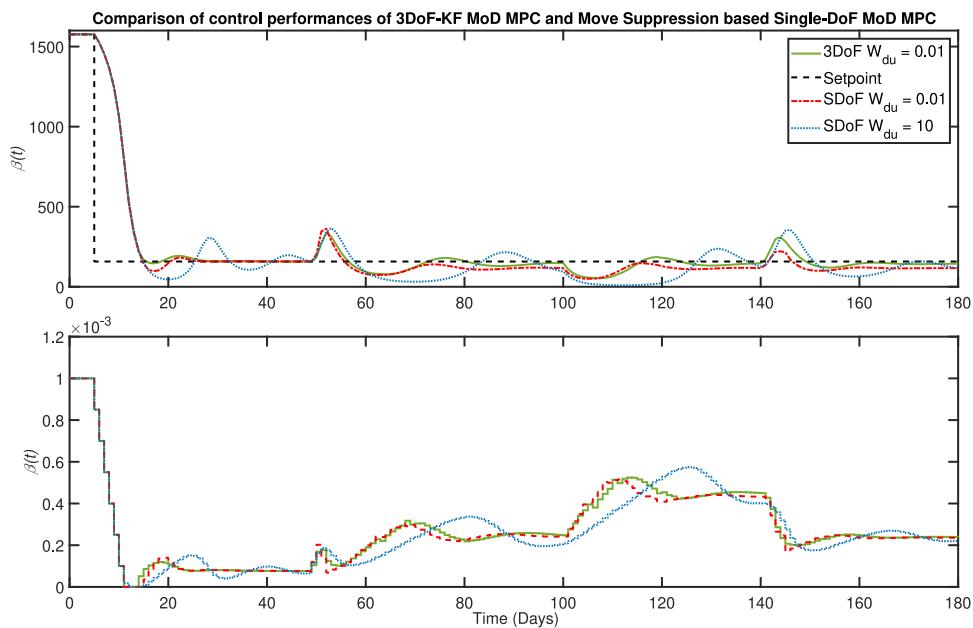


Figure 9. Plot illustrating the benefits of the 3DoF-KF MoD MPC ($W_{du} = 0.01$) over conventional move-suppression-based Single-DoF KF MoD MPC (SDoF; (a) $W_{du} = 0.01$ (b) $W_{du} = 10$), with improved set point tracking and effective disturbance rejection accompanied by low-variance manipulated variable response for the SIR model. Measured disturbance $\Delta k_v(t) = +0.3$ is introduced at $t = 50$ days, and unmeasured disturbances $\Delta \gamma(t) = +0.3$ and $\Delta \rho(t) = +0.2$ are introduced at $t = 120$ days and $t = 140$ days, respectively.

disturbance responses and, therefore, establish the benefits of the 3DoF framework.

4.2. The CSTR Model: A MIMO Case Study. The Continuous Stirred Tank Reactor (CSTR) model²⁶ is a commonly used unit operation in the chemical industry; in this paper, a perfectly mixed nonadiabatic reactor with a single irreversible exothermic reaction is considered, as illustrated in Figure 10. The kinetics considered are first-order, with the

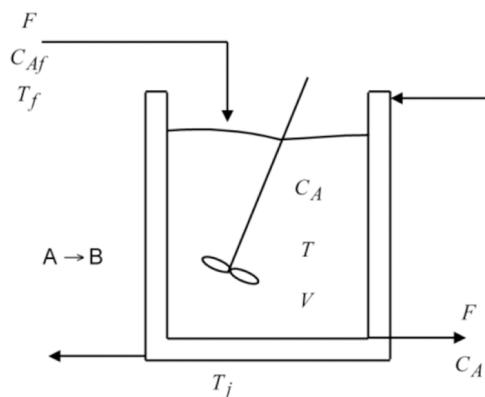


Figure 10. Schematic illustrating the MIMO CSTR considered in Case Study 2. C_A and T are controlled variables, T_j and F are manipulated, T_f and C_{A_f} are input disturbances.

overall system characterized through species accounting and energy balance equations²⁶ as follows:

$$\frac{dC_A(t)}{dt} = \frac{F(t)}{V}(C_{A_f}(t) - C_A(t)) - r_A(t) \quad (52)$$

$$\frac{dT(t)}{dt} = \frac{F(t)}{V}(T_f - T(t)) - \left(\frac{\Delta H}{C_p \rho} \right) r_A(t) - \frac{UA}{C_p \rho V} (T_j(t) - T(t)) \quad (53)$$

$$r_A(t) = k_o \exp\left(-\frac{\Delta E}{R_g T(t)}\right) C_A(t) \quad (54)$$

An inlet feed of the reactant A enters the reactor with concentration C_{A_f} at a flow rate F and temperature T_f . The product stream is removed from the reactor with a flow rate of F and possesses a concentration of C_A and an outlet temperature of T , which is the same as that of the reactor. The density of the fluids is assumed to be constant at ρ , and the volume of the reacting liquid is assumed to be constant at V , considering perfect level control. The temperature of the reactor jacket is denoted by T_j .

For control purposes, $T_j(t)$ and $F(t)$ are the manipulated variables. The primary control variable of interest is $C_A(t)$. However, considering the exothermic nature of the reaction, the reactor temperature $T(t)$ may vary significantly during operation, which is undesirable. Therefore, $T(t)$ is also controlled to minimize deviations from its steady-state value in the presence of disturbances arising in the feed temperature $T_f(t)$ (measured) and in feed concentration $C_{A_f}(t)$ (unmeasured). The remaining system parameters are noted in Table 5. The system is characterized by a high degree of nonlinearity involving ignition (rapid increase in reactor temperature and reactant conversion), extinction (the opposite phenomenon), and limit cycles (sustained oscillations). Additionally, there are strong interactions among outputs, making independent control of reactor concentration and reactor temperature challenging.

4.2.1. Input Signal Design. A multivariable input design is required for MIMO identification of the CSTR case study. This makes it imperative to obtain an understanding of gain directions of the system.²³ While it is easy to excite the high-gain direction, judicious design is required to obtain low-gain information relevant for control purposes. For the problem at hand, it is desired that a decrease in C_A (increase in yield) is accompanied by a decrease in reactor temperature T . However, this is difficult to achieve as it corresponds to the low-gain

Table 5. Plant Parameters for the MIMO CSTR Model

parameter	value	description
ΔH	-5960 kcal/kg mol	heat of reaction per mole
ΔE	11,843 kcal/kg mol	activation energy
ρC_p	500 kcal/m ³ °C	C_p : heat capacity
\bar{T}_f	298.15 K	steady-state feed temperature
\bar{C}_{A_f}	10 kg mol/m ³	steady-state feed concentration
UA	150 kcal/m ³ °C h	U : overall heat transfer coefficient A : area of heat exchange
\bar{T}_j	298.15 K	steady-state jacket temperature
\bar{C}_A	8.5695 kg mol/m ³	steady-state reactor concentration
\bar{T}	311.267 K	steady-state reactor temperature
R_g	1.98589 kcal/kg mol °C	universal gas constant

direction of the output. To properly excite both high and low-gain directions, Stec and Zhu²⁷ proposed an input signal design through concatenation of low-amplitude uncorrelated signals and high-amplitude correlated signals with user-specified relative amplitudes. The approach, while simple, runs the risk of pushing the plant actuators to their limits and involves long test durations. A suitable alternative is proposed in the frequency domain through a modified zippered approach using correlated and uncorrelated harmonics as depicted in Figure 11.^{23,25} The

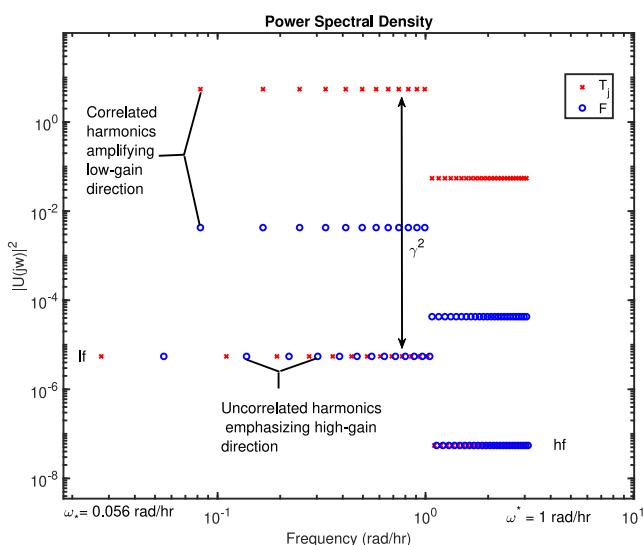


Figure 11. Plot illustrating the power spectrum corresponding to the modified zippered signals for T_j and F for the MIMO CSTR model, containing channels for both correlated and uncorrelated harmonics spread over the frequency grid, along with high (hf) and low-frequency (lf) contributions. The relative intensities are calculated from $\underline{\nu} = [-0.9996 \ -0.0279]^T$ and $\gamma = 100$; $lf = 0.5$, $hf = 0.1$.

modified zippered signal is realized in the time domain using the algorithm by Guillaume et al.,²⁴ providing a shorter, plant-friendly alternative to Stec and Zhu. For the 2-by-2 MIMO CSTR system, the modified zippered signals can be designed with three channels, two of which are associated with uncorrelated harmonics, and one is for the correlated. The uncorrelated harmonics are orthogonal in the frequency domain and correspond to the high-gain direction. The correlated harmonics can be utilized to provide the desired emphasis on the low-gain direction through a user-defined parameter γ and should be chosen in a manner that avoids high collinearity in the output as well as the input regressor space. A detailed discussion

on the modified zippered signal design can be found in Rivera et al.^{23,25} For the CSTR problem, directionality information can be obtained by considering the gain matrix of the linearized first-principles model (eqs 52–54) at steady-state conditions specified in Table 5:

$$y = G(0)u = \begin{bmatrix} -0.0565 & 3.2373 \\ 0.7490 & -26.6572 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (55)$$

$$\begin{aligned} SVD(G(0)) &= USV^T \\ &= \begin{bmatrix} 0.1205 & -0.9927 \\ -0.9927 & -0.1205 \end{bmatrix} \begin{bmatrix} 26.8635 & 0 \\ 0 & 0.0342 \end{bmatrix} \\ &\quad \begin{bmatrix} -0.0279 & -0.9996 \\ 0.9996 & -0.0279 \end{bmatrix}^T \end{aligned} \quad (56)$$

It is evident from eq 56 that it is easy to move the outputs in opposite directions; however, the designed input signal needs to be aligned with the low-gain direction for both C_A and T to decrease. The low-gain direction obtained from SVD pointed along $\underline{\nu} = [-0.9996 \ -0.0279]^T$ in the input regressor space. The corresponding amplification factor for the correlated harmonics was set to $\gamma = 1000$. This choice ensured that neither the C_A-T output pair nor the T_j-F input pair are highly correlated. Additionally, frequency support for high and low-frequency regions²⁸ was provided through the high-frequency ($hf = 0.1$) and the low-coefficient ($lf = 0.5$) magnitudes as illustrated in Figure 11.

The time-domain realizations of the signals were obtained through multisine signals implemented using the minimum crest-factor algorithm as per Guillaume et al.²⁴ with an amplitude ratio $\left(\frac{T_j}{F}\right)$ of $\frac{|u_1|}{|u_2|} = \left|\frac{-0.9996}{-0.0279}\right| = 35.79$, and a sampling time of $T = 1$ h. The actual amplitudes were chosen to be 25 K and $\frac{25}{35.79} \text{ m}^3/\text{hr} = 0.6985 \text{ m}^3/\text{hr}$ respectively. Furthermore, a band-limited white noise of amplitude 1 unit and an autoregressive filter $\frac{1.2}{z(z-0.3)}$ was used for generating data corresponding to the measured disturbance input. The signal parameters are provided in Table 6, and the power spectrum for the modified zippered design is presented in Figure 11.

Table 6. Signal Parameters for the Design of Guillaume-Phased Modified Zippered Signals²⁴

parameter	value	parameter	value
α_s	2	β_s	3
τ_{dom}^l	2	τ_{dom}^l	6
high-frequency ratio (hf)	0.1	low-frequency ratio (lf)	0.5
n_s	20	γ	1000

Different realizations of the signals with the same frequency content and amplitude were used for the estimation and the validation data, as shown in Figure 12. Figure 13 clearly demonstrates the benefit of the modified zippered design with correlated harmonics over a standard zippered design. The standard zippered approach generates data dominantly along the high-gain direction. However, the modified zippered input signals, when introduced into the system, spread the output data points widely across regressor space, covering both gain directions and reducing collinearity among outputs, thereby promoting well-conditioned estimation. This provides an

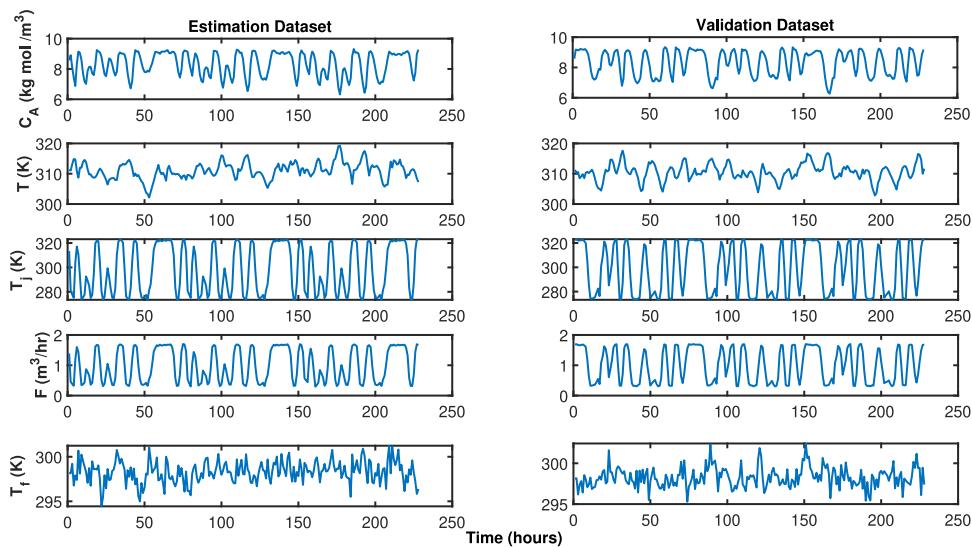


Figure 12. Plot illustrating the estimation and the validation database for the MIMO CSTR model using the modified zippered signal per Figure 11. The three inputs are T_j , F (manipulated), and T_f (measured disturbance), and the outputs are C_A and T . The input signals are different realizations with the same frequency contents and amplitudes.

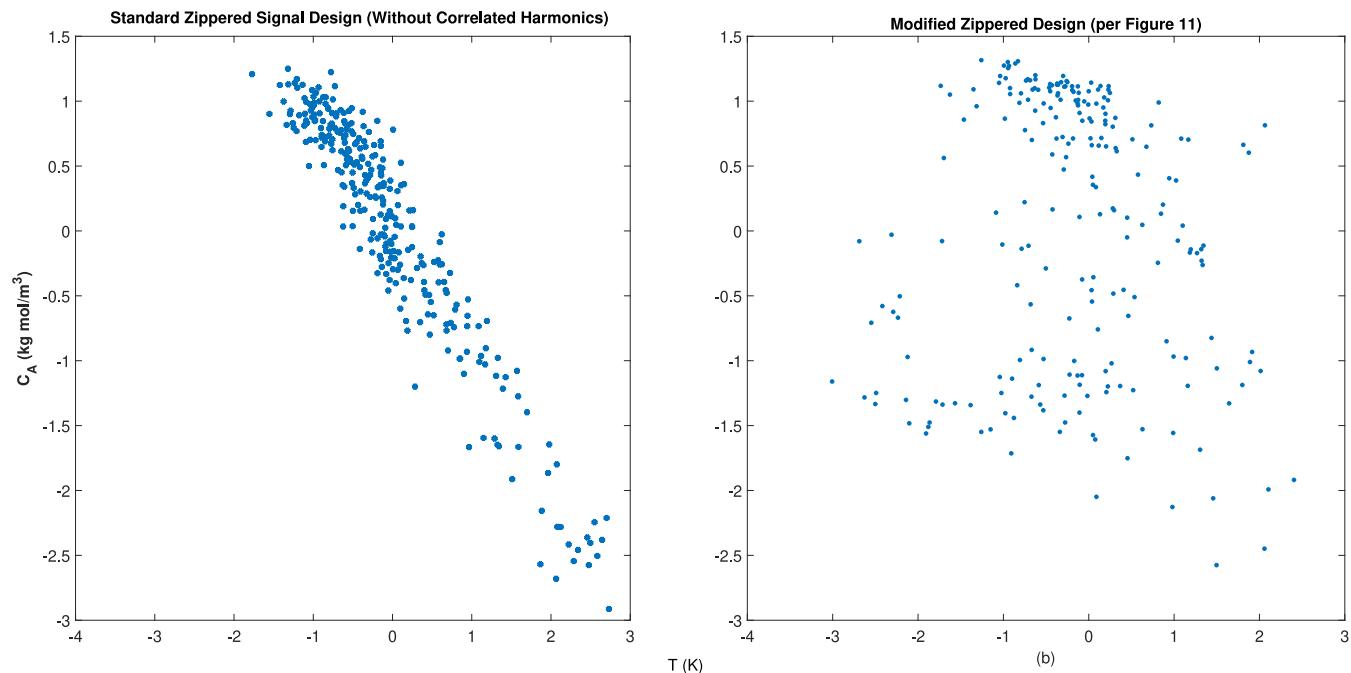


Figure 13. State-space plots illustrating the output regressor space for the MIMO CSTR model for (a) the standard and (b) the modified zippered input signal designs.

improved database that forms the heart of an MoD-based estimator. The overall regressor space is illustrated in Figure 14. Figure 15 further demonstrates the estimation database arising from the standard zippered signals for the inputs T_j and F . It struggles with issues of ignition and extinction, pushing the system to extremes in terms of C_A and T . Such signals are highly “plant-hostile” and consequently do not constitute a suitable database for MoD-based control. This was verified through closed-loop simulation, leading to an unstable response, the plot for which has not been included in this paper for the sake of brevity.

4.2.2. Open-Loop Model-on-Demand Estimation. The generated data set was standardized for well-conditioned

estimation. A locally linear polynomial was selected, with an ARX regressor structure of:

$$[n_a n_b n_k] = \begin{bmatrix} [22] & [222] & [111] \\ [22] & [222] & [111] \end{bmatrix} \quad (57)$$

giving rise to a model of the form:

$$\begin{aligned} C_{A,k} = & \theta_0 + \theta_1 C_{A,k-1} + \theta_2 C_{A,k-2} + \theta_3 T_{k-1} + \theta_4 T_{k-2} + \theta_5 T_{j,k-1} \\ & + \theta_6 T_{j,k-2} + \theta_7 F_{k-1} + \theta_8 F_{k-2} + \theta_9 T_{f,k-1} + \theta_{10} T_{f,k-2} \end{aligned} \quad (58)$$

with a similar model for the second output T_k . The k -based time indexing for the coefficients θ_i 's have been dropped for

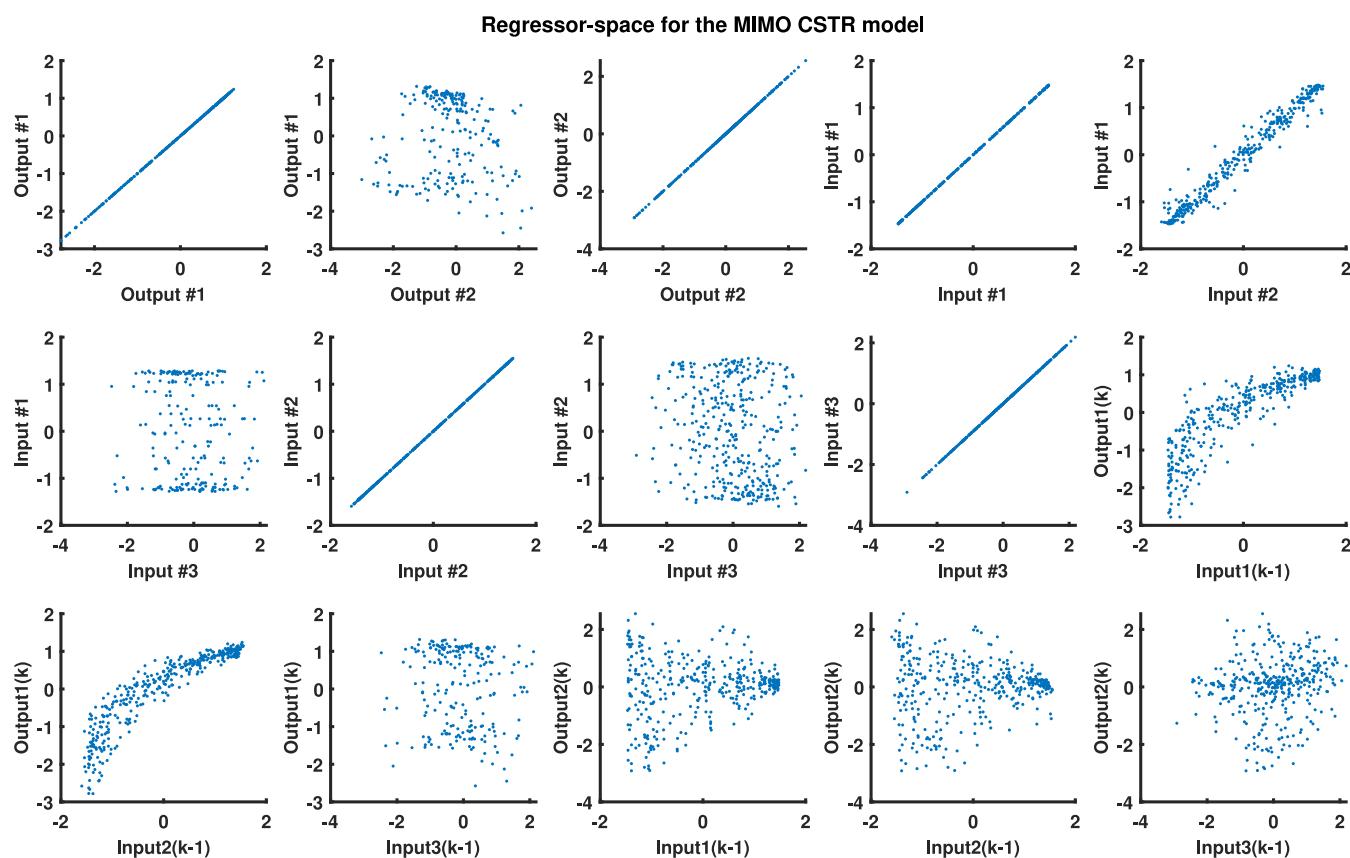


Figure 14. Plot illustrating the scaled input-output regressor space for the MIMO CSTR model from the modified zippered signal per Figure 11. The three inputs are T_p , F , and C_{A_f} (in that order), while the outputs are C_A and T .

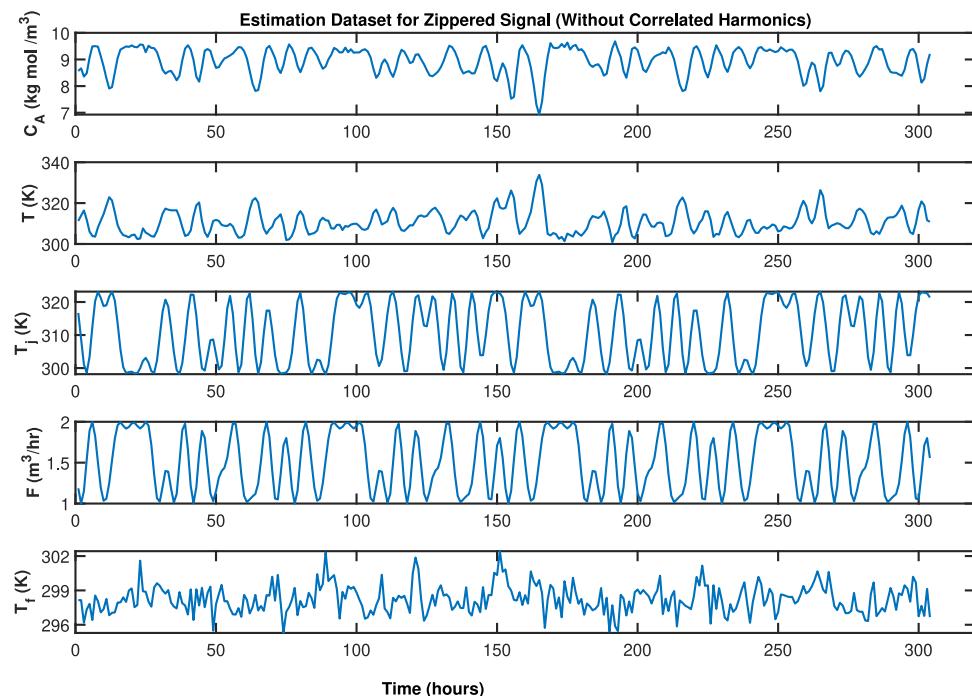


Figure 15. Plot illustrating the estimation database for the standard zippered signals for the input signals T_j and F for the MIMO CSTR model. These inputs feature ignition and extinction in the outputs and are not plant-friendly in nature.

notational convenience. A $[k_{\min}, k_{\max}]$ range of $[50, 100]$ was used to estimate the data. The localized Akaike Information Criterion (AIC) with a variance penalty of 3 was considered the

GOF criterion in the MoD estimation, while the weights were assigned through a tricube kernel function. Figure 16 compares the open-loop estimation results for the validation data for the

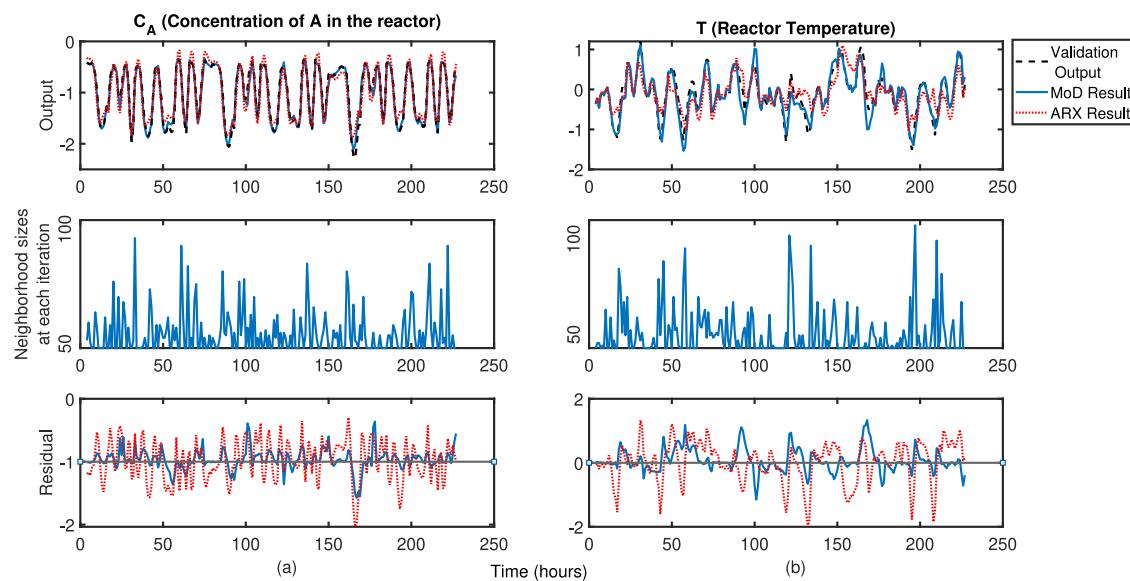


Figure 16. A comparative depiction of the open-loop identification performance of the MoD-based estimator and the ARX-based estimator for (a) concentration and (b) temperature in the MIMO CSTR problem. The middle plot shows neighborhood sizes for MoD estimates.

MoD and the ARX estimators under the same design parameters. Table 7 illustrates that the MoD-based estimation

Table 7. Open-Loop Estimation Results for MoD vs ARX-Based Estimators for C_A and T

method	C_A			T		
	NRMSE fit (%)	RMS error	max error	NRMSE fit (%)	RMS error	max error
MoD	84.12	0.16	0.63	63.45	0.37	1.33
ARX	68.87	0.32	1.03	38.28	0.64	1.94

(NRMSE fit: 84.12% for C_A and 63.45% for T) outperforms the conventional ARX estimator (NRMSE fit: 68.87% for C_A and 38.28% for T) in the aspects of fit percentages as well as the RMS

and maximum errors during open-loop estimation. This sets the ground for comparison of the two methods in the closed-loop setting.

4.2.3. Three-Degree-of-Freedom MPC based on the Model-on-Demand estimation. To evaluate the 3DoF-KF MoD MPC, a reference trajectory for C_A with multiple set point changes is evaluated, while the desired reactor temperature T set point is fixed at its steady-state value, as demonstrated in Figure 17. A step change in the measured disturbance variable $D_m = \Delta T_f = 2$ K at time $t = 120$ h, and the unmeasured disturbance $D_u = \Delta C_{A_f} = 0.1$ kg mol/m³ at $t = 180$ h are also introduced into the system. Results for the 3DoF-KF MoD MPC and the 3DoF-KF ARX MPC were generated using the plant parameters and input-output constraints provided in Table 8. The MoD MPC achieves precise set point tracking for C_A per the desired time-constant

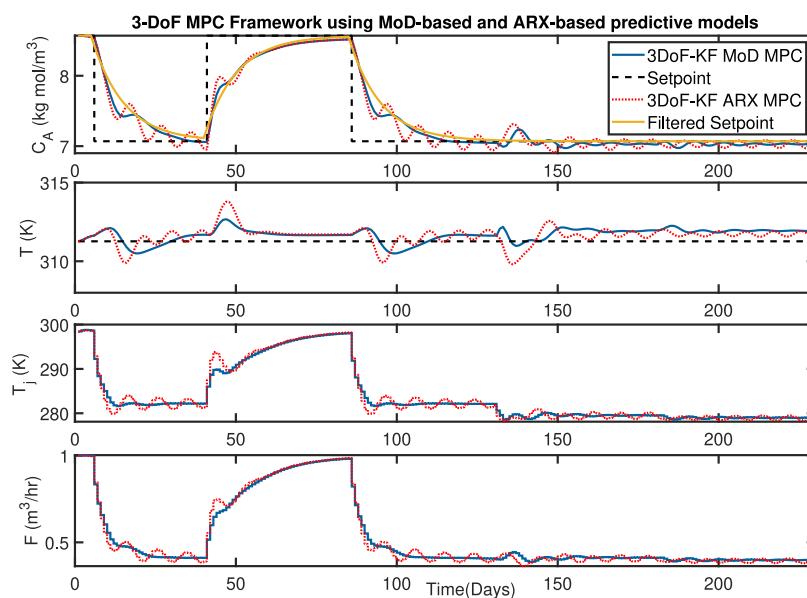


Figure 17. Closed-loop comparison between the 3DoF-KF MoD MPC and the ARX-based MPC for the MIMO CSTR model subject to set point changes, measured disturbance ($\Delta T_f = 2$ K at $t = 120$ h) and unmeasured disturbance ($\Delta C_{A_f} = 0.1$ kg mol/m³ at $t = 180$ h).

Table 8. Control Design Parameters for the MIMO CSTR Problem

parameter	value	parameter	value
p	50	y_{\min}	$[-\infty -\infty]$
m	20	y_{\max}	$[\infty \infty]$
τ_r	[10 10] h	Δu_{\min}	$[-\infty -\infty]$
τ_d	8 h	Δu_{\max}	$[\infty \infty]$
τ_u	[3.5 3.5] h	W_y	[1 1]
u_{\min}	[273.15 K 0 m ³ /h]	$W_{\Delta u}$	[0 0]
u_{\max}	[$\infty \infty$]	T_{sampling}	1 h

specifications, with no overshoot and faster response stabilization during abrupt set point changes. It also tracks the temperature set point with an acceptable worst-case deviation of 0.37 K from the desired steady-state value, as opposed to the 3DOF-KF ARX MPC with higher deviations in T (maximum deviation = 0.64 K) and a noticeably more oscillatory set point tracking for both outputs. This is a particularly challenging problem as the system tends to vary the temperature as the exothermic reaction progresses, and it highlights the benefits of MoD MPC over the ARX MPC. The ARX-based controller also exhibits prominent oscillations in manipulated variable response and pushes the system inputs to greater limits to meet the performance requirements that are easily satisfied by the MoD-based controller. Additionally, the 3DoF-KF MoD MPC effectively mitigates both measured and unmeasured disturbances, maintaining closer adherence to the set point trajectory. The ARX MPC, on the contrary, struggles more in the presence of these disturbances, leading to longer stabilization times and greater deviations from the set point. A quantitative comparison can also be made, as noted in Table 9. The MoD MPC exhibits

Table 9. Quantification of Control Performance Metrics for the 3DoF-KF MoD MPC and 3DoF ARX MPC, MIMO CSTR Model

method	ISE		max error		computational time
	C_A	T	C_A	T	
MoD	1.1063	81.0746	0.2461	1.4035	13 s
ARX	2.8037	103.5043	0.3740	2.5184	9 s

noticeably lower values for integral squared errors (ISE) and maximum deviations from filtered set point than the ARX counterpart, thereby showcasing enhanced plant friendliness, robustness, and stability in controlling the highly interactive MIMO system.

One way to better understand the results is to study the time-varying coefficients of the MoD-based models. The MoD-estimator-based controller enables effective data estimation in a local neighborhood of each operating point. Figure 18 demonstrates how changes in the time-varying coefficients for local linear models serve as a good indicator of model nonlinearities and interactions. While there are varying degrees of fluctuations among the coefficients associated with various time-lagged terms, their combined effects on the output predictions highlight the need for and the subsequent effectiveness of the MoD-based 3DoF algorithm in generating more contextually relevant control actions when compared to the ones arising from the ARX-based globally linear predictive model. A key consideration in this case study is the computational time. As noted in Table 9, the MoD-based MPC entails a computational time of 13 s, compared to the ARX

counterpart taking 9 s for the same simulation span of 250 days, when measured on an 11th Gen Intel(R) Core(TM) i7. While this extra effort is expected owing to the repeated estimation of the local models, the MoD-based controller does not involve insurmountable computational time and is well-suited for being implemented in situations where requirements are comparable to an ARX MPC.

Users can make design choices to further reduce computational time, such as selecting a smaller range for the neighborhood size [k_{\min}, k_{\max}], using a reduced-order regressor structure [n_a, n_b, n_k], or opting for a zeroth- or first-order local polynomial based on the complexity of the system. These decisions directly influence the design matrix used to calculate the coefficients of the local polynomial model.

5. SUMMARY AND CONCLUSIONS

This paper aims to bridge the gap between simplicity and performance of data-driven MIMO MPC algorithms for nonlinear process systems by developing the MIMO 3DoF-KF MoD MPC. This algorithm presents an appealing alternative to global nonlinear methods with noticeably better estimation and control while maintaining the ease and simplicity of fixed linear methods. The development of a successful MoD-based nonlinear estimation and control framework greatly relies on the development of a judicious database. This is attained through systematically integrating input signal design, parameter estimation, and 3DoF-KF predictive control. The results presented in this paper validate the advantages of the 3DoF-KF MoD MPC control scheme for nonlinear and highly interactive systems through two well-established case studies across domains of epidemiology and process systems engineering. The SIR model illustrates the workings of 3DoF-KF MPC for a nonlinear SISO problem and introduces many of the issues associated with the proposed methodology. The MIMO CSTR case study further demonstrates how MoD estimation can be combined with sensible input signal design and MIMO 3DoF-KF MPC to control highly interactive multivariable systems, which have long been considered demanding problems in process control.

One drawback of the current MoD approach includes discarding models generated at each time instant for MoD estimation. It can be envisioned that computational efficiency could be improved by reducing the frequency of local polynomial recalculations, updating these at longer time intervals; this may result in a performance loss that could be deemed not significant in many problems of practical significance.

Future research aims to better understand the effect of parameter regularization during MoD-based estimation and find better ways to measure the relevance of data during estimation beyond the current norm-based distance functions. In principle, the MoD database could be updated during the closed-loop operation, akin to closed-loop identification. This will be studied in the context of theoretical and practical implications.

Future and ongoing research also involves utilizing this framework in industrial settings, some initial efforts toward which have been performed in the context of a semi-industry-scale photobioreactor that performs combined wastewater treatment and sustainable microalgae synthesis.^{29,30} The work experimentally validates the 3DoF-KF MoD MPC framework in a real-world setting under disturbance and noise-prone nonlinear conditions with no computational delays. One must also consider that while the MoD-based approach provides a

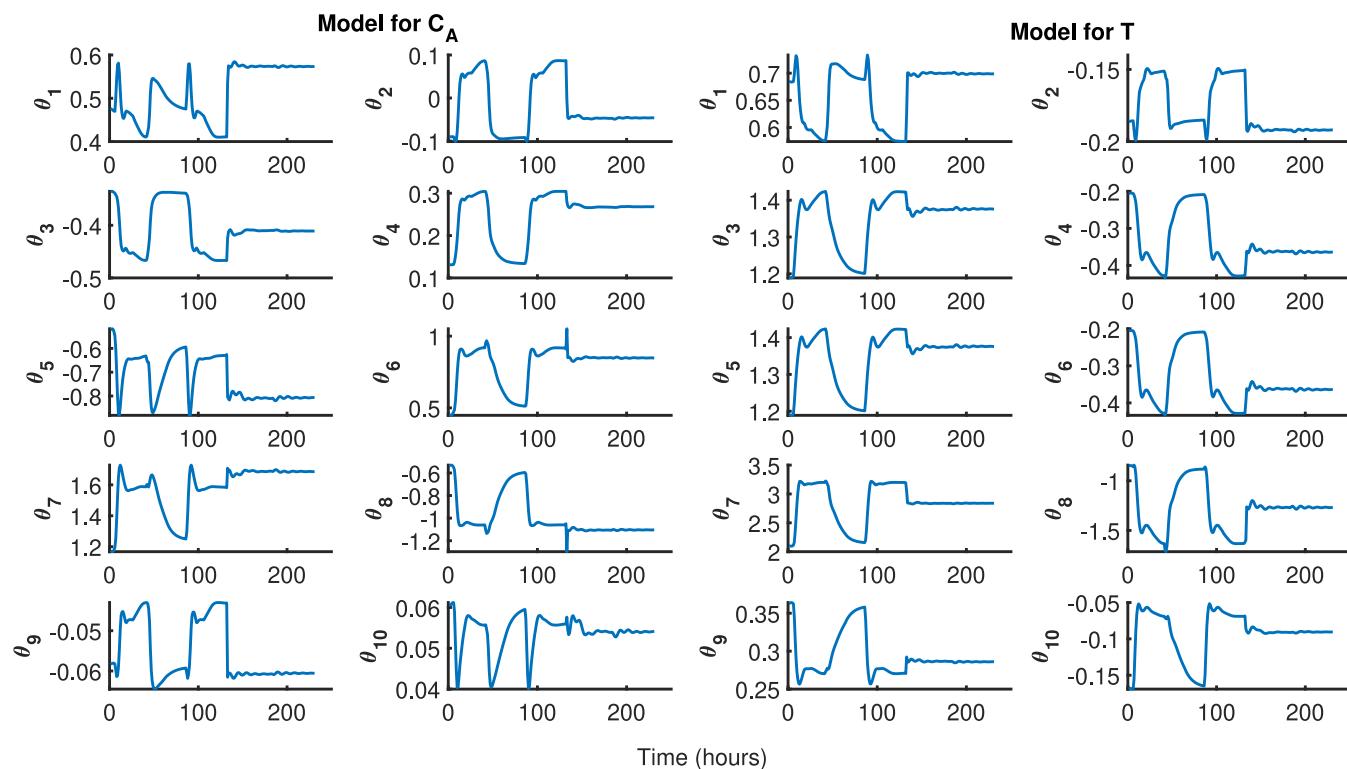


Figure 18. Time-varying coefficients of the locally linear models for C_A and T based on the 3DoF-KF MoD MPC estimation for the MIMO CSTR model. The coefficients correspond to the polynomial model generated using the $[n_a \ n_b \ n_k]$ regressor structure in eq S7. $\theta_1 - \theta_2$ are related to the time-lagged terms for C_A ; $\theta_3 - \theta_4$ are similarly associated with T ; $\theta_5 - \theta_6$, $\theta_7 - \theta_8$, $\theta_9 - \theta_{10}$ are coefficients for the inputs T_p , F , and C_A , respectively.

practical and accessible estimation algorithm, it presents a challenge for users due to the presence of additional tuning parameters. The heuristic nature of these parameters may lead to suboptimal solutions and even infeasibility in the computations if not properly understood. Additional work in examining alternative local polynomial modeling approaches such as the Direct Weight Optimization³¹ (which presents a compelling alternative by reducing the number of user decisions) seems warranted.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.iecr.4c04583>.

Coefficient matrices of the p -step ahead predictions of the output in eqs 26 and 27 (related to Section 3.3); Coefficient matrices of the objective function in eq 37 and constraints in eq 38 of the MPC problem (related to Section 3.3) ([PDF](#))

AUTHOR INFORMATION

Corresponding Authors

Sarasis Banerjee — Control Systems Engineering Laboratory, School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, Arizona 85287, United States; Email: sbaner74@asu.edu

Owais Khan — Control Systems Engineering Laboratory, School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, Arizona 85287, United States; Email: okhan5@asu.edu

Mohamed El Mistiri — Control Systems Engineering Laboratory, School for Engineering of Matter, Transport, and

Energy, Arizona State University, Tempe, Arizona 85287, United States; Email: melmistri@asu.edu

Naresh N. Nandola — Siemens Technology, Princeton, New Jersey 08540, United States; orcid.org/0000-0002-1195-3726; Email: nareshkumar.nandola@siemens.com

Eric Hekler — Center for Wireless & Population Health Systems, University of California, San Diego (UCSD), La Jolla, California 92093, United States; Email: ehekler@ucsd.edu

Daniel E. Rivera — Control Systems Engineering Laboratory, School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, Arizona 85287, United States; orcid.org/0000-0002-3141-0577; Email: daniel.rivera@asu.edu

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.iecr.4c04583>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Support for this research has been provided by NIH grants R01LM013107 and R01CA244777, and NSF grant CBET 2200161. The views presented in this paper are the authors' own and do not necessarily reflect the views of the NIH or the NSF.

REFERENCES

- (1) Carnerero, A. D.; Ramirez, D. R.; Limon, D.; Alamo, T. Kernel-Based State-Space Kriging for Predictive Control. *IEEE/CAC J. Autom. Syst.* **2023**, *10*, 1263–1275.
- (2) Piga, D.; Formentin, S.; Bemporad, A. Direct Data-Driven Control of Constrained Systems. *IEEE Trans. Control Syst. Technol.* **2018**, *26*, 1422–1429.

- (3) Berberich, J.; Köhler, J.; Müller, M. A.; Allgöwer, F. Data-Driven Model Predictive Control with Stability and Robustness Guarantees. *IEEE Trans. Autom. Control* **2021**, *66*, 1702–1717.
- (4) Torrente, G.; Kaufmann, E.; Föhn, P.; Scaramuzza, D. Data-Driven MPC for Quadrotors. *IEEE Rob. Autom. Lett.* **2021**, *6*, 3769–3776.
- (5) Ren, Y. M.; Alhajeri, M. S.; Luo, J.; Chen, S.; Abdullah, F.; Wu, Z.; Christofides, P. D. A tutorial review of neural network modeling approaches for model predictive control. *Comput. Chem. Eng.* **2022**, *165*, 107956.
- (6) Wu, Z.; Tran, A.; Rincon, D.; Christofides, P. D. Machine learning-based predictive control of nonlinear processes. Part I: Theory. *AIChE J.* **2019**, *65*, No. e16729.
- (7) Wu, Z.; Rincon, D.; Christofides, P. D. Real-Time Adaptive Machine-Learning-Based Predictive Control of Nonlinear Processes. *Ind. Eng. Chem. Res.* **2020**, *59*, 2275–2290.
- (8) Wu, Z.; Tran, A.; Rincon, D.; Christofides, P. D. Machine-learning-based predictive control of nonlinear processes. Part II: Computational implementation. *AIChE J.* **2019**, *65*, No. e16734.
- (9) Cybenko, G. Just-in-Time Learning and Estimation. *Identification, Adaptation, Learning* **1996**, 423–434.
- (10) Stenman, A. Model-on-Demand: Algorithms, Analysis and Applications. Ph.D. thesis, Linköping University, Sweden, 1999.
- (11) Stenman, A.; Gustafsson, F.; Ljung, L. Just in time models for dynamical systems. In *Proceedings of 35th IEEE Conference on Decision and Control* 1996; pp 1115–1120.
- (12) Braun, M. W.; Rivera, D. E.; Stenman, A. A Model-on-Demand identification methodology for nonlinear process systems. *Int. J. Control.* **2001**, *74*, 1708–1717.
- (13) Stenman, A. Model-free predictive control. In *Proceedings of the 38th IEEE Conference on Decision and Control* 1999; pp 3712–3717.
- (14) Braun, M. W. Model-On-Demand Nonlinear Estimation And Model Predictive Control: Novel Methodologies For Process Control And Supply Chain Management. Ph.D. thesis, Arizona State University, USA, 2001.
- (15) Braun, M.; Rivera, D. E.; Stenman, A. Model-on-Demand Model Predictive Control for Nonlinear Process Systems, AIChE 2000 Annual Meeting, Los Angeles, CA, paper 256h November 12–17, 2000; <https://labs.engineering.asu.edu/csel/wp-content/uploads/sites/43/2019/04/outline.pdf>.
- (16) Khan, O.; El Mistiri, M.; Banerjee, S.; Hekler, E.; Rivera, D. E. 3DoF-KF HMPC: A Kalman filter-based Hybrid Model Predictive Control Algorithm for Mixed Logical Dynamical Systems. *Control Eng. Pract.* **2025**, *154*, 106171.
- (17) Nandola, N. N.; Rivera, D. E. Model-on-Demand Predictive Control for Nonlinear Hybrid Systems With Application to Adaptive Behavioral Interventions. In *49th IEEE Conference on Decision and Control (CDC)* 2010; pp 6113–6118.
- (18) Banerjee, S.; Khan, O.; El Mistiri, M.; Nandola, N. N.; Rivera, D. E. Data-Driven Control of Highly Interactive Systems using 3DoF Model-On-Demand MPC: Application to a MIMO CSTR. *IFAC-PapersOnLine* **2024**, *58*, 420–425.
- (19) Simon, C. M. The SIR dynamic model of infectious disease transmission and its analogy with chemical kinetics. *PeerJ. Phys. Chem.* **2020**, *2*, e14.
- (20) Ljung, L. *System Identification: Theory for the User*; Pearson Education, 1998.
- (21) Morari, M.; Zafiriou, E. *Robust Process Control*; Prentice-Hall, Englewood Cliffs, NJ, 1989.
- (22) Lee, J. H.; Morari, M.; García, C. E. State-space Interpretation of Model Predictive Control. *Automatica* **1994**, *30*, 707–717.
- (23) Rivera, D. E.; Lee, H.; Mittelmann, H. D.; Braun, M. W. High-purity distillation. *IEEE Control Syst. Mag.* **2007**, *27*, 72–89.
- (24) Guillaume, P.; Schoukens, J.; Pintelon, R.; Kollar, I. Crest-Factor Minimization Using Nonlinear Chebyshev Approximation Methods. *IEEE Trans. Instrum. Meas.* **1991**, *40*, 982–989.
- (25) Rivera, D. E.; Lee, H.; Mittelmann, H. D.; Braun, M. W. Constrained multisine input signals for plant-friendly identification of chemical process systems. *J. Process Control* **2009**, *19*, 623–635.
- (26) Bequette, B. W. *Process Control: Modeling, Design, and Simulation*; Prentice Hall Professional, 2003.
- (27) Stec, P.; Zhu, Y. Some Study on Identification of Ill-Conditioned Processes for Control. In *Proceedings of the 2001 American Control Conference. (Catal. No. 01CH37148)* 2001; pp 1202–1207.
- (28) Lee, H. A plant-friendly multivariable system identification framework based on identification test monitoring. Ph.D. thesis, Arizona State University, 2006.
- (29) Banerjee, S.; Otálora, P.; El Mistiri, M.; Khan, O.; Guzmán, J. L.; Rivera, D. E. Control-Relevant Input Signal Design For Integrating Processes: Application to a Microalgae Raceway Reactor. *IFAC-PapersOnLine* **2024**, *58*, 360–365.
- (30) Otálora Berenguel, P. New control algorithms and artificial intelligence models to improve microalgae production systems. Ph.D. thesis, Universidad de Almería, 2024.
- (31) Roll, J.; Nazin, A.; Ljung, L. Nonlinear system identification via direct weight optimization. *Automatica* **2005**, *41*, 475–490.

Supporting Information for Data-Driven Control of Nonlinear Process Systems using a Three-Degree-of-Freedom Model-on-Demand Model Predictive Control Framework

Sarasij Banerjee,^{*,†} Owais Khan,^{*,†} Mohamed El Mistiri,^{*,†} Naresh N.
Nandola,^{*,‡} Eric Hekler,^{*,¶} and Daniel E. Rivera^{*,†}

[†]*Control Systems Engineering Laboratory, School for Engineering of Matter, Transport,
and Energy, Arizona State University, Tempe, AZ 85287 USA*

[‡]*Siemens Technology, Princeton, NJ 08540 USA*

[¶]*Center for Wireless & Population Health Systems, University of California, San Diego
(UCSD), La Jolla, CA 92093 USA*

E-mail: sbaner74@asu.edu; okhan5@asu.edu; melmist@asu.edu;
nareshkumar.nandola@siemens.com; ehekler@ucsd.edu; daniel.rivera@asu.edu

1 Coefficient Matrices of the p -step ahead predictions of the output in (28) and (29)

The matrices $\Phi, \mathcal{H}_u, \mathcal{H}_d, H_{i1} \forall i = u, d$ in Equations (28) and (29) from the main manuscript are given as follows:

$$\Phi = \begin{bmatrix} \mathcal{C}\mathcal{A} \\ \vdots \\ \mathcal{C}\mathcal{A}^p \end{bmatrix}, \quad H_{i1} = \begin{bmatrix} \mathcal{C}\mathcal{B}_i \\ \mathcal{C}\mathcal{A}\mathcal{B}_i \\ \vdots \\ \mathcal{C}\mathcal{A}^{p-1}\mathcal{B}_i \end{bmatrix}, \quad i = u, d \quad (1)$$

$$\mathcal{H}_u = \begin{bmatrix} \mathcal{C}\mathcal{B}_1 & 0 & \cdots & 0 & 0 \\ \mathcal{C}\mathcal{A}\mathcal{B}_1 - \mathcal{C}\mathcal{B}_1 & \mathcal{C}\mathcal{B}_1 & \cdots & 0 & 0 \\ \mathcal{C}\mathcal{A}^2\mathcal{B}_1 - \mathcal{C}\mathcal{A}\mathcal{B}_1 & \mathcal{C}\mathcal{A}\mathcal{B}_1 - \mathcal{C}\mathcal{B}_1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \mathcal{C}\mathcal{B}_1 & \vdots \\ \mathcal{C}\mathcal{A}^{m-1}\mathcal{B}_1 - \mathcal{C}\mathcal{A}^{m-2}\mathcal{B}_1 & \mathcal{C}\mathcal{A}^{m-2}\mathcal{B}_1 - \mathcal{C}\mathcal{A}^{m-1}\mathcal{B}_1 & \cdots & \mathcal{C}\mathcal{A}\mathcal{B}_1 - \mathcal{C}\mathcal{B}_1 & \mathcal{C}\mathcal{B}_1 \\ \mathcal{C}\mathcal{A}^m\mathcal{B}_1 - \mathcal{C}\mathcal{A}^{m-1}\mathcal{B}_1 & \mathcal{C}\mathcal{A}^{m-1}\mathcal{B}_1 - \mathcal{C}\mathcal{A}^{m-2}\mathcal{B}_1 & \cdots & \mathcal{C}\mathcal{A}^2\mathcal{B}_1 - \mathcal{C}\mathcal{A}\mathcal{B}_1 & \mathcal{C}\mathcal{A}\mathcal{B}_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathcal{C}\mathcal{A}^{p-1}\mathcal{B}_1 - \mathcal{C}\mathcal{A}^{p-2}\mathcal{B}_1 & \mathcal{C}\mathcal{A}^{p-2}\mathcal{B}_1 - \mathcal{C}\mathcal{A}^{p-1}\mathcal{B}_1 & \cdots & \mathcal{C}\mathcal{A}^{p-m+1}\mathcal{B}_1 - \mathcal{C}\mathcal{A}^{p-m}\mathcal{B}_1 & \mathcal{C}\mathcal{A}^{p-m}\mathcal{B}_1 \end{bmatrix} \quad (2)$$

$$\mathcal{H}_d = \begin{bmatrix} \mathcal{C}\mathcal{B}_i & 0 & \cdots & 0 & 0 \\ \mathcal{C}\mathcal{A}\mathcal{B}_i - \mathcal{C}\mathcal{B}_i & \mathcal{C}\mathcal{B}_i & \cdots & 0 & 0 \\ \mathcal{C}\mathcal{A}^2\mathcal{B}_i - \mathcal{C}\mathcal{A}\mathcal{B}_i & \mathcal{C}\mathcal{A}\mathcal{B}_i - \mathcal{C}\mathcal{B}_i & \cdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{C}\mathcal{A}^{p-1}\mathcal{B}_i - \mathcal{C}\mathcal{A}^{p-2}\mathcal{B}_i & \mathcal{C}\mathcal{A}^{p-2}\mathcal{B}_i - \mathcal{C}\mathcal{A}^{p-3}\mathcal{B}_i & \cdots & \mathcal{C}\mathcal{A}\mathcal{B}_i - \mathcal{C}\mathcal{B}_i & \mathcal{C}\mathcal{B}_i \end{bmatrix} \quad (3)$$

2 Coefficient matrices of the objective function in (42) and constraints in (43) of the MPC problem

The matrices \mathcal{H} and \mathcal{G} corresponding to the coefficient matrix for the quadratic and the linear terms, respectively, in Equation (42), and the coefficient matrices Γ and \mathcal{M} for constraints in Equation (43) are given as follows:

$$\mathcal{H} = 2(\mathcal{H}_u^T \widehat{W}_y \mathcal{H}_u + \widehat{W}_u + R_u^T \widehat{W}_{\Delta u} R_u), \quad \mathcal{G} = 2 \begin{bmatrix} g_u \end{bmatrix}^T, \quad \Gamma = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \end{bmatrix}, \quad \mathcal{M} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad (4)$$

$$g_u = \hat{X}_{flt,k}^T \Phi^T \widehat{W}_y \mathcal{H}_u - \mathcal{Y}_{r,flt,k}^T \widehat{W}_y \mathcal{H}_u + \mathcal{D}_{flt,k}^T \mathcal{H}_d^T \widehat{W}_y \mathcal{H}_u - d_{flt,k-1}^T H_{d1}^T \widehat{W}_y \mathcal{H}_u \\ - u_{k-1}^T (R_{u0}^T \widehat{W}_{\Delta u} R_u + H_{u1}^T \widehat{W}_y \mathcal{H}_u) - \mathcal{U}_{r,k}^T \widehat{W}_u \\ \Gamma_1 = \begin{bmatrix} \mathcal{H}_u \\ -\mathcal{H}_u \end{bmatrix}, \quad \Gamma_2 = \begin{bmatrix} \mathcal{I} \\ -\mathcal{I} \end{bmatrix}, \quad \Gamma_3 = \begin{bmatrix} R_u \\ -R_u \end{bmatrix} \quad (5)$$

$$b_1 = \begin{bmatrix} \mathcal{Y}_{max} - \Phi \hat{X}_k - \mathcal{H}_d \mathcal{D}_k + H_{u1} u_{k-1} + H_{d1} d_{k-1} \\ -\mathcal{Y}_{min} + \Phi \hat{X}_k + \mathcal{H}_d \mathcal{D}_k - H_{u1} u_{k-1} - H_{d1} d_{k-1} \end{bmatrix} \quad (6)$$

$$b_2 = \begin{bmatrix} \mathcal{U}_{max} \\ -\mathcal{U}_{min} \end{bmatrix}, \quad b_3 = \begin{bmatrix} \Delta \mathcal{U}_{max} + R_{u0} u_{k-1} \\ -\Delta \mathcal{U}_{min} - R_{u0} u_{k-1} \end{bmatrix} \quad (7)$$