Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-15 (2024) 360–365

Control-Relevant Input Signal Design For Integrating Processes: Application to a Microalgae Raceway Reactor ⋆

Sarasij Banerjee * **Pablo Otálora** ** **Mohamed El Mistiri** *
Owais Khan * **José Luis Guzmán** ** **Daniel E. Rivera** *

* *Control Systems Engineering Laboratory, School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, AZ 85287 USA (e-mail: {sbaner74, melmisti, daniel.rivera}@asu.edu).*
** *Department of Informatics, University of Almería, ceiA3, CIESOL, Ctra. Sacramento s/n, 04120 Almería, Spain (e-mail: {p.otalora, joguzman}@asu.edu).*

Abstract: Plants with integrators possess control-relevant modeling requirements that are typically ignored in the literature. Desired closed-loop speeds of response for such systems can differ significantly from their open-loop dynamics, causing conventional system identification guidelines to fail or be inaccurate. One such issue relates to experimental design. This paper presents guidelines for the design of excitation signals for system identification of plants with integrators, with application to the modeling and control of a microalgae raceway reactor. The concept is to excite the system through optimized test signals with control-relevant shaping of their power spectra. This facilitates a shift in emphasis of the identification objective from estimating a model with good open-loop performance to having a model possessing desired closed-loop characteristics. Such a consideration is particularly important when generating informative databases for estimating predictive models for closed-loop control. An illustration for this experimental design procedure is accomplished in this paper through the estimation of ARX-based models and, subsequently, model predictive control of the pH dynamics of an experimental raceway photobioreactor facility hosting sustainable microalgae production.

Keywords: identification for control, input and excitation design, robust control, model predictive control, process control

1. INTRODUCTION

Generating informative data forms a crucial step in system identification and subsequent control of any process system. However, many decisions in this regard, such as the input signal design, have to be undertaken based on limited *a priori* knowledge of the system. While guidelines for the input signal design of commonly encountered dynamics (e.g., overdamped, non-integrating plants) exist based on knowledge of their open-loop time constants (McFarlane and Rivera, 1992)) these fail or become impractical for cases where the desired closed-loop responses are significantly faster than their open-loop response. For such systems, the model emphasis has to be on the initial part of the response; an integrator provides a suitable approximation for the system dynamics in such a scenario. Nevertheless, conventional input signal design guidelines for such systems may result in excessively lengthy signals with improper time and frequency-domain emphasis. Control-relevant input signal design (and implementation) provides a fundamental and practical alternative, avoiding the need for long test durations.

This paper presents a comprehensive procedure for the design of custom input signals through the solution of a control-relevant parameter estimation problem (CRPEP) involving prediction error minimization (Rivera and Gaikwad, 1995). The power spectrum of the input signal forms a part of the control-relevant weight which can be designed to emphasize the frequency range interest while minimizing the prediction error term. We discuss a systematic procedure to design the weight based on control requirements, system type, and input signal type. The procedure is validated through the design of a plant-friendly multisine input signal that addresses the requirements of an integrating system. In particular, we demonstrate the generation of frequency-weighted multisine signals to identify the pH dynamics of an experimental photobioreactor facility. The control problem involves regulating the pH of a raceway reactor through the manipulation of CO_2 inflow into the reactor in the presence of varying solar radiation and water temperature (both measured) and other influencing factors (unmeasured). The dynamics can be modeled as first-order systems with dead time (FOPDT), with pronounced differences in the time constants of the open-loop process and

⋆ This work has been financed by the following projects: PID2020–112709RB-C21 project financed by the Spanish Ministry of Science and the Horizon Europe – the Framework Programme for Research and Innovation (2021–2027) under the agreement of grant no. 101060991 REALM.

10.1016/j.ifacol.2024.08.555

Fig. 1. Photobioreactor facility in Almería, Spain that forms the basis for this study. Figure adapted from Pataro et al. (2023).

the closed-loop pH regulation; these characteristics make it a suitable test case for validating our design procedure.

The paper is organized as follows: Section 2 describes the motivating problem involving a pilot-scale reactor facility. Section 3 deals with formulating power spectra as a control-relevant parameter estimation problem. Section 4 describes the reactor-specific control-relevant input signal design, the open-loop estimation, and the culminating closed-loop model-based control using a three-degree-of-freedom model predictive control (3DoF MPC) (Khan et al., 2022). Section 5 summarizes the advantages of this approach and notes the possible directions for future work.

2. MOTIVATING PROBLEM: A RACEWAY PHOTOBIOREACTOR

2.1 Description of the Raceway Reactor System

The current study involves a real-time raceway photobioreactor facility involved in the production of microalgae (Fernández et al., 2016), as depicted in Fig. 1. The strain of the cultivated microalgae belongs to the Scenedesmus almeriensis (CCAP 276/24) species. The reactor uses freshwater for the medium and additional nutrients necessary for the growth of microalgae are provided externally. For the current study, CO_2 *inflow*, *water temperature* and *radiation* are considered as inputs to the system, and the pH of the reactor is chosen to be the relevant output. The process model of the reactor can be conceptualized to be of the form:

$$pH(s) = \frac{K_{CO_2}}{\tau_{CO_2}s+1}e^{-t_{CO_2}s}CO_2(s) + \frac{K_{Temp}}{\tau_{Temp}s+1}Temp(s) + \frac{K_{Rad}}{\tau_{Rad}s+1}Rad(s) \tag{1}$$

The microalgae have a high growth rate and can tolerate pH from 3 to 10, the optimum value for their production being at 8 (Otálora et al., 2023). Typically, the pH of the reactor is kept around the optimal value through the application of an on-off controller. During the daytime (8 am to 8 pm), in the absence of CO_2 inflow, as the *radiation* and *temperature* increase, the culture grows through photosynthesis, subsequently increasing the level of pH in the medium. When the pH level crosses the optimal value of 8, a fixed amount of CO_2 is introduced into the reactor,

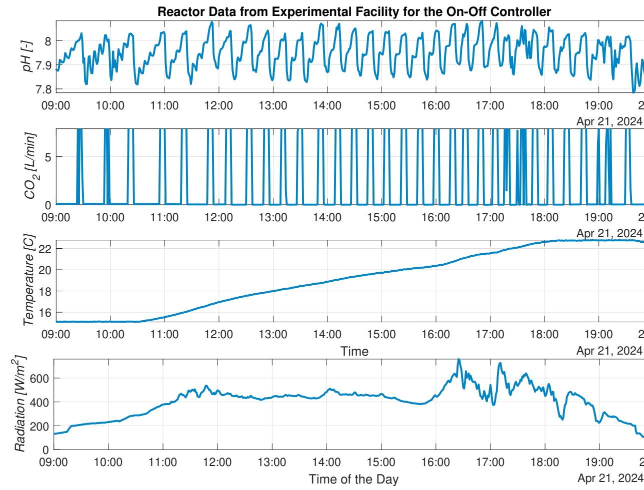

Fig. 2. Data collected from the raceway photobioreactor regulated by an on-off controller over a single day.

which dissolves in water to generate carbonic acid, thereby lowering the pH. This on-off mechanism keeps the pH of the medium within the system's limits; however, it does not constitute an efficient control architecture that allows tight setpoint tracking and disturbance rejection. More sophisticated control algorithms such as model predictive control (MPC) require system identification of models that suitably represent the relevant plant dynamics.

2.2 Identification of an Open-Loop Process Model

The data for the pH and the CO_2 signals are obtained from the reactor facility augmented with an on-off controller set to maintain the pH around 8. The radiation and water *temperature* signals for the simulation, along with other plant parameters, are obtained from real-time environmental data gathered from the actual reactor facility (Fernández et al., 2016). The time interval over which the data was obtained for identification corresponded to the presence of sunlight; this lies between 9 a.m. and 8 p.m., as demonstrated in Fig. 2. Once the input-output dataset is generated, first-order plus deadtime process models are estimated with a constraint of [-0.3 -0.05] on the gain of the $pH - CO_2$ dynamics. The estimated models are given as follows:

$$pH(s) = \frac{-0.08671}{15.61s+1}e^{-5s}CO_2(s) + \frac{0.004048}{320.96s+1}Temp(s) + \frac{0.009783}{136.20s+1}Rad(s) \tag{2}$$

where the time units are in minutes. The estimated open-loop time constants are noticeably slower than the desired closed-loop speeds of response; coupled with the "sawtooth" response in Fig. 2., these make an integrator a suitable approximation for the initial time dynamics of the problem.

At this point, one could use the estimated time constant of the $pH - CO_2$ dynamics in conjunction with the conventional guidelines specified in McFarlane and Rivera (1992) to generate PRBS signals covering the frequency range:

$$\frac{1}{\beta_s \tau^H_{dom}} \leq \omega \leq \frac{\alpha_s}{\tau^L_{dom}} \tag{3}$$

where $\tau_{dom}^H, \tau_{dom}^L$ denote the approximate values 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. Setting $\tau_{dom}^H = \tau_{dom}^L = 15.61$ mins according to (2) and choosing α_s =2 and $\beta_s = 3$ yields a signal of cycle length 296 minutes with a flat power spectrum over the frequency range [0.0214 0.1281] rad/min. This now presents a two-fold problem : (a) an excessively long test duration and (b) an equal emphasis on all frequencies of operation. The latter issue is particularly problematic for an integrator approximation which motivates the need for emphasis on the desired closed-loop dynamics and recognition of the integrating response. This emphasis is achieved in this paper through a control-relevant, plant-friendly input signal with a notch emphasis on the desired dynamics. Furthermore, the control-relevant design makes provisions for significantly shorter test durations as demonstrated in the subsequent sections, making their implementation physically realizable.

3. THE CONTROL-RELEVANT PARAMETER ESTIMATION PROBLEM

The Control-Relevant Parameter Estimation Problem (CRPEP) is an optimization problem that minimizes the closed-loop error by formulating it as a function of the weighted estimation error between the true plant and the identified model, where the weight, the error description, and the functional dependencies can be specified based on the end-use control problem for the model. The CRPEP problem relevant to the input signal design for integrating systems involves the evaluation of a control-relevant weight with a notch spectrum that de-emphasizes the low-frequency steady-state fit of the model and highlights the closed-loop dynamics.

3.1 Derivation of the CRPEP

For a unity feedback control structure, the objective of the CRPEP problem is to minimize the 2-norm of the control error $e_c(= r - y)$, given by:

$$||e_c||_2 = \left(\sum_{k=0}^{\infty} e_c^2(k) \right)^{1/2} \tag{4}$$

Assuming that the true plant $p(q)$ is estimated using the model $\tilde{p}(q)$, the nominal performance transfer functions of a feedback controller $c(q)$ can be written as:

$$\tilde{\eta}(q) = \tilde{p}c(1 + \tilde{p}c)^{-1} \quad \tilde{\epsilon}(q) = (1 + \tilde{p}c)^{-1} \tag{5}$$

where $\tilde{\eta}(q)$ and $\tilde{\epsilon}(q)(= 1-\tilde{\eta}(q))$ denote the complementary sensitivity and the sensitivity functions respectively. Considering the multiplicative error between $p(q)$ and $\tilde{p}(q)$ is given by $e_m = (p - \tilde{p})\tilde{p}^{-1}$, it can be shown via Linear Fractional Transformation that the open-loop modeling error affects the closed-loop error e_c as follows:

$$e_c = \frac{\tilde{\epsilon}(r - d)}{1 + \tilde{\eta}e_m} \tag{6}$$

where r and d denote changes in setpoint and disturbance respectively. Assuming that stability is enforced on the system via the Small Gain Theorem ($|\tilde{\eta}(e^{j\omega})e_m(e^{j\omega})| \leq 1 \; \forall -\pi \leq \omega \leq \pi$) and $|\tilde{\eta}(e^{j\omega})e_m(e^{j\omega})| << 1$ over the bandwidth of $\tilde{\epsilon}(r-d)$, the CRPEP can be reduced using (6) to a modified optimization problem (Rivera and Gaikwad, 1995), which is of the form:

$$\min_{\tilde{p}} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |\tilde{\epsilon}|^2 |\tilde{\eta}|^2 |r - d|^2 |e_m|^2 d\omega \right)^{1/2} \tag{7}$$

subject to

$$||\tilde{\eta}e_m||_\infty = \sup_{\omega} |1 - \tilde{\eta}e_m| < 1 \; \forall \omega \tag{8}$$

The merit of this formulation lies in the weight term augmented with multiplicative error in the overall expression. The frequency-weighted term $W_m(\omega) = |\tilde{\epsilon}\tilde{\eta}||r - d|$ can now be designed to satisfy the closed-loop requirements provided the information regarding the nominal performance requirements (specifications regarding $\tilde{\epsilon}$ and $\tilde{\eta}$), and the type of the control problem (setpoint tracking/ disturbance rejection) are available.

3.2 Effect of Input Signal Power on Prediction Error

To introduce the power spectrum of the input signal into this formulation, we consider prediction error estimation. The frequency domain representation of the prediction error method demonstrates the precise influence that the input signal power spectrum has on the parameter estimation problem. Consider a system given by:

$$y(t) = p(q)u(t) + \nu(t) \tag{9}$$

where $p(q)$ denotes the true plant. $\nu(t) = H(q)a(t)$ is a noise/unmeasured disturbance, and can be considered an output of a shaping filter $H(q)$ whose input is a white noise signal $a(t)$. Assuming $y(t)$, $u(t)$, and $\nu(t)$ to be stationary, the aim is to estimate the plant model $\tilde{p}(q)$ and the noise model $\tilde{p_e}(q)$ such that the error in our one-step ahead prediction corresponds to just the random unmeasured part of the data, or in other words, the prediction error is white. This is achieved by minimizing the prediction error. The one-step-ahead prediction error is given by:

$$e_P(t) = y(t) - \hat{y}(t|t-1) \tag{10}$$

$$= \tilde{p_e}^{-1}(q)[(p(q) - \tilde{p}(q))u(t) + \nu(t)] \tag{11}$$

where $\hat{y}(t|t-1) = \tilde{p_e}^{-1}(q)\tilde{p}(q)u(t) + (1 - \tilde{p_e}^{-1}(q))y(t)$ denotes the one-step-ahead prediction of the output. Using the formulation presented in Ljung (1999) the prediction error minimization problem can asymptotically be written as:

$$\min_{\tilde{p}, \tilde{p_e}} \lim_{N \to \infty} V_N = \min_{\tilde{p}, \tilde{p_e}} \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} [e_P(t)]^2 \tag{12}$$

$$= \min_{\tilde{p}, \tilde{p_e}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{e_P}(\omega) d\omega \tag{13}$$

$\Phi_{e_P}(\omega)$ is the power spectrum of the prediction error achieved from (11) using Parseval's theorem. Assuming y, u, and ν are uncorrelated, $\Phi_{e_P}(\omega)$ is given by:

$$\Phi_{e_P}(\omega) = \frac{1}{|\tilde{p_e}(e^{j\omega})|^2} [|p - \tilde{p}|^2 \Phi_U(\omega) + \Phi_\nu(\omega)] \tag{14}$$

where $\Phi_U(\omega)$ and $\Phi_\nu(\omega)$ are the power spectra of the input signal and the noise signal respectively. This is a multi-objective weighted optimization problem with $|\tilde{p_e}(e^{j\omega})|$ influencing the fit to the predictive model as well as the fit to noise model. To ensure that this does not introduce undesired bias in the situation, we assume an Output Error (OE) structure of the Prediction Error Model with

$|\tilde{p}_e(e^{j\omega})| = 1$. Keeping in mind that the aim is to design an input signal that has sufficient power over a particular frequency range, a *signal-to-noise* ratio of $\frac{\Phi_U(\omega)}{\Phi_\nu(\omega)} >> 1$ over the frequency band of interest is considered. The overall problem reduces to a frequency-weighted error minimization problem:

$$\min_{\tilde{p}, \tilde{p}_e} \frac{1}{2\pi} \int_{-\pi}^{\pi} |p - \tilde{p}|^2 \Phi_U(\omega) d\omega \tag{15}$$

with the power spectrum of the input signal $\Phi_U(\omega)$ acting as a frequency-dependent weight that determines the region of frequency over which the bias minimization occurs. Equating this with the modified CRPEP objective function from (7) results in:

$$|p - \tilde{p}|^2 \Phi_U(\omega) = |\tilde{\epsilon}|^2 |\tilde{\eta}|^2 |r - d|^2 |e_m|^2 \tag{16}$$

$$= |\tilde{\epsilon}|^2 |\tilde{\eta}|^2 |r - d|^2 |\tilde{p}^{-1}|^2 |p - \tilde{p}|^2 \tag{17}$$

which finally leaves us with a control-relevant power spectrum defined as:

$$\Phi_U(\omega) = A_R^2(\omega) = |\tilde{\epsilon}|^2 |\tilde{\eta}|^2 |r - d|^2 |\tilde{p}^{-1}|^2 \tag{18}$$

where $A_R(\omega) = |\tilde{\epsilon}||\tilde{\eta}||r - d||\tilde{p}^{-1}|$ denotes the relative magnitude of frequencies in the input signal. Equation (18) well illustrates the specifications that are needed *a priori* to design a control structure that caters to our needs:

- $\tilde{\eta}$ and $\tilde{\epsilon}$: These terms define the nominal performance of the closed-loop system. The desired speed of response dictates the frequency range over which the estimation error needs to be minimized. Faster the desired speed, greater the bandwidth over which $\tilde{p}(e^{j\omega})$ needs to match $p(e^{j\omega})$.
- r and d: The term $|r - d|$ defines the type of the control problem to be solved- setpoint tracking with $d = 0$ or disturbance rejection with $r = 0$. Furthermore, r and d also specify the type of input signal to be considered- step, ramp, or parabolic.
- $\tilde{p}(e^{j\omega})$: Considering the objective of the control-relevant identification is to minimize the multiplicative error generated from the estimated model, it forms a part of the power spectrum definition.

Now that a general methodology to generate a custom power spectrum based on the requirements of a control problem is presented, the case for an integrating system is demonstrated in the following subsection.

3.3 Input Signal Design for an Integrating System

The initial response dynamics of the pH to changes in CO_2 can be suitably approximated as a pure integrator with a delay. This approximation forms the basis for designing the control relevant power spectra as demonstrated in this section. Consider the approximation to be of the form:

$$\tilde{p}(s) \simeq \frac{e^{-\theta s}}{s} \tag{19}$$

Following the H_2-optimal IMC-based Q-parameterization from Morari and Zafiriou (1989), the nominal plant model can be factored into two parts:

$$\tilde{p}(s) = \tilde{p}_A(s)\tilde{p}_M(s) \tag{20}$$

where $\tilde{p}_A(s)$ and $\tilde{p}_M(s)$ comprise the all-pass and the minimum phase portion respectively. For the nominal plant model described in (19), $\tilde{p}_A = e^{-\theta s}$ and $\tilde{p}_M = \frac{1}{s}$.

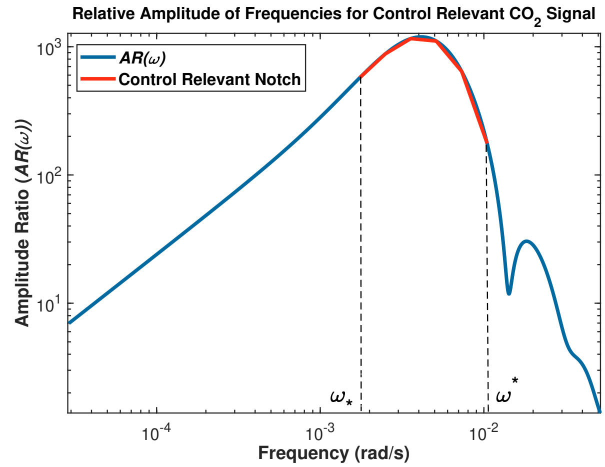

Fig. 3. Plot illustrating the relative amplitude of frequencies (red) over the control-relevant band $[\omega_*, \omega^*]$=[0.00095, 0.01050] rad/s.

With an integrator in the forward path, the corresponding control system demands a controller that rejects ramp output disturbance ($d = \frac{1}{s^2}$). This leads to an optimal Q-parameterized controller of the form:

$$\tilde{q} = \frac{\theta s + 1}{p_M} = s(\theta s + 1) \tag{21}$$

This setup further motivates the need for a Type-II filter $f(s)$ of the form $\frac{n_f \lambda s + 1}{(\lambda s + 1)^{n_f}}$ for a proper controller transfer function, we achieve:

$$q_{final}(s) = \tilde{q}(s) * f(s) = \frac{s(\theta s + 1)(n_f \lambda s + 1)}{(\lambda s + 1)^{n_f}} \tag{22}$$

where the filter time constant λ determines the desired closed-loop speed of response. This results in a nominal complementary sensitivity function η of the form:

$$\eta = \tilde{p}(s) * q_{final}(s) = \frac{(\theta s + 1)(n_f \lambda s + 1)e^{-\theta s}}{(\lambda s + 1)^{n_f}} \tag{23}$$

For our specific case of the photobioreactor, the desired closed-loop speed of response is set at 3 minutes, a transport delay of 5 minutes is considered, and $n_f = 4$ is chosen. This effectively provides notch emphasis on the relative amplitudes of the frequencies, as demonstrated in Fig. 3, and therefore on the power spectrum of the resultant input signal. The frequency band of interest is given by $[\omega_*, \omega^*]$ where: $\omega_* = \frac{2\pi}{N_s T_s}$, and $\omega^* = \frac{\pi}{T_s}$. N_s gives the sequence length and T_s denotes the sampling time of the signal. Considering that the length of the study was 11 hours (9 a.m. to 8 p.m.), plant-friendly signals with a cycle length (l_c) of 110 mins and a physically realizable sampling time of T_s= 5 min were designed. This resulted in the study being spanned by 6 cycles of CO_2 signals. The corresponding design parameters were N_s=l_c/Ts=(1*110/5) = 22, n_s=$N_s/2$= 11, and $[\omega_*, \omega^*]$= [0.00095, 0.01050] rad/s. The multisine input signal was designed through crest-factor minimization (Guillaume et al., 1991) according to the following equation:

$$u(k) = \lambda_s \sum_{j=1}^{n_s} \sqrt{2\alpha_{R_j}(\omega)}\, cos(\omega_j k T_s + \phi_j) \tag{24}$$

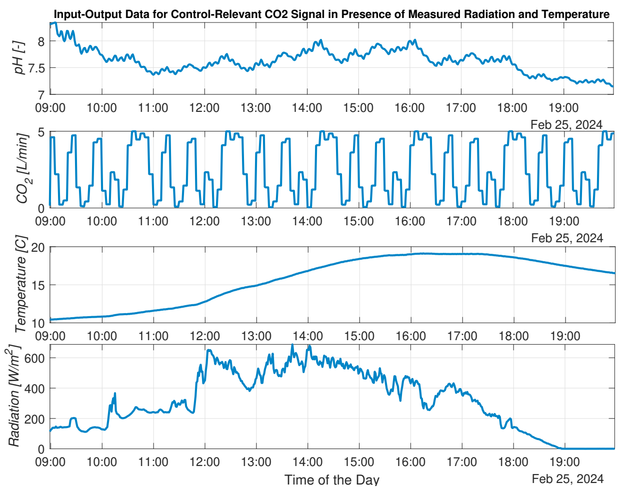

Fig. 4. Plot showing the control-relevant CO_2 signal and the corresponding input-output dataset for an illustrative day. The dataset arising from the on-off control for the same day is also provided for comparison.

where $n_s \leq N_s/2$ refers to the number of sinusoids, $\alpha_{R_j}(\omega) = \frac{A_R^2(\omega_j)}{\sum_{j=1}^{n_s} A_R^2(\omega_j)}$ gives the relative power of the frequencies. The harmonics present in the signal are given by $\omega_j = \frac{2\pi j}{N_s T_s}$, and their phases correspond to ϕ_j. λ_s denotes the scaling factor that meets amplitude requirements.

4. DATA-DRIVEN IDENTIFICATION AND CLOSED-LOOP CONTROL OF THE EXPERIMENTAL RACEWAY REACTOR

The data-driven estimation of the photobioreactor dynamics involves the identification of pH models as a function of CO_2 *inflow*, *radiation*, and *water temperature*. An important consideration in the design of estimation data involves the choice of amplitude of the CO_2 signal. A practical limitation in the use of pure CO_2 stream for the reactor operation is its cost-intensive nature. Therefore, it is crucial to minimize the overall consumption of CO_2 in addition to providing the required signal power to adequately excite the desired plant dynamics. Fig. 4 demonstrates the input-output data obtained for an illustrative day with an amplitude of 5 L/min for the control-relevant CO_2 signal. It can be observed that the control-relevant signal obtains an even spread of data over the range of operation of the pH dynamics in the open-loop settings. While the on-off controller pushes the system actuators to their limits with sharp changes of 10 L/min in the CO_2 inflow and overall consumption of 2435 L, the plant-friendly multisine signal with the control-relevant design facilitates noticeably smaller move sizes with an overall CO_2 consumption of 1589 L over the study day. This marks a reduction of 34.74% in daily average consumption of CO_2 during the period of system identification.

The estimation dataset is generated using input-output data collected over a single day of study. Similarly, for validation,input-output data from a separate day is used. System identification of the plant dynamics is achieved through the estimation of parsimonious ARX models with regressor structure of [2 [2 2 2] [6 1 1]]. The model estimated from the data illustrated in Fig. 4 generated an NRMSE fit percentage of 57.87% on the validation dataset demonstrated in Fig. 5, which can be considered to be an appropriate estimate accounting for several other influencing factors, such as dissolved oxygen level or biomass concentration, that are not included as a part of the model. Additional mismatch can also be ascribed to high-frequency dynamics arising from circulating water, which are unmodeled and whose influence in the problem is unclear.

Closed-loop control of the reactor dynamics is achieved through a 3DoF MPC architecture (Khan et al., 2022) involving the ARX-based predictive model. The MPC scheme involves solving a constrained optimization problem over a prediction horizon p and a move horizon m. The 3DoF philosophy entails independent tuning of parameters related to setpoint tracking (τ_r), measured disturbance rejection (τ_d), and noise rejection (τ_u). As a part of the 3DoF scheme, a filtered setpoint is provided to the controller after passing through a Type-I filter with a time constant τ_r. The higher the τ_r, the lower the speed of response to changes in reference trajectory. It is of the form, $f(q, \alpha_r) = \frac{(1-\alpha_r)q}{q-\alpha_r}$ where q^{-1} denotes the time lag operator, $\alpha_r = e^{-T_s/\tau_r}$ and T_s denotes the sampling time. A Type-II filter is implemented for the measured disturbance with a time constant τ_d due to the integrating nature of the system and the overdamped step nature of the input radiation disturbance. The filter is given by: $f(q, \alpha_d) = (\beta_0 + \beta_1 q^{-1} + \cdots + \beta_\omega q^{-\omega}) \times \frac{(1-\alpha_d)q}{q-\alpha_d}$ where $\alpha_d = e^{-T_s/\tau_d}$ and ω denotes the number of autoregressive lag terms. $\omega = 5$ for this study. The unmeasured disturbance rejection involves state estimation using Kalman filters, which uses a two-step process and the parameter $f_a = 1 - e^{-T_s/\tau_u}$ that directly influences the speed of unmeasured disturbance rejection ($0 \leq f_a < 1$). Higher values of f_a imply higher correction for prediction error arising from noise in the state estimation and, therefore, more aggressive control. The pH-control problem explored in this paper comprises tracking a pH-setpoint of 8 during the active hours of the day while rejecting measured disturbance arising from *radiation* and *temperature*, besides other time-varying unmeasured disturbances, as illustrated in Fig. 6. The controller parameters are summarized in Table 1. While the 3DoF structure provides a robust framework

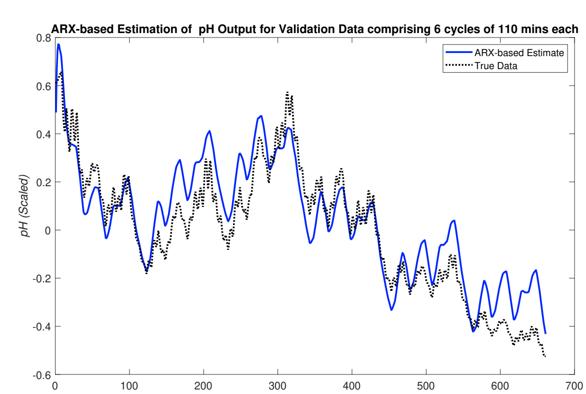

Fig. 5. Plot showing the open-loop simulation (NRMSE fit equal to 57.87%) of the control-relevant ARX model on a validation dataset with inputs CO_2, *temperature*, and *radiation*.

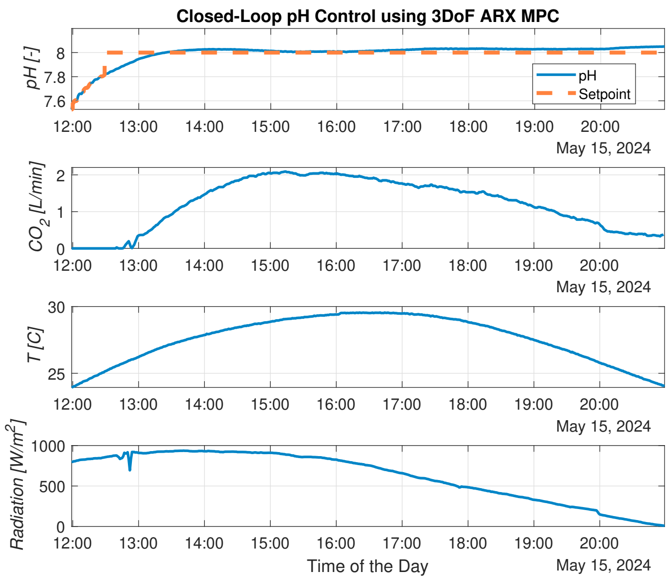

Fig. 6. Plot illustrating the pH-control through the 3DoF MPC based on the ARX model estimated from the control-relevant input signal design.

Table 1. Control design parameters for the 3-DoF ARX MPC for pH-control

Parameter	Value	Parameter	Value
p	150	y_{min}	$-\infty$
m	20	y_{max}	∞
τ_r	20 mins	Δu_{min}	-∞
τ_d	[1 1] min	Δu_{max}	∞
τ_u	1 min	W_y	1
u_{min}	0 L/min	$W_{\Delta u}$	0.5
u_{max}	15 L/min	T_s	1 min

in itself, the control performance primarily depends on the ability of the predictive model to capture the requisite closed-loop dynamics. Fig. 6 demonstrates that the MPC using the ARX predictive model arising from the control-relevant dataset successfully attains the pH setpoint with minimum overshoots and offsets during the active hours of the day. The controller is commissioned at 12:30 pm, before which the CO_2 flow is manually held constant at 0 L/min, and the setpoint is set to the current output value. A "bumpless transfer" mechanism is implemented at 12:30 pm, where the controller switches to automatic. It initially holds the CO_2 at minimum, facilitating the rise of the pH to sufficient levels, thereafter increasing CO_2 inflow to prevent overshoots arising from *radiation* and *temperature* changes. pH is then held at setpoint through effective rejection of measured and unmeasured disturbances over the course of the day. The manipulated variable response displays minimal CO_2 consumption of 637 L/min during the hours of operation and is characterized by a smooth, plant-friendly CO_2 response profile that is easy to implement in the real-time reactor facility.

5. SUMMARY AND CONCLUSIONS

This paper has examined, using an example of a raceway bioreactor, that considering control requirements forintegrating plants at the onset of experimental design can significantly reduce the length of the experiment, enables useful insights into estimation of various plant dynamics, and leads to a model achieving fast control actions along with robustness and reduced sensitivity to noise. MPC based on the predictive model generated from the control-relevant input signal design provides an attractive approach while minimizing design costs and maintaining simplicity in the estimation and control of the system dynamics. The methodology presented paves the way for future efforts in designing control-relevant input signals with additional considerations such as input and output constraints (Rivera et al., 2009). Further work may also involve control-relevant designs tailored towards more sophisticated data-driven estimation and control algorithms that can generate more contextually relevant models over varying operating conditions and, therefore, facilitate control of highly nonlinear systems in the future.

6. ACKNOWLEDGEMENT

The visit to the IFAPA reactor facility in Almería, Spain that led to the experimental results shown here was partially funded by an ASU Graduate and Professional Students Association (GPSA) Career Development Award.

REFERENCES

Fernández, I., Acién, F., Guzmán, J.L., Berenguel, M., and Mendoza, J. (2016). Dynamic model of an industrial raceway reactor for microalgae production. *Algal Research*, 17, 67–78.

Guillaume, P., Schoukens, J., Pintelon, R., and Kollar, I. (1991). Crest-factor minimization using nonlinear Chebyshev approximation methods. *IEEE Transactions on Instrumentation and Measurement*, 40(6), 982–989.

Khan, O., El Mistiri, M., Rivera, D.E., Martin, C.A., and Hekler, E. (2022). A Kalman filter-based Hybrid Model Predictive Control Algorithm for Mixed Logical Dynamical Systems: Application to Optimized Interventions for Physical Activity. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2586–2593. IEEE.

Ljung, L. (1999). *System Identification Theory For the User*. PTR Prentice Hall, New Jersey, second edition.

McFarlane, R. and Rivera, D. (1992). Identification of Distillation Systems. *Practical Distillation Control*, 96–139.

Morari, M. and Zafiriou, E. (1989). *Robust Process Control*. Prentice Hall Englewood Cliffs, NJ.

Otálora, P., Guzmán, J.L., Berenguel, M., and Acién, F.G. (2023). Data-Driven pH Model in Raceway Reactors for Freshwater and Wastewater Cultures. *Mathematics*, 11(7), 1614.

Pataro, I.M., Gil, J.D., Guzmán, J.L., Berenguel, M., and Lemos, J.M. (2023). A learning-based model predictive strategy for pH control in raceway photobioreactors with freshwater and wastewater cultivation media. *Control Engineering Practice*, 138, 105619.

Rivera, D.E. and Gaikwad, S.V. (1995). Systematic techniques for determining modelling requirements for SISO and MIMO feedback control. *Journal of Process Control*, 5(4), 213–224.

Rivera, D.E., Lee, H., Mittelmann, H.D., and Braun, M.W. (2009). Constrained multisine input signals for plant-friendly identification of chemical process systems. *Journal of Process Control*, 19(4), 623–635.