IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 14 August 2025, accepted 18 September 2025,
date of publication 24 September 2025, date of current version 30 September 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3614061

RESEARCH ARTICLE

End-to-End Process Optimization in Polymer Extrusion Lines Using Model Predictive Control and Multi-Task Learning

RAPHAEL HARTNER[1], MARTIN KOZEK[2], AND STEFAN JAKUBEK[2]
[1]Institute of Industrial Management, University of Applied Sciences FH JOANNEUM, 8605 Kapfenberg, Austria
[2]Institute of Mechanics and Mechatronics, Vienna University of Technology, 1040 Vienna, Austria

Corresponding author: Raphael Hartner (raphael.hartner2@fh-joanneum.at)

This work was in part by the Project Data Based Expert System for Process Efficiency (DEEPEN), funded by the Austrian Research Promotion Agency (FFG) through the 41st Call of the Initiative "Production of the Future," under Project 891247.

ABSTRACT Due to the inherent complexity of modern polymer extrusion lines caused by nonlinear, dynamic behavior with numerous influential factors, disturbances lead to higher scrap rates and downtimes. Unfortunately, conventional approaches to address this challenge rely on independent models and local control loops, neglecting the multistage characteristic of extrusion lines and therefore, cannot obtain a global optimum. Consequently, we propose a predictive control design based on a nonlinear and autoregressive multi-task learning model covering the entire extrusion line including local control loops and essential quality measurements. Training on an extended prediction horizon successfully addresses accumulating prediction errors and measurement noise. Due to its capability to predict state trajectories for up to 60 minutes, the proposed methodology enables effective model predictive control. Additionally, using an efficient method for adaptive error compensation based on previous error trajectories increases robustness substantially while the hierarchical control architecture supports global optimization and efficient local control loops. The proposed control design is validated on two polymer extrusion lines. The results show that changing operating points and optimizing process states, such as melt temperature, can be reliably achieved in spite of process disturbances. In comparison to baseline production periods, dominant oscillations are successfully damped by 62 %, reference values are closely followed and process variations are reduced by 41 % to 63 % improving product quality notably. A comparative analysis of disturbance rejection capabilities shows superior results while empirical analyses of stability and runtime further improve its closed-loop applicability. As a consequence, the proposed methodology for predictive control allows to steer extrusion processes accurately, address upcoming issues in advance and generally leads to more efficient processes.

INDEX TERMS Model predictive control, multistage production, multi-task learning, polymer extrusion, time series forecast.

I. INTRODUCTION

Polymer extrusion is one of the core process steps for a variety of production lines ranging from injection molding to a continuous production of profiles, pipes and plastic bottles. With regard to pipe production, modern extrusion lines consist of several process steps directly connected to the extruder, such as calibration, cooling, cutting and often specialized post processing [1]. Numerous influential factors at the extruder as well as upstream (material supply) and downstream lead to a complex nonlinear process [2], [3] with long transport delays [4], [5]. These factors include but are not limited to ambient conditions, variations in the feed, machine wear and oscillations in cooling systems. Consequently, the conditions and the behavior of the entire extrusion line are determining the resulting product quality of

The associate editor coordinating the review of this manuscript and approving it for publication was Mohsin Jamil.

polymer pipes requiring an end-to-end perspective for global optimization [6], [7].

To estimate the pressure in extrusion processes, shallow and deep neural networks are trained on experimental data to model the static dependencies on process settings, signals and recipes [8], [9]. Similarly, Tan et al. relied on deep neural networks to estimate the melt index as main quality indicator for polypropylene extrusion based on inputs from the extruder, gear pump and die [10]. Akram et al. focused on the same process to evaluate different machine learning methods to predict melt temperature and pressure showing that light gradient boosting machine and gradient boosting regression are well suited for the task [11]. In addition to that, Trifkovic et al. analyzed several process identification methods to model the dynamics of a twin screw extruder to predict motor load, melt temperature and pressure [12].

However, these static and dynamic approaches only consider the extruder itself while neglecting the resulting product quality. To address this point, Alhindawi and Altarazi investigated different machine learning algorithms to estimate the tensile strength of extrusion-blown high density polyethylene films based on features from the process and material [13]. Comparatively, Mulrennan et al. used tree-based algorithms to estimate the yield stress of extruded polyactide sheets to implement a soft sensor for real-time feedback [3]. Apart from mechanical properties, Garcia et al. as well as Bovo et al. showed that dimensional quality characteristics (diameter, wall thickness) can be reliably estimated with support vector machines (SVM) while Hartner et al. demonstrated that semi-supervised learning based on mass conservation can improve quality predictions in spite of data scarcity [14], [15], [16]. To predict the resulting product quality based on process parameters and material properties, Polychronopoulos et al. investigated three machine learning algorithms (random forest, XGBoost, support vector regression) and found that trial and error efforts for process optimization can be effectively reduced by utilizing model predictions [17]. Similarly, Takada et al. used random forest regression to predict and subsequently optimize the impact strength of polyphenylene sulfide blends based on process settings and material properties in high speed twin screw extrusion [18]. To take influential factors along the extrusion line into account and improve process yield, Jun et al. relied on a recurrent neural network (RNN) for quality predictions considering ambient conditions (temperature, humidity) and cooling water temperatures in addition to signals from the extruder itself [7].

Due to accumulating effects along multistage production lines, the specific process structure provides an important source of information for predictive models [19]. For this purpose, Ismail et al. relied on the cascade quality prediction method (CQPM) with individual sub-models separately trained to represent each production stage and improve quality predictions [20]. In addition to that, multi-task learning (MTL) allows to jointly train models for all outputs simultaneously within one end-to-end model, where the inputs and outputs are sequentially injected into and extracted from the model in accordance to their position and known causal dependencies in discrete and continuous production systems [21], [22], [23]. Furthermore, Hartner et al. extended the MTL approach to implement autoregressive end-to-end models based on physics-informed state propagation reliably capturing the dynamic and nonlinear behavior of entire extrusion lines. Training on an extended forecasting horizon allowed to predict the trajectory of relevant process states for a prolonged period of 60 to 120 minutes [24].

Nonetheless, optimizing an entire extrusion line requires an elaborate control design to steer the process and address upcoming disturbances effectively. Conventional methods of process control, such as proportional-integral-derivative (PID) control, fuzzy logic control (FLC) or extremum seeking control, are used to stabilize and optimize the extruder in terms of temperature, energy consumption and pressure [25], [26], [27], [28], [29]. While interesting results are achieved and hybrid approaches based on particle swarm optimization, fuzzy rules and PID control allow to address certain limitations [30], they cannot integrate future trajectories and still represent a reactive approach. On the other hand, model predictive control (MPC) allows to address this limitation while requiring an accurate process model. For this purpose, Grimard et al. designed a nonlinear MPC based on partial differential equations derived from mass and energy balance to control the extrusion output and the active pharmaceutical ingredient concentration in drug manufacturing [31]. Similarly, Celga et al. derived an analytical model for a reactive extrusion system and used a MPC to control the die pressure and throughput [32]. Additionally, Schwarzinger and Schlacher proposed a hierarchical control design with a MPC based on the heat transfer equation to control the melt temperature [33].

In contrast to MPCs based on analytical models, data-driven methods allow to model the dynamics based on actual measurements from the system. In this regard, Trifkovic et al. identified an autoregressive model with exogenous inputs (ARX) based on experimental data to control melt temperature and motor load in a twin-screw extruder for thermoplastic vulcanizate applications [34]. Moreover, Jiang et al. employed two control loops consisting of individual MPCs based on controlled autoregressive integrated moving average (CARIMA) structures to steer barrel temperatures and the melt pressure as proxy variable for product quality [4]. Apart from extrusion processes, Wu et al. relied on a RNN to model the nonlinear dynamics of chemical processes as the basis for a MPC. To account for known causal dependencies, the RNN was decoupled to represent the process structure [35], similar to aforementioned MTL methods.

In contrast to conventional MPC approaches based on forward models of the dynamics, Aschemann et al. and Lukas et al. relied on inverse models (neural networks) and process measurements to control rubber extrusion lines but

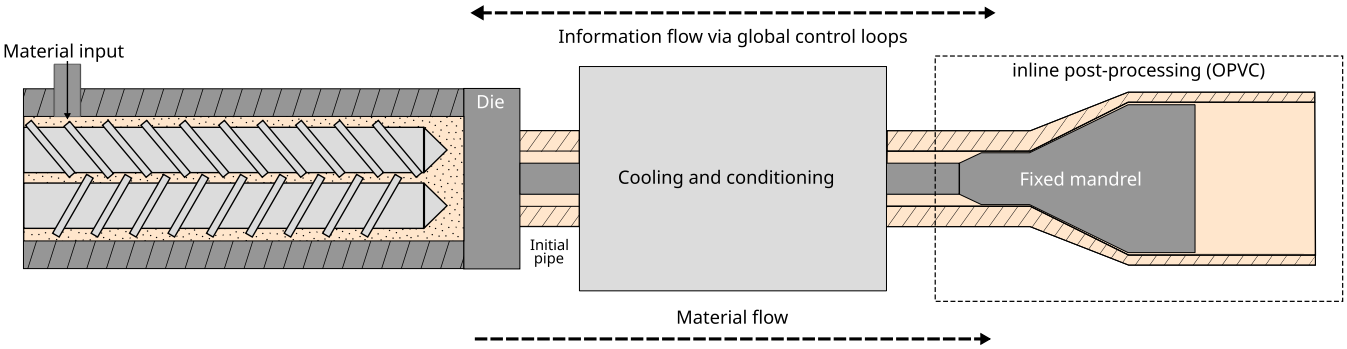


FIGURE 1. General setup of pipe extrusion processes with cooling and conditioning steps as well as a directly connected drawing section for bi-oriented PVC pipes (top view).

FIGURE 2. Setup of the bi-oriented PVC production line with gravimetric dosing system, extruder, adapter, die head, cooling zones and subsequent drawing section (left to right).

neglected the end-to-end dynamics of the process [36], [37]. Correspondingly, Adesanya et al. relied on an inverse model implemented as neural network and trained via the Levenberg Marquardt algorithm to predict suitable process parameters based on targeted quality characteristics in cable extrusion lines [38].

However, none of the existing work bridges the gap between end-to-end models and predictive control to steer entire extrusion lines for global optimization while accounting for downstream product quality characteristics. Additionally, conventional control approaches, static soft sensors and inverse models either completely neglect the multistage characteristics or the process dynamics, substantially limiting their applicability for global optimization. As a consequence, we propose to capture the nonlinear dynamic process behavior of extrusion lines with an autoregressive MTL model allowing to incorporate known causalities via physics-informed state propagation. The resulting model is used as the core of a nonlinear MPC that optimizes important state variables (e.g. melt temperature) and quality outputs (e.g. diameter) simultaneously in an end-to-end manner. Existing control loops along the production line are integrated in the MTL model so that relevant set points are updated as needed. To improve robustness of the resulting model and MPC design, an adaptive error compensation is proposed allowing to address changes in the dynamics or surroundings of the extrusion line. As a result, the main contributions of this work and advantages of the proposed methodology are:

- Physics-informed state propagation through autoregressive MTL to enforce causal dependencies and capture the end-to-end process dynamics.
- Training on an extended forecasting horizon to increase regularization, minimize accumulating prediction errors and improve robustness against measurement noise.
- End-to-end optimization of extrusion lines via MPC to enable closed-loop control and improve product quality while minimizing material and energy consumption.
- Hierarchical control design to support global optimization while distributing computational efforts and allowing fast responses to disturbances.
- Adaptive error compensation to efficiently address drifting process behavior improving the robustness of the control design.
- Validation on two actual polymer extrusion lines and empirical analyses of runtime and stability to ensure industrial applicability.

The remaining part of this paper is structured as follows. Section II describes the main process steps of two extrusion lines used for validation purposes. Section III elaborates on the proposed methodology and control design. Section IV provides an overview of the experimental setup. Section V and VI contain simulation and experimental results. Section VII includes a discussion of the results, limitations and future research directions whereas Section VIII concludes this paper.

II. PROCESS DESCRIPTION

Two PVC pipe extrusion lines are used for developing and validating the methodology of nonlinear MPC based on an autoregressive end-to-end MTL model. On the one hand, a continuous bi-oriented PVC (OPVC) production line with nine cooling and conditioning steps, one drawing section for pipe stretching and several inline quality measurements produces pipes with a diameter between 110 and 400 mm (see

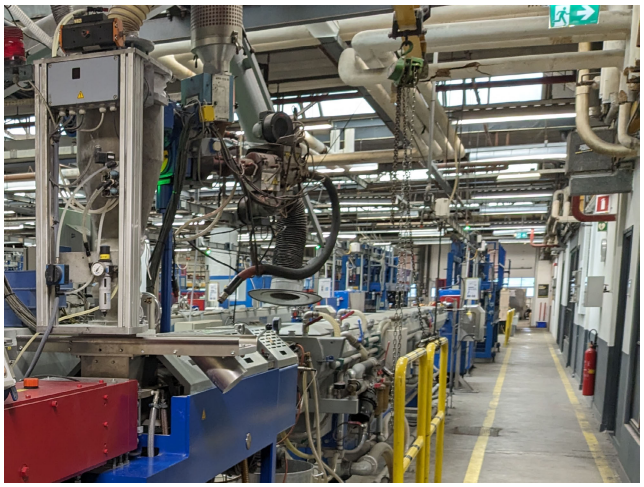


FIGURE 3. Setup of the conventional extrusion line (electro pipes) with feeding system, extruder, die head and cooling zones (left to right).

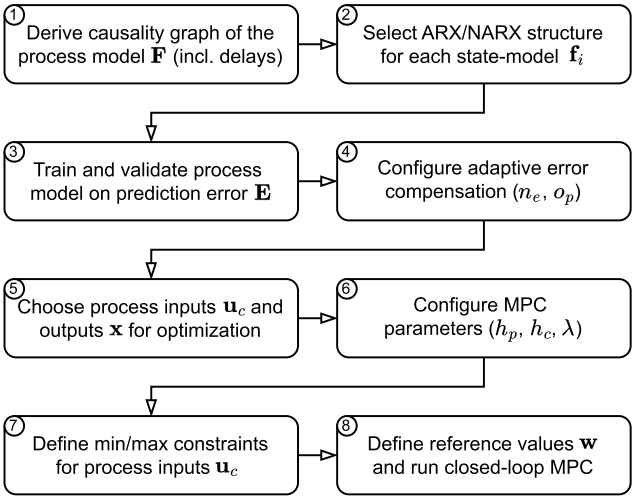


FIGURE 4. Overview of the main methodological steps to implement the proposed control design.

Fig. 1 and Fig. 2). In total, the OPVC line is approximately 55 m long with a lead time of 15 to 100 minutes from material input to the pipe cutter at the end of the line depending on the pipe dimension. The resulting pipes are characterized by oriented polymer chains in axial and circumferential direction, achieved through stretching the diameter and length, which results in improved mechanical properties, such as tensile strength and impact resistance.

On the other hand, a conventional extrusion line with three cooling and conditioning steps produces pipes used in electro installations with diameters between 16 and 21 mm (see Fig. 3). This electro line measures 25 m in length with a lead time of approximately 1 minute from material input to the pipe cutter. The conceptional setup of both lines is depicted in Fig. 1, whereas both use a twin-screw extruder which produces a melt that is pushed through a die head to form the initial (preform) pipe. In the case of the OPVC line, an additional post-processing step is built into the line which stretches the pipe in diameter and length. Furthermore, the two lines differ in terms of their material feeding system. While the OPVC line includes a gravimetric system with a dedicated feeding screw, the electro line relies purely on gravitational feeding introducing additional variations in material supply.

In both lines, influential factors, such as fluctuating properties of raw materials, ambient temperature and a strong temperature dependency of PVC lead to complex processes with dynamic, nonlinear behavior with transport delays [1], [5]. From a control perspective, all process components, for instance heating, cooling or stretching zones, have their own set points and local control loops which influence the dynamic behavior from this stage onward. These existing control loops are implemented as PID or on-off controller. Numerous sensors are placed along the extrusion lines and measure process states as well as quality metrics (e.g. wall thickness and diameter). As a result, set points, local control loops and process states must be considered in an end-to-end MPC design leading to a hierarchical architecture where the MPC determines the set points of existing controllers.

III. METHODOLOGY

The proposed methodology relies on an autoregressive MTL model of the entire extrusion line, an adaptive error compensation to increase robustness and a hierarchical MPC design for process optimization. The main methodological steps are summarized in Fig. 4 and described in this section.

A. END-TO-END MTL MODEL

To model the dynamic behavior of extrusion processes, an autoregressive MTL approach is employed [24]. For this purpose, known dependencies due to expert knowledge and physics, such as the effect of heating zones on the melt temperature and the propagating impact of melt temperature on the resulting wall thickness and diameter, are used to derive a causality graph of the system. This results in a model architecture representing the main sections of the process as exemplarily shown in Fig 5 (left) for three process sections of the OPVC line (Fig. 1). The resulting end-to-end model $\mathbf{F}(\cdot)$ (see Fig 5, right) relies on n_u system inputs $\mathbf{u} = [u_1, \ldots, u_{n_u}]$ and n_x states $\mathbf{x} = [x_1, \ldots, x_{n_x}]$ to predict the next time step of $\mathbf{x}$ while utilizing lagged values of order $\mathbf{l_u}$ and $\mathbf{l_x}$ respectively:

$$\hat{\mathbf{x}}(t+1) = \mathbf{F}(\mathbf{u}(t), \ldots, \mathbf{u}(t-\mathbf{l_u}), \mathbf{x}(t), \ldots, \mathbf{x}(t-\mathbf{l_x})). \quad (1)$$

Recursively feeding back $\hat{\mathbf{x}}(t+1)$ into $\mathbf{F}(\cdot)$ allows to forecast the trajectory of system states (e.g. melt temperature or pipe diameter) for an arbitrary horizon h. Each state is individually modeled and often depends on upstream states which are propagated through the model architecture as depicted by $\mathbf{f}_2(\cdot)$ and $\mathbf{f}_3(\cdot)$ allowing to represent physical dependencies along the extrusion line and to train all state-models simultaneously in accordance with the MTL approach.

Each state-model $\mathbf{f}_i(\cdot)$ covers the dynamic behavior of $n_{x_i} \geq 1$ states considered in the model-specific state vector

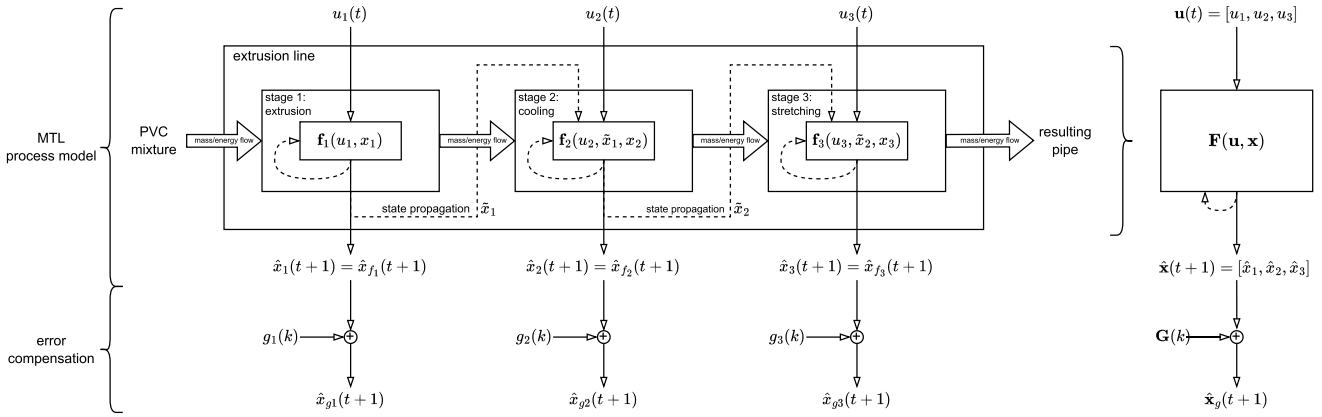


FIGURE 5. Proposed methodology to model extrusions lines from an end-to-end perspective based on (nonlinear) autoregressive state-models $\mathbf{f}_i$ and adaptive error compensation g_i while enforcing causal dependencies via physics-informed state propagation.

$\mathbf{x}_i$ representing a subset of all process states $\mathbf{x}$. To predict the trajectory of the state vector $\hat{\mathbf{x}}_i$ the state-model relies on lagged values of this state vector $\mathbf{x}_i$, relevant inputs $\mathbf{u}_i$ and propagated states from other upstream state-models $\tilde{\mathbf{x}}_i$:

$$\hat{\mathbf{x}}_i(t+1) = \mathbf{f}_i(\mathbf{x}_i, \mathbf{u}_i, \tilde{\mathbf{x}}_i) \tag{2}$$

The input orders for lagged values of $\mathbf{x}_i$, $\mathbf{u}_i$ and $\tilde{\mathbf{x}}_i$ are represented by one scalar l_i and two state-model specific vectors – $\mathbf{l}_{\mathbf{u}_i}$ and $\mathbf{l}_{\tilde{\mathbf{x}}_i}$ – individually defined for each state-model $\mathbf{f}_i$. Combining all input orders on a process level leads to a vector of orders $\mathbf{l}_f$,

$$\mathbf{l}_f = [l_1, \ldots, l_{n_f}] \in \mathbb{N}^{n_f}, \tag{3}$$

and two adjacency matrices – A_u, $A_{\tilde{x}}$ – containing the orders of exogenous $\mathbf{u}$ and propagated $\tilde{\mathbf{x}}$ inputs respectively determining the causality graph of the system:

$$A_u = \begin{bmatrix} l_{u_1,1} & \cdots & l_{u_1,n_f} \\ \vdots & l_{u_p,i} & \vdots \\ l_{u_{n_u},1} & \cdots & l_{u_{n_u},n_f} \end{bmatrix} \in \mathbb{N}^{n_u \times n_f}, \tag{4}$$

$$A_{\tilde{x}} = \begin{bmatrix} l_{\tilde{x}_1,1} & \cdots & l_{\tilde{x}_1,n_f} \\ \vdots & l_{\tilde{x}_p,i} & \vdots \\ l_{\tilde{x}_{n_x},1} & \cdots & l_{\tilde{x}_{n_x},n_f} \end{bmatrix} \in \mathbb{N}^{n_x \times n_f}. \tag{5}$$

For each state-model $\mathbf{f}_i$ an individual model structure is selected, ranging from linear ARX models with $\alpha_{j,k}$, $\beta_{j,k}$ and $\gamma_{j,k}$ as parameters:

$$\begin{aligned} \hat{x}_{i,j}(t+1) = \mathbf{f}_i(\cdot) = &\sum_{j=1}^{n_{u_i}} \sum_{k=0}^{\mathbf{l}_{\mathbf{u}_i,j}} \alpha_{j,k}\, u_{i,j}(t-k) \\ &+ \sum_{j=1}^{n_{\tilde{x}_i}} \sum_{k=0}^{\mathbf{l}_{\tilde{\mathbf{x}}_i,j}} \beta_{j,k}\, \tilde{x}_{i,j}(t-k) \\ &+ \sum_{j=1}^{n_{x_i}} \sum_{k=0}^{l_i} \gamma_{j,k}\, x_{i,j}(t-k), \end{aligned} \tag{6}$$

where n_{u_i}, $n_{\tilde{x}_i}$, n_{x_i} denote the number of exogenous, propagated and endogenous inputs for each state-model $\mathbf{f}_i(\cdot)$ respectively to nonlinear ARX (NARX) models of arbitrary complexity:

$$\begin{aligned} \hat{\mathbf{x}}_i(t+1) = \mathbf{f}_i(&\mathbf{x}_i(t), \ldots, \mathbf{x}_i(t-l_i), \\ &\mathbf{u}_i(t), \ldots, \mathbf{u}_i(t-\mathbf{l}_{\mathbf{u}_i}), \\ &\tilde{\mathbf{x}}_i(t), \ldots, \tilde{\mathbf{x}}_i(t-\mathbf{l}_{\tilde{\mathbf{x}}_i})). \end{aligned} \tag{7}$$

Essentially, all state-models – linear and nonlinear – are implemented as neural network and integrated into one end-to-end process model covering the entire extrusion line. This allows to use focused model structures while considering physical dependencies via state propagation. Additionally, training all state-models simultaneously improves robustness due to a strong regularization and mitigates the effect of spurious correlations. Importantly, transport delays in the system are not learned by the MTL model but explicitly considered in the state-models. For this purpose, delays are identified based on the relative position and production speed as well as a nonlinear impulse response estimation [5].

To train the overall process model with all its state-models, a three-dimensional forecasting tensor $\hat{\mathbf{T}} \in \mathbb{R}^{n_t \times n_x \times h_t}$ is created by recursively predicting the state trajectory of all n_x states over the train horizon h_t for n_t time instances in the training set. The resulting forecasting tensor is compared to the ground truth of actual process measurements $\mathbf{T} \in \mathbb{R}^{n_t \times n_x \times h_t}$ to calculate the prediction error $\mathbf{E}$ of the process model over the trajectory of length h_t:

$$\mathbf{E} = \mathbf{T} - \hat{\mathbf{T}} \in \mathbb{R}^{n_t \times n_x \times h_t}. \tag{8}$$

Consequently, the prediction error $\mathbf{E}$ includes propagated errors in space (between state-models) and time (along the trajectory) and is used to calculate the mean squared error (MSE) over the entire error tensor:

$$\mathrm{MSE}(\mathbf{E}) = \frac{1}{n_t n_x h_t} \sum_{j=1}^{n_t} \sum_{i=1}^{n_x} \sum_{k=1}^{h_t} \mathbf{E}_{j,i,k}^2 \,. \tag{9}$$

Based on this MSE a gradient descent algorithm optimizes the parameters θ of the process model (weights and biases of the neural networks) via backpropagating gradients through

space (model architecture) and time effectively minimizing the prediction error:

$$\theta^* = \arg\min_{\theta} \text{MSE}(\mathbf{E}). \tag{10}$$

Notably, θ contains all parameters of the process model including $\alpha_{j,k}$, $\beta_{j,k}$ and $\gamma_{j,k}$ for linear state-models. However, due to its modular architecture, the parameters of linear ARX models might be identified in advance via the computationally efficient least squares algorithm. As shown by Hartner et al., using autoregressive MTL to model entire extrusion lines allows to capture the dynamic behavior and to accurately predict the trajectory of process states for a prolonged period of 60 to 120 minutes [24].

B. ADAPTIVE ERROR COMPENSATION

To increase the robustness of the resulting process model $\mathbf{F}(\cdot)$ with optimized parameters θ, an efficient and adaptive error compensation is employed extending the applicability of the model in spite of drifting changes in extrusion lines caused by machine wear and ambient conditions. For each forecast at t_0 the matrix of past errors E_i is calculated as a subset of the three-dimensional error tensor $\mathbf{E}$ based on the available ground truth $\mathbf{T}$ and past predictions $\hat{\mathbf{T}}$:

$$E_i = \mathbf{E}_{:t_0,i,:} = \mathbf{T}_{:t_0,i,:} - \hat{\mathbf{T}}_{:t_0,i,:}\,. \tag{11}$$

From the resulting error matrix n_e most recent error trajectories $\mathbf{e}_{-a}$ are extracted as conceptionally shown in Fig. 6. The shifted values of past forecasting errors are used to form the error trajectories $\mathbf{e}_{-a} \in \mathbb{R}^{h_p}$ over the entire prediction horizon h_p. These trajectories are then used as basis to fit a polynomial function $g_i(k)$ of order o_p as depicted in Fig. 6b, effectively constituting a nonparametric approximation of the residual dynamics. To prioritize most recent error trajectories (e.g. $\mathbf{e}_{-1}$ over $\mathbf{e}_{-3}$) a linearly decreasing weight is applied during fitting the polynomial function.

As shown in Fig. 5 (bottom), the resulting polynomial functions $g_i(k)$ are used as adaptive offset for each forecasted state in the process model:

$$\hat{x}_{gi}(t+k) = \hat{x}_i(t+k) + g_i(k). \tag{12}$$

As a result, drifting behavior in actual extrusion lines are adaptively addressed, increasing the robustness of the process model and prolonging its applicability as long as the main dynamics (captured by the process model $\mathbf{F}$) remains intact and the error trajectories are stationary within the considered time frame.

C. CONTROLLER DESIGN

A nonlinear, hierarchical MPC is proposed which uses the autoregressive MTL model extended with adaptive error compensation for process predictions to optimize the operation of entire extrusion lines. In this regard, the MPC represents a global control loop focusing on strategic objectives (process efficiency and product quality) while adjusting the relevant set points $\mathbf{u}_c$ of existing controllers (PID, on-off) used to steer individual process components locally and independently, such as heating or cooling zones. For this purpose, reference values $\mathbf{w}$ are defined for n_c system states in $\mathbf{x}$ where $n_c \leq n_x$. To steer the selected number of states, relevant system inputs $\mathbf{u}_c$ of length $n_{u_c} \leq n_u$ are identified via a sensitivity analysis [24] or expert knowledge. Other states and inputs which are not selected for the optimization procedure but are included in the process model are left unchanged by the MPC. Due to its hierarchical and distributed characteristics, the proposed MPC approach allows to distribute resources and enables fast responses by local control loops while the overall process is steered and optimized by the global MPC.

To identify optimal set point changes in $\mathbf{u}_c$, the process model $\mathbf{F}(\mathbf{u}, \mathbf{x})$ is recursively applied to predict the trajectory of all states over the prediction horizon h_p. Additionally, an adaptive offset based on past error trajectories is used to compensate for drifting behaviors in the extrusion line. The resulting state trajectories $\hat{x}_{gi}(t+k)$ are compared to the corresponding reference value w_i and combined with a penalty term for set point changes to form the loss function L for process optimization:

$$L(\mathbf{u_c}) = \sum_{k=1}^{h_p}\sum_{i=1}^{n_w}[w_i - \hat{x}_{gi}(t+k)]^2 + \sum_{k=1}^{h_c}\sum_{i=1}^{n_{u_c}}[\lambda \Delta u_{ci}(t+k)]^2, \tag{13}$$

which is used by the MPC to steer the process effectively. To ensure efficient control sequences, a penalty term is added for each set point change,

$$\Delta u_{ci}(t+k) = u_{ci}(t+k-1) - u_{ci}(t+k)), \tag{14}$$

in considered system inputs $\mathbf{u}_c$ over the control horizon $h_c \leq h_p$ and scaled via the penalty factor λ. The resulting loss function is used in an iterative optimization procedure to determine the optimal set points U_c over h_c:

$$U_c^* = \arg\min_{U_c} L(U_c), \tag{15}$$

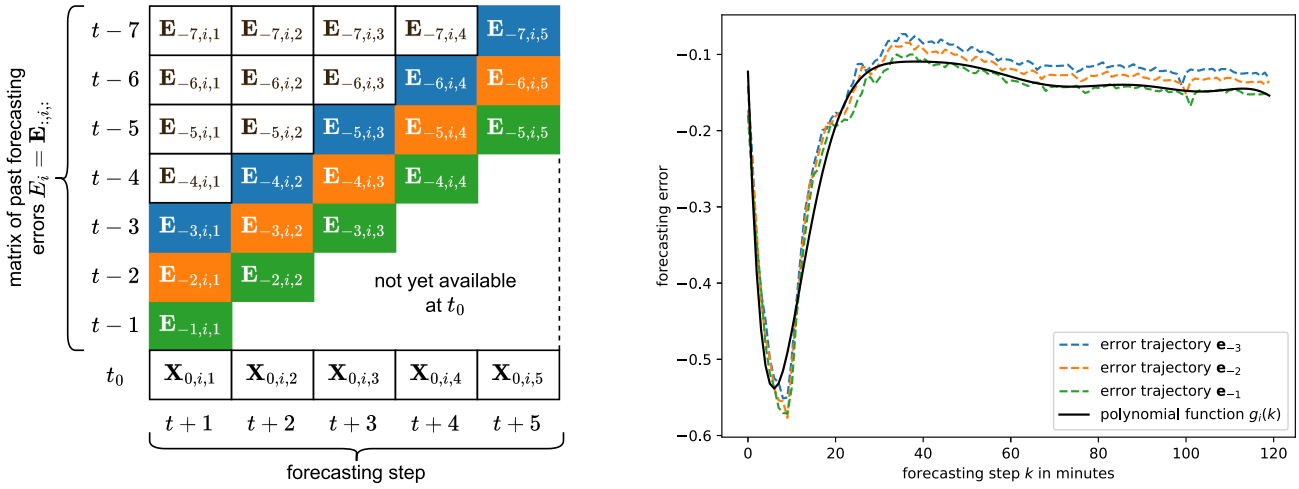
subject to

$$u_{ci_{min}} \leq u_{ci} \leq u_{ci_{max}}, \tag{16}$$

where the lower $u_{ci_{min}}$ and upper boundaries $u_{ci_{max}}$ for each set point u_{ci} in $\mathbf{u}_c$ are enforced as hard constraints and each column vector in U_c represents the values of all relevant set points $\mathbf{u}_c$ at time $t+k$:

$$U_c = \begin{bmatrix} | & | & | \\ \mathbf{u}_c(t+1) & \dots & \mathbf{u}_c(t+h_c) \\ | & | & | \end{bmatrix} \in \mathbb{R}^{n_{u_c} \times h_c} \tag{17}$$

To solve this nonlinear minimization problem iteratively, the trust region reflective (TRR) algorithm is used due to its native consideration of parameter constraints [39] ensuring safety and feasibility provided that the chosen input constraints are not mutually exclusive. Additionally, to increase



(a) Conceptional error matrix E_i

(b) Example of polynomial function $g_i(k)$

FIGURE 6. **Conceptional example for an error matrix E_i (left) with selected error trajectories (green, orange, blue) and the resulting polynomial function $g_i(k)$ fitted on these trajectories for a prediction horizon of $h_p = 120$ (right).**

the conservativeness of the optimization procedure, the constraints can be narrowed or the penalty factor λ increased. While the optimal set points in $\mathbf{u}_c$ are identified over the entire control horizon h_c, only the values at the first time step $\mathbf{u}_c(t+1)$ are applied to the process. The procedure of identifying and applying optimal set points is then repeatedly performed at each sampling instance to steer the process accordingly and enable closed-loop optimization.

IV. EXPERIMENTAL SETUP

For developing and validating the proposed methodology of nonlinear MPC, two extrusion lines are used – a large diameter line for bi-oriented PVC pipes (OPVC) and a small diameter line for pipes used in electrical installations (see Fig. 2 and Fig. 3). For the OPVC line two autoregressive MTL models were trained and used for the MPC validation (see Tab. 1). On the one hand, a model (OPVC-1) solely focused on the extruder as first part of an extrusion line was trained with $n_u = 4$ and $n_x = 5$ system inputs and states respectively where 4 heating zones with local control loops as well as the resulting melt temperature are considered. The OPVC-1 process model with its inner structure based on the causality graph of the extruder requires $n_f = 5$ state-models and 3353 parameters. On the other hand, a second model (OPVC-2) covers the entire extrusion line in an end-to-end manner relying on $n_u = 24$ inputs and $n_x = 29$ states to capture and forecast the nonlinear system dynamics with $n_f = 15$ state-models and 13987 parameters. Consequently, the OPVC-2 model extends the OPVC-1 model and includes causal dependencies along the extrusion line, for instance the effect of melt temperature on wall thickness and diameter or the impact of intermediate heating zones on the drawing process. Additionally, to ease the training procedure, a known day and night oscillation was explicitly included as sinusoidal

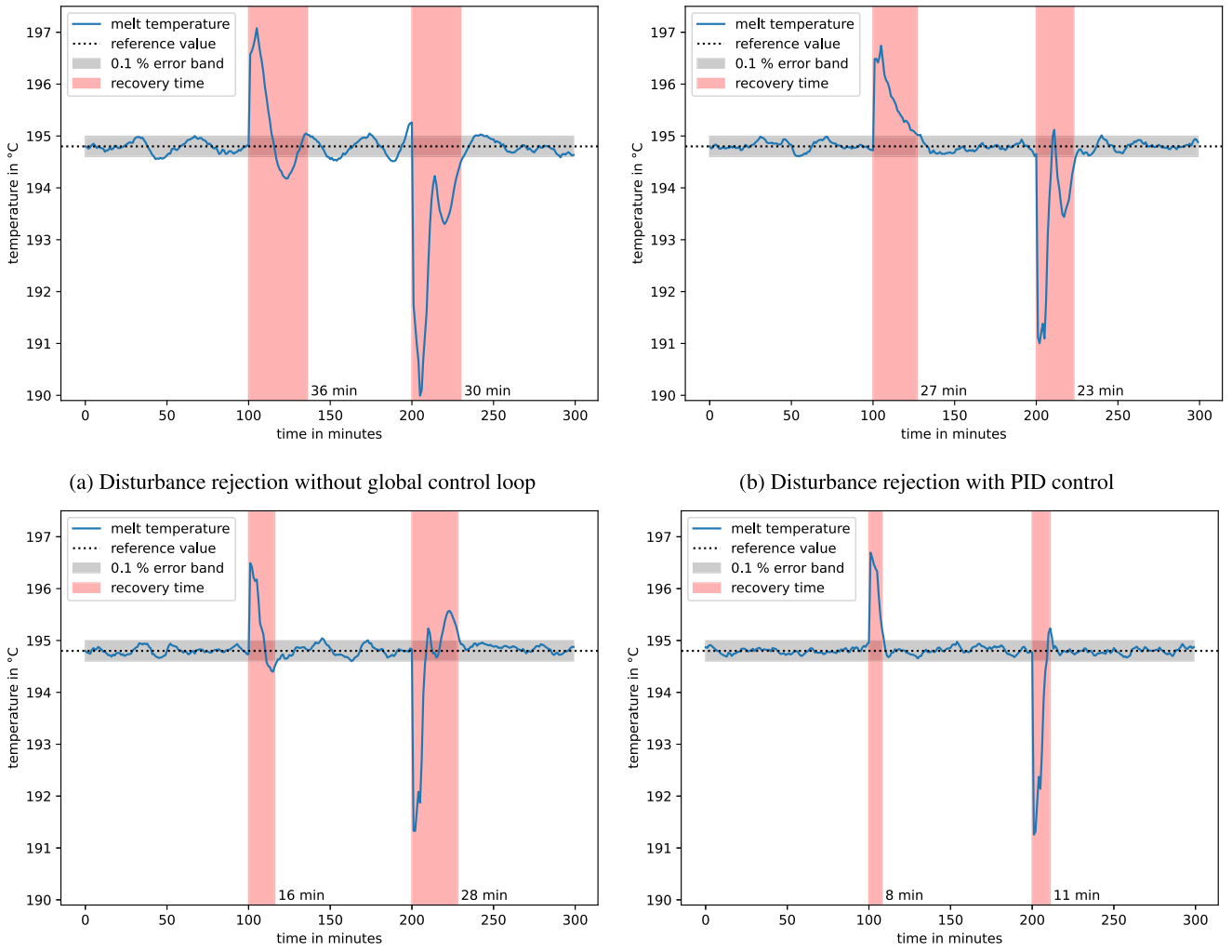
TABLE 1. **Process models used for validation.**

model name	parameters	n_u	n_x	n_f	h_t	n_{u_c}	n_w	λ
OPVC-1	3353	4	5	5	32	2	1	0.5
OPVC-2	13987	24	29	15	16	1	1	0.5
ELECTRO	2233	5	9	9	16	3	2	0.5

signal with a period of 24 hours. The train horizon h_t was experimentally selected as 16 and 32 for training the OPVC-1 and OPVC-2 process model. For the electro line, a process model with $n_u = 5$ inputs and $n_x = 9$ states was used where the inner structure consists of $n_f = 9$ state-models (see Tab. 1) allowing to forecast the trajectory of relevant heating zones, throughput, melt temperature and pressure. Similar to the OPVC-1 model, the causal dependencies include the effect of the heating zones on the melt temperature while also considering varying screw speeds and the impact on throughput and pressure. In total 2233 parameters were optimized during the training procedure based on a train horizon of $h_t = 16$.

All models (OPVC-1, OPVC-2, ELECTRO) were trained on data collected over a production period of approximately 14 days sampled at an interval of 1 minute. Additionally, the leaky rectified linear unit was used to represent the nonlinearity of the system while the Adam [40] gradient descent algorithm with a cosine learning rate [41] and a batch size of 512 was used to optimize the weights of the models. For this purpose the initial learning rate of 10^{-3} was iteratively lowered to 10^{-5} over 500 epochs. Additionally, the data was normalized via min-max scaling to improve the convergence during training.

While validating the MPC on the extrusion lines, adaptive error compensation was used for the OPVC-2 and ELECTRO models. The polynomial function $g_i(k)$ for approximating the error trajectory is of order $o_p = 10$ and fitted on



(a) Disturbance rejection without global control loop

(b) Disturbance rejection with PID control

(c) Disturbance rejection with LQI control

(d) Disturbance rejection with proposed nonlinear MPC

FIGURE 7. Comparative analysis of disturbance rejection capabilities (red area) based on impulse disturbances (+1.8 °C at $t = 100$, −3.6 °C at $t = 200$) applied to the melt temperature while different control approaches are active (a-d).

$n_e = 10$ most recent trajectories. Due to its restricted scope, the error compensation is not required for the OPVC-1 model. The prediction and control horizon were defined as $h_p = 60$ and $h_c = 1$ respectively for all process models while the boundaries for $\mathbf{u}_c$ are equal to the minimum and maximum values encountered in the training data which ensures feasibility and safety. Moreover, the penalty factor for set point changes was defined as $\lambda = 0.5$.

The number (n_{u_c}, n_w) and type of manipulated as well as controlled variables were chosen to allow a broad range of experiments. The OPVC-1 model is used to manipulate two heating zones in the extruder to control the melt temperature. While the OPVC-2 model captures the entire extrusion line to forecast the trajectory, it is only used to control the diameter as main quality characteristic via the air pressure at the drawing process (see Fig. 1). The ELECTRO model relies on two heating zones and the screw speed as manipulated variables to optimize the melt temperature and throughput simultaneously.

V. SIMULATION RESULTS

To investigate important properties of the proposed control design, a comparative analysis of the disturbance rejection capabilities, an empirical stability analysis and a runtime analysis are conducted and described in this section.

A. ANALYSIS OF DISTURBANCE REJECTION

During normal production periods a controller is usually operated as compensator to address disturbances in the extrusion line and keep the system within its tolerances. Therefore, the capabilities to reject disturbances are investigated for the proposed control design in comparison to alternative control approaches. For this purpose, the OPVC-1 process model (see Tab. 1) is used to iteratively simulate the extrusion process while Gaussian noise is added to all states at each simulation step. A signal-to-noise ratio (SNR) of 10 is used to generate the additive noise based on the standard deviation measured for each process state.

The simulated melt temperature at the extruder outlet of the OPVC line is mainly influenced by four heating zones which are individually regulated via simple on-off controllers. Consequently, the set points for the heating zones must be adjusted by a global controller to keep the melt temperature within the tolerance of $\pm$ 0.1 %. For testing purposes, two impulse disturbances with a magnitude of $+1.8\,^{\circ}\mathrm{C}$ and $-3.6\,^{\circ}\mathrm{C}$ (20 % and 40 % of the melt temperature range) are applied to the melt temperature at $t = 100$ and $t = 200$ respectively (see Fig. 7). These disturbances represent actual scenarios often encountered in polymer extrusion, such as issues in the feeding system, fluctuating degrees of moisture in the powder or mechanical issues in the extruder.

To evaluate the proposed control design, the prediction and control horizon of the nonlinear MPC are defined as $h_p = h_c = 15$ while the OPVC-1 model is used to predict the state trajectories based on the simulated data with added noise. Additionally, three alternative control approaches are investigated. First, the current setup of the OPVC line is used as baseline, where all heating zones are individually steered by on-off controllers without any global control loop. Second, a conventional PID controller ($K_p = 0.5$, $K_d = 1$) is implemented where the melt temperature represents the output regulated by adjusting the temperature set point of the last heating zone in the extruder. Third, a linear-quadratic-integral (LQI) controller is employed as an alternative model-based approach. The underlying state space model is derived from the OPVC-1 model via linearization around the operating point. The set points of all four heating zones are used as system inputs to steer the melt temperature while the weights of the cost matrices are selected to ensure that the inputs remain feasible.

The results of the comparative analysis are shown in Fig. 7 where the simulated melt temperature with added noise is represented by the blue line and the gray area shows the 0.1 % error band around the reference value of 194.8 °C. The recovery time is defined as settling time after an impulse disturbance occurs and is highlighted as red area. As depicted in Fig. 7a, without any additional global control loop, the melt temperature oscillates around the reference value and occasionally violates the tolerances. After an impulse disturbance occurs the system requires 30 to 36 minutes to recover. Using a PID controller (Fig. 7b) to steer the melt temperature reduces the recovery time to 23-27 minutes while damping the system oscillations in general. Similarly, implementing a LQI controller reduces the system oscillations while efficiently addressing smaller disturbances as shown in Fig. 7c at $t = 100$ (recovery time of 16 minutes). However, large disturbances lead to notably longer recovery times (28 minutes) with significant overshoots caused by nonlinearities outside the normal operating range. In contrast to that, relying on a nonlinear MPC based on an autoregressive MTL model allows to significantly reduce the recovery time to 8-11 minutes while keeping the melt temperature close to the reference value during normal periods as demonstrated in Fig. 7d. Additionally, while this comparative analysis only focused on the extruder outlet, the proposed methodology based on an end-to-end process model allows to automatically adjust downstream process components (e.g. cooling or drawing sections) to compensate for disturbances ensuring overall process efficiency.

B. EMPIRICAL STABILITY ANALYSIS

To evaluate the empirical stability of the proposed control design, a Monte Carlo simulation with 100 repetitions is performed. While the reference value for the melt temperature remains at 194.8 °C, each repetition starts with randomized initial conditions where system inputs and states are drawn uniformly from the range of feasible values. During the simulation based on the OPVC-1 model, Gaussian noise with a SNR of 10 is added to all states in each simulation step. As shown in Fig. 8 for all 100 melt temperature trajectories, the closed-loop MPC with a prediction and control horizon of $h_p = h_c = 15$ reaches the reference value within the 0.1 % error band for all repetitions after 15-20 minutes. Even in rare cases where the melt temperature drops below 190 °C due to extreme initial conditions, the MPC is able to stabilize the system successfully.

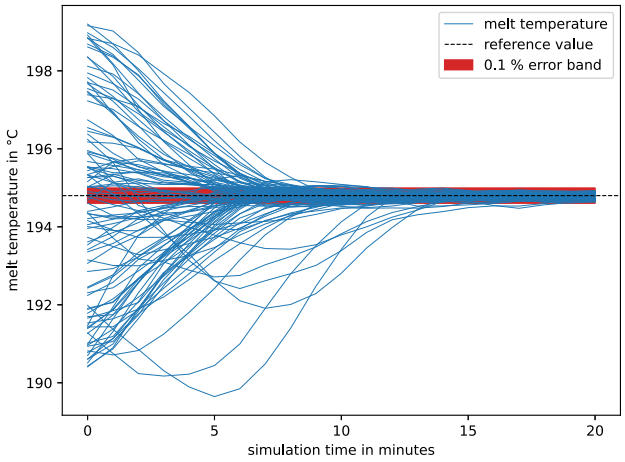


FIGURE 8. Empirical stability analysis based on a Monte Carlo simulation with 100 repetitions and randomized initial conditions.

C. EMPIRICAL RUNTIME ANALYSIS

To ensure the proposed control approach is applicable for online usage in actual production lines, the expected runtime of each component (individual prediction, error compensation, optimization) is empirically analyzed in this section. For this purpose, the main drivers for prediction and optimization time (model complexity, prediction horizon) are investigated based on the OPVC-1 and OPVC-2 models with 3353 and 13987 parameters respectively (see Tab. 1). Since the adaptive error compensation method is computationally independent from the preceding process model, the number of considered past error trajectories n_e is analyzed instead. All measurements were repeated 100 times and performed on a notebook with Windows 11, 32 GB memory and an i7

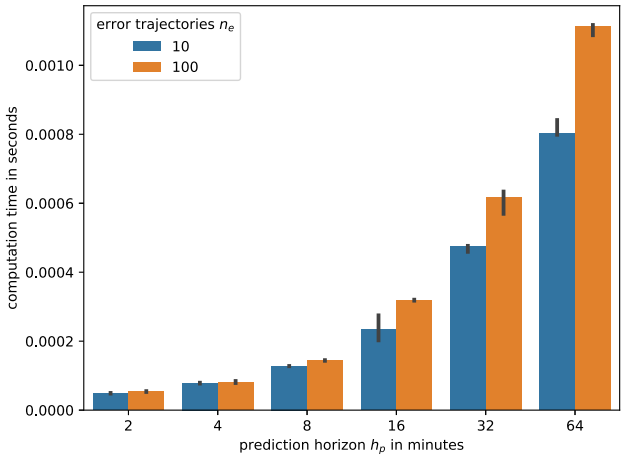


FIGURE 9. Median runtime including the 95 % confidence interval for computing the adaptive error compensation offsets over an increasing prediction horizon h_p and two levels of considered error trajectories n_e.

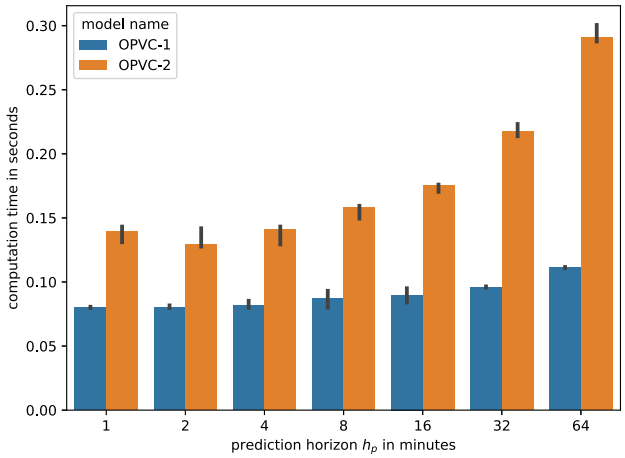


FIGURE 10. Median runtime including the 95 % confidence interval for predicting state trajectories based on two process models over an increasing prediction horizon h_p.

1365U CPU. Additionally, Python 3.10.13 with TensorFlow 2.9.1, NumPy 1.23.4 and SciPy 1.13.1 was used for the implementation.

As shown in Fig. 9 for computing the adaptive error compensation offset, the median computation time increases steadily with the prediction horizon h_p, while the number of considered past error trajectories n_e has less influence on the runtime. The results for $h_p = 1$ are not included in this analysis because the error matrix E_i does not contain any trajectories in this case.

Similarly, the results for individual predictions (see Fig. 10) and full control steps (see Fig. 11) show a strong dependence on the prediction horizon h_p. Moreover, the model complexity significantly influences the median runtime. In this regard, a control step includes the iterative optimization procedure (repeated state predictions) to identify new process inputs to minimize the cost function. Even though the error compensation offset must be computed once per control step for all relevant states separately, the overall runtime (max. 1 millisecond per state, see Fig. 9) is negligible in comparison to the control step runtime (>1 seconds).

The results of the empirical runtime analysis show, that considering a sampling interval of 1 minute, the median runtime of 5-6 seconds for one full control step is sufficient to ensure its online applicability in actual extrusion lines.

VI. EXPERIMENTAL RESULTS

In addition to the simulation results, the effectiveness of the proposed methodology is industrially validated in the context of process optimization and changing operating points. For this purpose, both extrusion lines are used to demonstrate the general applicability in extrusion processes.

A. SMALL DIAMETER ELECTRO LINE

To investigate the capabilities of the proposed approach for extrusion lines in general, a conventional electro line is used since it contains the basic process steps often found in

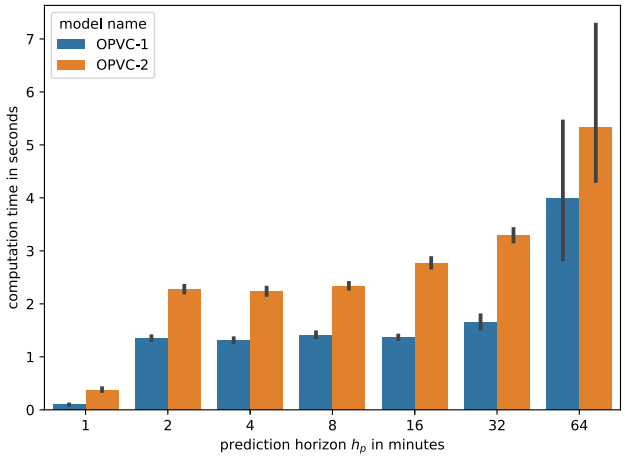


FIGURE 11. Median runtime including the 95 % confidence interval for a full control step (iterative optimization to identify process inputs) based on two process models over an increasing prediction horizon h_p.

extrusion processes. The aforementioned ELECTRO model is employed as nonlinear autoregressive model by the MPC to steer the process in a closed-loop manner. As shown by the examples in Fig. 12, the MPC with an update interval of 1 minute (equal to the sampling interval) ensures that the melt temperature closely follows the reference values (dotted lines) which are repeatedly changed by the operator. To control the melt temperature three relevant set points (two heating zones and the screw speed) are used as $\mathbf{u}_c$ by the MPC while the lower and upper boundaries are defined as minimum and maximum values encountered in the training data.

The vertical dashed lines indicate an arbitrary point in time where the prediction of the state trajectory starts for demonstration purposes, whereas the actual forecast is iteratively performed as receding horizon in closed-loop control. Even though the actual melt temperature signal is obscured by short-term fluctuations and noise, the underlying model accurately captures the main trend due to its strong

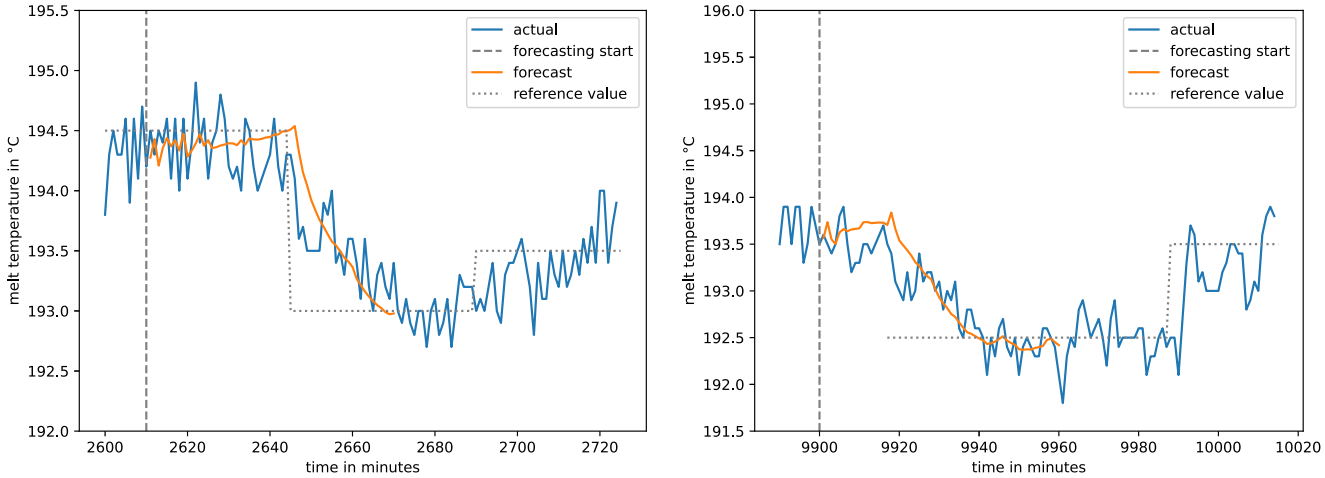


FIGURE 12. Examples of closed-loop control via the proposed MPC based on MTL predictions (orange line) for repeatedly changing the reference value of the melt temperature in the electro pipe production line.

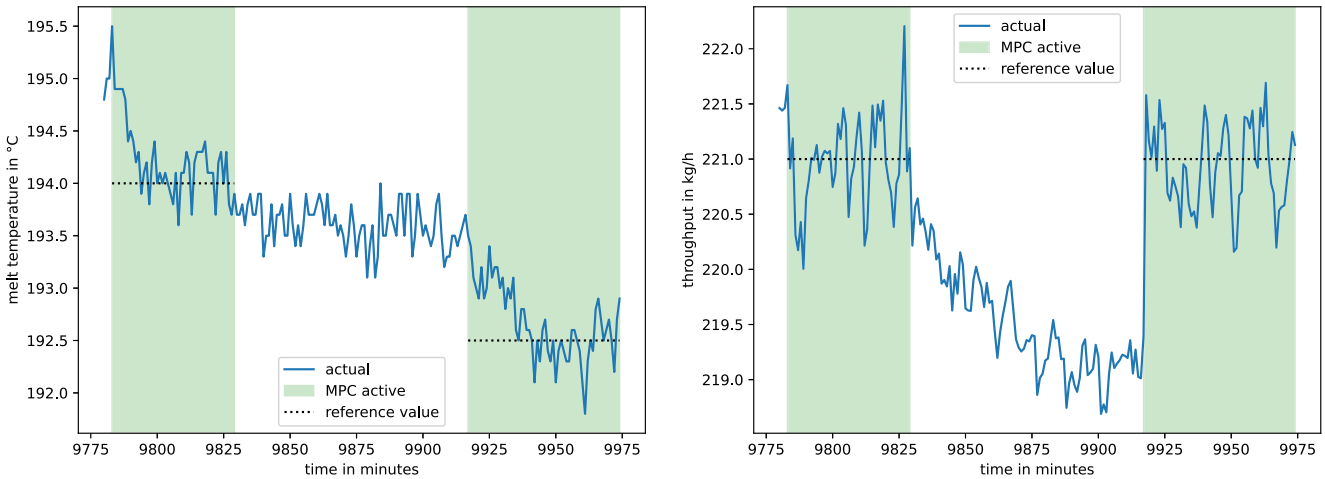


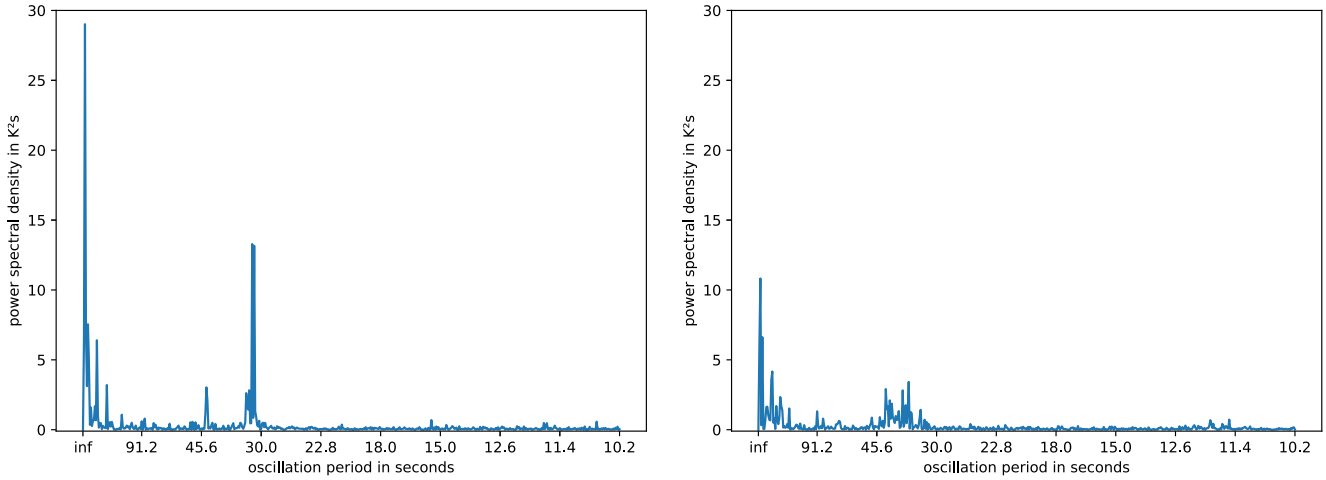
FIGURE 13. Examples of closed-loop control (green area) via the proposed MPC for simultaneously optimizing melt temperature (left) and throughput (right).

regularization enforced by the MTL approach as well as the extended forecasting horizon during the training process.

To analyze the effectiveness of the MPC approach when multiple process states are simultaneously controlled, the ELECTRO model was used to steer melt temperature and throughput via manipulating the same set points $\mathbf{u}_c$ as before. As shown in Fig. 13, the MPC was switched on twice with different reference values for the melt temperature while the throughput was kept on the same level (dotted lines).

The period between the closed-loop modes represents normal production without any interference. In this period of normal production the melt temperature stabilizes itself between 193.5 °C and 194 °C. However, the throughput changes significantly due to the system dynamics showing the need for a global control loop. Interestingly, increasing the throughput again for the second period of closed-loop control (around $t = 9920$) requires that the screw speed is raised inducing more energy into the extruder barrel. Normally, that would lead to an increase in melt temperature as well. Nevertheless, due to the MPC it is possible to reduce the melt temperature to a lower reference value (192.5 °C) via manipulating the set points of the heating zones accordingly. As a result, the melt temperature and throughput closely follow the reference values allowing a counterintuitive control of the main process states, while the short-term fluctuations around the target throughput are mainly caused by measurement delays and technical constraints in the existing screw speed controller.

To determine the effect of the proposed MPC on process oscillations, the power spectral density of the melt temperature is analyzed through a periodogram. For this purpose, the signal is sampled at an interval of 5 instead of 60 seconds allowing improved insights into the signal. Additionally, since some dominant oscillations in the extruder system are expected to last several minutes, the spectral density is shown in bins related to period length in seconds instead of frequency in Hz. As depicted in Fig. 14a, the baseline signal captured during normal production shows multiple dominant oscillations, most notably with a period of 45 minutes and 30 seconds. After enabling closed-loop control (Fig. 14b)



(a) Baseline periodogram

(b) Closed-loop periodogram

FIGURE 14. **Periodogram of the melt temperature in the electro pipe production line during normal operating conditions without any global control loop (baseline, left) and closed-loop mode via the proposed MPC (right).**

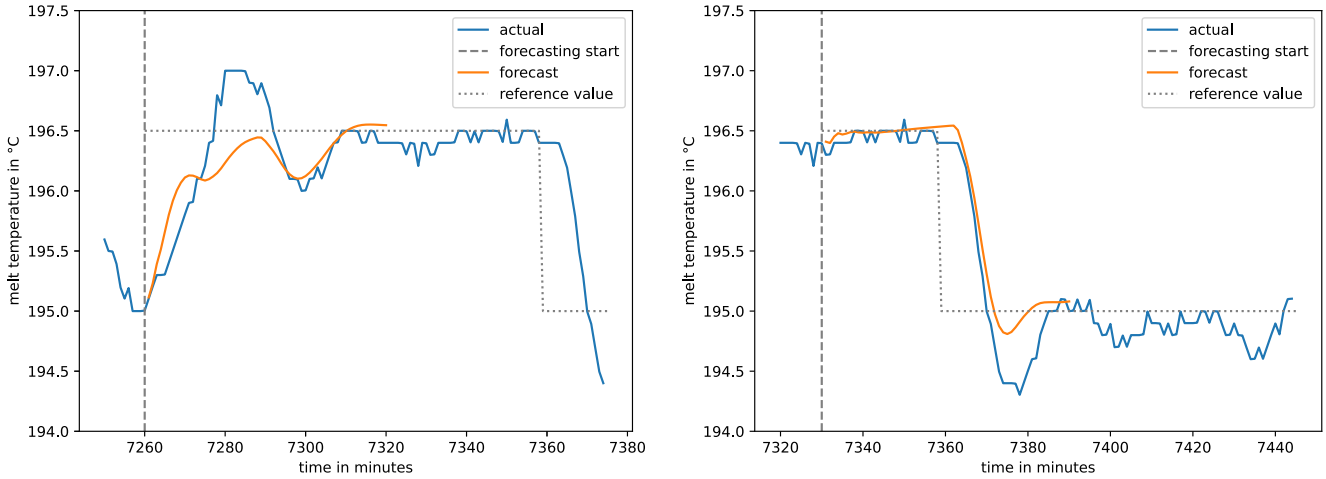


FIGURE 15. **Examples of closed-loop control (human-in-the-loop) via the proposed MPC based on MTL predictions (orange line) for repeatedly changing the reference value of the melt temperature in the bi-oriented pipe production line.**

with the proposed MPC, the 45 minute oscillation, representing a limit cycle of the system, is significantly damped by 62 % from 29 to 11 K^2s while the 30 seconds oscillation is almost entirely removed showing that process oscillations are effectively addressed.

B. LARGE DIAMETER OPVC LINE

Due to the setup of the OPVC line, an automated control loop cannot be implemented directly, however, a human-in-the-loop approach is a viable option to analyze the effectiveness of the proposed MPC solution. The human operator blindly executes set point changes without interference guaranteeing unbiased results. Nevertheless, to ensure that long-term tests over several hours can be performed, the update interval of corresponding set points is defined as 10 to 15 minutes (instead of the sampling interval of 1 minute). Additionally, to show the advantage of a long prediction horizon $h_p = 60$ multiple tests are conducted for the pipe diameter, where the set points are changed only once by the operator.

Due to its significant impact on the entire extrusion line, the extruder is investigated first based on the OPVC-1 model. For this purpose, the last two heating zones in the extruder are defined as relevant inputs $\mathbf{u}_c$ for controlling the melt temperature measured at the extruder outlet via a thermocouple sensor. As shown in Fig. 15 (left), the reference value for the melt temperature was set to 196.5 °C at $t = 7260$ and recommended set point changes were immediately applied by the operator leading to a sharp increase in the melt temperature. Even though the overshoot was not captured by the forecast of the MTL model at $t = 7260$, the settling time was accurately predicted and the process settled at the reference value of 196.5 °C. Additionally, the step response at $t = 7359$ in Fig. 15 (right) shows a similar result. However, the process settles faster at the new reference value of 195 °C

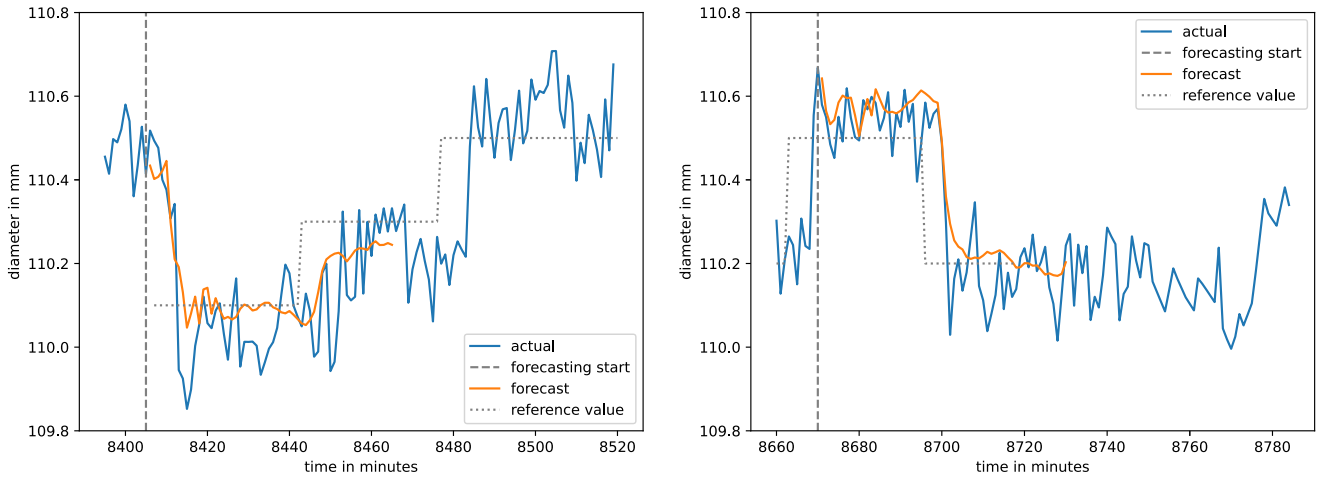


FIGURE 16. Examples of closed-loop control (human-in-the-loop) via the proposed MPC based on MTL predictions (orange line) for repeatedly changing the reference value of the resulting pipe diameter in the bi-oriented pipe production line.

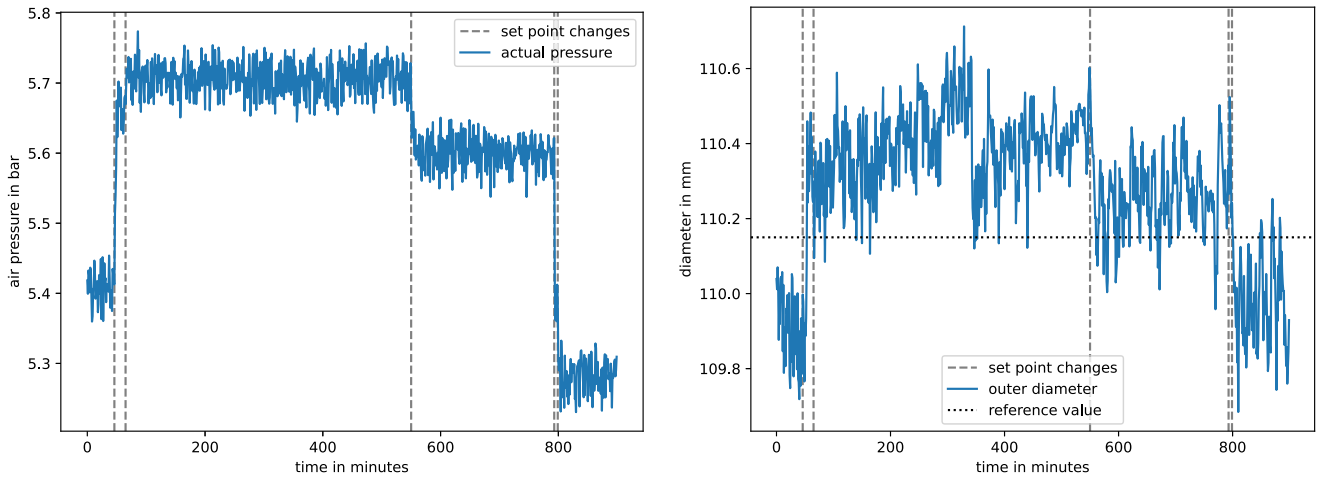


FIGURE 17. Baseline of 15 hours of normal production (without model-based recommendations) where the operator manually adjusts the air pressure of the stretching unit (left) to steer the outer diameter (right) based on experience alone.

and the overall trajectory of the melt temperature is accurately predicted by the OPVC-1 model starting from $t = 7330$.

In addition to the melt temperature, the OPVC-2 model is used to change operating points of the resulting outer diameter while considering the dynamic behavior of the entire extrusion line. For this purpose, the air pressure at the drawing mandrel (see Fig. 1) is used as relevant system input $\mathbf{u}_c$ while the penalty factor for set point changes is defined as $\lambda = 0.5$. As shown by repeatedly performed step tests of the outer diameter in Fig. 16, the reference values are closely followed due to accurate predictions of the trajectory. Furthermore, due to a strong regularization implicitly enforced during training the model via the MTL approach over an extended forecasting horizon, the effect of noisy measurements is mitigated. As depicted in Fig. 16, the reference values were changed by the operator in an ad hoc manner, therefore, the transport delays in the system cannot be avoided.

To evaluate the effectiveness of the proposed methodology in comparison to a baseline, a normal production period of 15 hours is shown in Fig. 17. On the left side, the actual air pressure at the drawing mandrel is depicted, which is regularly changed by the operator to adjust the outer diameter of the resulting product (right). The reference value for the diameter was 110.15 as indicated by the dotted line in Fig. 17 (right). In this period, the operator tried to optimize the diameter five times (dashed lines) while it is difficult to determine a proper value of the air pressure on experience alone without any model-based recommendation. This often results in an over- or underestimation of the effect of set point changes (air pressure) leading to long periods, where the actual diameter is far from the targeted reference value.

TABLE 2. Estimated process improvements with regard to variance, and 95 % production tolerances for different cutoff periods in minutes in comparison to the baseline.

operating mode	cutoff period	variance	95 % tolerance
unsupported baseline	-	0.037	± 0.375 mm
human-in-the-loop	120 min	0.013	± 0.220 mm
	90 min	0.011	± 0.204 mm
⇳	60 min	0.009	± 0.190 mm
	30 min	0.008	± 0.171 mm
closed-loop control	15 min	0.005	± 0.139 mm

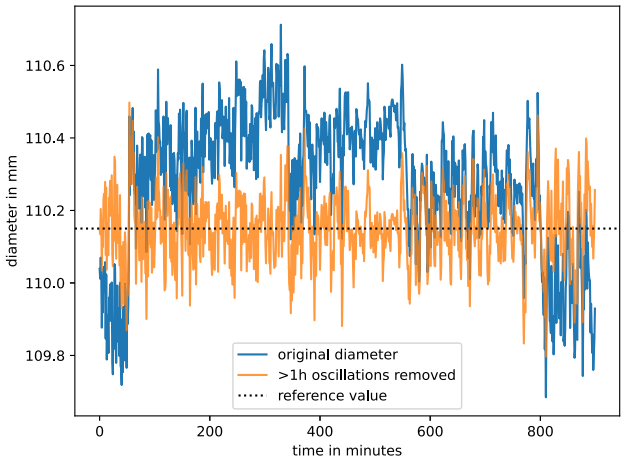


FIGURE 18. Estimated improvements in terms of reduced diameter variations due to removed oscillations with a period of 1 hour or more based on the proposed control design.

This becomes evident when the periods from $t = 46$ to $t = 65$ and $t = 793$ to $t = 799$ are considered, where the operator changed the set point again after the initial change did not show the expected result. As a consequence, based on the results in Fig. 16, significant improvements are expected due to human-in-the-loop and closed-loop control facilitated by the proposed nonlinear MPC.

To calculate an estimate of expected process improvements, it is assumed that oscillations with a period longer than a certain threshold are reliably removed by the MPC. For simulation purposes, a high-pass filter is applied to the original diameter signal to remove oscillations which are below a certain cutoff frequency, as shown in Fig. 18 with a threshold period of 1 hour (1/3600 Hz). As a result, a large portion of diameter deviations is removed and the product quality is substantially improved. Depending on the specific organizational and technical implementation (human-in-the-loop or closed-loop), the proposed MPC removes or damps oscillations with shorter or longer periods (cutoff frequency is increased or decreased) leading to narrower or wider tolerances respectively. Consequently, a range of cutoff periods is investigated in Tab. 2 showing the impact on diameter variance and therefore, on the achievable 95 % production tolerance as a measure of process variation. As clearly demonstrated by cutoff periods $\geq$ 60 minutes, removing slow oscillations already leads to significant improvements, reducing the process variation by 41 % to 49 % when compared to the baseline production. Implementing a fully automated closed-loop control is also expected to remove and damp oscillations with shorter periods, thus reducing the process variation by up to 63 % when a cutoff period of 15 minutes is assumed. As a result, significant improvements in terms of process optimization, material consumption and product quality are expected when the proposed method of nonlinear MPC is employed for polymer extrusion lines.

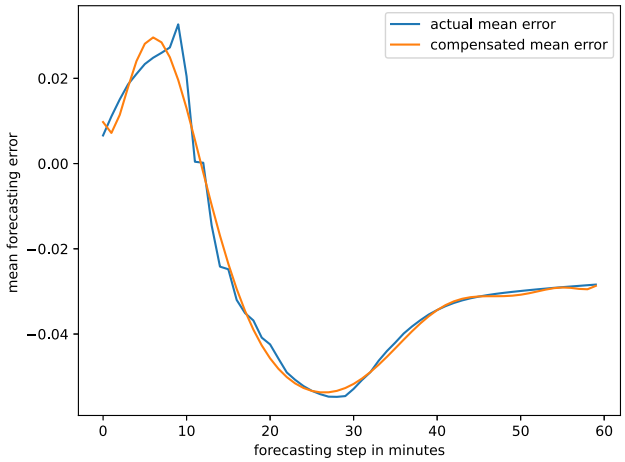


FIGURE 19. Comparison of actual error trajectories (blue) and associated compensation functions (orange) averaged over several days of production for the pipe diameter.

C. EFFECT OF ADAPTIVE ERROR COMPENSATION

To assess the effect of the proposed method for adaptive error compensation, the actual error trajectory (blue) and associated adaptive compensation function $g_i(k)$ (orange) are averaged over several days of production and shown in Fig. 19 for the diameter as target variable. As depicted, the actual mean error trajectory increases in the first 10 minutes ($k = 10$) of forecasting, drops sharply afterwards and settles at approximately -0.028. Since the averaged results of the polynomial compensation function (orange line), which is adaptively fitted onto the most recent error trajectories, align well with the actual mean trajectory (blue line), the forecasting error can be significantly reduced. Consequently, the model robustness is increased and drifting behavior in the extrusion line is effectively addressed.

Additionally, Fig. 20 includes an arbitrary example at time t from the underlying production run. As illustrated, the $n_e = 10$ most recent error trajectories (dashed lines) from $t-1$ to $t-n_e$ are used to fit the coefficients of the polynomial compensation function $g_i(k)$ (solid black line). Consequently, the main pattern of the uncompensated future error trajectory at t (solid red line) is captured well allowing to decrease the resulting forecasting error while enforcing a sufficient degree of regularization.

VII. DISCUSSION

The complexity of polymer extrusion lines with its sensitivity to ambient conditions and material inputs as well as the nonlinear and complex dynamic behavior with long transport

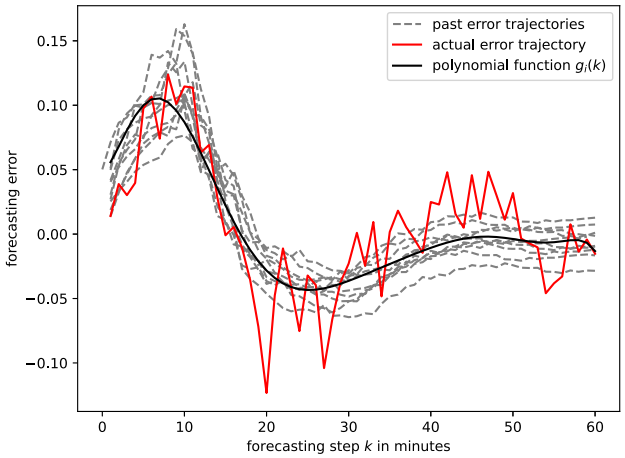


FIGURE 20. Example of the resulting adaptive error compensation function (black) fitted on $n_e = 10$ past error trajectories (gray dashed) in comparison with the actual future error trajectory (red).

delays requires an end-to-end perspective for predictive control to globally optimize the process. Since existing works in this domain either focus on optimizing individual process sections or neglect the dynamic behavior, a methodology for global optimization based on autoregressive MTL and nonlinear MPC is proposed in this work.

As shown by the results with two actual extrusion lines, where process oscillations and step responses were investigated, the underlying predictive models capture the nonlinear dynamics well and allow to accurately forecast state trajectories. Due to a strong regularization implicitly enforced due to training the MTL model on an extended forecasting horizon, the effect of noisy measurements and short-term process fluctuations are mitigated. Additionally, the robustness against process drifts, caused by machine wear and unmeasured external factors, is reliably improved via an adaptive error compensation based on past error trajectories. As a consequence, multiple process states can be simultaneously optimized while counterintuitive objectives can be achieved as demonstrated by several step response tests. Importantly, local control loops are naturally built into the autoregressive MTL model allowing to rely on existing controllers in a hierarchical control design. Furthermore, investigating the periodogram of closed-loop control revealed that dominant oscillations are successfully damped leading to an optimized process behavior. Analyzing the capabilities to reject impulse disturbances showed that the proposed control design outperforms alternative control approaches in terms of rejection time and maximum overshoot. The investigated empirical runtime with regard to model complexity and prediction horizon demonstrated that the nonlinear MPC with a median runtime of 5-6 seconds per control step is suitable for online application in extrusion processes.

The proposed methodology relies on an autoregressive MTL process model [24], builds on the ideas of CPQM [19], [20] and other end-to-end modeling approaches [21], [22], [23] and uses its forecasting capabilities for MPC in continuous extrusion processes. In contrast to analytical models for MPC [31], [32], [33] and conventional PID and fuzzy logic control [25], [26], [27], [28], [29], a data-driven approach based on a causality graph of the system allows to depend on actual measurements to model and control the process dynamics. As shown by the results of the OPVC line, the methodology in this work extends the ideas of data-driven MPC [4], [34], [35] to the entire extrusion line and offers a modular framework to model any process based on individual state-models jointly trained on an extended horizon. Due to its generic setup, the idea of using causal dependencies to generate a structured model architecture with individual dynamic state-models that are used for process control within the MPC framework is process agnostic. Therefore, the proposed approach is equally applicable to continuous processes with multistage characteristics outside the polymer extrusion domain where traditional control approaches fail to consider the end-to-end process dynamics. Examples include but are not limited to food processing or pulp and paper production, provided that the knowledge of causal dependencies and sufficient data with adequate quality are available.

Nevertheless, certain limitations and aspects must be considered. First, measurement delays and other technical constraints in the electro line, particular in the material feeding and screw speed control system, limit the possibilities of predictive control leading to short-term fluctuations when steering the throughput. Second, since the reference values are set by the operator in an ad hoc manner, the transport delays cannot be addressed by the MPC. However, the delays are explicitly built into the MTL models, hence, it is possible to consider them when upcoming reference changes are known in advance. Third, due to the technical setup of the OPVC line, automatically changing set points via a global control loop is not possible. Therefore, a human-in-the-loop approach was chosen for validation purposes where the operator blindly executes the recommendations by the MPC restricting the update interval to 10 to 15 minutes. Consequently, further improvements are expected when the MPC is fully enabled for closed-loop control. Fourth, while relying on a causality graph improves the model robustness in general, incorrect causal assumptions lead to an increased bias or variance of the process model, thus limiting the applicability to processes where the dependencies are known in advance. Lastly, several design choices (model structures, MPC and error compensation parameters, etc.) must be experimentally made when employing the proposed methodology requiring an iterative validation procedure to ensure proper configurations are used.

To further improve the methodology, multiple directions of future work can be identified. On the one hand, other optimization methods, for instance heuristic or genetic optimization algorithms might improve the quality of closed-loop control and should be evaluated in the future. On the other hand, other approaches for adaptive error compensation, such as an autoregressive moving average model, might

improve the robustness of the model against process drifts even further. Moreover, while open-loop process stability and the empirical stability analysis are strong indicators for closed-loop stability, a full proof was not provided yet and must be addressed in future work. Even though the computational efficiency (median runtime of 5-6 seconds) is certainly sufficient for a sampling interval of 60 seconds, other approaches, such as reinforcement learning [42], [43] and imitation learning [44], [45], allow to significantly reduce the runtime and represent future research directions. Similarly, the proposed methodology might be extended by event-triggered aspects to extend its applicability in domains with restricted communication and computation resources [46]. Moreover, the underlying autoregressive MTL model can be used for other applications in the context of process optimization, such as fault detection and soft sensing, which represents future work.

VIII. CONCLUSION

Due to the complexity and associated challenges of modern extrusion lines, existing approaches focusing on local controllers or static models are insufficient for global process optimization. To address this research gap, a methodology of nonlinear model predictive control was proposed in this work relying on the idea of autoregressive multi-task learning on an extended forecasting horizon. Additionally, known physical dependencies and expert knowledge are built into the process model through a causality graph of the system allowing to predict the future state trajectory for up to 60 minutes. Combined with an adaptive error compensation based on a polynomial function representing the most recent error trajectories, the process model can be used robustly for global optimization of multiple states simultaneously.

As shown by the experimental results on two actual extrusion lines, the state trajectories are accurately forecasted allowing to steer the process to closely follow the reference values. Moreover, analyzing the frequency spectrum after enabling closed-loop control revealed that the dominant natural oscillations in the system are effectively addressed by the proposed approach. In addition to that, comparing the results with normal production, supervised by a human operator, over an extended period showed significant potential when autoregressive end-to-end models are used for nonlinear model predictive control in extrusion lines.

Furthermore, the analysis of disturbance rejection capabilities showed that the proposed control design produces superior results when compared to alternative control approaches. The empirical stability analysis based on a Monte Carlo simulation with randomized initial conditions provided a strong indication for closed-loop stability of the nonlinear MPC. Apart from this, an empirical runtime analysis verified its applicability for online process control in terms of computational requirements. Nevertheless, to further improve the proposed methodology, limitations and promising research directions were discussed to facilitate future work.

ACKNOWLEDGMENT

This work was supported by Pipelife International GmbH and Pipelife Nederland B.V. as industrial project partners.

REFERENCES

[1] R. Hartner, M. Kozek, and S. Jakubek, "Addressing multi-step inverse heat transfer problems via reduced order models in a cooling process for polymer pipes with sparse measurements," *Appl. Thermal Eng.*, vol. 257, Dec. 2024, Art. no. 124483.

[2] J. Grimard, L. Dewasme, and A. Vande Wouwer, "Nonlinear model predictive control of a twin-screw extruder," in *Proc. 20th Int. Conf. Syst. Theory, Control Comput. (ICSTCC)*, Oct. 2016, pp. 204–209.

[3] K. Mulrennan, J. Donovan, L. Creedon, I. Rogers, J. G. Lyons, and M. McAfee, "A soft sensor for prediction of mechanical properties of extruded PLA sheet using an instrumented slit die and machine learning algorithms," *Polym. Test.*, vol. 69, pp. 462–469, Aug. 2018.

[4] Z. Jiang, Y. Yang, S. Mo, K. Yao, and F. Gao, "Polymer extrusion: From control system design to product quality," *Ind. Eng. Chem. Res.*, vol. 51, no. 45, pp. 14759–14770, Nov. 2012.

[5] R. Hartner, M. Kozek, S. Jakubek, and B. Mayer, "Gradient boosting regression trees for nonlinear delay identification in a polymer extrusion process," in *Proc. IEEE 21st Int. Ccnference Sci. Techn. Autom. Control Comput. Eng. (STA)*, Dec. 2022, pp. 192–197.

[6] E. M. del Pilar Noriega and C. Rauwendaal, *Troubleshooting the Extrusion Process*, 3rd ed., Munich, Germany: Hanser Publications, Oct. 2019.

[7] J.-H. Jun, T.-W. Chang, and S. Jun, "Quality prediction and yield improvement in process manufacturing based on data analytics," *Processes*, vol. 8, no. 9, p. 1068, Sep. 2020.

[8] Y. Y. Li and J. Bridgwater, "Prediction of extrusion pressure using an artificial neural network," *Powder Technol.*, vol. 108, no. 1, pp. 65–73, Mar. 2000.

[9] Y. S. Perera, J. Li, A. L. Kelly, and C. Abeykoon, "Melt pressure prediction in polymer extrusion processes with deep learning," in *Proc. Eur. Control Conf. (ECC)*, Jun. 2023, pp. 1–6.

[10] C. H. Tan, K. M. Yusof, and S. R. W. Alwi, "Quality prediction for polypropylene extrusion based on neural networks," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 1257, Oct. 2022, Art. no. 012034.

[11] M. B. Akram, C. Gilarno, K. McCoy, S. Osten, M. Rumps, U. Thota, D. Campbell, and G. Balasubramanian, "Machine learned and explainable prediction of melt pressure and temperature for polypropylene extrusion," *J. Intell. Manuf.*, pp. 1–14, Aug. 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10845-025-02647-9#citeas

[12] M. Trifkovic, M. Sheikhzadeh, K. Choo, and S. Rohani, "Model identification of a twin screw extruder for thermoplastic vulcanizate (TPV) applications," *Polym. Eng. Sci.*, vol. 50, no. 6, pp. 1168–1177, Dec. 2009.

[13] F. Alhindawi and S. Altarazi, "Predicting the tensile strength of extrusion-blown high density polyethylene film using machine learning algorithms," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage. (IEEM)*, Dec. 2018, pp. 715–719.

[14] V. García, J. S. Sánchez, L. A. Rodríguez-Picón, L. C. Méndez-González, and H. D. J. Ochoa-Domínguez, "Using regression models for predicting the product quality in a tubing extrusion process," *J. Intell. Manuf.*, vol. 30, no. 6, pp. 2535–2544, Aug. 2019.

[15] E. Bovo, M. Sorgato, and G. Lucchetta, "Using analytical and data-driven methods to develop a soft-sensor for flow rate monitoring in tube extrusion," *Proc. Comput. Sci.*, vol. 217, pp. 114–125, Jan. 2023.

[16] R. Hartner, M. Kozek, S. Jakubek, and F. Zittmayr, "Semi-supervised learning based on mass conservation to improve quality predictions in a polymer pipe production process," in *Proc. 4th Int. Symp. Artif. Intell. Intell. Manuf. (AIIM)*, Dec. 2024, pp. 272–277.

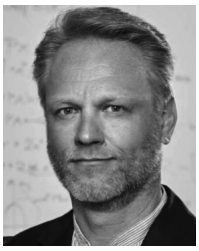
[17] N. D. Polychronopoulos, I. Sarris, and J. Vlachopoulos, "Implementation of machine learning in flat die extrusion of polymers," *Molecules*, vol. 30, no. 9, p. 1879, Apr. 2025.

[18] S. Takada, T. Suzuki, Y. Takebayashi, T. Ono, and S. Yoda, "Machine learning assisted optimization of blending process of polyphenylene sulfide with elastomer using high speed twin screw extruder," *Sci. Rep.*, vol. 11, no. 1, Dec. 2021, Art. no. 24079. [Online]. Available: https://www.nature.com/articles/s41598-021-03513-3#citeas

[19] F. Arif, N. Suryana, and B. Hussin, "Cascade quality prediction method using multiple PCA+ID3 for multi-stage manufacturing system," *IERI Proc.*, vol. 4, pp. 201–207, Jan. 2013.

[20] M. Ismail, N. A. Mostafa, and A. El-Assal, "Quality monitoring in multistage manufacturing systems by using machine learning techniques," *J. Intell. Manuf.*, vol. 33, no. 8, pp. 2471–2486, Dec. 2022.
[21] H. Yan, N. D. Sergin, W. A. Brenneman, S. J. Lange, and S. Ba, "Deep multistage multi-task learning for quality prediction of multistage manufacturing systems," *J. Qual. Technol.*, vol. 53, no. 5, pp. 526–544, Oct. 2021.
[22] H. Cho, K. Kim, K. Yoon, J. Chun, J. Kim, K. Lee, J. Lee, and C. Lim, "MMP Net: A feedforward neural network model with sequential inputs for representing continuous multistage manufacturing processes without intermediate outputs," *IISE Trans.*, vol. 56, no. 10, pp. 1–12, Oct. 2024.
[23] P. Wang, Q. Zhang, H. Qu, X. Xu, and S. Yang, "Time series prediction for production quality in a machining system using spatial–temporal multi-task graph learning," *J. Manuf. Syst.*, vol. 74, pp. 157–179, Jun. 2024.
[24] R. Hartner, M. Kozek, and S. Jakubek, "Multi-task learning with state propagation for quality forecasts in polymer extrusion lines," *J. Intell. Manuf.*, pp. 1–15, May 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10845-025-02616-2#citeas
[25] I. Yusuf, N. Iksan, and N. S. Herman, "Weight-feeder control for plastic extruder using fuzzy genetic algorithms," in *Proc. 2nd Int. Conf. Comput. Autom. Eng. (ICCAE)*, Feb. 2010, pp. 145–149.
[26] J. Deng, K. Li, E. Harkin-Jones, M. Price, N. Karnachi, A. Kelly, J. Vera-Sorroche, P. Coates, E. Brown, and M. Fei, "Energy monitoring and quality control of a single screw extruder," *Appl. Energy*, vol. 113, pp. 1775–1785, Jan. 2014.
[27] C. Abeykoon, "Single screw extrusion control: A comprehensive review and directions for improvements," *Control Eng. Pract.*, vol. 51, pp. 69–80, Jun. 2016.
[28] Y. S. Perera, J. Li, and C. Abeykoon, "Adaptive neuro-fuzzy controller for real-time melt pressure control in polymer extrusion processes," in *Proc. Eur. Control Conf. (ECC)*, Jun. 2024, pp. 2023–2028.
[29] Y. S. Perera, J. Li, and C. Abeykoon, "Adaptive control of melt pressure in polymer extrusion processes using extremum-seeking control," in *Proc. 10th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Jul. 2024, pp. 371–376.
[30] J. Meng, W. Sun, Q. F. Pan, and M. X. Ruan, "Research and application of improved particle swarm fuzzy PID algorithm based on self-disturbance rejection in temperature control system of plastic extruder," *IEEE Access*, vol. 12, pp. 41620–41630, 2024.
[31] J. Grimard, L. Dewasme, and A. V. Wouwer, "Dynamic model reduction and predictive control of hot-melt extrusion applied to drug manufacturing," *IEEE Trans. Control Syst. Technol.*, vol. 29, no. 6, pp. 2366–2378, Nov. 2021.
[32] M. Cegla and S. Engell, "Application of model predictive control to the reactive extrusion of e-Caprolactone in a twin-screw extruder," *IFAC-PapersOnLine*, vol. 54, no. 3, pp. 225–230, 2021.
[33] K. Schwarzinger and K. Schlacher, "Melt temperature control for an extruder," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 3223–3228, 2023.
[34] M. Trifkovic, M. Sheikhzadeh, K. Choo, and S. Rohani, "Model predictive control of a twin-screw extruder for thermoplastic vulcanizate (TPV) applications," *Comput. Chem. Eng.*, vol. 36, pp. 247–254, Jan. 2012.
[35] Z. Wu, D. Rincon, and P. D. Christofides, "Process structure-based recurrent neural network modeling for model predictive control of nonlinear processes," *J. Process Control*, vol. 89, pp. 74–84, May 2020.
[36] M. Lukas, S. Leineweber, B. Reitz, and L. Overmeyer, "Real-time temperature control in rubber extrusion lines: A neural network approach," *Int. J. Adv. Manuf. Technol.*, vol. 133, nos. 11–12, pp. 5233–5241, Jul. 2024.
[37] A. Aschemann, P.-F. Hagen, S. Albers, R. Rofallski, S. Schwabe, M. Dagher, M. Lukas, S. Leineweber, B. Klie, P. Schneider, H. Bossemeyer, L. Hinz, M. Kästner, B. Reitz, E. Reithmeier, T. Luhmann, H. Wackerbarth, L. Overmeyer, and U. Giese, "Smart rubber extrusion line combining multiple sensor techniques for AI-based process control," *Adv. Eng. Mater.*, vol. 27, no. 8, Oct. 2024, Art. no. 2401316. [Online]. Available: https://advanced.onlinelibrary.wiley.com/action/showCitFormats?doi=10.1002%2Fadem.202401316
[38] A. Adesanya, A. Abdulkareem, and L. M. Adesina, "Predicting extrusion process parameters in Nigeria cable manufacturing industry using artificial neural network," *Heliyon*, vol. 6, no. 7, Jul. 2020, Art. no. e04289.
[39] M. A. Branch, T. F. Coleman, and Y. Li, "A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems," *SIAM J. Sci. Comput.*, vol. 21, no. 1, pp. 1–23, Jan. 1999.
[40] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017, *arXiv:1412.6980*.
[41] I. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," 2017, *arXiv:1608.03983*.
[42] H. Yoo, H. E. Byun, D. Han, and J. H. Lee, "Reinforcement learning for batch process control: Review and perspectives," *Annu. Rev. Control*, vol. 52, pp. 108–119, Jan. 2021.
[43] H. Zhang, C. Zhao, and J. Ding, "Robust safe reinforcement learning control of unknown continuous-time nonlinear systems with state constraints and disturbances," *J. Process Control*, vol. 128, Aug. 2023, Art. no. 103028.
[44] P. Kumar, J. B. Rawlings, and S. J. Wright, "Industrial, large-scale model predictive control with structured neural networks," *Comput. Chem. Eng.*, vol. 150, Jul. 2021, Art. no. 107291.
[45] A. Pozzi, A. Incremona, and D. Toti, "Neural network-based imitation learning for approximating stochastic battery management systems," *IEEE Access*, vol. 13, pp. 71041–71052, 2025.
[46] J. Sun, X. Meng, and J. Qiao, "Event-based data-driven adaptive model predictive control for nonlinear dynamic processes," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 54, no. 4, pp. 1982–1994, Apr. 2024.

RAPHAEL HARTNER received the dual M.Sc. degree in IT and mobile security and international industrial management from the University of Applied Sciences FH JOANNEUM, Kapfenberg, Austria. He is currently pursuing the Ph.D. degree with TU Wien, Vienna, Austria. Since 2019, he has been with the University of Applied Sciences FH JOANNEUM, where he is a Lecturer and the Deputy Head of the Smart Production Laboratory. His current research interests include modeling of nonlinear processes, model predictive control, hybrid modeling approaches, and physics-informed machine learning.



MARTIN KOZEK received the M.S. degree in mechanical engineering, and the Ph.D. and Habilitation degrees from TU Wien, Vienna, Austria, in 1994, 2000, and 2009, respectively. He is currently a Professor with the Institute of Mechanics and Mechatronics, TU Wien. His research interests include nonlinear systems modeling and identification, model predictive process control, and active control of structural vibrations.

STEFAN JAKUBEK received the M.Sc. degree in mechanical engineering, the Ph.D. degree in technical sciences, and the Habilitation (Professorial Qualification) degree in control theory and system dynamics from TU Wien, Vienna, Austria, in 1997, 2000, and 2007, respectively. From 2006 to 2009, he was the Head of Development of the Hybrid Powertrain Calibration and Battery Testing Technology, Automotive Industry Company AVL List GmbH, Graz, Austria. He is currently a Professor and the Head of the Institute of Mechanics and Mechatronics, TU Wien. His research interests include fault diagnosis, nonlinear system identification, and simulation technology.