Silesian University of Technology

Faculty of Automatic Control, Electronics and Computer Science

Department of Automatic Control and Robotics

Patryk Grelewicz

Application of fusion of chosen control performance assessment methods for automatic diagnostics and performance improvement of control systems

Supervisor:
prof. dr hab. inż. Jacek Czeczot

Co-supervisor:
dr hab. inż., prof. Pol. Śl. Tomasz Kłopot

Gliwice 2022

Acknowledgments

First and foremost I would like to express my deepest appreciation to my supervisors, Prof. Jacek Czeczot and Prof. Tomasz Kłopot for their invaluable support and guidance during my PhD study. Their knowledge and plentiful experience have encouraged me in all the time of my academic research.

I would like to express my deepest gratitude to Prof. Bogdan Gabryś and PhD Thanh Tung Khuat for their commitment to our cooperation, insightful comments and suggestions.

Sincere thanks to all members in the Department of Automatic Control and Robotics for all the collaborative research and constructive discussions.

This thesis was supported by the European Union through the European Social Fund under grant POWR.03.02.00-00-I029.

Mojej najbliższej Rodzinie

Żonie, Rodzicom i Dziadkom

za ciągłe wsparcie, cierpliwość i troskę.

Dziękuję za wszystko.

Abstract

In modern, industrial control systems it is crucial to monitor control performance in continuous manner to detect any degradations. Maintaining control performance at a satisfactory level ensures high energy efficiency, final product quality and lifetime of control equipment with low post-production wastages. Considering Industry 4.0 transformation, control performance assessment (CPA) algorithms are a point of interest for academic and industrial researchers.

This PhD dissertation describes synthesis of CPA system dedicated to industrial, PID closed loop systems based on machine learning (ML) approach. General concept of suggested system is to assess the control performance based on the rejection response of the system to an intentionally applied additive disturbance and compare it with the so-called reference response. For practical application, binary assessment is suggested (OK or NOK) based on degree of difference with reference response. For this purpose, ML classification algorithms were applied. Training and validating datasets consists of feature vector of thirty control performance indices (CPIs) calculated based on the closed loop response data with the final assessment (OK or NOK) as a label. Automatic labelling method was developed, based on frequency-based indices: gain and phase margins supplemented with normalized distance from reference response. Generated datasets were used for training of ML algorithms, achieving high classification accuracy (higher than 95% for selected algorithms). The performance of CPA system was compared with other existing methods and verified based on simulations and experimental studies. For the latter, cloud-based implementation of the CPA system was prepared and it was verified on a real laboratory setup. Both simulation and experimental validation confirm high accuracy of control assessment.

Generated dataset was investigated for potential correlations between CPIs and then, forward feature selection method was used to determine the universal subset of features, reducing the number of thirty to seven features. It drastically decreased computational and memory resources required for CPA system and allowed to prepare PLC-based implementation. General purpose function block was prepared for Siemens S7-1200/1500 PLC, implemented in TIA Portal software. It consists of several functionalities, i.e. process model identification, CPIs calculations and classification of control performance. Its operation has been verified on real laboratory setup and again, obtained results indicate high accuracy of performance assessment.

In conclusion, suggested CPA system ensures high accuracy of performance assessment for considered class of closed loop systems. Thus, the following thesis statement was confirmed*: the developed CPA system can explicitly assess control performance based on load disturbance rejection response data for closed loop systems in process automation, while clearly and objectively considering the predefined assumptions and limitations*.

Streszczenie

W nowoczesnych, przemysłowych systemach sterowania kluczowe jest ciągłe monitorowanie jakości regulacji w celu wykrywania wszelkich nieprawidłowości. Utrzymywanie jakości regulacji na wysokim poziomie zapewnia wysoką efektywność energetyczną, jakość produktu końcowego oraz żywotność urządzeń wykonawczych przy niskich odpadach poprodukcyjnych. Biorąc pod uwagę transformację do przemysłu 4.0, algorytmy oceny jakości regulacji są punktem zainteresowania badaczy akademickich i przemysłowych.

Niniejsza rozprawa doktorska opisuje syntezę systemu do oceny jakości regulacji dedykowanego dla przemysłowych, zamkniętych układów regulacji PID, opartego na wykorzystaniu metod uczenia maszynowego. Ogólna koncepcja proponowanego systemu polega na ocenie jakości regulacji na podstawie odpowiedzi układu na celowo wprowadzone zakłócenie obciążeniowe i porównaniu jej z tzw. przebiegiem referencyjnym. Do praktycznego zastosowania, zasugerowano ocenę binarną (OK lub NOK) na podstawie stopnia podobieństwa z przebiegiem referencyjnym. W tym celu, zastosowano algorytmy uczenia maszynowego. Zbiory uczący i walidacyjny składają się z wektora trzydziestu wskaźników oceny jakości regulacji obliczonych na podstawie odpowiedzi układu na zakłócenie obciążeniowe wraz z oceną (OK lub NOK) jako etykietą. Opracowano automatyczną metodę etykietowania opartą na wskaźnikach częstotliwościowych: zapasie amplitudy i fazy uzupełnionych o znormalizowaną odległość od przebiegu referencyjnego. Wygenerowane zbiory wykorzystano do treningu metod uczenia maszynowego, uzyskując wysoką dokładność klasyfikacji (powyżej 95% dla wybranych algorytmów). Działanie zaproponowanego systemu zostało porównane z innymi istniejącymi metodami i zweryfikowane na podstawie symulacji i badań eksperymentalnych. W tym celu, przygotowano implementację systemu z wykorzystaniem infrastruktury chmurowej oraz zweryfikowano ją na rzeczywistej instalacji laboratoryjnej. Zarówno symulacja jak i walidacja eksperymentalna potwierdzają wysoką dokładność oceny jakości regulacji.

Wygenerowany zbiór został przebadany pod kątem potencjalnych korelacji pomiędzy wybranymi wskaźnikami oceny jakości regulacji, a następnie metodą *forward feature selection* wyznaczono uniwersalny podzbiór cech, redukując liczbę trzydziestu do siedmiu cech. Zmniejszyło to drastycznie zasoby obliczeniowe i pamięciowe wymagane dla systemu i pozwoliło na przygotowanie implementacji dla sterowników PLC. Blok funkcyjny ogólnego przeznaczenia został przygotowany dla sterowników Siemens S7-1200/1500, zaimplementowany w oprogramowaniu TIA Portal. Składa się on z kilku funkcjonalności, m.in. identyfikacji modelu procesu, obliczania wskaźników oceny jakości oraz klasyfikacji jakości regulacji. Jego działanie zostało zweryfikowane na rzeczywistej instalacji laboratoryjnej i ponownie uzyskane wyniki wskazują wysoką dokładność oceny regulacji.

Podsumowując, zaproponowany system zapewnia wysoką dokładność oceny jakości regulacji dla rozważanej klasy układów zamkniętych. Tym samym potwierdzono następującą tezę: *opracowany system może jednoznacznie ocenić jakość regulacji na podstawie odpowiedzi na zakłócenie obciążeniowe dla zamkniętych układów w automatyce procesowej, przy jasno zdefiniowanych założeniach i ograniczeniach.*

Table of Contents

1 Introduction

In recent years, the global manufacturing industry has been subject to dynamic transformations in digital technology. To maintain the country's position as a leading supplier of manufacturing equipment [1], the German government introduced a high-tech strategy known as Industry 4.0 (I4.0) in 2011. This smart manufacturing concept was necessitated by growing demand for high quality, customized products with fast delivery times [2]. The key principle of I4.0 is the incorporation of machinery, warehousing systems, and production facilities within a global network [3]. Such a combination is intended to optimize production not just for a single machine or factory, but for an entire enterprise. The realization of this concept requires the introduction of novel technologies such as cyber-physical systems, the internet of things, big data, and cloud computing to the manufacturing industry.

The final report of the I4.0 working group [3] provides explicit recommendations for the implementation of the I4.0 strategy, as prepared by the National Academy of Science and Engineering. Resource efficiency is a priority area. The authors of the report note that a reduction in both the energy and resources consumed during manufacturing processes is critical, while maintaining the stability of such processes and avoiding faulty products. This position is convergent with the European Union's Energy Efficiency Directive, which obliges EU member states to use energy more effectively [4]. The increased prevalence of global environmental movements suggests that resource efficiency will be a key issue in coming years. Such efficiency can be increased on many different levels. From a high-level enterprise-wide perspective, supply operations, batch and continuous manufacturing processes, and other activities can be optimized to reduce costs [5]. Even a 1% decrease in resource consumption can substantially affect the entire enterprise [6]. On the next level, resource consumption can be decreased by the optimization of individual processes. In [7], authors suggest improvements for industrial ovens, increasing both energy efficiency and process performance. The application of these suggestions reduces gas consumption by 20–30% and cooling time by 87.5% for a 1 MW festoon oven. The improvement of process performance, such as by increasing energy efficiency, often has a positive effect on additional manufacturing factors. In [8], an internal roller burnishing process is optimized to reduce energy requirements. This optimization also led to improved surface roughness. Finally, low-level aspects of a manufacturing line can also be optimized, such as control systems [6], [9]. Such systems are rarely developed with the purpose of minimizing resource utilization [10]. Despite this, the performance of these systems has a substantial effect on factors such as energy consumption [11], [12]. The satisfactory performance of a control system does not only ensure high energy efficiency, but can also decrease the usage of media such as chilled water or steam, reduce post-production wastage, and increase the final product quality and the lifetime of control equipment such as pumps and valves [13].

In the vast majority of cases, control systems are optimized during the commissioning stage—that is, during the start-up of new systems or the retrofitting of systems that are already in operation. During this stage, experienced engineers adapt the system hardware to conform to requirements, install modern control equipment such as sensors, actuators, or programmable logic controllers (PLCs), and select an appropriate structure for the control algorithm and tune its parameters, based on the dynamic behavior of the system. Each of these steps provide continuous improvement in the control performance, until optimal performance is reached. Fig. 1.1 shows the progression of control performance over time, from the commissioning period through normal operation. During normal operation, the monitoring of control performance is typically discontinued, and performance degrades over time. The half-life of good control performance is only approximately 6 months [14]. The degradation of control performance is associated with fluctuations in process dynamics caused by effects such as slow fouling, the wearing of control equipment such as sensors and actuators without periodic maintenance, hardware modifications made to the system, and changes in the type of manufactured product [15]. Hence, monitoring of the control performance is crucial—either online or during periodical inspections. Continuous monitoring allows any degradation in control performance to be immediately detected, then remedied following deeper inspection. As such, the demand for control performance assessment (CPA) algorithms is constantly increasing. Optimal performance can only be achieved following a diligent commissioning period, carried out by experienced process engineers. In practice, the vast majority of control systems enter normal operation while displaying suboptimal (only fair or poor) performance. CPA algorithms can indicate such situations.

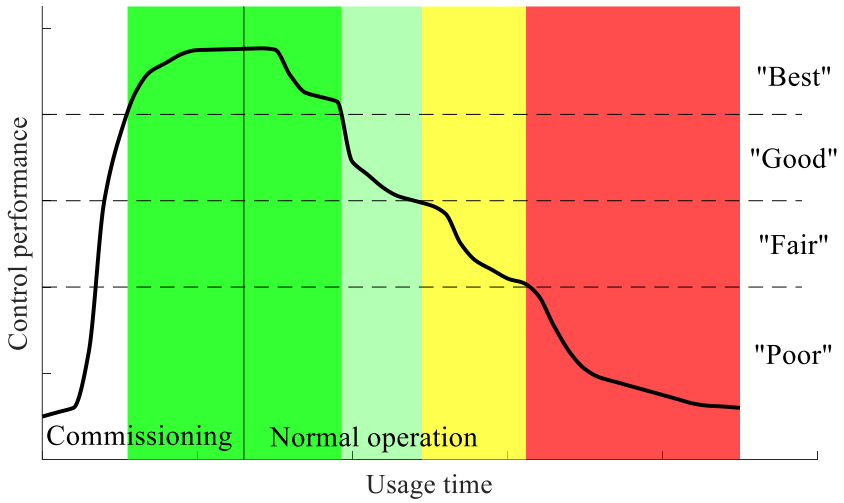

Fig. 1.1 The control performance of a closed loop system over time, showing both the commissioning and normal operation periods. Figure based on [15]**.**

CPA algorithms partially fall within the scope of fault detection techniques, with both model-based and signal-based approaches [16], [17]. Faults are typically associated with malfunctioning actuators or sensors, which directly affect the control system and lead to degradation of the control performance. CPA algorithms lack the ability to perform complex diagnoses, such as determining the location of

a faulty component. However, such algorithms can be incorporated as part of a fault detection system. Upon identification of a control system with poor (faulty) performance using a CPA algorithm, further investigation can be undertaken.

Commercially available solutions for the assessment of control performance [14] invariably take the form of additional software for analyzing process data. This approach is very convenient when the process data is archived and is available directly from a database. However, if the hardware and software infrastructure does not allow for the archiving of data, then further modifications are required, with concomitant costs, when implementing the CPA system. In such a case, costs can be reduced substantially by implementing the CPA system directly in the control layer, that is, the PLC. However, PLCs typically feature low computational and memory resources. As such, lack of professional implementations of CPA systems exist as ready-to-use general purpose function blocks for PLCs.

Fig. 1.2 shows the typical structure of an industrial control system with feedback control. The main purpose of a feedback controller is to maintain the process value *y(t)*, as measured by the sensor, at the desired setpoint *sp(t)*. This is accomplished by calculating the control signal *u(t)* required to properly manipulate the actuator.

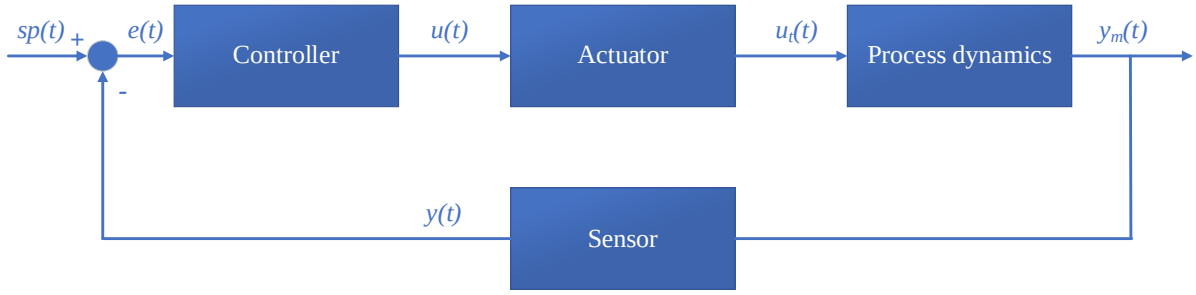

Fig. 1.2 The typical structure of a closed loop system with feedback control.

The large majority of industrial control systems use proportional integral derivative (PID) controllers. Such controllers are used in more than 90% of all industrial control systems [15], [18]. The form most commonly used has a transfer function $K_r(s)$ of

$$K_r(s) = k_r\left(1 + \frac{1}{sT_i} + \frac{sT_d}{1+s\alpha T_d}\right), \tag{1.1}$$

where s is the Laplace operator, k_r is the proportional gain, T_i is the integral constant, T_d is the derivative constant, and α is the filter coefficient. The derivative part of the PID controller is typically filtered ($\alpha > 0$) to ensure sufficient resistance to measurement noise. The high popularity of PID controllers in industrial systems is due to their simple, but complex structure, which contains both static (proportional) and dynamic (integral and derivative) parts. Both parts show deterministic behavior with respect to changes in the controller parameters; a highly desirable property for practical

applications. However, PID controllers are linear, thus any variations in the process parameters cause behavioral changes in the closed loop system.

Many different manufacturers provide their own implementations of industrial PID algorithms, in the form of general-purpose PLC function blocks. Notable manufacturers include Siemens, Mitsubishi, and Allen Bradley. In practical applications, PID algorithms are supplemented with additional functionalities, including manual control, anti-windup, and the scaling of input and output signals. For the efficient commissioning of applications, each PID function block should include an autotuning algorithm. The large majority of autotuning algorithms calculate the controller parameters based on the response for relay signal [19]–[21]. However, autotuning experiments can rarely be conducted during the commissioning of new or newly modernized control systems for industrial applications, due to the lack of necessary functionality. Moreover, the operating conditions during commissioning procedures can vary from those during normal operation of a control system. Typically, the control parameters are not retuned following commissioning, leading to fair or poor performance during normal operation (Fig. 1.1). Thus, even if autotuning is successfully applied, the obtained PID parameters should be considered as initial, and their optimality evaluated. This is particularly true for strongly nonlinear and nonstationary processes [22].

Further to the above, the optimality of PID tunings for industrial control parameters should always be considered in relation to *the best achievable performance* of the control system, which is strictly dependent on the structure and parameters of the system. During the design of control systems, many technological requirements and constraints must be considered, such as the combination of desired closed loop system dynamics and appropriate level of robustness. Thus, from a technological perspective, there always exists a trade-off between ensuring the best possible closed loop performance and providing stable conditions. As such, *the best possible performance* cannot be achieved in practice; rather, designers should focus on producing *the best achievable performance*.

In practice, an estimated 80% of control loops can be assessed as poor due to bad tuning [23]. Almost 30% have received no tuning and operate using default tunings. There currently exist hundreds of ready-made tuning rules for PID controllers [24]. These rules can be applied to many different conditions, such as desired control system time response [25], frequency response [26], or robustness level [27], [28]. In the vast majority of cases tuning rules are calculated based on the simplified process model. Various methods exist to simplify the modelling of processes. One of the most popular is the first-order plus dead-time (FOPDT) process model:

$$K(s) = \frac{k}{(1+s\tau)} e^{-s\tau_0}, \tag{1.2}$$

where k is the gain of the linear transfer function, τ is the time constant, and τ_0 is the transportation delay time. The model parameters describe the three most important features of the modelled process:

the strength of the relationship between the control signal and the process output (gain), the speed of the process (time constant), and the time lag between the control signal and the process output (transportation delay time). A key practical advantage of the application of FOPDT modelling is the simplicity of parameter identification. The parameters can be calculated based on the step response of the process using, for example, graphical area methods [29] or the least-square method [30]. Parameter identification for closed loop systems is also possible, based on the response to step setpoint change [31], or using a specific excitation signal [32], for example. Fig 1.3 shows the accuracy with which the FOPDT model can approximate the step response of a higher-order process.

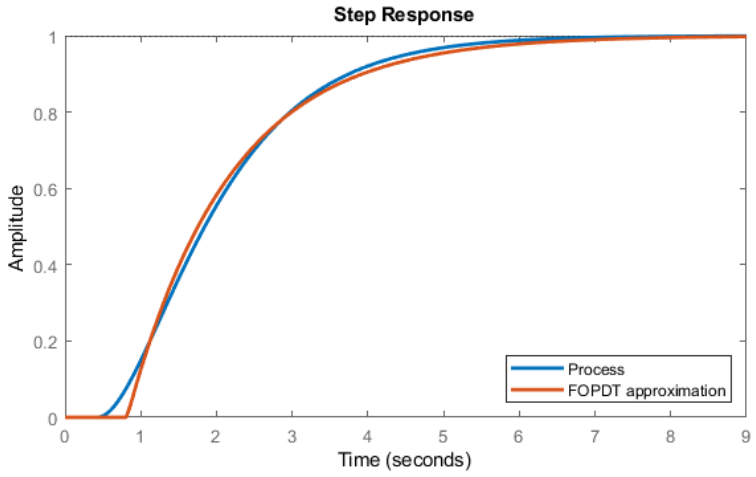

Fig. 1.3 A comparison of the step response of a higher-order process and its FOPDT approximation.

Despite its advantages, the accuracy of FOPDT modelling decreases substantially when applied to higher-order dynamical systems. For more accurate modelling of higher-order systems, the following model can be applied:

$$K(s) = \frac{k}{\prod_i (1+s\tau_i)} e^{-s\tau_0}, \tag{1.3}$$

where the τ_i, (i = 1, …, n) are time constants ($\tau_1 \geq \tau_2 \geq \ldots \geq \tau_n$) and n is the model order. Note that increasing the model order n considerably increases the difficulty of parameter identification. In practice, a second-order plus dead-time (SOPDT) model (n = 2) is often applied, as it provides a trade-off between modelling accuracy and the difficulty of parameter identification.

Control performance can be improved by modifying the structure of the controller. Given that PID controllers are linear (1.1), any changes in process dynamics can cause a significant reduction in control performance. As such, adaptation mechanisms should be included to ensure correct operation. When using gain scheduling [33]–[35], the controller parameters are dependent upon external disturbances, which can affect the dynamic behavior of the control system. PID controllers do not ensure proper performance when applied to delay dominated processes. However, the controller structure can be enhanced with a Smith predictor, which compensates for transportation time [36]–

[38]. Additional information about the process dynamics is required when using either gain scheduling or a Smith predictor. Such information can include the process model, the estimated delay time, or possible variations of the process dynamics.

In recent years, many advanced control algorithms have been developed and prepared for practical applications. Model predictive control (MPC) is an optimization-based control approach, which predicts the future behavior of a process by using a dynamic model [39]. Algorithms based on MPC are the second most commonly used controllers in industrial applications, after PID controllers [15]. MPC-based algorithms provides superior control performance to PID controllers, particularly for delay dominated processes [40], [41]. However, MPC-based algorithms do require additional modelling of the process dynamics.

Another highly promising group of advanced algorithms are based on active disturbance rejection control (ADRC) [42]. For algorithms of this type, the internal dynamics and external disturbances are modelled by a parameter known as the total disturbance, which is ultimately estimated online using an extended state observer [43]. Many different extensions of the ADRC algorithm have been developed in recent years [44], [45]. These extensions have provided more convenient implementations of the algorithm, and improved its operation with delay dominated processes.

Despite their performance advantage over PID controllers, advanced controllers lack universal tuning rules, autotuning methods and PLC-based implementations, which are crucial for industrial applications. These issues are the subject of current investigation by many researchers [46]–[50] with highly promising results. However, the industry remains to be convinced of the usefulness of advanced controllers, and work remains to develop general purpose function blocks for MPC or ADRC controllers in Siemens or Mitsubishi PLCs with all necessary functionalities.

1.1 Control performance assessment—the state of the art

CPA algorithms have been a point of interest for academic researchers since 1989, when Harris [51] introduced the very first CPA index. The Harris index is based on stochastic performance criteria, which typically includes the variance of the process variable or the control signal:

$$\sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y(i) - \bar{y})^2, \tag{1.4}$$

where σ_y^2 is the variance of the process variable, $\bar{y}$ is the mean value of the process variable, and N is the number of investigated samples. The index can be applied to control systems that operate under stochastic disturbance; such disturbance is assumed to be filtered white noise. The performance criteria are directly associated with energy consumption, product quality, and overall process performance [15]. In addition to typically increasing product quality, a reduction in variance can also reduce energy consumption, as the system is allowed to operate more closely to the constraints of the

technological process. This is demonstrated in Fig. 1.4. The assessment of variance is further justified by the fact that the vast majority of control systems in process automation operate with a constant setpoint. Hence, proper disturbance rejection is more important than tracking properties [52].

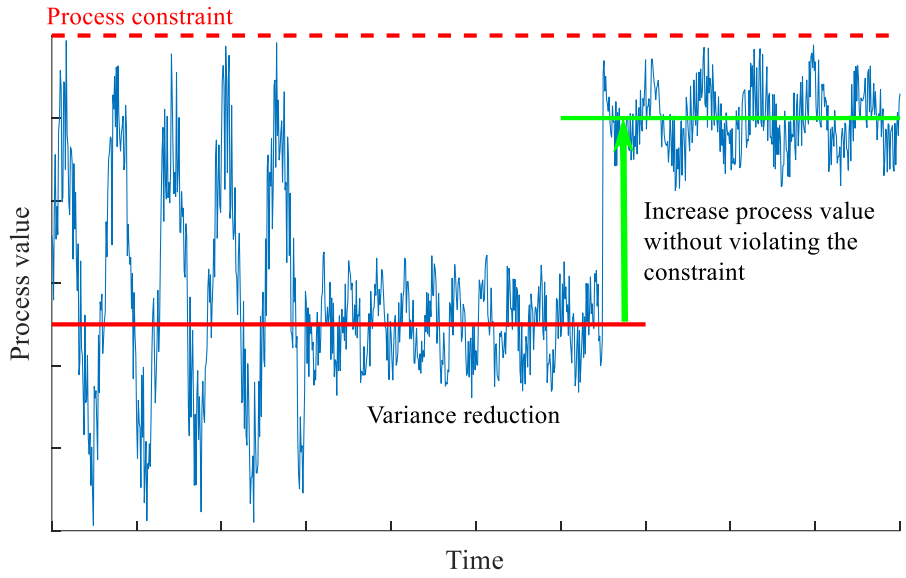

Fig. 1.4 The effect of reducing variance on the operating conditions of a closed loop system.

Many CPA algorithms compare actual control performance with a benchmark or reference value. However, the Harris index compares the actual variance of the process variable with the variance that could be achieved by implementing a minimum variance (MV) controller [53]:

$$\eta_{MV} = \frac{\sigma_{MV}^2}{\sigma_y^2}, \tag{1.5}$$

where η_{MV} is the Harris index and σ_{MV}^2 is the variance obtained with an MV controller. The index itself is normalized and bounded, and hence can be interpreted relatively easily by practitioners. The value $\eta_{MV} \to 1$, thus $\sigma_y^2 \to \sigma_{MV}^2$ indicates perfect minimum variance control, with $\eta_{MV} \to 0$ indicating the worst possible performance. The practical implementation of an MV controller is difficult, due to unrealistic variations of the control signal [54]. However, an online implementation is not required for the controller to be effectively used as a benchmark for performance assessment. Calculation of the Harris index does not require a precise model of the control system; only the delay time is required [55]. However, the index can be effectively estimated using real process data [50]. For more convenient calculations, and to avoid solving the Diophantine equation, the Harris index can be estimated using autoregressive model fitting [56]. Although it is normalized, the Harris index does not indicate potential improvements in control performance. However, the index can be used to detect the degradation of control performance. Fig. 1.5 demonstrates a possible application of the Harris index. In this case, the index itself is used to detect the increased variance of the process value, which results

in a degradation of control performance. However, this example requires that the nominal value of the Harris index (the best possible control) is first determined by rule of thumb.

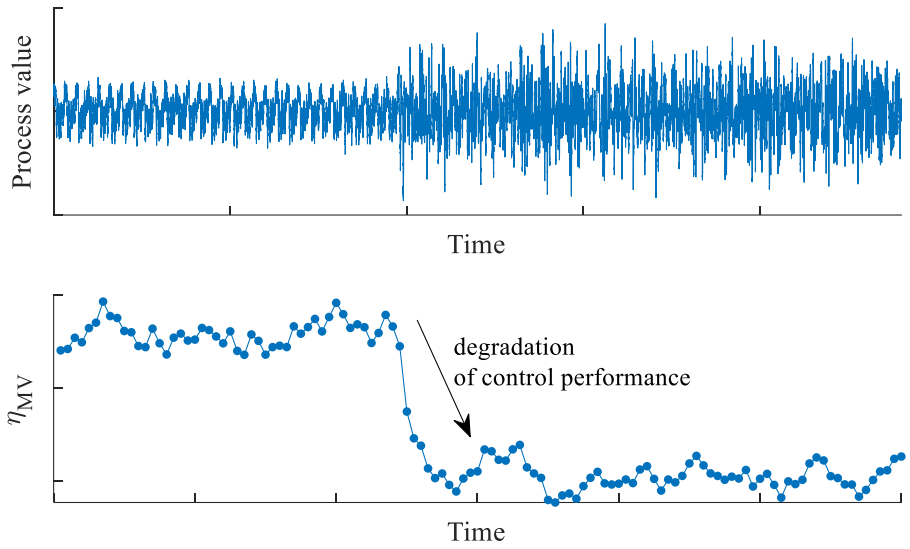

Fig. 1.5 The application of the Harris index to the detection of control performance degradation.

The Harris index is based on the theoretically achievable minimum variance. However, as discussed previously, this minimum level is impossible to achieve in practice, particularly when an appropriate level of robustness must be maintained. The generalized minimum variance index (η_{GMV}) can be introduced to limit the control activity. This index is based on the general minimum variance (GMV) control law [57], in which a generalized output ϕ_{GMV} is minimized as

$$\phi_{GMV}(t) = P_c e(t) + F_c u(t), \tag{1.6}$$

where P_c and F_c are weighting functions. The GMV index has a very similar form to the Harris index, but compares the actual variance of the generalized signal (1.6) to that which could be achieved by a GMV controller [58]. The introduction of weighting functions to both the control error and the control signal provides a more realistic benchmark of control performance. An appropriate choice of weighting functions allows the desired dynamical behavior of a closed loop system to be determined.

Such a choice of weighting functions P_c and F_c is critical for proper control and assessment using the GMV index. The function P_c should include an integral term, so that the integral action is present within the controller, and the function F_c should be constant or feature a lead term to ensure that the controller rolls-off at high frequencies [59]:

$$P_c = 1 + \frac{\omega_p}{s}, \tag{1.7}$$

$$F_c = -\rho_c \left(1 + \frac{s}{\omega_f}\right), \tag{1.8}$$

where ω_p and ω_f are the cut-off frequencies and $\rho_c > 0$ is the tuning parameter. The frequency ω_p should be chosen to remove the integral action at the lowest possible frequency. Typically, it is chosen as $\omega_p = \omega_c/10$, where ω_c is the desired unity-gain crossover frequency of the system [59]. The frequency ω_f should be chosen to reduce the control signal at high frequencies, with an initial value of $\omega_f = 10\omega_c$. Finally, the tuning parameter ρ_c can be used to determine the speed of the response. Fig. 1.6 presents the frequency responses of example weighting functions, together with the response of the process model. Notably, the high frequencies of the control error (measurement noise) and the low frequencies of the control signal (constant term) are efficiently filtered out by the P_c and F_c functions, respectively, and thus do not affect the generalized output (1.6). It should also be noted that in the presence of transportation delay within the process model, the control signal affects the output with a time delay. Hence, the value of τ_0 should be included within the weighting function F_c. This emphasizes the necessity of estimating the delay time with the highest possible accuracy.

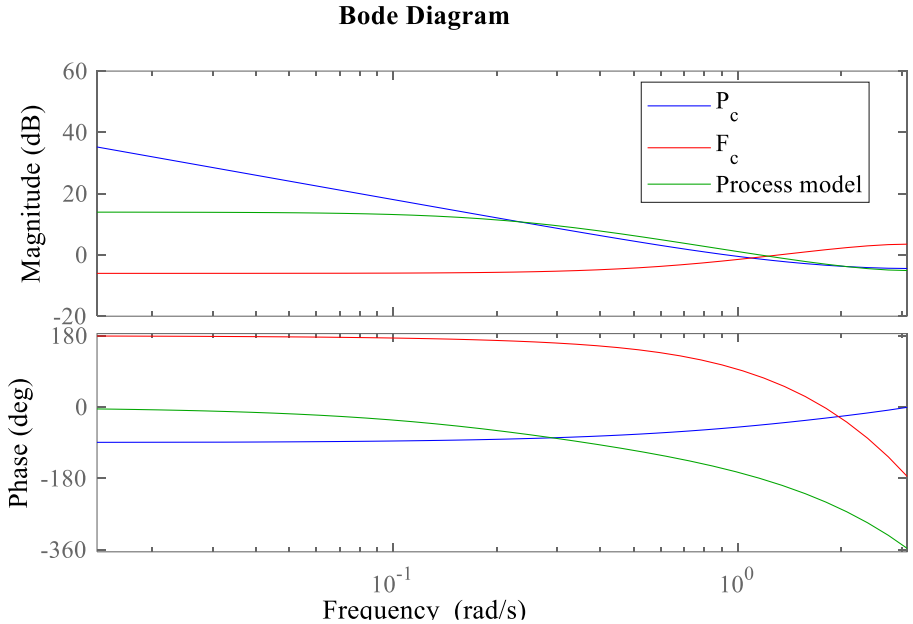

Fig. 1.6 The frequency responses of example weighting functions and the process model.

A choice of weighting functions as described above results in GMV control with very similar characteristics to linear-quadratic-Gaussian (LQG) control [58]. The latter can also be used as a CPA method by obtaining the performance limit curve, which defines the achievable region of control performance. This curve can be generated by solving the LQG problem defined by the objective function J_{LQG} [60]:

$$J_{LQG} = \sigma_e^2 + \rho_{LQG}\sigma_u^2, \tag{1.9}$$

where σ_e^2 and σ_u^2 denote variances of control error and signal respectively, and ρ_{LQG} is the tuning parameter. By varying the tuning parameter ρ_{LQG} the objective function J_{LQG} can be optimized

iteratively. The obtained results can be presented as a performance limit curve. Fig. 1.7 shows an example of such a curve.

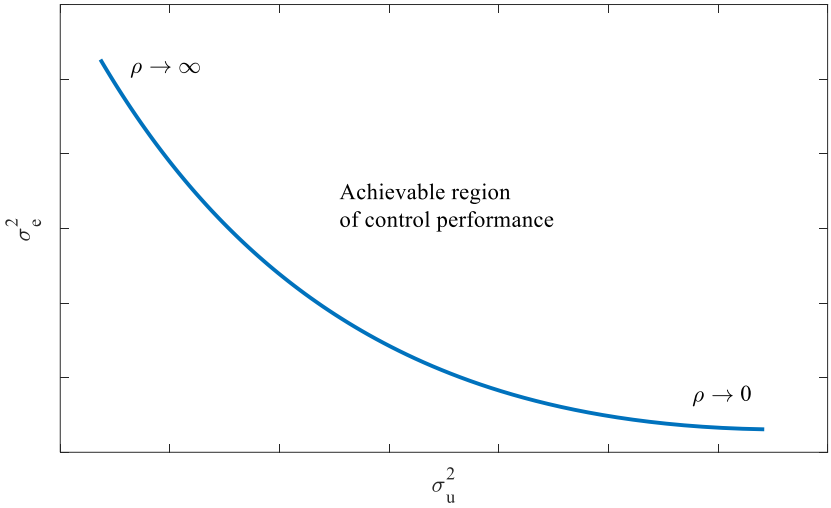

Fig. 1.7 An example performance limit curve obtained for an LQG objective function (1.9).

In this example, the curve represents the best achievable performance, which is dependent on the tuning parameter ρ_{LQG}. From this the possible performance improvement can be determined, in terms of the control error and signal variances. Two types of improvement can be then analyzed: the degree to which the variance of the control signal can be decreased by the same variance of the control error, and the degree to which the variance of the control error can be decreased by the same variance of the control signal. Clearly, the limit curve can be only achieved by applying the LQG controller. In practice, calculation of the performance limit curve can involve unnecessary computation and produce an unideal regression effect. A numerical, recursive, or analytical algorithm can be applied to obtain the equigrid LQG benchmark, thus improving the regression effect and reducing computing costs [61], [62].

The vast majority of industrial control systems use PID controllers, which cannot achieve the performances displayed by MV, GMV, or LQG controllers. The application of CPA indices to PID control systems does not provide explicit information concerning the potential improvement of control performance, due to the limitations of the PID algorithm. Indices designed for use with PID systems can be applied for a more explicit assessment in practice. In [63], a numerical approach is used to estimate the achievable control when using a proportional integral (PI) controller. This approach uses the variance of the output signal as the performance measure. To calculate the PI index, only the process model and delay time are required, and must be estimated with sufficient accuracy. The literature presents two approaches to determine the minimum variance of PI controllers: with a disturbance model and without a disturbance model. In [64], an iterative procedure is used to obtain the minimum variance and performance index of PID controllers. These parameters are calculated based on the step response coefficients of the process model.

All of the presented approaches assume that the disturbances affecting the control system are stochastic, and that the introduction of additional excitation is not required. However, in practice, control systems are not always sufficiently affected by disturbances that can be considered as stochastic and that meet these theoretical assumptions. As such, the methods described above cannot be used for the assessment of absolute control performance. In these cases, additional excitation is applied to the control system in the form of a deterministic disturbance such as a setpoint change *sp(t)* or a load disturbance *d(t)*. The time response samples are then used for performance assessment. These can be manually introduced to the closed loop system by a process operator, or their occurrence can be the result of a production process that can be automatically detected, such as changes in operating point or product demand [65]. If a deterministic disturbance is manually applied to the control system, the amplitude and precise time of application can be adjusted by the process operator or the master system according to the actual operating conditions. Obviously, the assessment and monitoring of performance based on a deterministic disturbance response cannot be conducted continuously. Rather, this type of disturbance should be applied periodically to a control system to provide a periodically accessible assessment result.

For a load disturbance, the disturbance and process dynamics are similar. Hence, the block diagram shown in Fig. 1.2 can be simplified to the form presented in Fig. 1.8.

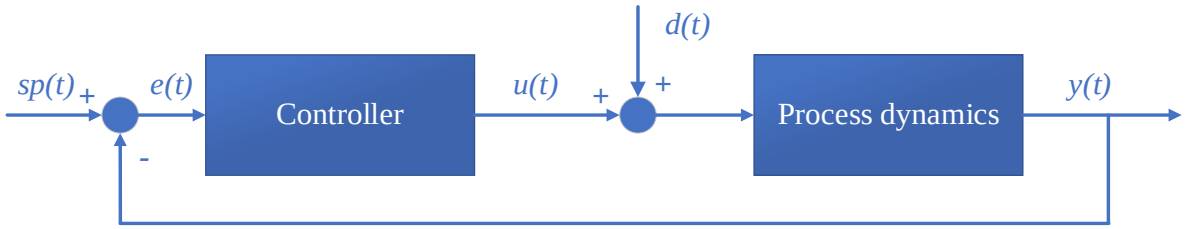

Fig. 1.8 A simplified block diagram of a closed loop system subject to both types of deterministic disturbances.

Many well-known control performance indices (CPIs) exist, such as settling time or maximum overshoot, or integral indices such as integral absolute error (IAE). The latter group are typically used as measures of economic performance [66]. They do not require a high degree of computational resources, as the vast majority of them can be calculated directly from closed loop response data. These approaches are widely used to compare the control performance of various tuning methods or control algorithms [22], [40], [46], [67]–[69], but do not provide an explicit assessment or allow possible improvements to be determined. The possible values of CPIs are strictly dependent upon the closed loop system parameters. For example, the settling time of a system with slow closed loop dynamics is typically much higher than for a system with fast closed loop dynamics, even if the performance of both control systems is satisfactory. Moreover, CPIs lack reference values, and thus can only be used for the monitoring of control degradation, with reference to the historical trends of

the chosen CPI. Finally, individual CPIs typically provide much less information concerning control performance than stochastic indices such as the Harris or GMV indices.

Another approach to determine control performance when subject to a deterministic disturbance is the idle index, proposed by Hägglund [70]. The idle index is particularly effective for the detection of sluggish control loops. The index itself compares the total time during which there is a positive correlation between the control signal Δu and process signal Δy increments with the total time during which there is a negative such correlation:

$$J_{IdleIndex} = \frac{t_+ - t_-}{t_+ + t_-}, \tag{1.10}$$

where t_+ and t_- are the total times during which the product of $\Delta u\ \Delta y$ has a positive or negative sign, respectively. An idle index with a value close to 1 indicates sluggish control, while a value close to 0 indicates good performance. However, the idle index should be interpreted jointly with an oscillation-detection procedure; a value close to -1 does not provide an explicit assessment of performance, and can indicate either a well-tuned controller or oscillatory behavior [15].

Deterministic disturbance rejection response data, gathered from a PID control system, can also be used to retune the controller via calculation of the process parameters. This approach is presented in [71] for setpoint change and in [72] for load disturbance rejection. In [73], a data-driven approach is used to obtain a performance limit curve with a PID controller subject to setpoint change. A PID controller performance benchmark based on the min-max principle is presented in [74]. This approach is particularly effective for nonlinear processes. Finally, [76] presents the frequency and time domain assessment and retuning of an internal model control PI controller [75].

1.2 Summary and identification of research gaps

To summarize, the continuous monitoring of control performance is critical; control performance directly affects economic factors such as energy efficiency, product quality, and material usage. Hence, CPA algorithms are under investigation by both academia and industry [77], [78].

As demonstrated by the presented state-of-the-art review, several issues remain unresolved and additional questions are posed when considering the practical application of CPA algorithms.

1. The vast majority of CPA algorithms require that the transportation delay time is estimated with the highest possible accuracy. The novel increment count method (ICM) was suggested by author as a PLC-based implementation of such algorithms [50]. The method ensures that the transportation delay time is estimated using steady-state detection, and produces results that are comparable with other existing methods. Moreover, the ICM algorithm has low

computational and memory resource requirements, and is thus suitable for implementation on a PLC. Appendix 1 provides further details of this approach.

2. Stochastic indices, such as the Harris or GMV indices, provide highly accurate assessment of control performance. However, this accuracy comes at the cost of high demand for computational and memory resources. Alternatively, CPIs are calculated based on the response to a deterministic disturbance, and can be obtained and calculated with ease, as the vast majority of them do not require complex mathematical functions. Thus, a worthwhile line of investigation is to compare the performance assessment provided by CPIs with that provided by stochastic indices, to determine the existence of any correlation between the two.
3. Several different CPIs can be calculated to assess control performance based on the system response to a deterministic disturbance. These indices are widely used to compare different control strategies or tunings for a given control system, but do not provide an explicit assessment of control performance. This is because the provided values are relative, and depend upon the process dynamics. This suggests a possible line of investigation is to determine the ability of CPIs to explicitly assess control performance.
4. Finally, implementation issues remain to be resolved when considering the practical applications of a CPA system. The state-of-the-art provides many examples of CPA implementations that take the form of external or master systems, but there is a lack of examples of implementations directly in the control layer. A PLC-based implementation of a CPA system could provide a substantial reduction in cost, as additional software and hardware modifications would not be required.

1.3 Aim and scope of the work

The main goal of this thesis is to develop a system for the explicit control performance assessment of PID-based closed loop systems. Having reviewed the state-of-the-art and the practical issues concerning such systems, the following general specification of a CPA system is suggested:

1. The system should be as general as possible. That is, it should be capable of assessing control performance for the widest possible range of processes that are typically found in automated industrial closed loop systems.
2. Typically, the control systems of automated processes operate using a constant setpoint. Thus, the assessment procedure should be prepared for load disturbance rejection. This type of disturbance can be manually introduced to any considered closed loop system. A load disturbance of predefined amplitude should be introduced periodically. Note that the investigated closed loop system should not be disturbed significantly.
3. The reference behavior of a closed loop system should be explicitly predefined. Moreover, the behavior should be adjusted according to technological requirements and conditions.

4. For the majority of convenient practical applications, the assessment of the considered closed loop system should be explicit and binary. For example, an assessment of OK could indicate satisfactory performance, and an assessment of NOK could indicate poor performance. Thus, the control assessment will be clear even for inexperienced users.
5. The proposed system should assess control performance based on predefined indices calculated from closed loop response data.
6. The system should be capable of offline design without requiring additional experiments to be conducted on a real control system. Only the load disturbance rejection response can be obtained from the control system for performance assessment.
7. The structure of the CPA system should be suitable for implementation on a PLC. That is, its design should consider the computational and memory resource limits of PLCs.

This thesis describes the development of the proposed CPA system, in addition to its verification via simulation and experiment. The thesis author suggests the following thesis statement: *the developed CPA system can explicitly assess control performance based on load disturbance rejection response data for closed loop systems in process automation, while clearly and objectively considering the predefined assumptions and limitations*.

The general concept and detailed developmental steps of the CPA system are presented in Section II, together with a robustness analysis and simulated comparison to other CPA methods. Section II concludes with an experimental verification of the developed system using a cloud-based implementation.

Section III presents the implementation of the CPA system on a PLC. The reduction of the CPA system structure is fully described, together with implementation issues. A reduced set of indices are also discussed.

Finally, Section IV concludes the thesis and indicates possible future directions of CPA system development.

2 Synthesis of the CPA system

This section describes the methodology that is used in this thesis to develop a CPA system that meets the requirements defined in Section 1.3.

2.1 General concept

The general concept of the proposed CPA system is to assess the control performance of the considered class of closed loop systems based on the rejection response of the system to an intentionally applied additive disturbance. The data that is obtained from closed loop responses can be used to calculate various indices with low computational complexity. Separately, such indices describe only limited properties of the performance of the considered systems. The proposed CPA system is based on the assumption that the capture of various key features of closed loop responses by a sufficiently large number of indices could provide useful information to assess the overall control performance. However, the major difficulty with this assumption is that the process dynamics and dynamical parameters strongly affect the closed loop response.

As defined in Section 1.3, a key requirement of the proposed CPA system is the potential for its application to the widest possible range of processes. Thus, a general process modelling method should be used to model the system, thereby ensuring high modelling accuracy combined with a relatively straightforward procedure to identify the model parameters. Moreover, the process dynamics should be normalized appropriately, given that various processes can display similar dynamic behaviors across different time scales.

Potential system users will require a convenient assessment system. Among the most convenient approaches is binary assessment, with the control performance being judged as either satisfactory (OK) or poor (NOK). This approach ensures that the performance assessment is explicit and can be easily interpreted by users such as process operators, who are typically inexperienced with process automation. Such a binary classification problem could be solved with a supervised machine learning (ML) approach. Typically, ML methods are used for the assessment of explicit technological process, such as the smelting process of an electro-fused magnesium furnace [79]. An ML system with a more generalized approach is presented in [80], wherein the performance of a PID-based control system is assessed with use of a k-nearest neighbors (KNN) method. Finally, multi-class SVM is used in [81] to indicate potential problems with control systems based on an autocorrelation function coefficient and statistical features calculated from time response data.

The use of a functional classifier requires the generation of training and validation datasets that consist of feature vectors with the expected output of the classifier (label). However, offline preparation is a requirement of the CPA system, and therefore the use of additional online training experiments using

real control setups is not possible. Thus, the training and validation datasets should be generated via simulation, before their usefulness for real control systems is evaluated.

For use with the ML algorithms, a dataset of closed loop systems and corresponding performance evaluations is required. To generate such a dataset, a reference response can be defined for each considered closed loop system, based on predefined criteria, resulting in reference PID tunings. Then, these tunings can be randomly modified to provide various closed loop responses, and thus various closed loop systems. Basing on predefined criteria, the randomly generated closed loop responses can then be labelled as displaying satisfactory (OK) or poor (NOK) performance. Typically, the control performance should be assessed as satisfactory (OK) if the considered closed loop response is relatively close to the corresponding reference response. Else, if the dynamic behavior of the considered control system differs from the predefined reference, the control performance should be assessed as poor (NOK).

Various classification algorithms can be investigated using the generated training and validation datasets, including tree-based, kernel-based, or even fuzzy methods. The following sections present further details concerning the discussed issues.

2.2 Process modelling and model normalization

One of the key requirements of the proposed CPA system is applicability to the widest possible range of dynamic industrial processes. Use of the SOPDT model is appropriate for this purpose, as it provides relatively high modelling accuracy in the presence of higher order dynamics. The model parameters can be identified based on either open or closed loop response data, using, for example, the least squares algorithm.

Assuming a unitary gain, the SOPDT model consists of three key dynamical parameters: two time constants $\tau_1 \geq \tau_2$ and the transportation delay time τ_0. For a more convenient classification of the process dynamics, these parameters can be reduced to two normalized dynamical parameters:

$$L_1 = \frac{\tau_0}{\tau_1+\tau_0} \in [0.1, 0.6], \tag{2.1}$$

and

$$L_2 = \frac{\tau_2}{\tau_1} \in [0.1, 1]. \tag{2.2}$$

The ranges of L_1 and L_2 cover all of the dynamic processes that can be reasonably controlled by the PID algorithm. The transportation delay dominates processes for which $L_1 > 0.6$. In this case, more advanced control algorithms should be implemented (e.g., MPC) to achieve satisfactory performance. Conversely, for the case in which L_1, $L_2 < 0.1$, the application and tuning of a conventional PI controller is sufficient.

The L_1 and L_2 parameters are dimensionless. However, sets of SOPDT model parameters are dependent upon the process dynamics; that is, the time constants and transportation delay time increase for slower processes. For example, for given values of L_1 and L_2 parameters, an infinite number of equivalent SOPDT processes can be determined, each with a different set of parameters (τ_1, τ_2, and τ_0). These equivalent processes have different time responses with similar shapes. Thus, a comparison of the time responses requires that they are first normalized. The problem of dynamics normalization is discussed in [82], where the authors use gain and time scales to normalize the process and controller transfer functions.

Applying a similar approach to [82], the closed loop response can be normalized using the process gain k (along the y axis) and the dominating time constant τ_1 (along the time axis). The purpose of normalization is to obtain similar dynamic responses for various sets of equivalent SOPDT parameters that are all associated with a specific set of values L_1 and L_2.

For illustration, let us consider two dynamic processes:

$$K_1(s) = \frac{1}{(1+s)(1+0.2s)} e^{-0.428s}, \tag{2.3}$$

and

$$K_2(s) = \frac{1}{(1+2s)(1+0.4s)} e^{-0.857s}. \tag{2.4}$$

It should be noted that both the time constants τ_1 and τ_2 and the transportation delay τ_0 are twice as high in the second process K_2 as in the first process K_1. However, in both cases L_1 = 0.3 and L_2 = 0.2. Tunings for PID controllers were suggested so as to preserve the similar gain of both controllers, with the integral and derivative constants twice as high for the second process K_2 as for the first K_1. Thus, the controller parameters for both processes are presented below:

$$\begin{aligned} k_{r1} &= 1.47, \\ T_{i1} &= 1.086, \\ T_{d1} &= 0.22, \end{aligned} \tag{2.5}$$

and

$$\begin{aligned} k_{r2} &= 1.47, \\ T_{i2} &= 2.17, \\ T_{d2} &= 0.45. \end{aligned} \tag{2.6}$$

Fig. 2.1 presents the responses of each closed loop system to a step load disturbance.

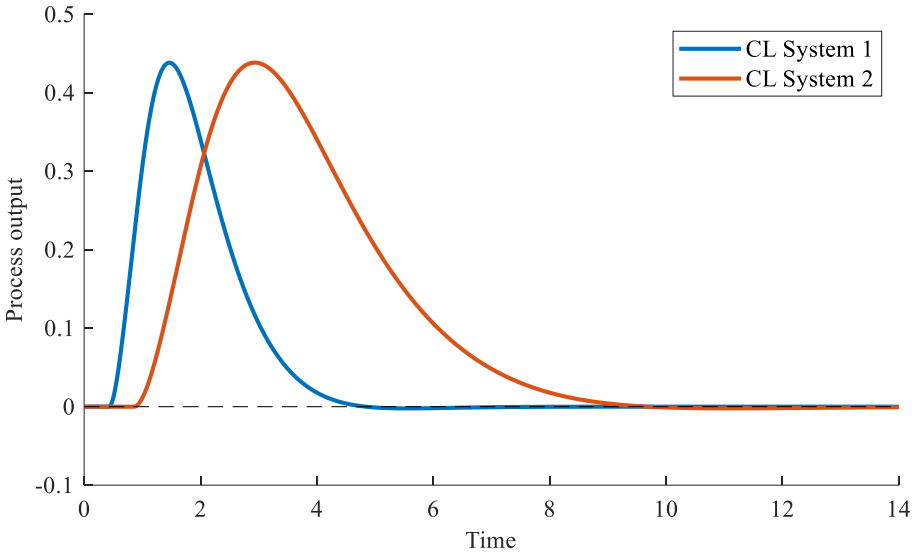

Fig. 2.1 The responses of each of the two example closed loop systems to a step load disturbance, prior to normalization of the time axis.

The dynamic behavior of both closed loop systems is identical, but the second system (blue plot) is apparently slower. However, when considering that the ratio of the dominant time constants of the K_2 (2.4) and K_1 (2.3) processes is equal to 2, the time axis of the K_2 closed loop response can be normalized by dividing it by 2 to obtain the two identical time responses presented in Fig. 2.2.

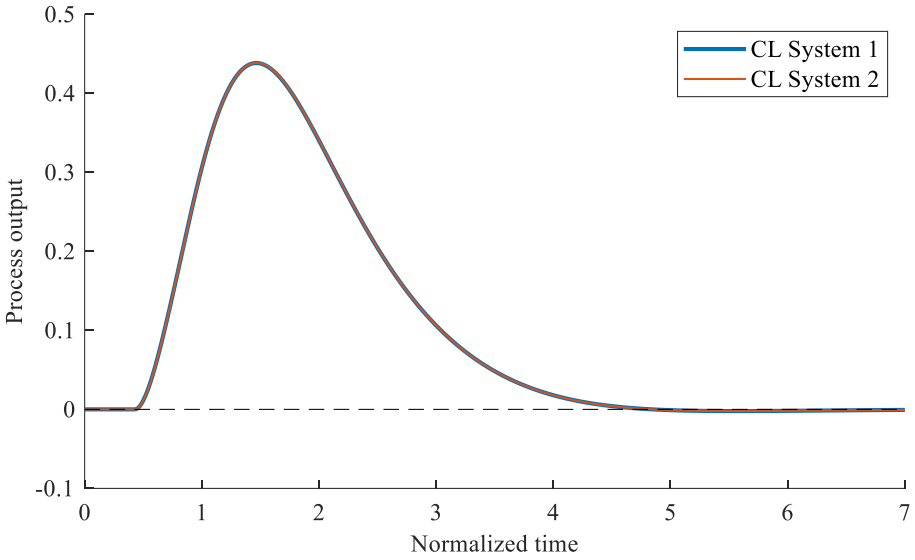

Fig. 2.2 The responses of each of the two example closed loop systems to a step load disturbance, following normalization of the time axis.

The responses of equivalent closed loop systems with all possible sets of parameters (time constants and transportation delay time) can be normalized for further analysis in a general manner. This is accomplished by normalizing the time axis with the dominant time constant τ_1, that is, by dividing the time vector samples by the dominant time constant τ_1.

A similar effect to the normalization of the time axis can be achieved by the normalization of the process (both time constants and transportation delay time) and controller parameters (integral and derivative constants) with the dominant time constant τ_1:

$$\begin{matrix}\tau_1, \tau_2 \tau_0 \\ T_i, T_d\end{matrix} \xrightarrow{normalization} \begin{matrix}\tau_{1,n} = 1, \tau_{2,n} = \frac{\tau_2}{\tau_1}, \tau_{0,n} = \frac{\tau_0}{\tau_1} \\ T_{i,n} = \frac{T_i}{\tau_1}, T_{d,n} = \frac{T_d}{\tau_1}\end{matrix} \tag{2.7}$$

where $\tau_{1,n}$, $\tau_{2,n}$, and $\tau_{0,n}$ are the new, normalized process parameters, and $T_{i,n}$ and $T_{d,n}$ are the new, normalized controller parameters. Hence, by introducing normalized L_1 and L_2 parameters, and by normalizing the time axis, the entire range of dynamic processes for a given L_1 and L_2 can be compared with one other, regardless of the SOPDT time-based parameters:

$$\tau_1, \tau_2 \tau_0 \in (0, +\infty) \xrightarrow{normalization} \tau_{1,n} = 1 \; and \; \tau_{2,n} \tau_{0,n} \in (0, 1] \tag{2.8}$$

For the normalization of process gain *k*, let us consider two dynamic processes:

$$K_3(s) = \frac{1}{(1+s)(1+0.2s)} e^{-0.428s}, \tag{2.9}$$

and

$$K_4(s) = \frac{2}{(1+s)(1+0.2s)} e^{-0.428s}. \tag{2.10}$$

Both processes have the same time constants and transportation delay time. However, the gain of the K_4 process is twice as high as the gain of the K_3 process. Again, PID tunings were suggested for both processes to preserve the similar dynamic behavior of the closed loop systems. The selected integral and derivative constants are the same for both processes, but the controller gain for the K_3 process is twice as high as the controller gain for the K_4 process:

$$\begin{matrix} k_{r3} = 1.47, \\ T_{i3} = 1.086, \\ T_{d3} = 0.22, \end{matrix} \tag{2.11}$$

and

$$\begin{matrix} k_{r4} = 0.735, \\ T_{i4} = 1.086, \\ T_{d4} = 0.22. \end{matrix} \tag{2.12}$$

Fig 2.3 presents the responses of each closed loop system to a step load disturbance.

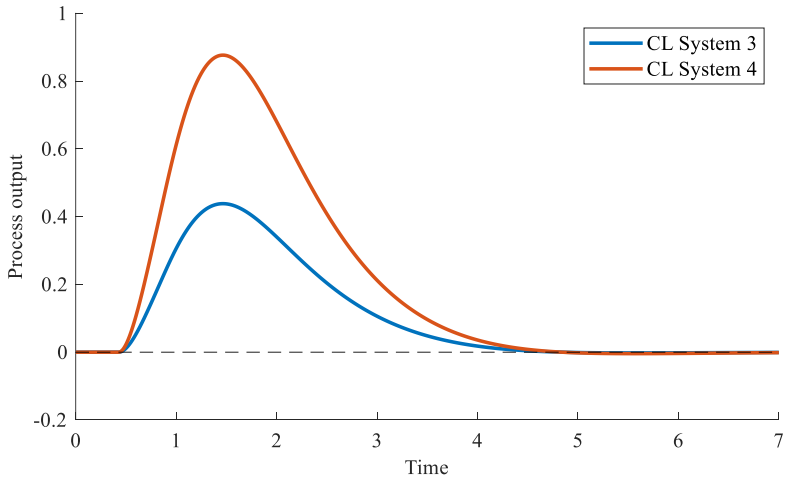

Fig. 2.3 The responses of each of the two example closed loop systems to a step load disturbance, prior to normalization of the *y* axis.

Again, the dynamic behavior of both closed loop systems is identical. However, the response of CL System 4 is twice as high as the response of CL System 3. It should be noted that the dynamic behaviors of the two systems are similar because the product of process and controller gain is equal for the K_3 (2.9, 2.11) and K_4 (2.10, 2.12) processes. To normalize the closed loop responses, each response sample must be divided by the gain of system K_4. This produces equivalent closed loop responses for both processes, as shown in Fig. 2.4.

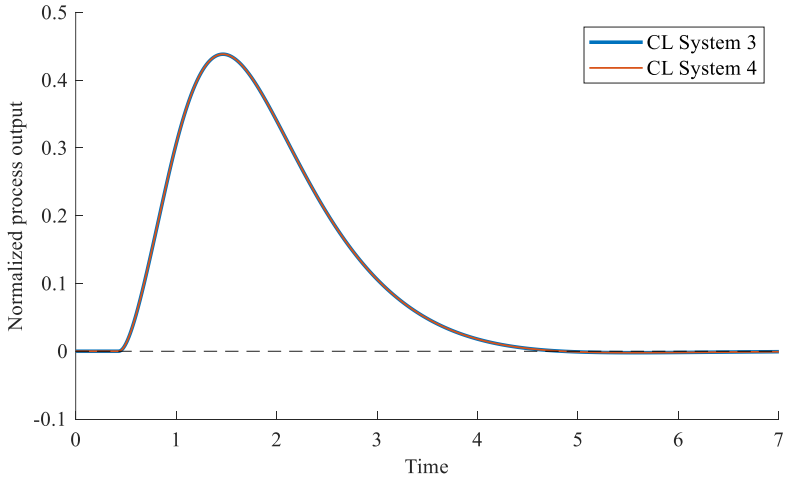

Fig. 2.4 The responses of each of the two example closed loop systems to a step load disturbance, following normalization of the *y* axis.

Generally, each closed loop system can be normalized by dividing the process gain k by itself and multiplying the controller gain k_r by the process gain. Following this normalization process, the gain has a unitary value:

$$\begin{matrix} k \\ k_r \end{matrix} \xrightarrow{normalization} \begin{matrix} k_n = 1 \\ k_{r,n} = k_r k \end{matrix} \tag{2.13}$$

where k_n is unitary process gain and $k_{r,n}$ is the controller gain following normalization.

The process gain k and the dominant time constant τ_1 can be used to normalize the response along the y axis and the time axis, respectively. Thus, the normalized dynamic behaviors of two processes that are described by the same L_1 and L_2 parameters are identical, independent of the SOPDT parameters. However, it should be noted that the SOPDT parameters (gain k and dominant time constant τ_1) are required for normalization, and hence must be identified for the considered closed loop system.

2.3 The selection of closed loop response features

This thesis suggests calculating various CPIs based on the closed loop disturbance rejection response to capture the key features of the response and generate the feature vector. The proposed CPA system requires the generation of training and validation datasets for the testing of different ML algorithms. Hence, each feature vector (sample) should be supplemented with the expected classifier output (label), corresponding to either OK or NOK performance. The selected CPIs should not require computationally intensive calculations, but should provide useful information about control performance when evaluated together.

2.3.1 Correlation between CPIs and stochastic indices

Initial studies were conducted to verify the possibility of using relatively simple CPIs to achieve a similar accuracy of control performance assessment to that obtained by stochastic indices. For this purpose, several CPIs and stochastic indices were selected. Many different CPIs are defined in the literature. The most popular were chosen for these studies:

- the *MaxPeak* is defined as the maximum value of control error during the closed loop response,
- the *SettlingTime* is defined as the period of time that is required for the output signal to reach its setpoint value following the introduction of an additional excitation,
- the *integral absolute error* (*IAE*) is defined as

$$IAE = \int_0^\infty |e(t)|dt, \tag{2.14}$$

- the *integral time absolute error* (*ITAE*) is defined as

$$ITAE = \int_0^\infty t|e(t)|dt. \tag{2.15}$$

The Harris and GMV indices were chosen as stochastic indices. The potential correlation between CPIs and stochastic indices was investigated via simulation. The model of an electric flow heater was simulated, as part of the heat exchange and distribution system described in Appendix 4. The heater had a constant flow rate of F = 4 L/min. For increased realism, measurement noise was added to the simulated output temperature. The noise had a variance of σ^2_a = 0.005, and was filtered using a first order filter with a time constant τ_f = 2 s. The tuning of the temperature control system was based on an identified FOPDT process model (k = 0.41, τ = 27.1 s, and τ_0 = 17.8 s). A total of 13 PI and PID tuning

methods were used (see Appendix 5) to generate 13 different temperature control systems. The PI and PID parameters of the control systems are summarized in Table 2.1.

	PI		PID		
	k_r	T_i	k_r	T_i	T_d
ZN	*3.26*	*59.33*	*4.35*	*35.6*	*8.9*
SIMC	*1.81*	*27.1*	-	-	-
CHR 20% reg	*2.53*	*41.47*	*4.35*	*35.6*	*7.47*
CHR 20% serv	*2.17*	*27.1*	*3.44*	*36.86*	*8.37*
CHR 0% reg	*2.17*	*71.2*	*3.44*	*42.36*	*7.47*
CHR 0% serv	*1.27*	*31.71*	*2.17*	*27.1*	*8.9*
AMIGO	*0.76*	*52.89*	*2.11*	*25*	*7.43*

Table 2.1 The PI and PID parameters calculated for the considered process.

Two experiments were conducted to compare the CPIs and stochastic indices. To calculate the CPIs, a deterministic disturbance in the form of an additive load disturbance was introduced to the control system as ΔP_h = 10%, with the constant setpoint T_{SP} = 35 °C. The responses of the PI and PID control systems are presented in Fig. 2.5 and Fig. 2.6, respectively.

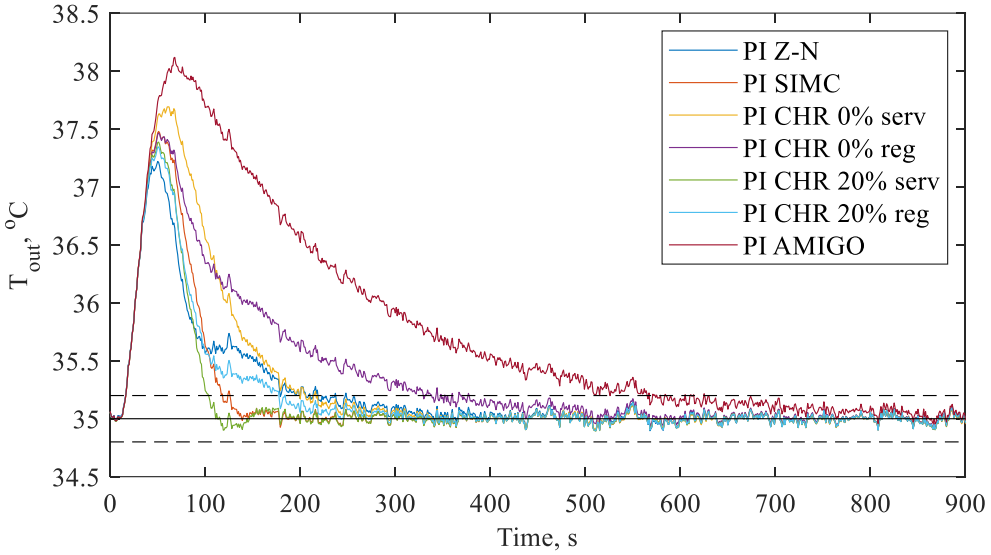

Fig. 2.5 The closed loop responses of the considered PI controllers.

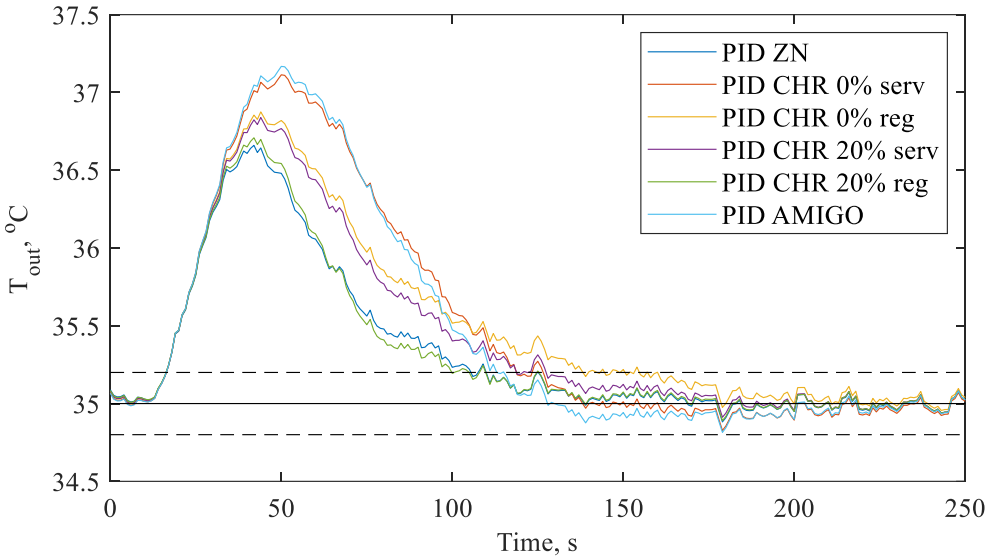

Fig. 2.6 The closed loop responses of the considered PID controllers.

Based on each response, four CPIs were calculated and normalized with respect to their minimum values, with a unitary value indicating best performance. The normalized values of the studied CPIs are presented in Table 2.2.

The stochastic indices required an additional form of excitation. This was provided in the form of an introduced load disturbance. The load disturbance ΔP_h consisted of two sinusoidal and two square wave signals with different amplitudes and frequencies, and subject to Gaussian noise. The transportation delay time required to calculate the stochastic indices was estimated by a correlation function of the input and output signals [83]. Using this approach, the transportation delay time was calculated as two samples with a sampling time of τ_p = 5 s. To standardize testing conditions, the same stochastic disturbances were applied during every experiment. Although this consistency is not achievable in practice, it provides a more reliable comparison between the considered indices.

Fig. 2.7 presents three example responses to an applied excitation for an example control system. To calculate the stochastic indices, 6000 samples were obtained with the same constant sampling time of τ_p = 5 s.

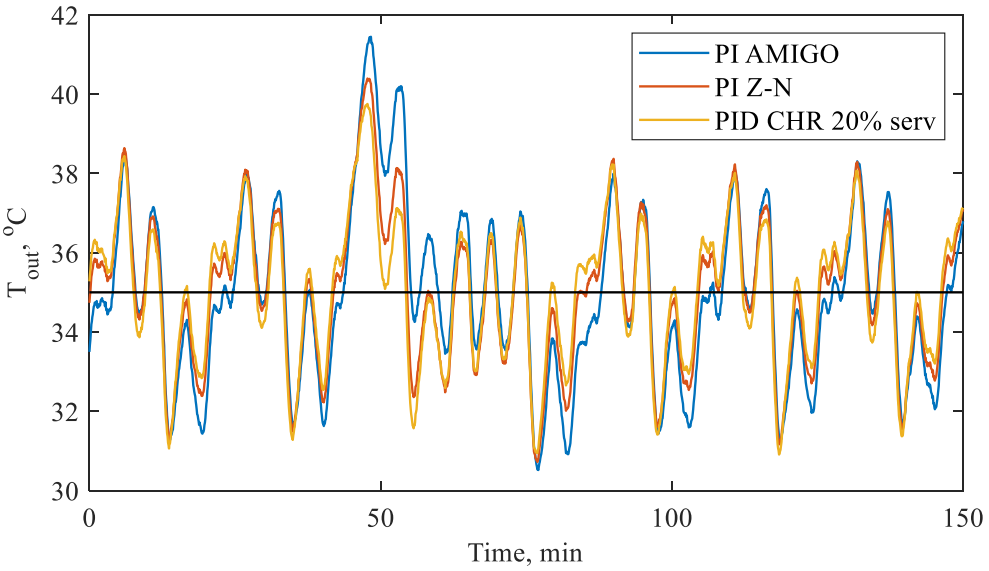

Fig. 2.7 Three example responses of closed loop systems to the suggested excitation.

A moving window of N = 1000 samples was used to calculate the stochastic indies. The window was shifted successively by 200 samples. Thus, during each iteration, N = 1000 samples were used to calculate consecutive values of the stochastic indices. Both the Harris and GMV indices were calculated using an autoregressive model with order m = 30 [56]. The calculated Harris indices are presented in Fig. 2.8 and 2.9 for the PI and PID systems, respectively. The calculated GMV indices are presented in Fig. 2.10 and 2.11 for the PI and PID systems, respectively. In each case, additional filtering of the indices is included for clarity of presentation.

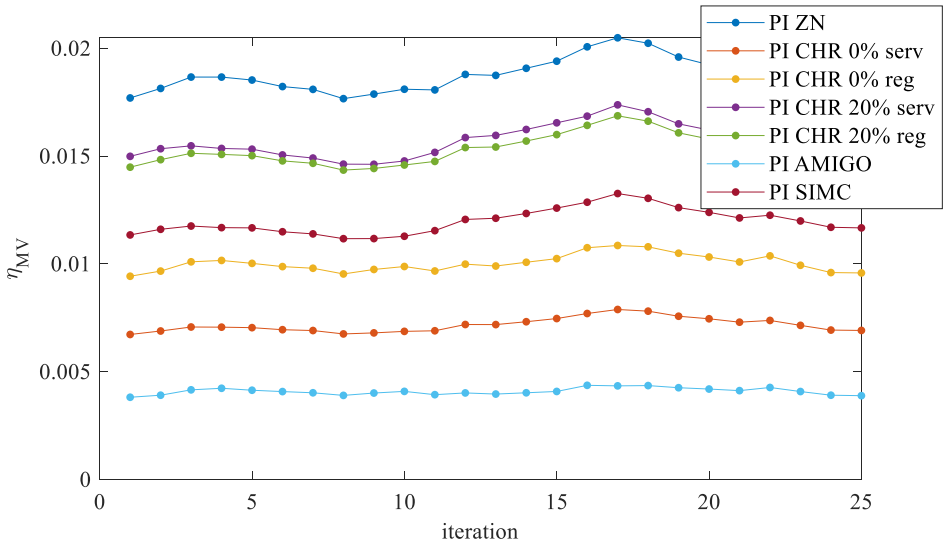

Fig. 2.8 The calculated Harris indices for each of the considered PI closed loop systems.

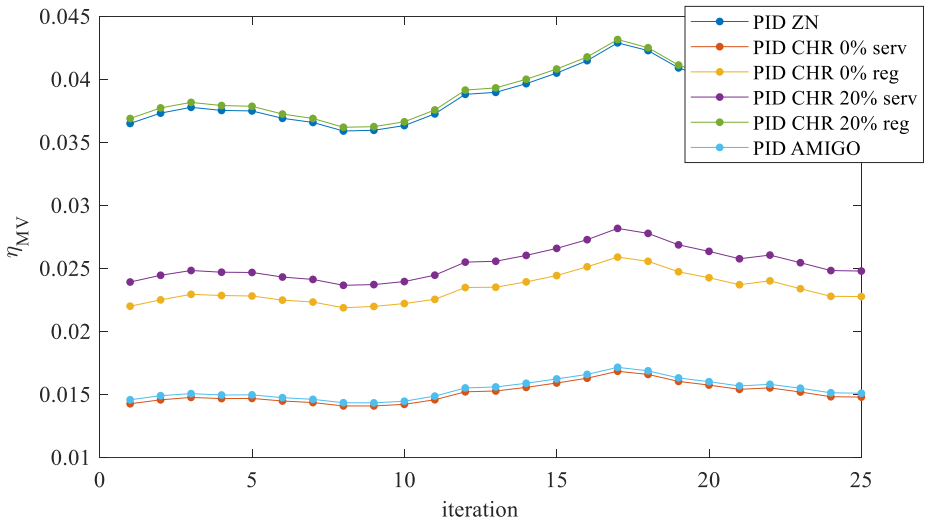

Fig. 2.9 The calculated Harris indices for each of the considered PID closed loop systems.

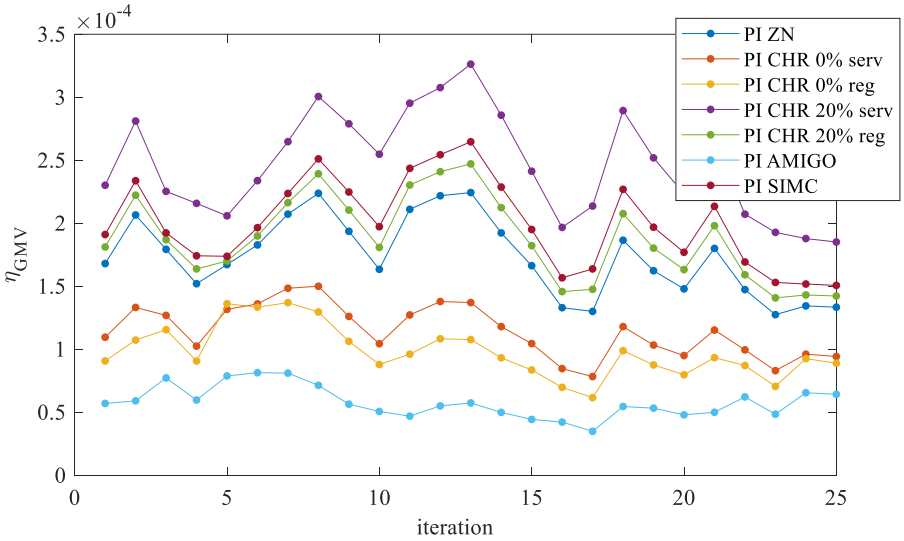

Fig. 2.10 The calculated GMV indices for each of the considered PI closed loop systems.

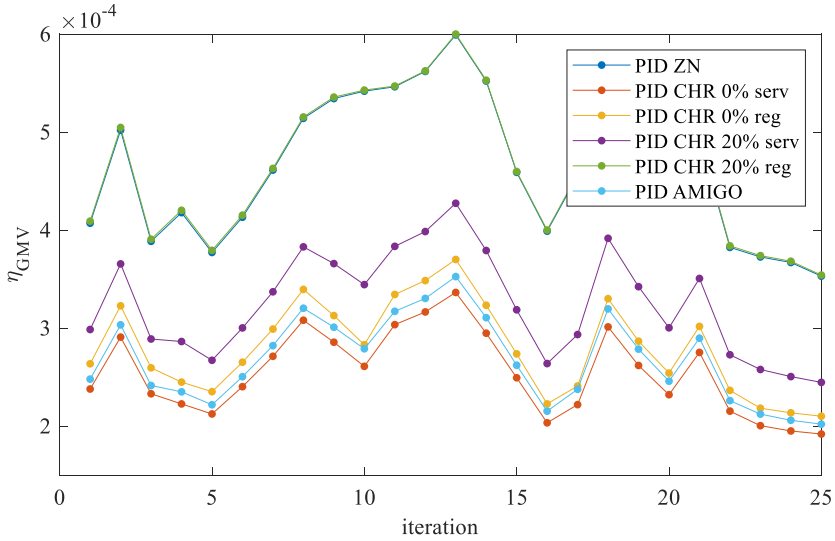

Fig. 2.11 The calculated GMV indices for each of the considered PID closed loop systems.

To provide a comparison with the CPIs, the mean value of each stochastic index was calculated and normalized with respect to the minimal value. In this case, a unitary value indicates worst performance. Table 2.2 presents the normalized values of the Harris and GMV indices as calculated for the studied control systems. The results are presented in descending order according to the GMV index.

		$MaxPeak_{norm}$	$SettlingTime_{norm}$	IAE_{norm}	$ITAE_{norm}$	$\eta_{MV,norm}$	$\eta_{GMV,norm}$
PID	**CHR 20% reg**	*1.04*	*1.06*	*1.00*	*1.00*	*9.49*	*8.03*
	ZN	*1.00*	*1.06*	*1.00*	*1.00*	*9.41*	*8.02*
	CHR 20% serv	*1.12*	*1.23*	*1.17*	*1.01*	*6.20*	*5.59*
	CHR 0% reg	*1.13*	*1.54*	*1.27*	*1.04*	*5.71*	*4.83*
	AMIGO	*1.32*	*1.11*	*1.32*	*1.02*	*3.76*	*4.6*
	CHR 0% serv	*1.28*	*1.22*	*1.34*	*1.02*	*3.68*	*4.38*
PI	**CHR 20% serv**	*1.44*	*1.00*	*1.32*	*1.04*	*3.83*	*4.24*
	SIMC	*1.50*	*1.21*	*1.47*	*1.05*	*2.93*	*3.45*
	CHR 20% reg	*1.42*	*1.71*	*1.55*	*1.10*	*3.73*	*3.24*
	ZN	*1.35*	*2.40*	*1.66*	*1.18*	*4.56*	*3.00*
	CHR 0% serv	*1.63*	*2.12*	*2.20*	*1.22*	*1.76*	*1.97*
	CHR 0% reg	*1.49*	*3.65*	*2.66*	*1.56*	*2.44*	*1.69*
	AMIGO	*1.88*	*6.01*	*5.26*	*3.02*	*1.00*	*1.00*

Table 2.2 The normalized values of the selected indices, calculated for the considered closed loop systems.

The results show that the assessment of the control performance made by the GMV index almost perfectly aligns with the assessment made by the integral *IAE* and *ITAE* indices. Slight differences in classification exist between the *IAE* index and the PI CHR 20% serv tunings, and between the *ITAE* index and the PID CHR 0% reg tunings. However, the classifications made by the *MaxPeak* and *SettlingTime* indices show no correlation with those of the stochastic indices.

To provide a deeper analysis, the correlation coefficients *r(X, Y)* between each pair of indices *X, Y* were calculated:

$$r(X,Y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-y)^2}} \tag{2.16}$$

where *x* and *y* represent the two values of the studied indices and $\bar{x}$ and $\bar{y}$ are their mean values. The calculated correlation coefficients are presented in Table 2.3. Highly correlated values are highlighted in green ($|r| \geq 0.9$) and red ($|r| \geq 0.95$).

	MaxPeak	*SettlingTime*	*IAE*	*ITAE*	η_{MV}	η_{GMV}
MaxPeak	*1.00*	*0.74*	*0.82*	*0.73*	*-0.91*	*-0.91*
SettlingTime		*1.00*	*0.97*	*0.96*	*-0.57*	*-0.70*
IAE			*1.00*	*0.98*	*-0.64*	*-0.71*
ITAE				*1.00*	*-0.52*	*-0.59*
η_{MV}					*1.00*	*0.95*
η_{GMV}						*1.00*

Table 2.3 The calculated correlation coefficients $r(X,Y)$ between considered indices. Highly correlated values are highlighted in green ($|r| \geq 0.9$) and red ($|r| \geq 0.95$).

The results indicate that the *MaxPeak* index is strongly correlated with the stochastic indices. These indices are presented in Fig. 2.12 in the form of a scatter plot. The graphical representation confirms the high degree of correlation between these indices. All of the studied CPIs are more strongly correlated with the GMV index than with the Harris index. A strong correlation is also noticeable between the *SettlingTime*, *IAE*, *ITAE*, and stochastic indices.

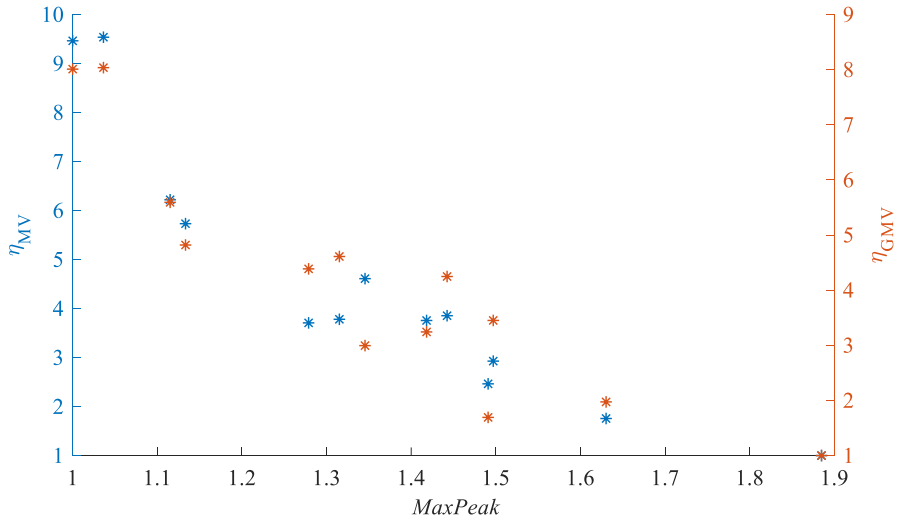

Fig. 2.12 A scatter plot of the *MaxPeak* and stochastic indices, showing a strong correlation.

The relatively low degree of correlation between the other CPIs and stochastic indices can indicate nonlinear dependency between these indices. To determine the presence of inverse proportionality between the different indices, additional $r(X, 1/Y)$ correlation coefficients were calculated. The results are presented in Table 2.4. Highly correlated values are highlighted in green ($|r| \geq 0.9$) and red ($|r| \geq 0.95$). The results indicate inverse proportionality between the *SettlingTime* and GMV indices, between the *IAE* index and the MV and GMV indices, and between the GMV and *IAE* indices. The highly correlated pairs of indices ($|r| \geq 0.95$) are presented as scatter plots in Fig. 2.13.

	MaxPeak	*SettlingTime*	*IAE*	*ITAE*	η_{MV}	η_{GMV}
1 / MaxPeak	*-0.98*	*-0.64*	*-0.72*	*-0.61*	*0.96*	*0.94*
1 / SettlingTime	*-0.68*	*-0.88*	*-0.80*	*-0.73*	*0.62*	*0.81*
1 / IAE	*-0.92*	*-0.87*	*-0.88*	*-0.79*	*0.87*	*0.95*
1 / ITAE	*-0.79*	*-0.99*	*-0.98*	*-0.96*	*0.61*	*0.73*
$1 / \eta_{MV}$	*0.91*	*0.87*	*0.95*	*0.91*	*-0.75*	*-0.77*
$1 / \eta_{GMV}$	*0.88*	*0.96*	*0.98*	*0.94*	*-0.72*	*-0.81*

Table 2.4 The calculated correlation coefficients $r(X,1/Y)$ between considered indices. Highly correlated values are highlighted in green ($|r| \geq 0.9$) and red ($|r| \geq 0.95$).

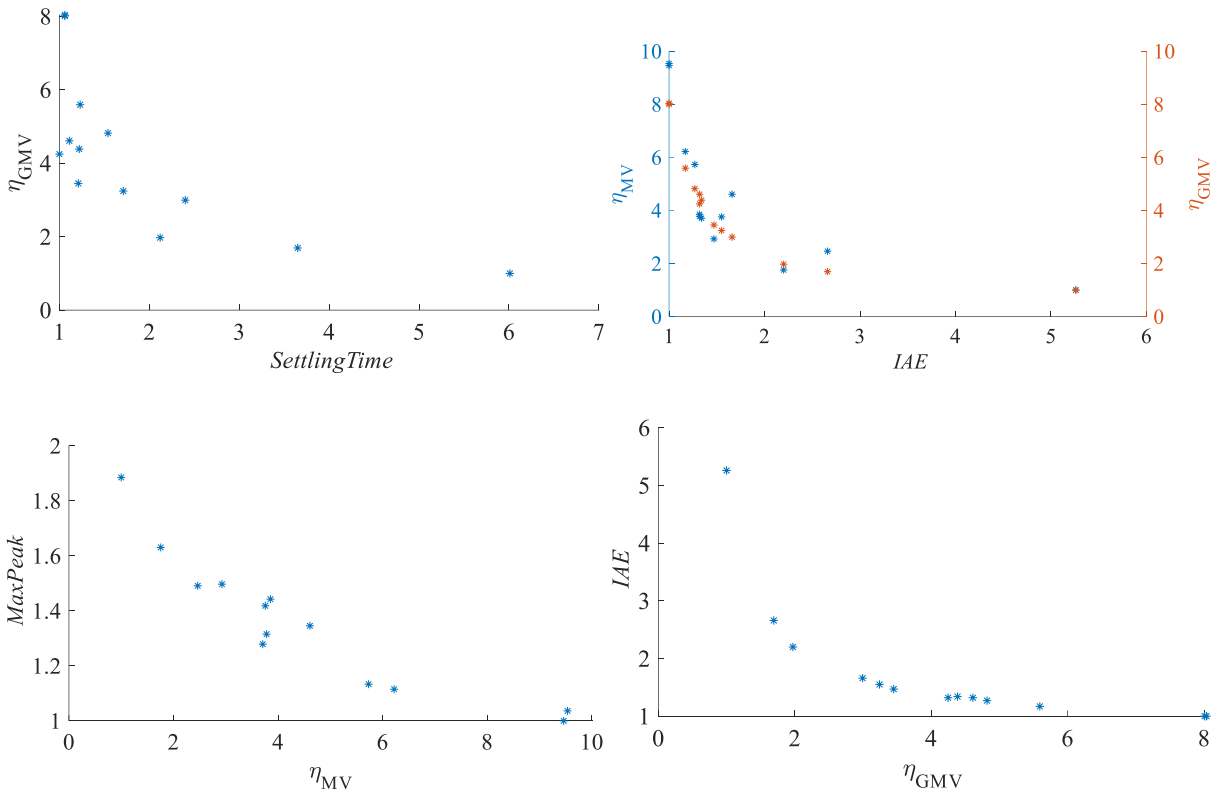

Fig. 2.13 Scatter plots of index pairs with high inverse correlation.

The Harris index correlates strongly with both *MaxPeak* and 1/*MaxPeak*. Additional results are required to confirm the presence of either a direct or inverse dependency. However, the key result for this study is that such a dependency exists.

The obtained results show that for the considered case, the quantification of control performance using the *IAE* or *ITAE* indices provides similar results to the use of the GMV index. These results are particularly promising, given that the GMV index is more suitable for industrial applications than the Harris index. Thus, in some cases control performance can be quantified using relatively simple integral CPIs, which have low computational and memory resource requirements. The results obtained

by these CPIs should be similar to those obtained by the GMV index. However, these conclusions were obtained from a specific case study; their generalization requires additional studies.

It should be noted that the calculations of CPIs and stochastic indices require different types of excitations. Typically, industrial control systems operate with a constant setpoint and act to compensate for existing disturbances. In such cases, the control system is already affected by a stochastic disturbance, and the introduction of an additional stochastic disturbance is not required. However, the calculation of a CPI requires the intentional introduction of a load disturbance to the control system. For some industrial applications, this could occur incidentally from the result of a technological process or additional requirement, such as an increase in demand for a product. In this case, only automatic detection is required. Further details on these studies can be found in [84].

2.3.2 CPIs–discussion and selection

As demonstrated in the previous section, control performance can be assessed using CPIs that are relatively easy to calculate. The obtained results are similar to those that can be produced using more complex stochastic indices. Thus, the suggested CPA system could potentially utilize simple CPIs and conduct assessments that are similar to those obtained using complex CPA algorithms.

However, although CPIs are widely used to compare control tunings or strategies, lack of CPI applications exist which are used for the explicit assessment of control performance. This is because the CPI values depend on the process dynamics, and hence control performance cannot be assessed explicitly using only a single CPI. That is, the use of just a single CPI can provide an inaccurate assessment of control performance. This problem is illustrated in Fig. 2.14, which presents the responses of two different closed loop systems for the same step load disturbance. Both responses display a similar overshoot (*MaxPeak*) of approximately 0.35, however, CL System 1 substantially outperforms CL System 2, which has a significantly longer *SettlingTime* with oscillatory behavior.

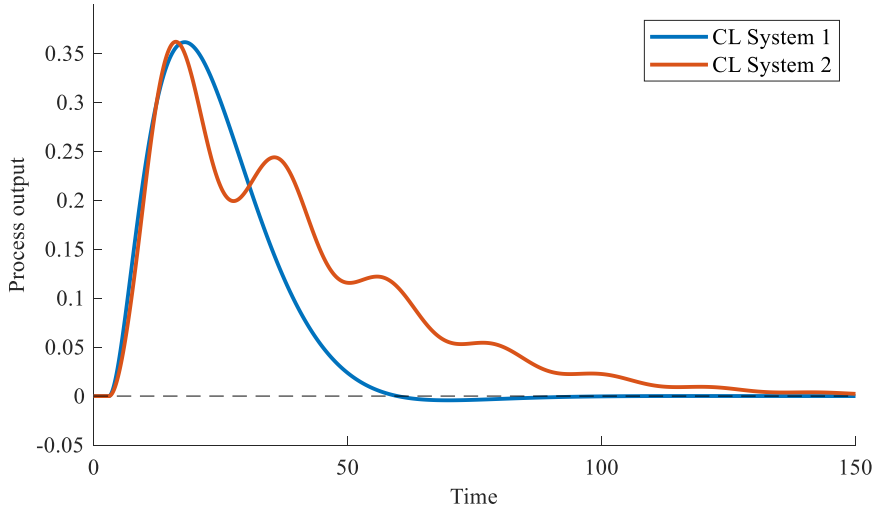

Fig. 2.14 The first comparison between the responses of two example closed loop systems to a step load disturbance.

Fig. 2.15 presents a second comparison between two different systems. The responses of CL Systems 1 and 2 display significantly different values of maximum overshoot (*MaxPeak*). As determined by this CPI, CL System 2 outperforms CL System 1. However, the dynamics of CL System 2 are dominated by oscillatory behavior, which is detrimental from a practical perspective. Again, focus on just a single CPI can provide incorrect results.

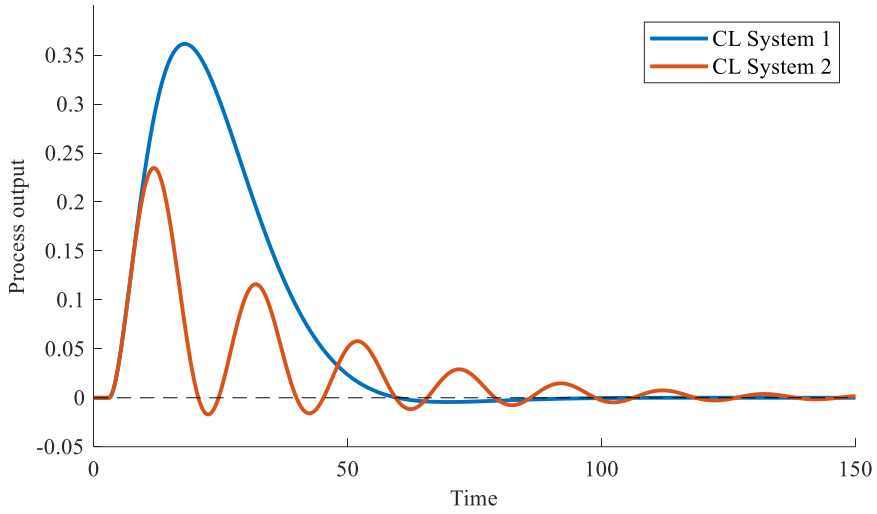

Fig. 2.15 The second comparison between the responses of two example closed loop systems to a step load disturbance.

Fig. 2.16 presents a third and final comparison. CL Systems 1 and 2 display a similar *SettlingTime* of approximately 70, but different dynamic behavior. CL System 1 does not undershoot, but reaches a higher value of overshoot (*MaxPeak*) than CL System 2. A more detailed analysis is required to determine which CL System provides better performance. The choice of system is also dependent on the precise technological requirements.

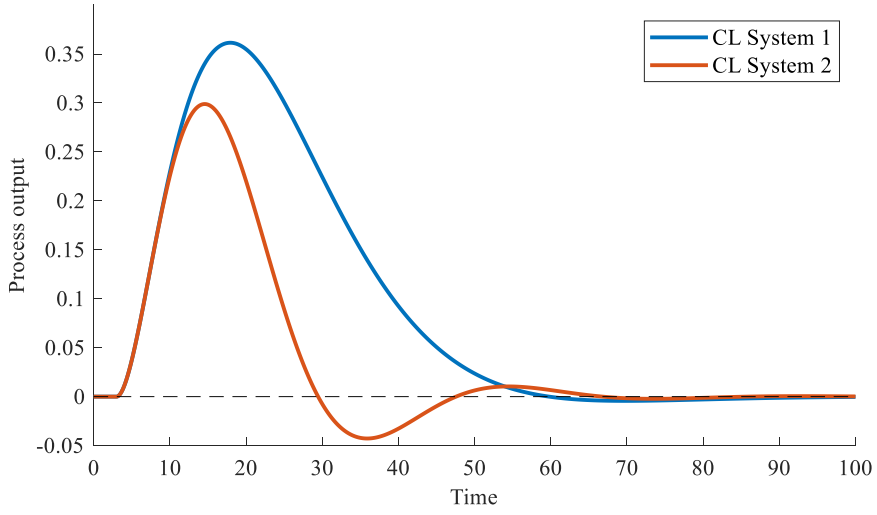

Fig. 2.16 The third comparison between the responses of two example closed loop systems to a step load disturbance.

Two major conclusions can be drawn from the above analysis:

1. It is impossible to explicitly assess control performance using only a single CPI, as none of them can fully describe the dynamic behavior of the control system. Rather, CPIs describe only chosen features of the closed loop response. Analysis using multiple CPIs can provide a more accurate description of dynamic behavior, and thus a more accurate assessment of performance. Hence, the number and type of CPIs that are required to effectively capture the key features of a closed loop response should be investigated.
2. The definition of satisfactory closed loop response—that is, control performance—depends on the precise technological requirements and constraints. For example, one technological process may require that a setpoint trajectory is followed as quickly as possible, with the existence of overshoots and undershoots being permissible. However, for a different application a lack of overshoots may be mandatory, and thus the behavior of the control system should be more conservative. As such, the reference dynamic behavior can be different for every application.

At this initial stage of the study, the greatest possible number of CPIs should be defined to ensure that all key features are captured and no important information is omitted. To systematize selection of CPIs, the closed loop response is divided into three suggested stages, as shown in Fig. 2.17. The first stage is defined as the time between the introduction of a load disturbance to the closed loop system and the maximum value of the response being reached. This duration of this stage depends on process dynamics, transportation time delay, and the initial action (aggressivity) of the controller. The second stage encompasses the damping of the maximum peak. Finally, the third stage describes the period in which the system reaches a steady state. To fully describe the dynamic behavior of the control system, the suggested CPIs should capture features of each distinct stage and of the entire response.

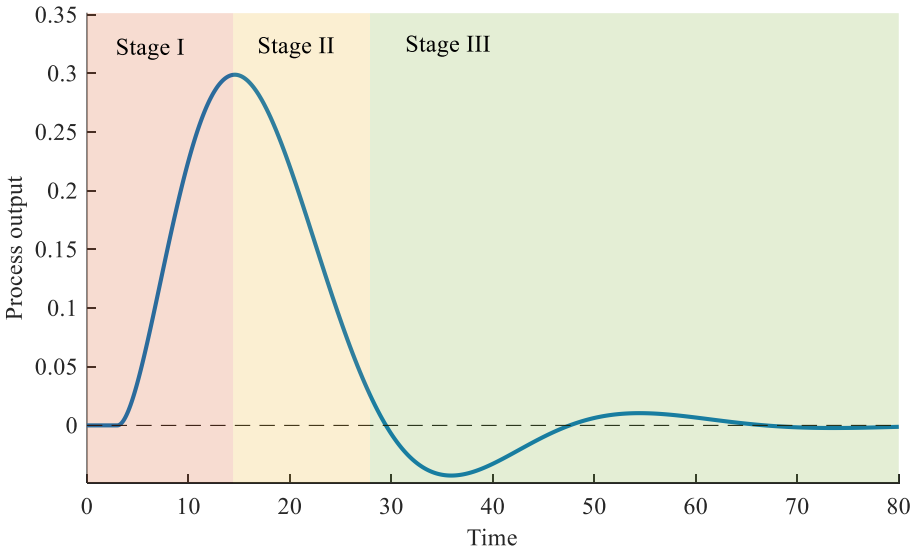

Fig. 2.17 The division of a closed loop response into three suggested stages.

Considering the above specification, a total of thirty different CPIs were selected for further analysis. Table 2.5 presents an exhaustive list of these CPIs, together with short descriptions and appropriate acronyms. Appendix 2 contains graphical representations of the selected CPIs. The 12 CPIs that are highlighted in grey in Table 2.5 are those that are widely known and used by practitioners: overshoot (F1), undershoot (F3), the ratio of overshoot and undershoot (F5), settling time (F7), integral indices (F8–F11), decay ratio (F15, F16), and indices calculated using the response derivative (F28, F29). The first stage of the response is described by indices F1 and F28, as they evaluate the maximum peak value and the rate at which the output signal reaches this value. The second stage of the response is directly described by F29, which captures the maximum gradient of the output signal during of its decrease from the maximum peak value. The key features of third stage are indirectly captured by the decay ratio indices F15 and F16 and by the undershoot indices F3 and F5, as they describe the effectiveness with which the close loop response damps oscillations. Features of the overall response are assessed by indices F7–F11, which are based on the transient time in correlation with the control error *e*.

CPI	**Short description**	**Acronym**
MaxPeak	The maximum value of the dynamic system response	F1
MaxPeakTime	The time at which the maximum peak occurs	F2
MinPeak	The minimum value of the dynamic system response (absolute value)	F3
MinPeakTime	The time at which the minimum peak occurs	F4
MinToMax	The ratio of the minimum and maximum peaks	F5
MaxToMinTime	The difference in time between the occurrence of the maximum and minimum peaks $MaxToMinTime = MinPeakTime - MaxPeakTime$	F6
SettlingTime	The time at which the system response returns to within the range of 1% of its steady state value $\|e\| < 0.01$	F7
IAE	The integral absolute error $IAE = \int \|e\| dt$	F8
ISE	The integral square error $ISE = \int e^2 dt$	F9
ITAE	The integral time absolute error $ITAE = \int t\|e\| dt$	F10
IT2AE	The integral time square absolute error $IT2AE = \int t^2\|e\| dt$	F11
IAEPos	The integral absolute error calculated for positive values of the system response $IAEPos = \int \|e\| dt, e > 0$	F12
IAENeg	The integral absolute error calculated for negative values of the system response $IAENeg = \int \|e\| dt, e < 0$	F13
IAENegToPos	The ratio of *IAENeg* and *IAEPos*	F14
DecayRatio	The ratio of the maximum peak value to the second positive peak value $DecayRatio = 2^{nd} Peak/MaxPeak$	F15
DecayRatioTime	The difference in time between the occurrence of the maximum and second peaks $DecayRatioTime = 2^{nd} PeakTime - MaxPeakTime$	F16
PeakSettlingTime	The difference between *SettlingTime* and *MaxPeakTime*	F17
TimePos	The total time during which the response of the system is positive $TimePos = \int dt, e > 0$	F18
TimeNeg	The total time during which the response of the system is negative $TimeNeg = \int dt, e < 0$	F19
TimeNegToPos	The ratio of *TimeNeg* and *TimePos*	F20
RisingTime	The rising time of the maximum peak, calculated as the time required for the system response to move from 5% to 95% of *MaxPeak*	F21

FallingTime	The falling time of the maximum peak, calculated as the time the time required for the system response to move from 95% to 5% of *MaxPeak*	F22
RisingToFallingTime	The ratio of *RisingTime* and *FallingTime*	F23
25%DistRejected	The time at which the response of system reaches 25% of *MaxPeak*, $\|e\| < 25\%*MaxPeak$	F24
50%DistRejected	The time at which the response of system reaches 50% of *MaxPeak*, $\|e\| < 50\%*MaxPeak$	F25
75%DistRejected	The time at which the response of system reaches 75% of *MaxPeak*, $\|e\| < 75\%*MaxPeak$	F26
ZeroCrossingTime	The first time at which the response of the system reaches zero	F27
MaxDiff	The maximum value of the derivative of the dynamic response	F28
MinDiff	The minimum value of the derivative of the dynamic response (absolute value)	F29
DiffMaxToMin	The ratio of *MaxDiff* and *MinDiff*	F30

Table 2.5 The exhaustive list of CPIs used for the synthesis of the described CPA system, including short descriptions and acronyms.

The remaining 18 CPIs are novel, and were proposed for this study. The introduction of new CPIs was motivated by the desire to capture more dynamic features of the closed loop response. Thus, the large majority of the novel CPIs are intended to supplement the 12 commonly used indices described above. Well known CPIs F1, F3, and F5 are supplemented with F2, F4, and F6 respectively to capture time domain features. They describe the time required to reach the maximum and minimum peaks, in addition to the ratio of the peak values. The widely used *IAE* index is supplemented by indices F12–F14, which describe changes in the sign of the control error *e*. Based on this sign, additional integral time-based indices were suggested (F18–F20). These indices more explicitly assess the overall closed loop response. Key points during the closed loop response in the time domain are detected by indices F17 and F24–F27, which supplement the settling time index F7. The exceptions to this supplementary behavior are the rising and falling time indices F21–F23, which were introduced to fully capture key features of the maximum peak, and thus describe the first and second stages of the closed loop response.

After considering all of the suggested CPIs, some preliminary conclusions can be drawn:

1. Calculation of the suggested CPIs does not require a high degree of computational and memory resources. Only basic mathematical functions are required for the calculations, such as addition, multiplication, and comparison. However, indices F28–F30 are calculated based on the derivative of the output signal, which can be problematic in the presence of measurement noise. In this case, additional filtering can be required.
2. Potentially, CPIs that capture similar key features of the closed loop response are correlated. For example, a long maximum peak time (F2) is likely to result in a long rising time (F21). Alternately, a high ratio of maximum and minimum peak values (F5) is likely to indicate oscillations in the closed loop response and a possible increase in decay ratio (F15). Moreover, many CPIs are defined as the ratio between other CPIs. Thus, a high correlation between some CPIs is to be expected and will be studied further.

2.4 Generating closed loop reference responses

A central requirement of the suggested CPA system is that information from multiple CPIs is combined to fully describe the dynamic behavior of the control system. This behavior should then be compared with a predefined reference response. Such a comparison is a general assumption of many CPA algorithms, including the Harris, GMV, LQG, and optimal PID indices, among others. The reference response can be described by reference PID tunings that are defined for a considered SOPDT process described by fixed L_1 and L_2 parameters. However, the definition of the reference behavior can vary depending on the differing technological requirements and constraints. Thus, the reference PID tunings can also differ, even for the same SOPDT process. Generally, as a reference, any suitable PID tuning rule can be chosen [24], or appropriate tunings can be generated using any other method, such as optimization-based tuning. In this work, the reference PID tunings are calculated by minimization of the *IAE* index for a closed loop response to a load disturbance. Following this, for fixed L_1, L_2 parameters and load disturbance, and for a fixed filter coefficient for the derivative part α = 0.3, the *IAE* value depends only on the PID tunings and can be calculated as

$$J = IAE(k_r, T_i, T_d) = \int_0^{t_{max}} |e(t)| dt, \tag{2.17}$$

where t_{max} is the settling time after introducing the load disturbance to the control system. However, the direct minimization of (2.17) without any constraints results in very aggressive tunings. In practice, this approach is not acceptable because constraints are required to ensure proper robustness of the control system. Robustness is defined by general frequency-based limitations: the gain A_m and phase ϕ_m margins. Introducing the gain and phase margins as additional constraints results in the following three-dimensional constrained optimization problem:

$$\begin{aligned} &\underset{k_r, T_i, T_d \epsilon \mathcal{R}^+}{minimize}\ IAE(k_r, T_i, T_d) \\ &\quad subject\ to \qquad A_m \geq A_{m_{min}} \\ &\qquad\qquad\qquad\ \ \phi_m \geq \phi_{m_{min}} \end{aligned} \tag{2.18}$$

where the limit values for the gain and phase margins are denoted as $A_{m_{min}}$ and $\phi_{m_{min}}$, respectively. This work assumes that $A_{m_{min}}$ = 2.5 and $\phi_{m_{min}}$ = 60°, which results in conservative tunings. However, this method can provide different tunings by adjusting the thresholds. For example, lower threshold values will produce more aggressive closed loop behavior.

The reference closed loop responses were computed within discretized L_1, L_2 space consisting of a mesh of equidistant points with ΔL = 0.1. This produced 60 SOPDT processes, evenly distributed throughout the considered ranges of L_1 and L_2. Note that a denser mesh can be used if greater accuracy is required. Then, for each considered process the optimization problem (2.18) was solved numerically using an interior point approach [85]. Each optimization problem was solved multiple times, with each

solution starting from a different initial point to avoid reaching local minima. This approach generated reference PID tunings $k_{r,ref}$, $T_{i,ref}$, and $T_{d,ref}$ for each of the considered SOPDT processes. The tunings are presented in Fig. 2.18–2.20 and Table 2.6–2.8.

$k_{r,ref}$		L_2									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
L_1	**0.1**	4.051	3.862	3.623	3.551	3.532	3.559	3.607	3.672	3.752	3.833
	0.2	2.189	2.177	2.223	2.281	2.338	2.393	2.442	2.487	2.526	2.586
	0.3	1.484	1.471	1.506	1.558	1.617	1.680	1.743	1.806	1.871	1.934
	0.4	1.089	1.076	1.093	1.127	1.169	1.217	1.267	1.318	1.369	1.421
	0.5	0.823	0.813	0.822	0.842	0.869	0.901	0.935	0.972	1.011	1.050
	0.6	0.634	0.629	0.633	0.644	0.660	0.680	0.702	0.726	0.752	0.779

Table 2.6 Obtained reference values of the controller gain.

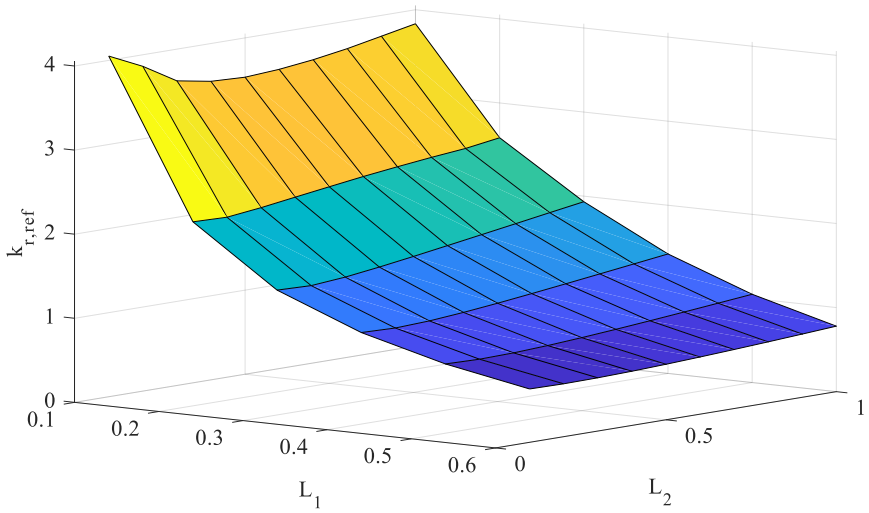

Fig. 2.18 A graphical representation of the obtained reference values of controller gain.

$T_{i,ref}$		L_2									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
L_1	**0.1**	0.577	0.604	0.632	0.725	0.801	0.874	0.939	0.999	1.058	1.107
	0.2	0.838	0.919	0.986	1.041	1.082	1.116	1.139	1.159	1.176	1.226
	0.3	1.004	1.087	1.169	1.249	1.324	1.394	1.458	1.518	1.574	1.624
	0.4	1.137	1.209	1.287	1.369	1.452	1.536	1.617	1.697	1.774	1.849
	0.5	1.267	1.327	1.397	1.472	1.552	1.635	1.719	1.805	1.891	1.976
	0.6	1.413	1.464	1.534	1.611	1.694	1.781	1.865	1.943	2.021	2.103

Table 2.7 Obtained reference values of the integral constant.

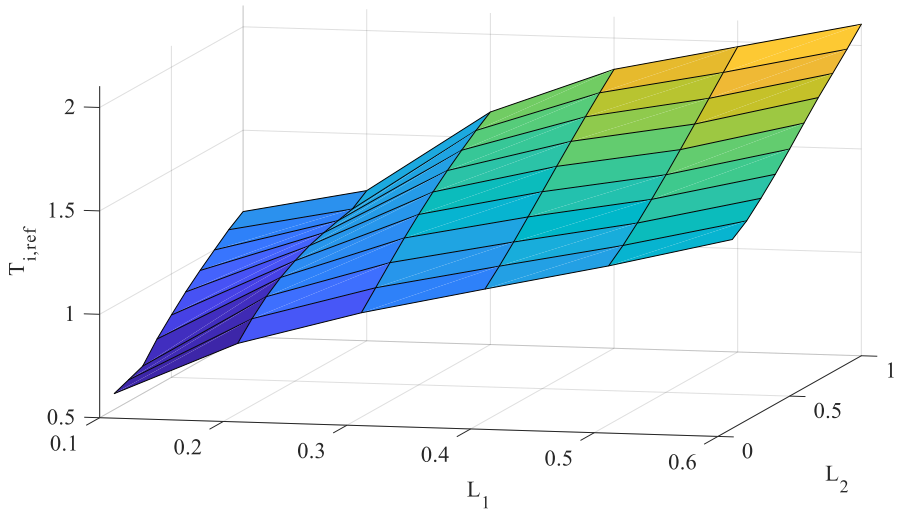

Fig. 2.19 A graphical representation of the obtained reference values of the integral constant.

$T_{d,ref}$		L_2									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
L_1	**0.1**	0.125	0.217	0.309	0.353	0.392	0.425	0.455	0.481	0.504	0.527
	0.2	0.150	0.211	0.272	0.332	0.390	0.448	0.506	0.563	0.619	0.644
	0.3	0.181	0.227	0.274	0.320	0.365	0.407	0.447	0.485	0.520	0.554
	0.4	0.217	0.251	0.289	0.328	0.365	0.400	0.434	0.466	0.496	0.525
	0.5	0.261	0.286	0.316	0.348	0.381	0.413	0.443	0.471	0.498	0.524
	0.6	0.360	0.387	0.408	0.428	0.445	0.461	0.478	0.499	0.518	0.538

Table 2.8 Obtained reference values of the derivative constant.

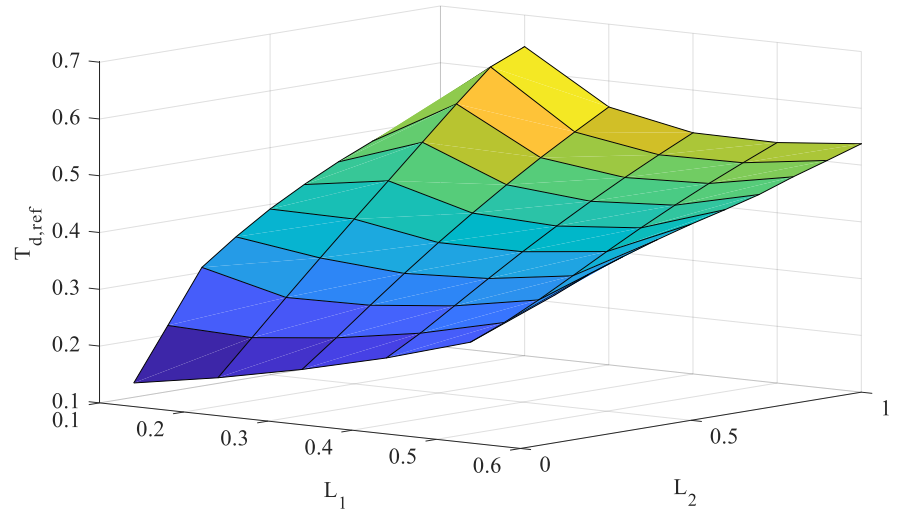

Fig. 2.20 A graphical representation of the obtained reference values of the derivative constant.

For each of the considered closed loop systems defined by the SOPDT process, in addition to the reference PID tunings computed as described above, corresponding values of gain and phase margins can be calculated. Given that they are strictly correlated with the reference PID tunings, these variables

are referred to as reference gain $A_{m,ref}$ and phase $\phi_{m,ref}$ margins. Fig. 2.21 and Fig. 2.22 present the values of the reference gain and phase margins, respectively. Note that for the vast majority of considered closed loop systems, the limiting values of the gain and phase margins as predefined for the optimization problem (2.18) and preserved as equality constraints, with $A_{m,ref} = 2.5$ and $\phi_{m,ref} = 60^o$. Only for the case $L_1 = 0.1$ and $L_2 > 0.1$ is the reference gain margin $A_{m,ref} > 2.5$, and only for $L_1 > 0.5$ is the reference phase margin $\phi_{m,ref} > 60^o$. In each case, the constraints assumed for (2.18) are not violated.

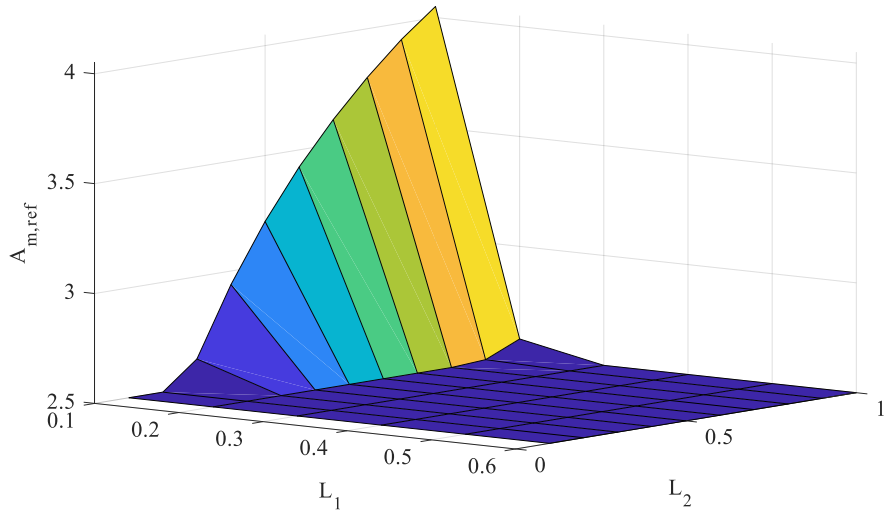

Fig. 2.21 A graphical representation of the obtained reference values of gain margin.

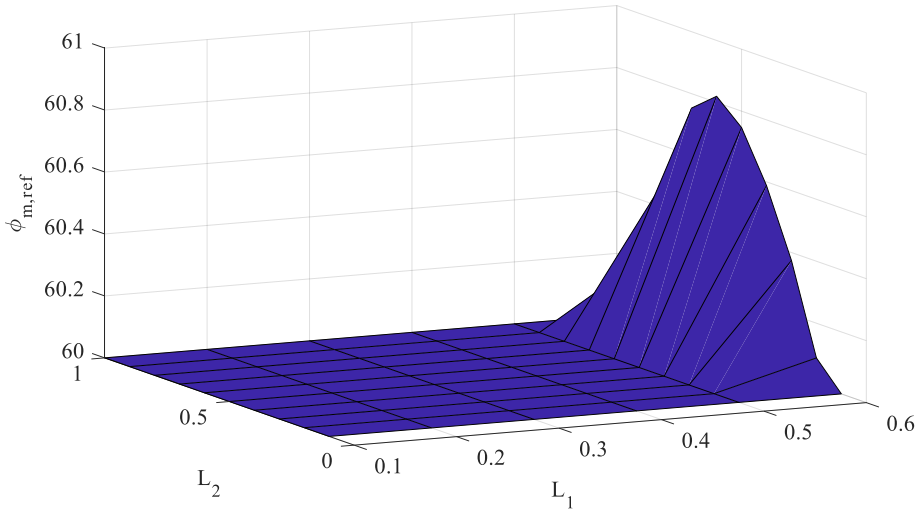

Fig. 2.22 A graphical representation of the obtained reference values of phase margin.

Fig. 2.23 presents examples of the reference closed loop responses corresponding to selected SOPDT processes. Note that despite differences in the values of the process parameters L_1 and L_2, each of the presented responses display similar behavior.

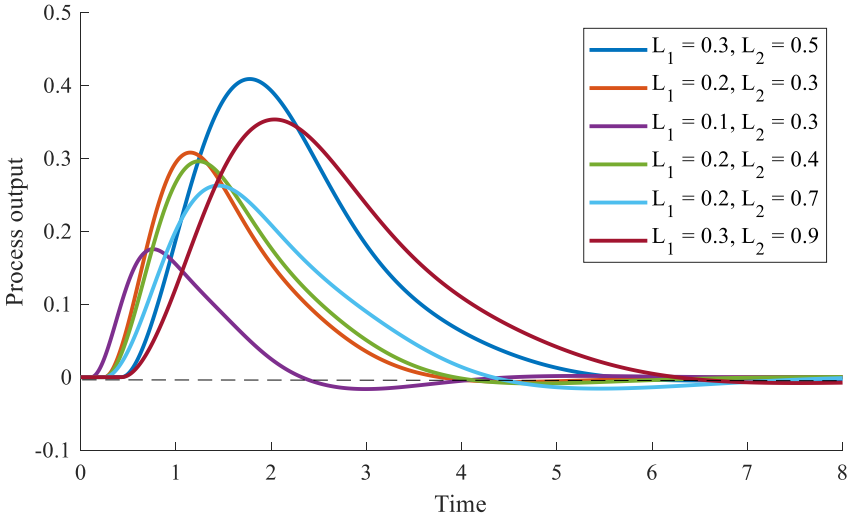

Fig. 2.23 Example reference responses obtained for the selected closed loop systems.

The optimization-based approach that was used to obtain reference PID tunings was practically justified. However, if necessary, the approach could be modified to meet specific technological requirements. Additionally, the assumed ranges of the L_1 and L_2 parameters can be divided into smaller, local sub regions within the L_1, L_2 plane, and the reference PID tunings then computed for each sub region individually. This approach could be used to compensate for the nonlinear nature of the process.

The computation of the PID reference tunings as described above assumed that the operation of the closed loop system processes can be described by SOPDT dynamics. However, industrial processes are very often characterized by higher order dynamics. As such, it is reasonable to investigate the difference between the reference PID tunings calculated directly for a closed loop system running higher order processes, and the tunings calculated for their corresponding SOPDT approximations. If the difference proves to be insignificant, the proposed method of computing the reference PID tunings can be directly used for higher order processes based only on their SOPDT approximation.

This investigation studies a closed loop system that controls a fourth order (4OP) process with the following dynamics:

$$K(s) = \frac{1}{(1+s)(1+sP_1)(1+sP_2)(1+sP_3)}, \tag{2.19}$$

where P_1, P_2, and P_3 represent the relative time constants. The ranges of P_1, P_2, and P_3 were chosen to cover the entire landscape of possible dynamics, and discretized with mesh size 0.1. Thus, $P_1 \, \epsilon \, < 0.1 : 0.1 : \, 1.0 >$, $P_2 \, \epsilon \, < 0 : 0.1 : \, P_1 >$, and $P_3 \, \epsilon \, < 0 : 0.1 : \, P_2 >$. Following this, all of the selected 4OP processes were approximated by the SOPDT model. For each 4OP process and its SOPDT approximation, the optimization problem (2.18) was solved, resulting in two separately

computed sets of reference PID tunings: one for the closed loop system with the 4OP process itself, and one for the closed loop system with the respective SOPDT process approximation. The relative differences between these reference PID tunings are given by

$$\Delta x_{Ref} = \frac{x_{Ref,4OP} - x_{Ref,SOPDT}}{x_{Ref,4OP}} * 100\%, \tag{2.20}$$

where Δx_{Ref} is the relative difference between the reference tuning parameter calculated for the closed loop system with 4OP process $x_{Ref,4OP}$ and for the closed loop system with the respective SOPDT approximation $x_{Ref,SOPDT.}$ These differences were calculated for all PID tuning parameters, and the results are presented in Fig. 2.24–Fig. 2.26 in both the P_1, P_2, P_3 (left figures) and L_1, L_2 (right figures) subspaces.

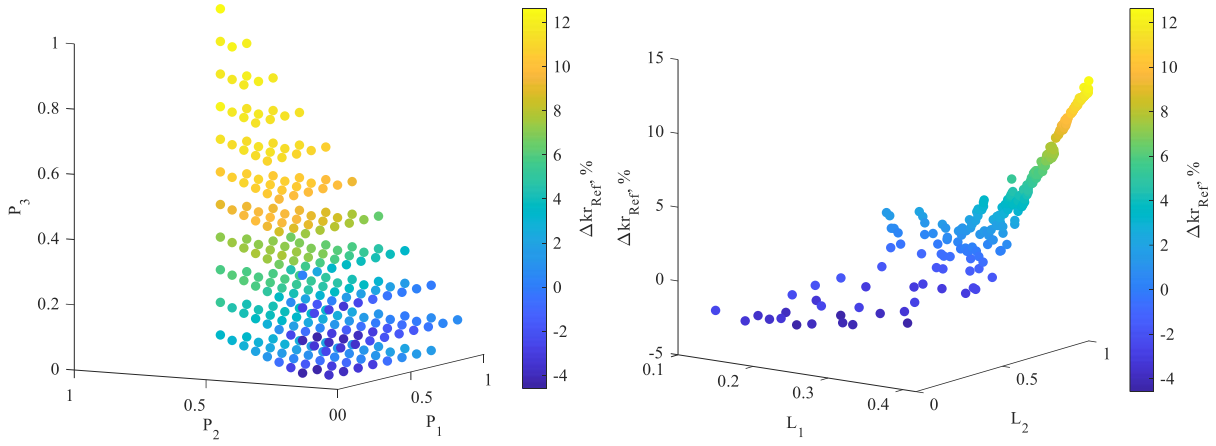

Fig. 2.24 The relative difference between the reference controller gain calculated for a 4OP process and its SOPDT approximation in the P_1, P_2, P_3 (left) and L_1, L_2 subspace (right).

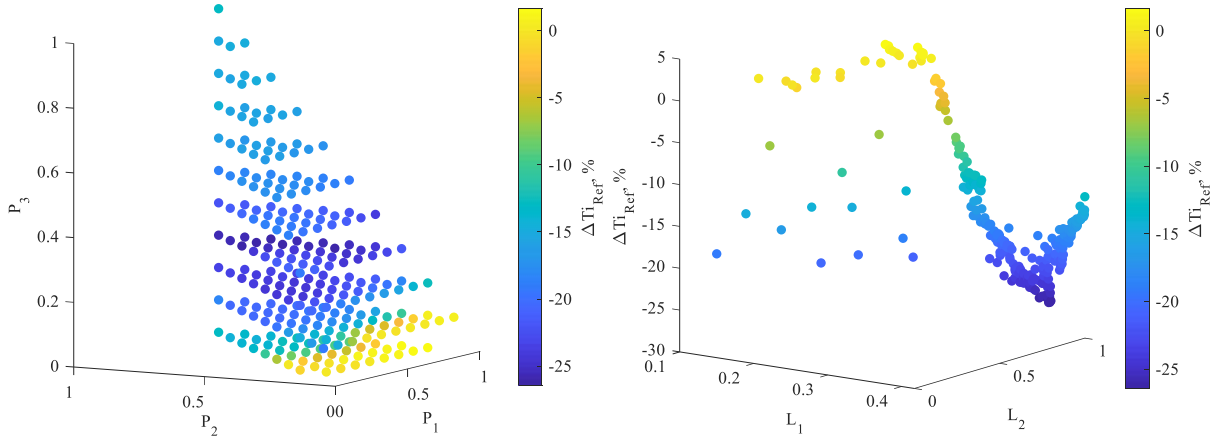

Fig. 2.25 The relative difference between the reference integral constant calculated for a 4OP process and its SOPDT approximation in P_1, P_2, P_3 subspace (left) and L_1, L_2 subspace (right).

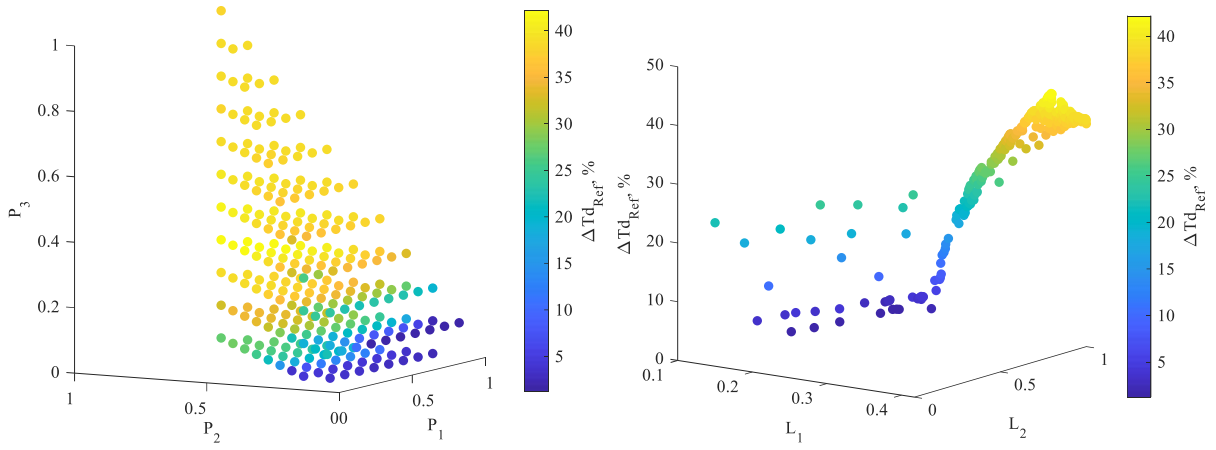

Fig. 2.26 The relative difference between the reference derivative constant calculated for a 4OP process and its SOPDT approximation in P_1, P_2, P_3 subspace (left) and L_1, L_2 subspace (right).

For further investigation, Fig. 2.27 presents the relative difference between the reference value of the *IAE* index computed using (2.17) for a closed loop system with a 4OP process, and the same index for a closed loop system with the corresponding SOPDT approximation. The graph indicates that the maximum relative difference of the *IAE* index is close to 16%. This difference is positive for some processes, and thus the *IAE* value is higher for the closed loop system with a 4OP process than for the closed loop system with the corresponding SOPDT approximation. However, both reference responses are very similar in terms of *IAE*.

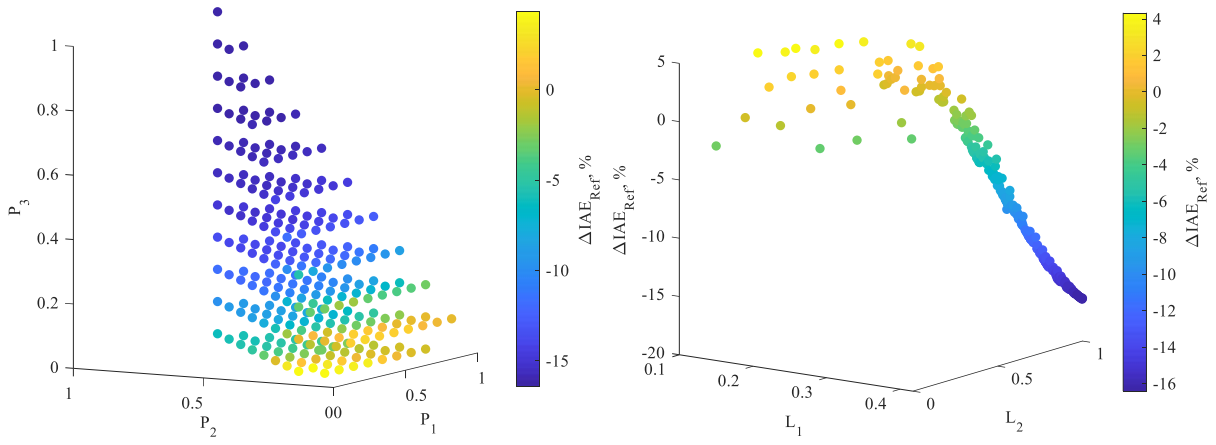

Fig. 2.27 The relative difference between the reference cost function (*IAE*) calculated a 4OP process and its SOPDT approximation in P_1, P_2, P_3 subspace (left) and L_1, L_2 subspace (right).

The smallest difference between the reference value of the *IAE* index was 0.0076 %, obtained for P_1 = 0.6, P_2 = 0.4, and P_3 = 0. Fig. 2.28 shows a comparison of the reference responses for both the 4OP process and the corresponding SOPDT approximation at this point in P_1, P_2, P_3 subspace. Conversely, the largest difference of -16.4 % was obtained for P_1 = 0.9, P_2 = 0.9, and P_3 = 0.9. Fig. 2.29 provides a corresponding comparison of the reference responses.

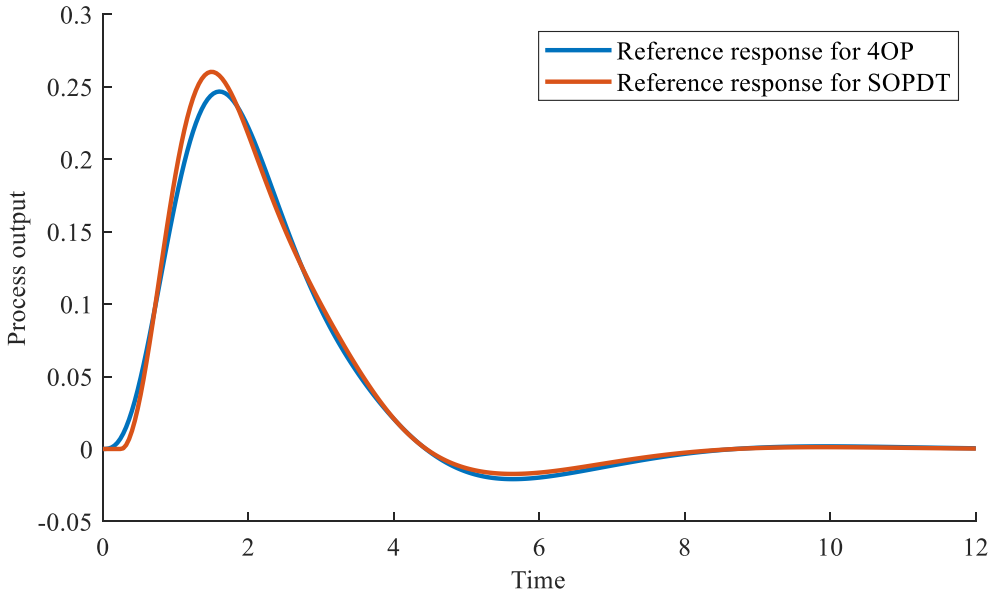

Fig. 2.28 A comparison of the reference responses for a 4OP process and its corresponding SOPDT approximation for $P_1 = 0.6$, $P_2 = 0.4$, $P_3 = 0$.

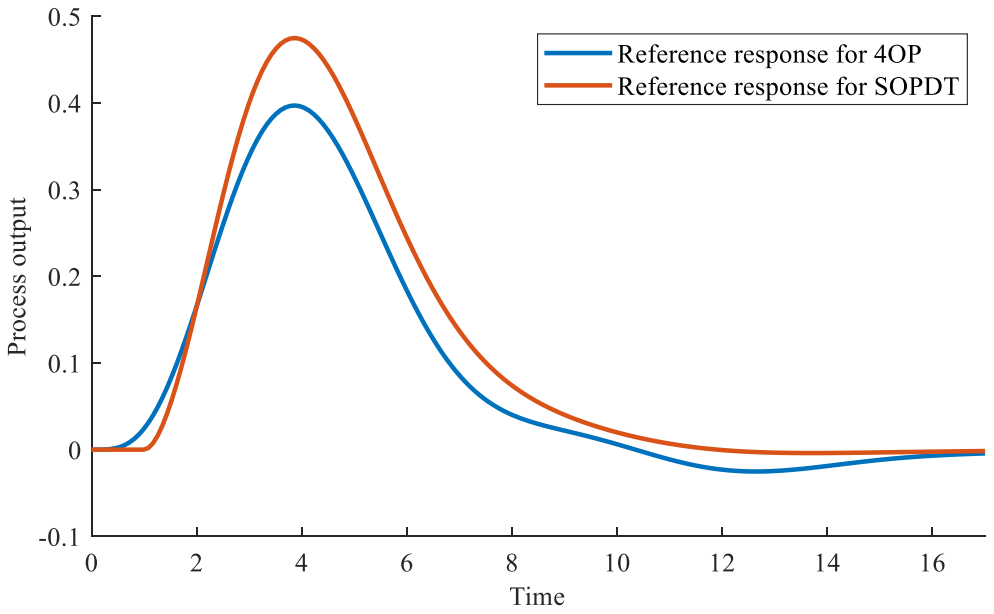

Fig. 2.29 A comparison of the reference responses for a 4OP process and its corresponding SOPDT approximation for $P_1 = 0.9$, $P_2 = 0.9$, $P_3 = 0.9$.

The robustness of a closed loop system with a 4OP process is predefined for the optimization problem (2.18). The degree to which a reference PID tuning based on a SOPDT approximation preserves this robustness should be investigated. To this end, for each considered closed loop system with a 4OP process, a PID controller was tuned using reference tunings obtained for the corresponding SOPDT process approximation. Then, based on the responses of simulated closed loop systems with 4OP processes and with 4OP and SOPDT reference tunings, the gain and phase margins were calculated, in an analogous manner to (2.20). The relative differences between the gain and phase margins are presented in Fig. 2.30 and Fig. 2.31, respectively.

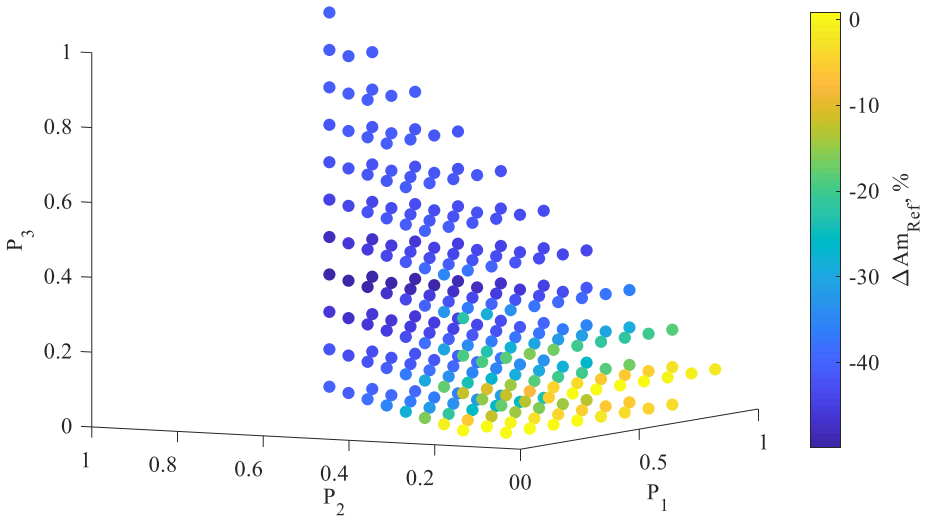

Fig. 2.30 The relative difference between the reference gain margins as calculated for a 4OP process using 4OP tunings and the corresponding SOPDT approximation tunings.

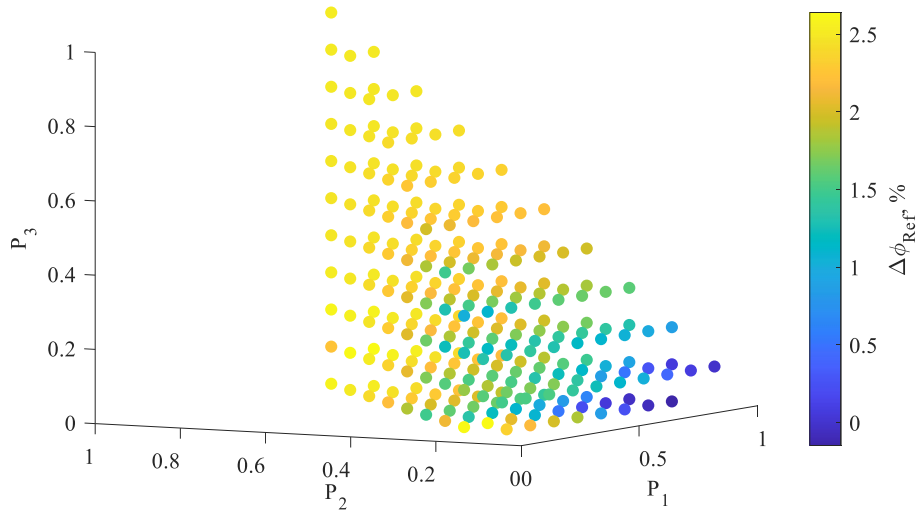

Fig. 2.31 The relative difference between the reference phase margins as calculated for a 4OP process using 4OP tunings and the corresponding SOPDT approximation tunings.

For the vast majority of closed loop systems with 4OP processes, the relative differences between the gain margins are negative. For these cases, the gain margins obtained using the SOPDT reference tunings are higher, and thus the system has a greater robustness. However, the relative differences for the phase margins are negative, and thus the system is less robust. However, the maximum difference is only 2.5%, which is acceptable from a practical perspective. To provide a complete overview, the relative difference of the *IAE* index was also calculated, and is presented in Fig. 2.32. The results show that the direct transfer of closed loop reference tunings computed using an SOPDT process approximation to closed loop systems with a 4OP process does not lead to a substantial decrease in performance, with a maximum difference of approximately 15%.

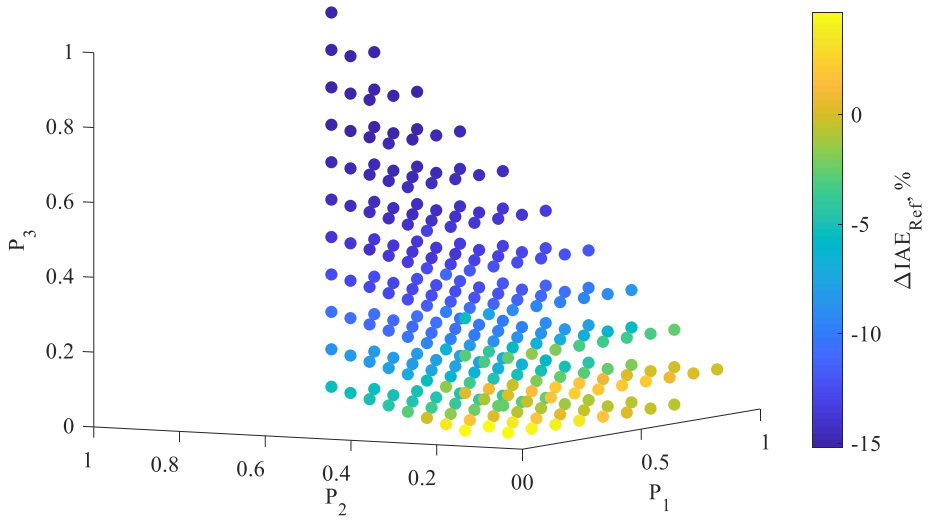

Fig. 2.32 The relative difference between the reference cost function (*IAE*) as calculated for a 4OP process using 40P tunings and the corresponding SOPDT approximation tunings.

To summarize, reference PID tunings calculated using the proposed method for SOPDT processes can be applied as *close to reference* for a relatively wide range of higher order processes. This was demonstrated using 4OP process as an example (2.19), with the approximated tunings showing no substantial degradation of performance when compared to reference PID tunings calculated directly for the considered 4OP processes. This result is of great practical importance, given that the practical determination of higher order dynamics ($n \geq 3$) is difficult when using real process data. Thus, the use of SOPDT approximations for deriving reference PID tunings is much more convenient from a practical perspective.

2.5 The generation of training and validation datasets

Following the definition of the general process model, the reference closed loop response, and a set of features, the next stage is to propose the general methodology for generating training and validation datasets that can be used to train and validate the proposed CPA system.

Both datasets contain a large set of different SOPDT processes defined by the L_1 and L_2 parameters, in addition to closed loop reference trajectories obtained for the PID reference tunings computed for each considered SOPDT process. To generate the datasets, a number of SOPDT processes were randomly drawn within the predefined ranges $L_1 \in \langle 0.1, 0.6 \rangle$ and $L_2 \in \langle 0.1, 1.0 \rangle$. For each generated process, PID reference tunings were computed to produce the closed loop reference response. As described in Section 2.4, the reference tunings were computed only for SOPDT processes defined by combinations of the parameters L_1 and L_2 that represent mesh points within the space $L_1 \in \langle 0.1 : 0.1 : 0.6 \rangle$, $L_2 \in \langle 0.1 : 0.1 : 1.0 \rangle$. Thus, for generated SOPDT processes defined by combinations of the parameters L_1 and L_2 that were between mesh points, the reference PID tunings were calculated by spline interpolation between reference tunings. The interpolation was calculated by solving the

optimization problem (2.18) for neighboring mesh points $L_1 \in <0.1 : 0.1 : 0.6>$, $L_2 \in <0.1 : 0.1 : 1.0>$. This approach substantially reduced the number of times that the optimization problem (2.18) required solving during the generation of the training and validation datasets. Moreover, it ensured that the reference tunings for each considered closed loop system were sufficiently accurate. However, this approach can produce PID reference tunings that are not optimal with respect to the optimization problem (2.18), and can violate the problem constraints slightly. If this limited accuracy is not acceptable, then the PID reference tunings can be calculated by solving the optimization problem (2.18) for each randomly generated SOPDT process, or by increasing the density of the L_1, L_2 subspace mesh to increase the accuracy of spline interpolation.

For each considered closed loop system, the control performance was assessed as satisfactory (OK) if the closed loop response was relatively similar to the corresponding reference response. Thus, for each closed loop system defined by a generated SOPDT process and a corresponding set of PID reference tunings, a very large number of different closed loop responses were generated to densely cover the region of satisfactory control performance. To this end, for the closed loop system defined by each randomly generated SOPDT process, the corresponding PID reference tunings were modified as follows:

$$\begin{aligned} k_{r,mod} &= a_1 k_{r,ref}, \\ T_{i,mod} &= a_2 T_{i,ref}, \\ T_{d,mod} &= a_3 T_{d,ref}, \end{aligned} \tag{2.21}$$

where $k_{r,mod}$, $T_{i,mod}$, and $T_{d,mod}$ are the modified PID tunings and a_1, a_2, and a_3 are random numbers that are taken from a normal distribution with $N(1, 0.0225)$. For PID tunings modified in this manner, the corresponding closed loop responses are generated by simulation. The degree by which the PID tunings are modified depends on the random values of a_1, a_2, and a_3. If the corresponding simulated closed loop response is then similar to the predefined reference response, then the response is assessed as OK. Otherwise, the generated closed loop response is assessed as NOK.

The similarity between the reference closed loop response and the corresponding response simulated using the modified PID reference tunings must be assessed. This can be accomplished by comparing the gain and phase margins for each response: $A_{m,ref}$ and $\phi_{m,ref}$ for the former and $A_{m,mod}$ and $\phi_{m,mod}$ for the latter. Hence, the similarity between the responses can be quantified by the distances $A_{m,dist}$ and $\phi_{m,dist}$, calculated as

$$A_{m,dist} = \frac{|A_{m,ref} - A_{m,mod}|}{A_{m,ref}}, \tag{2.22}$$

and

$$\phi_{m,dist} = \frac{|\phi_{m,ref} - \phi_{m,mod}|}{\phi_{m,ref}}. \tag{2.23}$$

A 10% deviation from the reference gain and phase margins is considered allowable. Hence, the control performance of a closed loop response generated using modified PID reference tunings (2.21) is assessed as OK if the distance $A_{m,dist} < 10\%$ and $\phi_{m,dist} < 10\%$.

However, initial studies indicate that an assessment based only upon the distances $A_{m,dist}$ and $\phi_{m,dist}$ provides incorrect results. There exist certain closed loop responses with a shape that diverges substantially from the reference response, but which preserve the distances $A_{m,dist} < 10\%$ and $\phi_{m,dist} < 10\%$. This problem can be illustrated using the normalized distance e_{dist} as a measure of the difference in shape between the closed loop control error samples generated the for modified PID tunings (2.21) e_{mod} and those generated for the PID reference tunings e_{ref}:

$$e_{dist} = \frac{\int |e_{ref} - e_{mod}| dt}{\int |e_{ref}| dt}. \tag{2.24}$$

For illustration, 60000 different SOPDT processes were randomly generated, with the modified PID tunings (2.21) ensuring stability of the closed loop system. Following this, 60000 corresponding closed loop responses were generated by simulation. Fig. 2.33 shows a graphical representation of this dataset, with each dot representing a generated closed loop response. To improve clarity, the amplitude $A_{m,dist}$ and phase $\phi_{m,dist}$ distances were normalized:

$$A_{m,norm} = \frac{A_{m,dist}}{10\%}, \tag{2.25}$$

$$\phi_{m,norm} = \frac{\phi_{m,dist}}{10\%}. \tag{2.26}$$

Thus, $|A_{m,norm}| \leq 1$ and $|\phi_{m,norm}| \leq 1$ indicate satisfactory (OK) control performance. This region is highlighted with a green box in Fig. 2.33.

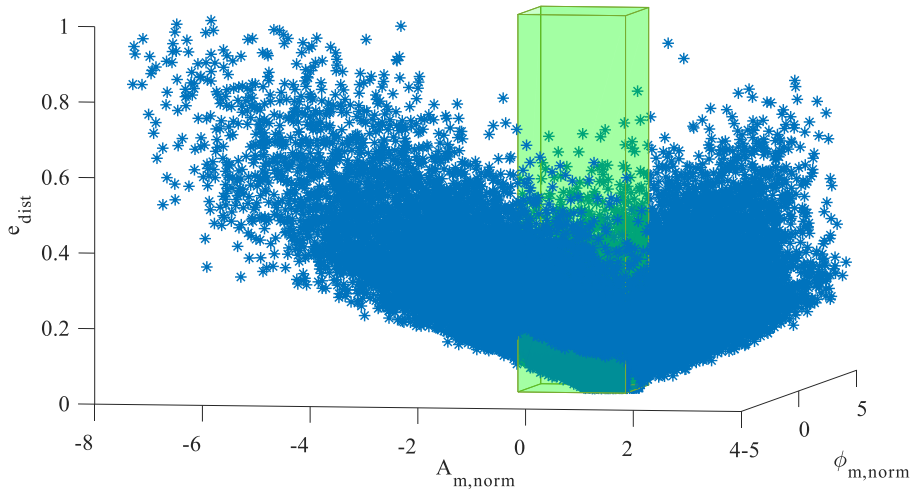

Fig. 2.33 A graphical representation of the generated training dataset, with the green box highlighting the region of satisfactory (OK) control performance.

For further analysis, the closed loop systems for which the gain and phase margins were preserved ($|A_{m,norm}| \leq 1$ and $|\phi_{m,norm}| \leq 1$) were selected from the generated dataset. Fig. 2.34 presents a histogram of the normalized distances e_{dist} of these closed loop systems. The histogram shows that the closed loop responses of the large majority of these systems are shaped very similarly to the corresponding reference response ($e_{dist} \approx 0$). However, there also exist many closed loop systems with preserved gain and phase margins that display a relatively high value of e_{dist}; the largest value of e_{dist} within the generated dataset is 0.51. Thus, despite the preservation of assumed robustness, the dynamic behavior of the closed loop systems with modified PID tunings is different from that of the corresponding reference systems.

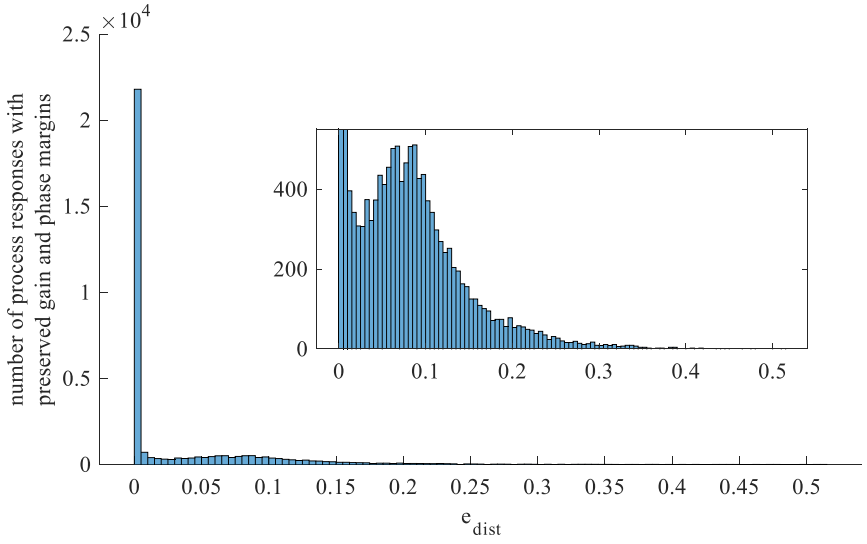

Fig. 2.34 A histogram of the normalized distances of the closed loop responses with preserved gain and phase margins.

As such, relying solely on the deviation of the gain and phase margins from their reference values is insufficient for the explicit assessment of control performance. Thus an additional condition must be introduced: for the control performance of a closed loop system to be assessed as satisfactory (OK), the normalized distance should be less than a predefined threshold value $e_{dist} < e_{dist,thresh}$. The value taken by the threshold $e_{dist,thresh}$ can be adjusted by rule of thumb, depending on the acceptable difference in shape between the closed loop response and the corresponding reference. The threshold value could also be determined by preliminary studies. By manipulating the distance threshold $e_{dist,thresh}$, only those closed loop systems that meet the condition $e_{dist} < e_{dist,thresh}$ will be extracted. From this, the percentage of closed loop systems that preserve the gain and phase margins can be determined, with respect to all such extracted systems. This approach was applied using threshold values $e_{dist,thresh} \in (0; 1.0]$, with the results presented in Fig. 2.35. For $e_{dist,thresh} < 0.05$, almost 100% of the closed loop systems are assessed as providing satisfactory (OK) control performance in terms of preserving the gain and phase margins. However, as the value of the distance threshold increases, this percentage decreases. These results suggest that a value of $e_{dist,thresh} = 0.1$ is appropriate for further studies, given that for this value almost 96% of all closed loop systems preserve the gain and phase margins.

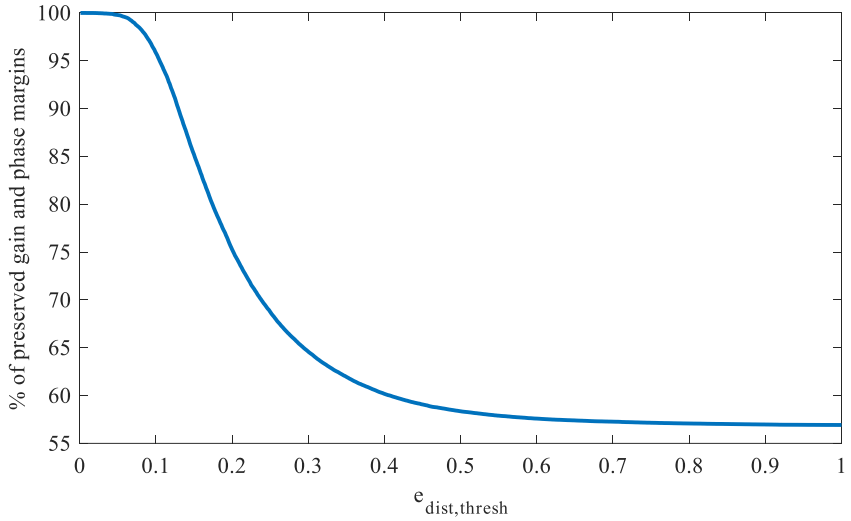

Fig. 2.35 The percentage of closed loop responses that preserve the gain and phase margins for various values of distance threshold $e_{dist,thresh}$.

One can notice, that basing on normalized distance e_{dist} with predefined threshold value $e_{dist,thresh} = 0.1$, the closed loop response can be assessed in terms of deviations of gain and phase margins with accuracy of almost 96 %. Process parameters L_1 and L_2 must be determined at initial stage of assessment for normalization of closed loop response (see Section 2.2), thus they could be potentially used for generating reference response (based on reference tunings, see Section 2.4) and normalized distance (2.24) could be calculated. Then, the control performance assessment can be determined based on threshold value $e_{dist,thresh} = 0.1$ defined based on the above considerations. However, in this

thesis, this possibility was rejected due to the fact that even if the presented results are complete, they should be treated as an important but also initial stage of developing the final CPA system. It is assumed that this final version will not require identification of the process model and the classification will be based only on a number of features computed directly from measurement data of closed loop disturbance rejection response. Then, any assessment basing only on the value of the normalized distance is not possible.

In summary, a generated closed loop response can be assessed as satisfactory (OK) if the following three conditions are met:

1. $|A_{m,dist}| < 10\%$,
2. $|\phi_{m,dist}| < 10\%$,
3. $e_{dist} < 0.1$.

Using these criteria, the generated dataset can be easily labelled. Fig. 2.36 presents example results of this process. The green and red dots represent closed loop responses with satisfactory (OK) and poor (NOK) control performance, respectively. For clarity, the region of satisfactory control performance is highlighted with a green box.

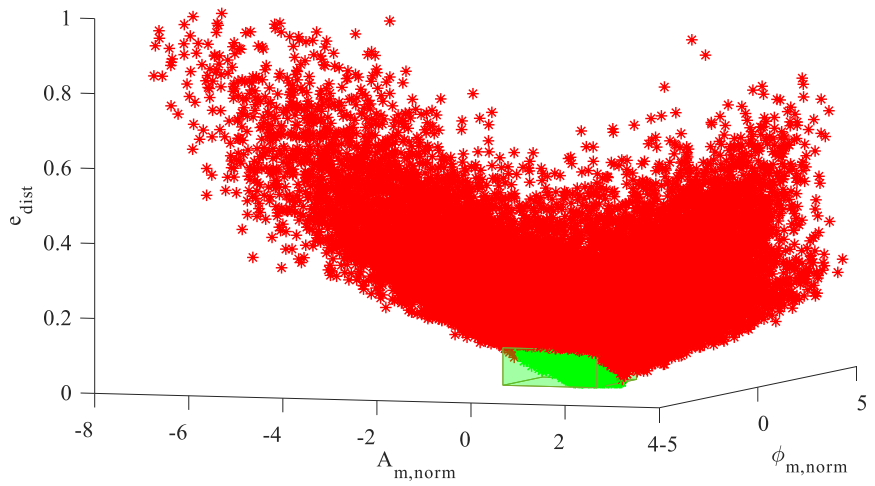

Fig. 2.36 A graphical representation of the generated training dataset, labelled according to the suggested criteria. Green and red dots represent satisfactory (OK) and poor (NOK) control performance, respectively. The region of satisfactory control performance is highlighted with a green box.

Finally, using the proposed methodology, a dataset of 60000 closed loop systems was generated. Half of the systems were labelled as OK and the other half as NOK, in accordance with the suggested criteria. Then, for each closed loop response, the thirty CPIs defined above were calculated forming a corresponding feature vector.

The robustness of the suggested labelling method was verified for higher order processes (2.19). For each considered 4OP process, a SOPDT approximation was used to randomly generate 200 sets of PID

tunings. Based on these PID tunings 200 closed loop responses were generated, with half of them labelled as OK and the other half as NOK, in accordance with the suggested criteria. Following this, the PID tunings were directly transferred to closed loop systems with 4OP process. Another set of closed loop responses were then generated and labelled; now based upon the closed loop reference responses of the 4OP processes. The two sets of labels produced by this approach for each closed loop response were then compared. The degree of compliance between them is presented in Fig. 2.37.

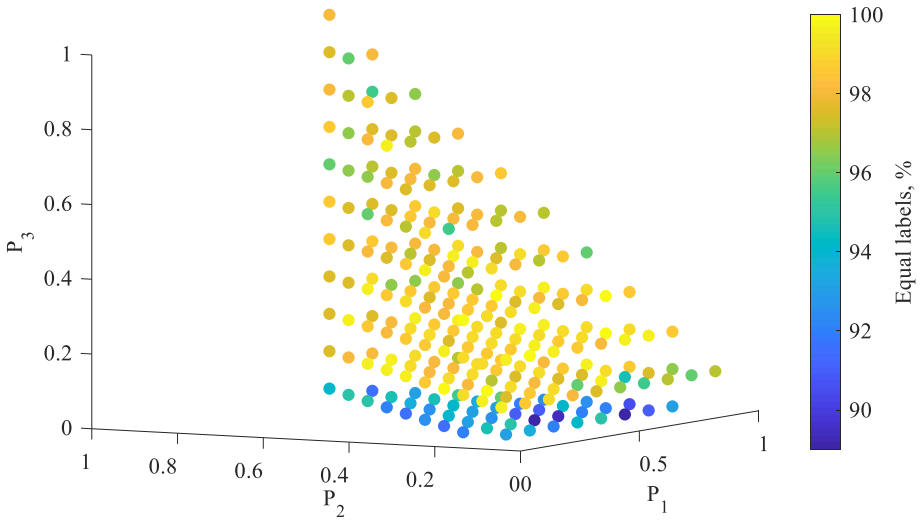

Fig. 2.37 The compliance of labels obtained using a 4OP process model and those obtained using an SOPDT approximation.

The results show very high compliance between labels generated directly for closed loop systems with 4OP processes and those generated for systems using SOPDT approximations. The minimum compliance level is 89%, although it is greater than 98% for the large majority of process. These results demonstrate that using SOPDT approximations to assess closed loop systems with higher order processes produces similar results to an assessment that directly considers higher order processes in the closed loop system. As described previously, the identification of higher order models using real measurement data is difficult. The ability to use SOPDT approximations that are relatively easy to identify significantly simplifies the assessment procedure. Thus, although the training dataset of the presented CPA systems was generated for assumed SOPDT processes, it can be effectively used for assessment of closed loop systems with higher order processes.

2.6 Selection of the classification algorithm

Various classification methods can be used to assess control performance as satisfactory or poor, based on the dataset described in the previous section. For this purpose, the following classification algorithms were selected, with varying degrees of complexity: Gaussian Naïve Bayes (GNB) [86], Linear Discriminant Analysis (LDA) [87], K-nearest Neighbors (KNN) [88], Decision Tree (DT) [89], and General Fuzzy Min-Max Neural Network trained by either an online learning algorithm (Onln-GFMM) [90] or an agglomerative learning algorithm (AGGLO-2) [91]. In addition, less transparent

but more powerful classification algorithms were selected, including kernel-based methods: Support Vector Machines (SVM) [92] and tree-based ensembles such as Light Gradient Boosted Machine (Light GBM) [93], Extreme Gradient Boosting (XGBoost) [94], Adaptive Boosting (AdaBoost) [95], Extremely Randomized Trees (Extra Trees) [96], and Random Forest (RF) [97]. Hyper-parameters of the selected models (except GNB and LDA) were obtained using random search within given ranges. This ensured the highest possible classification accuracy, as calculated based on 5-fold cross validation. Optimal values of the individual parameters together with their assumed ranges are presented in Table 2.9.

Classification algorithm	**Parameter**	**Range**	**Optimal value**
Decision Trees	Max depth	[4, 20]	19
	Min samples per leaf	[4, 30]	4
Light GBM	Max depth	[4, 20]	20
	Min samples per leaf	[4, 30]	12
	Sampling rate	{0.3, 0.4, 0.5, 0.6, 0.7}	0.4
	% features used	{20%, 30%, 40%, 50%, 60%, 70%}	70%
	Learning rate	{0.025, 0.05, 0.1, 0.2, 0.3}	0.3
	No. of estimators	{30, 50, 70, 100, 150, 200}	200
XGBoost	Max depth	[4, 20]	8
	Sampling rate	{0.3, 0.4, 0.5, 0.6, 0.7}	0.7
	% features used	{20%, 30%, 40%, 50%, 60%, 70%}	70%
	Learning rate	{0.025, 0.05, 0.1, 0.2, 0.3}	0.2
	Gamma	{0, 0.1, 0.2, 0.3, 0.4, 1, 1.5, 2}	1
	No. of estimators	{30, 50, 70, 100, 150, 200}	200
Extra Trees	Max depth	[4, 20]	20
	Min samples per leaf	[4, 30]	6
	% features used	{20%, 30%, 40%, 50%, 60%, 70%}	40%
	Sampling rate	{0.3, 0.4, 0.5, 0.6, 0.7}	0.7
	No. of estimators	{30, 50, 70, 100, 150, 200}	50
Random Forest	Max depth	[4, 20]	20
	Min samples per leaf	[4, 30]	6
	% features used	{20%, 30%, 40%, 50%, 60%, 70%}	40%
	Sampling rate	{0.3, 0.4, 0.5, 0.6, 0.7}	0.7
	No. of estimators	{30, 50, 70, 100, 150, 200}	50
AdaBoost	Max depth	[4, 20]	11
	Min samples per leaf	[4, 30]	12
	No. of estimators	{30, 50, 70, 100, 150, 200}	150
	Learning rate	{0.001, 0.01, 0.1, 0.2, 0.5, 1}	0.1
Support Vector Machines	Kernel	{'rbf', 'sigmoid', 'linear'}	rbf
	Gamma	{2^-15, 2^-13, …, 2^3}	8
	C	{2^-5, 2^-3, …, 2^15}	512
K-nearest Neighbor	K	{1, 3, …, 29}	5
Onln-GFMM	Maximum hyperbox size θ	{0.1, 0.15, …, 0.55, 0.6}	0.1
AGGLO-2	Maximum hyperbox size θ	{0.1, 0.15, …, 0.55, 0.6}	0.4

Table 2.9 Individual parameters together with their assumed ranges for the considered classification algorithms, obtained via a hyper-parameter optimization procedure.

The generated dataset of 60000 samples was used to train the selected classifiers. An additional validation dataset, consisting of 10000 samples, was generated using similar criteria to the generation of the training dataset. Fig. 2.38 presents the classification accuracy of the selected algorithms when applied to all 30 considered CPIs (red bars). The majority of the selected algorithms provide a classification accuracy of greater than 91%, and are thus suitable for the CPA classification problem. The highest accuracy, of more than 96%, was obtained when using the SVM classifier. The classification accuracy was also verified for training and validation datasets consisting of the 12 commonly used CPIs (those highlighted in grey in Table 2.5). The results are also presented in Fig. 2.38 (the blue bars). The results show that when restricting the CPIs to only those that are commonly used, the classification accuracy significantly decreases for the large majority of classifiers. Only the SVM classifier displays substantial robustness, as its accuracy does not significantly decrease. It can be concluded that the use of all 30 CPIs rather than the most common 12 positively affects the classification algorithms, as more nuanced features of the closed loop response can be captured.

Furthermore, the simple linear models (GDA or LDA) provide a classification accuracy of less than 80%—an insufficient level of accuracy. Significantly higher performance is obtained by the nonlinear models. This suggests that the decision boundary between the OK and NOK classes is strongly nonlinear, and cannot be effectively captured by linear models, such as GDA or LDA. Thus, nonlinear algorithms are the most appropriate for the CPA classification problem. A relatively high classification accuracy was obtained for models such as DT, KNN, or AGGLO-2, which are complex but more interpretable than the black-box models, such as SVM or tree-based ensemble models. The best performance was obtained by powerful nonlinear classifiers such as SVM, in addition to the following nonlinear kernel and boosted ensemble classifiers: Light GBM, AdaBoost, and XGBoost.

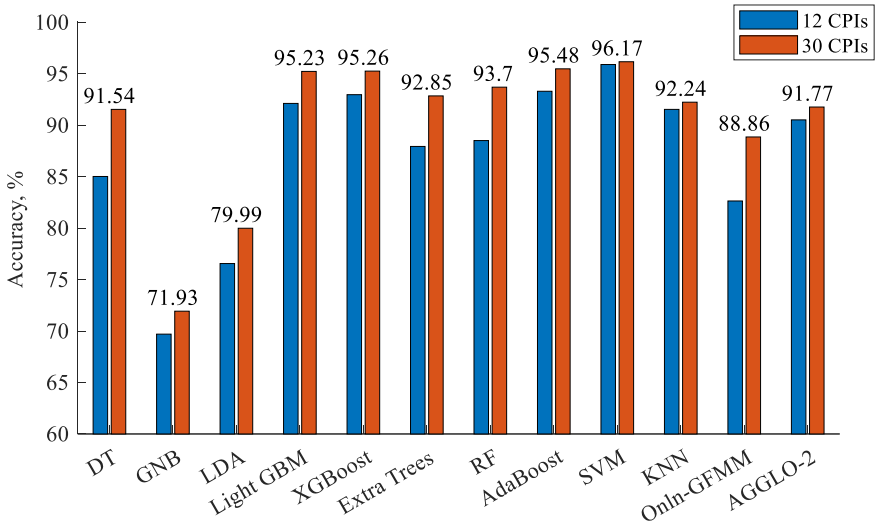

Fig. 2.38 The classification accuracy obtained by each of the considered algorithms, when trained on the 12 commonly used CPIs (blue bar plot) and all 30 CPIs (red bar plot).

Although the classification accuracy of the fuzzy models (Onln-GFMM, AGGLO-2) was lower than that of the SVM classifier, the fuzzy-model membership functions can be used to provide a more precise assessment of control performance, by determining the distance between the assessed closed loop response and the OK and NOK control performance boundary. The membership functions for the OK and NOK classes were calculated for the validation dataset. The histograms for correctly assessed closed loop responses are presented in Fig. 2.39 and Fig. 2.40. The histograms for incorrectly assessed closed loop responses are presented in Fig. 2.41 and Fig. 2.42. The results show that for correctly assessed closed loop responses the membership functions of the correct class (OK for Fig. 2.39 and NOK for Fig. 2.40) dominates the incorrect class, with a value of close to 1. Moreover, the distribution of the membership functions of the incorrect classes has a somewhat high deviation.

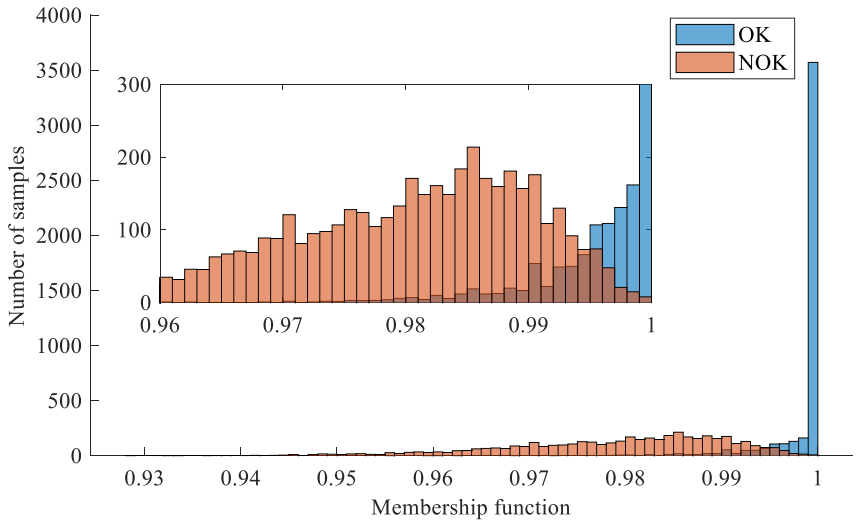

Fig. 2.39 Histograms of the membership functions of the OK and NOK classes, calculated for closed loop responses that are correctly assessed as OK.

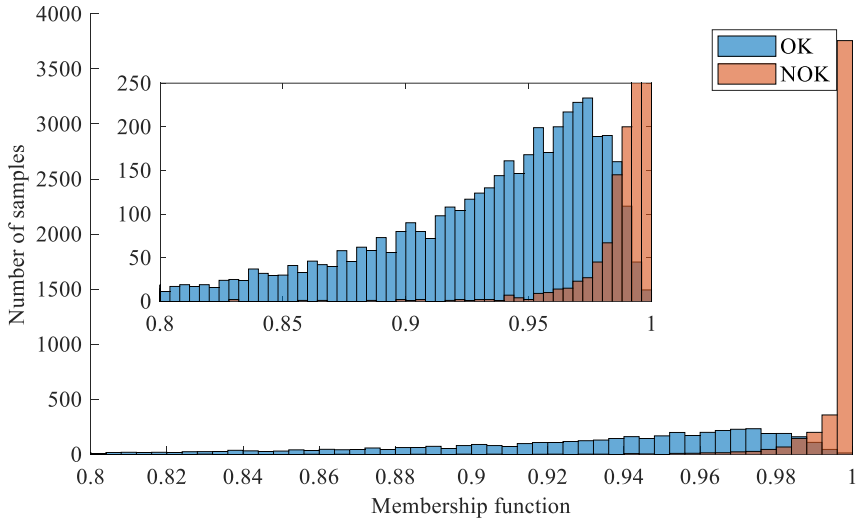

Fig. 2.40 Histograms of the membership functions of the OK and NOK classes, calculated for closed loop responses that are correctly assessed as NOK.

As expected, the incorrect class dominates in the case of incorrectly assessed control performance (OK for Fig. 2.41 and NOK for Fig. 2.42). However, in this case the membership function of the correct class is also very close to unitary value. The distributions of both membership functions are clustered around unitary value, and the deviation is small.

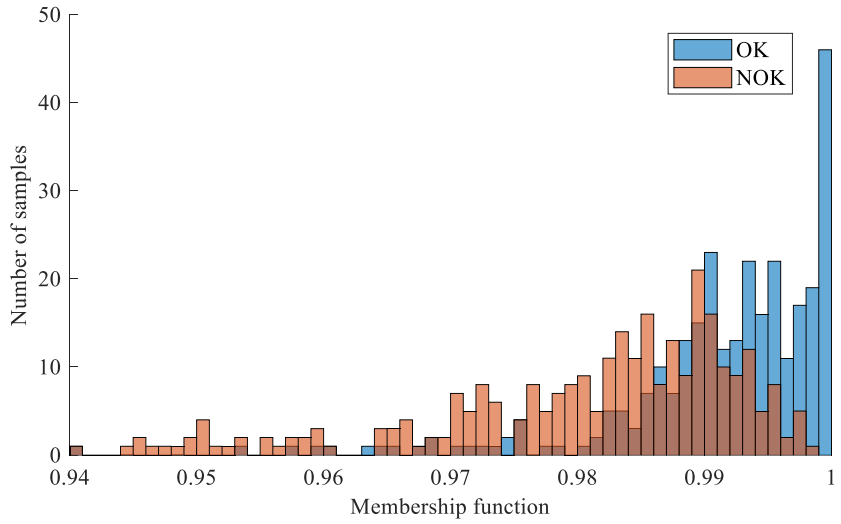

Fig. 2.41 Histograms of the membership functions of the OK and NOK classes, calculated for closed loop responses that are incorrectly assessed as OK.

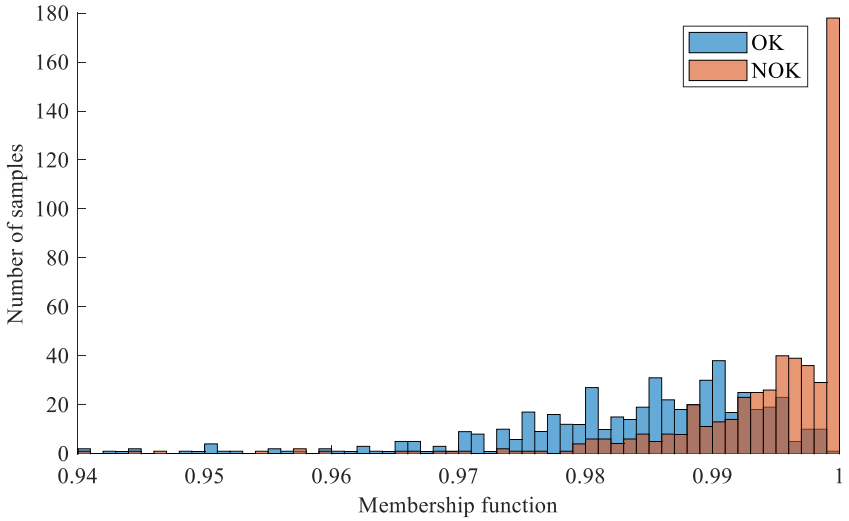

Fig. 2.42 Histograms of the membership functions of the OK and NOK classes, calculated for closed loop responses that are incorrectly assessed as NOK.

Finally, the differences between the membership functions for the OK and NOK classes were calculated for both correctly and incorrectly assessed closed loop responses. The normalized histogram of calculated differences is presented in Fig. 2.43. The histogram shows that the differences between the membership functions in the case of incorrectly assessed closed loop responses are close to zero. In these cases, the closed loop response is close to the boundary between OK and NOK performance,

and thus explicit assessment can be more difficult. In the case of correct assessment, the calculated differences are higher, and thus membership of the given class is more obvious.

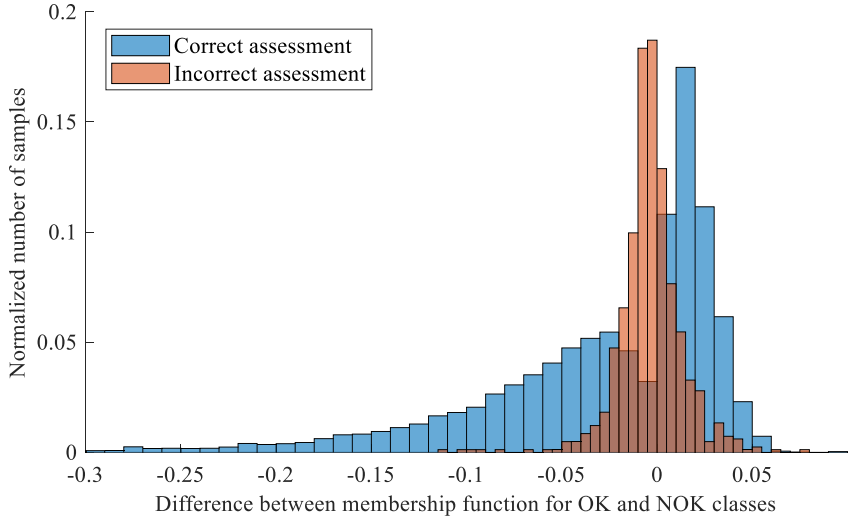

Fig. 2.43 Normalized histograms of the calculated differences between membership functions for OK and NOK classes, for correctly (blue) and incorrectly (red) assessed closed loop responses.

Information concerning the membership functions, such as the difference between the functions for the OK and NOK classes, can be used to determine the effectiveness of the CPA algorithms. This approach can potentially be used to monitor the degradation of control performance. If this occurs, a gradual decrease in the OK membership function and corresponding increase in the NOK membership function should be noticed. Such an investigation provides an interesting direction for further studies, but is beyond the scope of this thesis.

2.7 Simulated and experimental verification

The practical utility of the suggested CPA system was verified based on simulations and experimental studies. All of the presented results were obtained using the SVM classifier, as this was shown to achieve the highest possible classification accuracy amongst all investigated classifiers (see Section 2.7). The classifier was trained off-line using the methodology described previously.

2.7.1 Simulated validation of the system with SOPDT processes

The system was initially validated using simulations. Two different SOPDT processes were selected, defined by $L_1 = 0.4$, $L_2 = 0.5$ and $L_1 = 0.3$, $L_2 = 0.9$. For each process, 35 different PID tunings were obtained based on a FOPDT approximation of the process step response. These tunings are shown in Appendix 5. Following this, a load disturbance was introduced to each control system, and the closed loop response data was collected. The CPIs were then calculated, to produce a new validation dataset consisting of 35 different samples.

Fig. 2.44 presents the classification accuracy in the form of a confusion matrix for the SOPDT process with L_1 = 0.4, L_2 = 0.5. The matrix shows perfect classification accuracy, with 100% of responses being assessed correctly. As such, control performance was assessed well. Fig. 2.44 also shows the validation dataset, with the green and red dots representing OK and NOK closed loop responses, respectively. The green box indicates the region of satisfactory (OK) control performance.

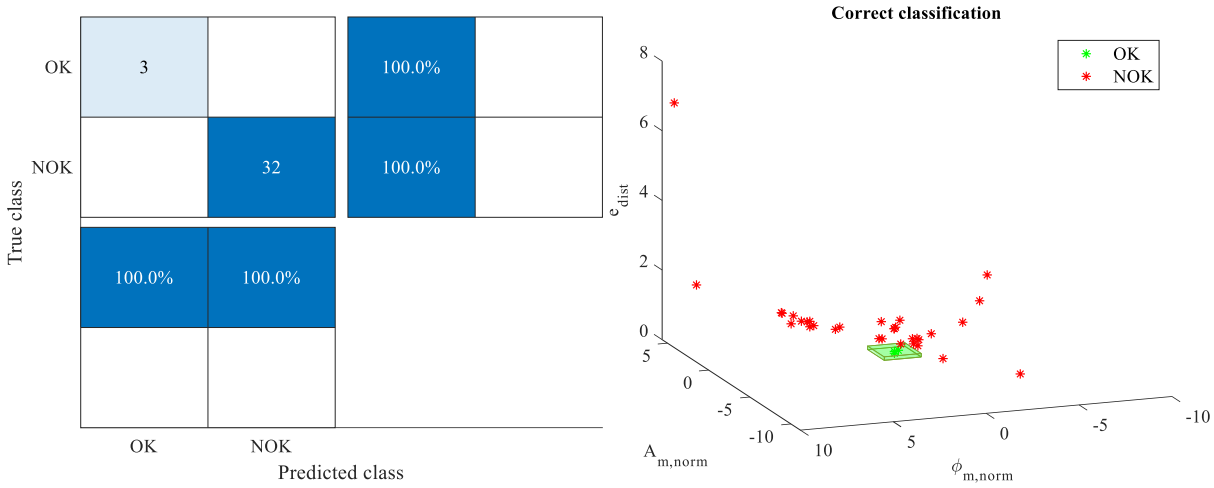

Fig. 2.44 The confusion matrix obtained for the simulation dataset with L_1 = 0.4, L_2 = 0.5 (left). A graphical representation of the simulation dataset with L_1 = 0.4, L_2 = 0.5 (right).

The closed loop responses from the generated validation dataset are presented in Fig. 2.45. The green and red responses represent OK and NOK closed loop responses respectively, and the black line represents the reference response. Those closed loop responses that are assessed as OK are relatively close to the corresponding reference response. However, those responses for which the dynamic behavior differs substantially from the predefined reference are assessed as NOK.

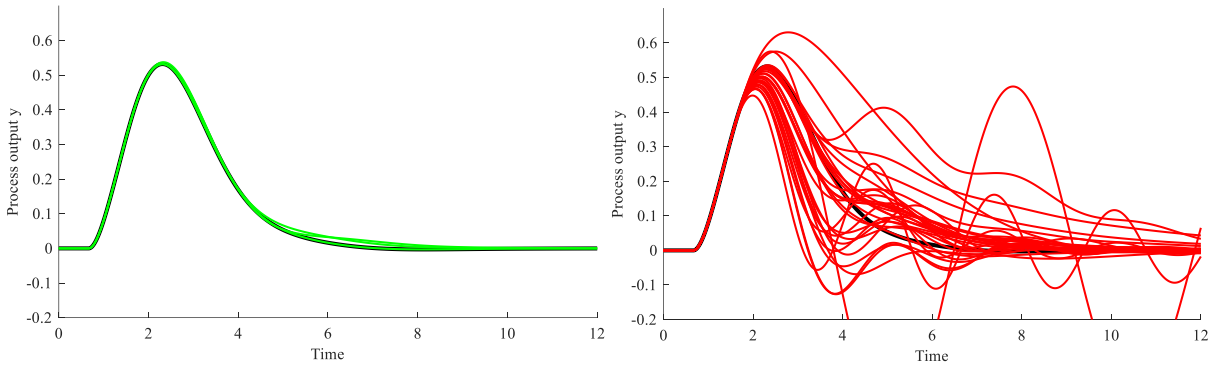

Fig. 2.45 The closed loop responses assessed as OK (left) and NOK (right), with the reference response indicated (black line), for the simulation dataset with L_1 = 0.4, L_2 = 0.5.

Using the same methodology, a validation dataset was generated for the second considered SOPDT process with L_1 = 0.3, L_2 = 0.9. However, for this process, one set of PID tunings lead to unstable behavior, and thus the corresponding closed loop system was excluded from the dataset.

The classification accuracy is presented in Fig. 2.46. The accuracy remains high, but is not perfect. The correctly classified samples are presented in Fig. 2.47. One NOK closed loop response was incorrectly classified as OK according to the assumed assessment criteria. However, this particular sample is very close to the decision boundary between OK and NOK performance, and thus is particularly difficult to assess accurately.

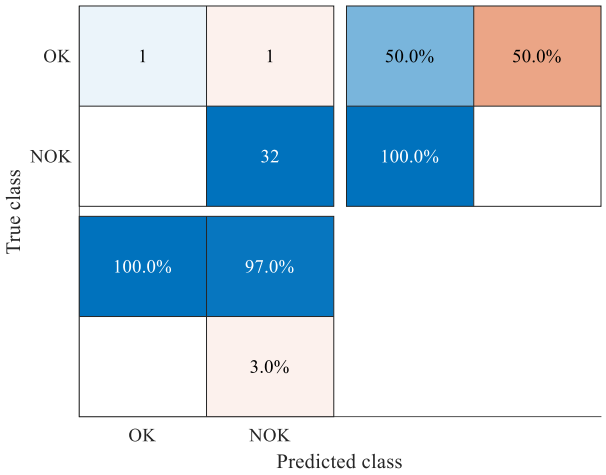

Fig. 2.46 The confusion matrix obtained for the simulation dataset with L_1 = 0.3, L_2 = 0.9.

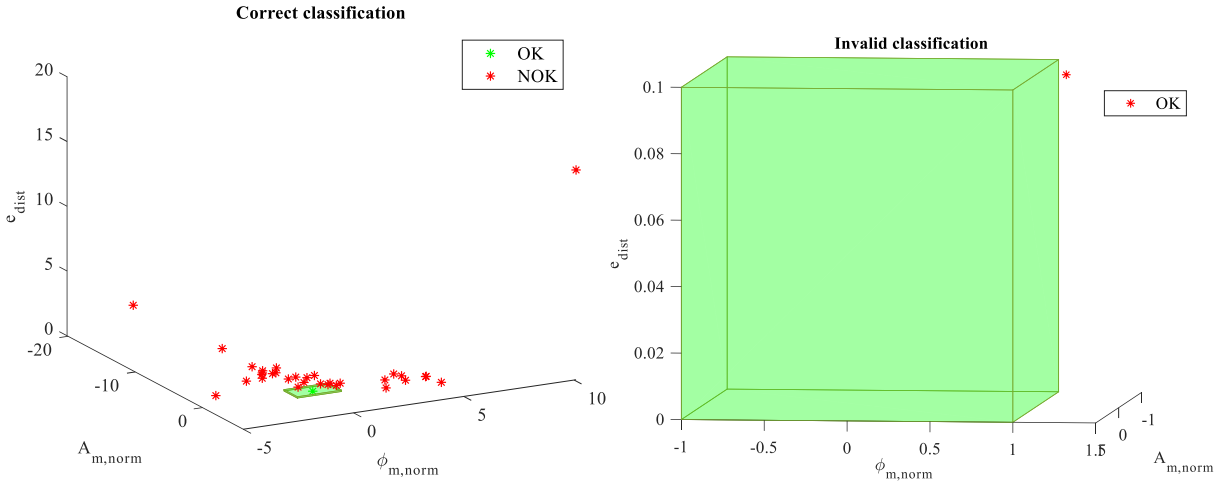

Fig. 2.47 A graphical representation of the correctly (left) and incorrectly (right) assessed samples of the simulation dataset with L_1 = 0.3, L_2 = 0.9.

Fig. 2.48 presents the closed loop responses for the second SOPDT process that were correctly assessed as either OK or NOK. The single response that was incorrectly assessed (as NOK) is presented in Fig. 2.49. This response is relatively close to the corresponding reference response, indicated in black. Thus, from a practical perspective this misclassification is insignificant. As described earlier, the use of fuzzy models and their related membership functions could potentially be helpful for such cases.

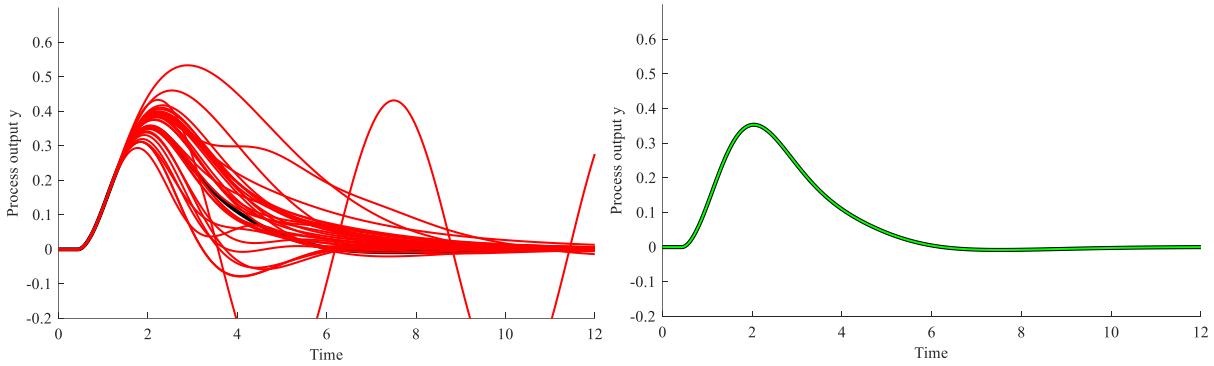

Fig. 2.48 The closed loop responses correctly assessed as OK (left) and NOK (right) with the reference response indicated (black line), for the simulation dataset with L_1 = 0.3, L_2 = 0.9.

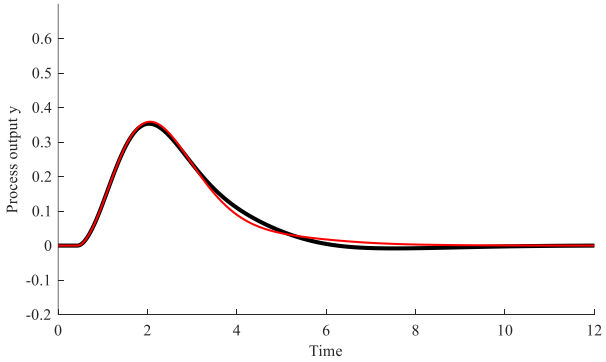

Fig. 2.49 The closed loop response incorrectly assessed as NOK, with the reference response indicated (black line), for the simulation dataset with L_1 = 0.3, L_2 = 0.9.

For completeness, the classification results, classification accuracies, and confusion matrices for the generated datasets are shown in Table 2.10 for the other considered classifiers. The results show that other classifiers, such as Random Forest or AdaBoost, provide a similar accuracy to that of the selected SVM classifier. However, the simulation datasets are less representative than the validation dataset described in Section 2.6, and thus these results cannot be used for an explicit comparison of the involved classifiers.

Classification algorithm	**Simulation dataset L_1 = 0.4, L_2 = 0.5**		**Simulation dataset L_1 = 0.3, L_2 = 0.9**	
	Confusion matrix	**Accuracy, %**	**Confusion matrix**	**Accuracy, %**
Decision Trees	$\begin{bmatrix} 3 & 0 \\ 0 & 32 \end{bmatrix}$	100	$\begin{bmatrix} 1 & 1 \\ 2 & 30 \end{bmatrix}$	91.17
Gaussian Naïve Bayes	$\begin{bmatrix} 2 & 1 \\ 3 & 29 \end{bmatrix}$	88.57	$\begin{bmatrix} 1 & 1 \\ 5 & 27 \end{bmatrix}$	82.35
Linear Discriminant Analysis	$\begin{bmatrix} 1 & 2 \\ 3 & 29 \end{bmatrix}$	85.71	$\begin{bmatrix} 1 & 1 \\ 1 & 31 \end{bmatrix}$	94.11
Light GBM	$\begin{bmatrix} 3 & 0 \\ 1 & 31 \end{bmatrix}$	97.14	$\begin{bmatrix} 1 & 1 \\ 0 & 32 \end{bmatrix}$	97.05
XGBoost	$\begin{bmatrix} 3 & 0 \\ 1 & 31 \end{bmatrix}$	97.14	$\begin{bmatrix} 1 & 1 \\ 0 & 32 \end{bmatrix}$	97.05
Extra tree	$\begin{bmatrix} 2 & 1 \\ 0 & 32 \end{bmatrix}$	97.14	$\begin{bmatrix} 1 & 1 \\ 1 & 31 \end{bmatrix}$	94.11
Random Forest	$\begin{bmatrix} 3 & 0 \\ 0 & 32 \end{bmatrix}$	100	$\begin{bmatrix} 1 & 1 \\ 0 & 32 \end{bmatrix}$	97.05

AdaBoost	$\begin{bmatrix} 3 & 0 \\ 0 & 32 \end{bmatrix}$	100	$\begin{bmatrix} 1 & 1 \\ 0 & 32 \end{bmatrix}$	97.05
Support Vector Machine	$\begin{bmatrix} 3 & 0 \\ 0 & 32 \end{bmatrix}$	100	$\begin{bmatrix} 1 & 1 \\ 0 & 32 \end{bmatrix}$	97.05
k-Nearest Neighbor	$\begin{bmatrix} 3 & 0 \\ 1 & 31 \end{bmatrix}$	97.14	$\begin{bmatrix} 1 & 1 \\ 0 & 32 \end{bmatrix}$	97.05
Onln-GFMM	$\begin{bmatrix} 3 & 0 \\ 1 & 31 \end{bmatrix}$	97.14	$\begin{bmatrix} 1 & 1 \\ 1 & 31 \end{bmatrix}$	94.11
AGGLO-2	$\begin{bmatrix} 2 & 1 \\ 0 & 32 \end{bmatrix}$	97.14	$\begin{bmatrix} 1 & 1 \\ 0 & 32 \end{bmatrix}$	97.05

Table 2.10 The classification results obtained for other classifiers as applied to both simulation datasets.

This validation shows that the proposed CPA system can accurately assesses the control performance of closed loop systems with SOPDT processes.

2.7.2 Comparison with existing CPA algorithms

The accuracy with which the suggested CPA system can assess performance was compared with other existing CPA methods. Based on the closed loop response data, the following CPA indices were calculated: the R index [98], the idle index [70], the area index [99], the load disturbance rejection performance (LDR) index [72], and the Harris index [51]. Note that the assessment procedure for each of these CPA algorithms, other than the Harris index, is similar to that required by the suggested CPA system. That is, each index is calculated based on the load disturbance rejection response of the closed loop system. However, the Harris index is a more complex CPA algorithm, and requires stochastic disturbances. For this purpose, load disturbances with different amplitudes were introduced to the assessed closed control system across multiple consecutive steps. The LDR index was calculated based on the reference PID tunings used for the synthesis of the suggested CPA system.

Explicit assessment using the Harris index is impossible. However, control performance can be assessed using the other indices based on the corresponding index values presented in Table 2.11.

R Index		**Idle Index**		**Area Index**		**LRP Index**	
NOK (oscillatory)	1.0	NOK (sluggish)	1.0	NOK (sluggish)	1.0	NOK	> 1.4
OK	0.5	OK / NOK (oscillatory)	-1.0	OK	0.5	OK	1.0
NOK (sluggish)	0.0			NOK (oscillatory)	0.0	NOK	< 0.6

Table 2.11 Values of the selected CPA algorithms that assessed performance as OK or NOK.

The comparison was conducted based on the generated closed loop responses of the datasets described in the previous section, for the same two SOPDT processes with L_1 = 0.4, L_2 = 0.5 and L_1 = 0.3, L_2 = 0.9, and the same PID tunings. The selected CPA indices were calculated for each of the processes, with the results presented in Table 2.12 andTable 2.13. Score Expert is the expected assessment based on the criteria suggested in Section 2.5, and Score SVM is the assessment made by the suggested CPA system. The results are color-coded: green and red denote OK and NOK closed loop systems, respectively. The results are also presented in graphical form in Fig. 2.50 and Fig. 2.51.

Number of response	Score Expert	Score SVM	R index	Idle index	Area index	LRP index	Harris index
1	OK	OK	0.6270	-0.1939	1.0000	0.9887	0.3871
2	NOK	NOK	0.5342	-0.6137	0.2740	1.1142	0.3767
3	NOK	NOK	0.3147	0.8930	1.0000	0.6107	0.4650
4	NOK	NOK	0.4462	0.6386	0.6642	0.9179	0.4429
5	NOK	NOK	0.5301	0.1070	1.0000	0.6856	0.3742
6	NOK	NOK	0.4820	0.1329	1.0000	0.8692	0.4328
7	NOK	NOK	0.6383	-0.7357	0.4337	1.2327	0.3649
8	NOK	NOK	0.3403	0.7884	0.0872	0.7650	0.4613
9	NOK	NOK	0.2313	0.8772	1.0000	0.5100	0.5009
10	NOK	NOK	1.0690	-0.4078	0.1180	0.1339	0.0329
11	NOK	NOK	0.9147	-0.6850	0.4050	1.3557	0.3552
12	NOK	NOK	0.4462	0.6386	0.6642	0.9179	0.4429
13	NOK	NOK	0.4379	0.1609	0.6630	0.9179	0.4471
14	NOK	NOK	0.4355	0.1929	1.0000	0.7915	0.4462
15	NOK	NOK	1.0494	-0.6838	0.1124	0.6052	0.1742
16	NOK	NOK	0.4937	-0.7260	0.3717	1.1246	0.3749
17	NOK	NOK	0.9086	-0.8995	0.3734	1.1001	0.3160
18	NOK	NOK	0.1959	0.1500	0.2548	0.4301	0.4384
19	NOK	NOK	0.9472	-0.6311	0.6057	1.4707	0.3411
20	NOK	NOK	0.9246	-0.6128	0.8035	1.3816	0.3294
21	NOK	NOK	1.0352	-0.6692	0.4962	1.3538	0.2849
22	NOK	NOK	1.0208	-0.6713	0.5252	1.3639	0.2860
23	NOK	NOK	0.4554	0.8496	0.0518	0.8383	0.4256
24	NOK	NOK	0.4467	0.8626	0.0153	0.8171	0.4278
25	OK	OK	0.5878	-0.1430	0.6564	0.9554	0.3938
26	NOK	NOK	0.5049	0.8188	1.0000	0.8451	0.4099
27	NOK	NOK	0.5658	-0.0262	0.5255	1.0327	0.4055
28	NOK	NOK	0.6176	-0.1033	0.4705	1.0622	0.4084
29	NOK	NOK	0.4086	-0.2740	0.5649	0.9235	0.4510
30	NOK	NOK	0.6576	-0.6575	0.5248	1.2462	0.3826
31	NOK	NOK	0.5083	0.1647	0.0177	0.8810	0.4185
32	NOK	NOK	0.6269	-0.6490	0.4588	1.2396	0.3962
33	NOK	NOK	0.5287	0.6664	0.3074	0.9263	0.4096
34	NOK	NOK	0.4567	0.3150	1.0000	0.4865	0.3590
35	OK	OK	0.6012	0.6574	0.6598	0.9449	0.3804

Table 2.12 The assessment of the 35 samples from the simulation dataset with L_1 = 0.4, L_2 = 0.5 by selected indices, and comparison with the expected assessment (Score Expert), and the assessment of the suggested CPA system (Score SVM).

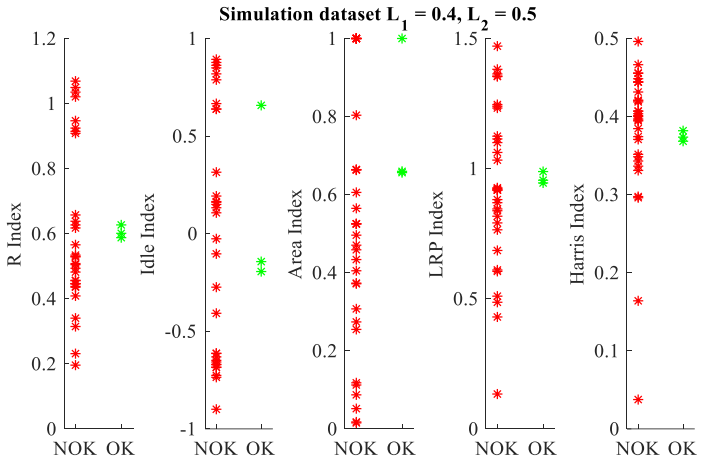

Fig. 2.50 A graphical representation of the assessment by selected CPA indices of the 35 samples from the simulation dataset with L_1 = 0.4, L_2 = 0.5.

Number of response	Score Expert	Score SVM	R index	Idle index	Area index	LRP index	Harris index
1	OK	OK	0.5564	-0.3342	1.0000	0.9480	0.0632
2	NOK	NOK	0.5452	-0.4309	0.6530	1.2876	0.0762
3	NOK	NOK	0.4469	0.8697	1.0000	0.6439	0.0562
4	NOK	NOK	0.5761	0.4196	0.8982	0.9679	0.0597
5	NOK	NOK	0.5807	-0.1245	1.0000	0.5419	0.0431
6	NOK	NOK	0.4487	0.8836	1.0000	0.6629	0.0581
7	NOK	NOK	0.6380	-0.0386	0.7204	1.1844	0.0603
8	NOK	NOK	0.4374	0.8596	0.4473	0.8065	0.0650
9	NOK	NOK	0.3185	0.8895	1.0000	0.5377	0.0655
10	NOK	NOK	0.8151	-0.5854	0.7924	1.5144	0.0618
11	NOK	NOK	0.5761	0.4196	0.8982	0.9679	0.0597
12	OK	NOK	0.5638	-0.1041	0.9153	0.9679	0.0610
13	NOK	NOK	0.5907	-0.2695	1.0000	0.7664	0.0525
14	NOK	NOK	0.5100	-0.6416	0.4977	1.3877	0.0772
15	NOK	NOK	0.4191	0.8592	0.2641	0.7031	0.0600
16	NOK	NOK	0.4369	0.8116	0.3438	0.7031	0.0566
17	NOK	NOK	0.2460	0.3929	1.0000	0.4205	0.0682
18	NOK	NOK	0.8555	-0.8055	0.8662	1.2802	0.0517
19	NOK	NOK	0.8346	-0.8716	0.8545	1.1781	0.0488
20	NOK	NOK	0.9194	-0.8441	0.8137	1.3401	0.0511
21	NOK	NOK	0.9224	-0.8046	0.8174	1.3444	0.0514
22	NOK	NOK	0.4275	0.8580	0.0375	0.6380	0.0566
23	NOK	NOK	0.4141	0.8624	1.0000	0.6095	0.0566
24	NOK	NOK	0.5613	0.8012	0.7716	0.7111	0.0493
25	NOK	NOK	0.4979	0.8528	0.0097	0.6292	0.0507
26	NOK	NOK	0.5161	0.8494	0.6189	0.7714	0.0559
27	NOK	NOK	0.5946	-0.4869	1.0000	0.8774	0.0556
28	NOK	NOK	0.3504	0.8927	0.0716	0.6902	0.0714
29	NOK	NOK	0.5251	0.7878	0.7306	0.9491	0.0623
30	NOK	NOK	0.4769	0.8548	0.0406	0.6482	0.0532
31	NOK	NOK	0.5032	0.7324	0.7265	0.9491	0.0650
32	NOK	NOK	0.5027	0.8475	0.4358	0.7030	0.0533
33	NOK	NOK	0.5263	0.1553	1.0000	0.3727	0.0382
34	NOK	NOK	0.5712	0.7623	0.7224	0.7171	0.0480

Table 2.13 The assessment of the 35 samples from the simulation dataset with L_1 = 0.3, L_2 = 0.9 by selected indices, and comparison with the expected assessment (Score Expert), and the assessment of the suggested CPA system (Score SVM).

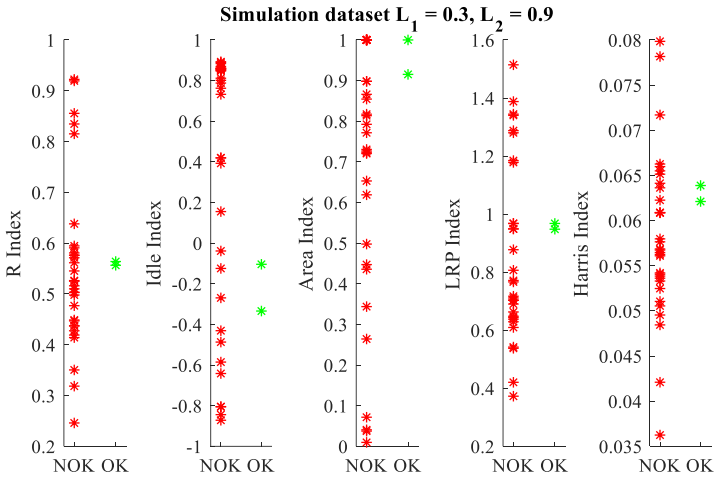

Fig. 2.51 A graphical representation of the assessment by selected CPA indices of the 35 samples from the simulation dataset with L_1 = 0.3, L_2 = 0.9.

The results indicate that OK and NOK samples cannot be distinguished by individual CPA indices. Thus, individual CPA indices are insufficient for explicit control assessment. In [99], a combination of the idle index and the area index are suggested for more precise assessment. However, the results here indicate that a combination of all of the selected CPA indices does not ensure accurate assessment.

Potentially, the selected CPA indices could be used to supplement the feature vector of the derived classification system, as they are more complex than simple CPIs and potentially provide more information. The supplemented feature vector could capture additional key features of the closed loop response, thus potentially improving the classification accuracy. This approach is beyond the scope of this thesis, but will be investigated in future works.

Closed loop responses numbers 16, 27, 28, and 29 were assessed as OK by all of the selected CPA indices, but as poor (NOK) by the predefined criteria for the suggested CPA system. These closed loop responses are presented in Fig. 2.52, together with the corresponding reference response. The dynamic behavior of the selected closed loop system is noticeably different from the predefined reference. Moreover, closed loop responses numbers 16, 27, and 29 display oscillatory behavior, which is not acceptable in industrial control systems.

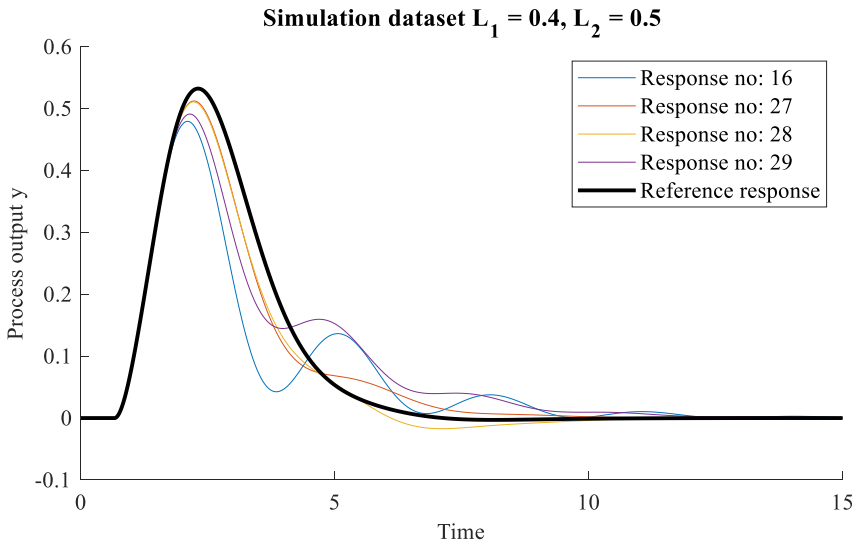

Fig. 2.52 Closed loop responses that were assessed as NOK by predefined criteria (Score Expert) and OK by selected CPA algorithms.

As described above, assessment using the Harris index requires a more aggressive excitation of the closed loop system, when compared to the single step load disturbance required for the other CPA indices and for the proposed CPA system. In practice, the achievable performance of a PID-based control system is always limited when compared to a minimum variance controller, for example. Thus, the Harris index reference value in such cases is unknown. This ambiguity means that explicit assessment using the Harris index can be challenging task.

2.7.3 Simulated validation of the system with higher order processes

The previous section validated the proposed CPA system for the assessment of systems with second order dynamics. The purpose of this section is to validate the CPA system for the assessment of systems with higher order dynamics. Such systems differ substantially from those modelled by SOPDT, which was used to generate the training dataset for the CPA classification algorithm. Two benchmark transfer functions were chosen [100] with an additional scalable time delay for $K_2(s)$:

$$K_1(s) = \frac{1}{(1+s)^{\mu}}, \tag{2.27}$$

$$K_2(s) = \frac{1}{(1+s)(1+\mu s)(1+\mu^2 s)(1+\mu^3 s)} e^{-\mu s}. \tag{2.28}$$

These transfer functions can be parametrized with the μ coefficient, resulting in different dynamic behavior. Table 2.14 presents the transfer functions that were selected to validate the proposed CPA system. Notably, the SOPDT approximations of these higher order systems cover a wide range of assumed L_1 and L_2 parameters.

Process Acronym	Transfer Function	L_1	L_2
P1	$K_1, \mu = 3$	0.27	1.0
P2	$K_1, \mu = 4$	0.41	1.0
P3	$K_2, \mu = 0.25$	0.24	0.28
P4	$K_2, \mu = 0.3$	0.28	0.33
P5	$K_2, \mu = 0.4$	0.37	0.5
P6	$K_2, \mu = 0.5$	0.49	1.0
P7	$K_2, \mu = 0.6$	0.53	1.0

Table 2.14 The selected higher order processes used to validate the proposed CPA system, together with the calculated L_1 and L_2 parameters based on the corresponding SOPDT approximations.

For each closed loop system with selected higher order process, 20 different sets of PID tunings were selected, resulting in 20 different control systems. Some of these were chosen from Appendix 5. Others were generated by using a trial-and-error method to manipulate the reference tunings, in an attempt to ensure the most satisfactory control performance.

The applied methodology together with the obtained results is presented below, using the *P1* process as an example. The results of the same experiments for processes *P2–P7* are presented in Appendix 3 At the outset of the experiment, the step response of *P1* was used to identify its SOPDT approximation (Fig. 2.53). Following this, the reference PID tunings were calculated from the identified parameters L_1, L_2 and the reference closed loop response was generated, both for a closed loop system with *P1* and a system with the corresponding SOPDT approximation (Fig. 2.53). The results show that the direct transfer of the reference PID tunings calculated using the SOPDT approximation to the closed loop systems with higher order *P1* processes does not cause a substantial difference in the dynamic behavior of the control systems, given that both reference responses are very similar.

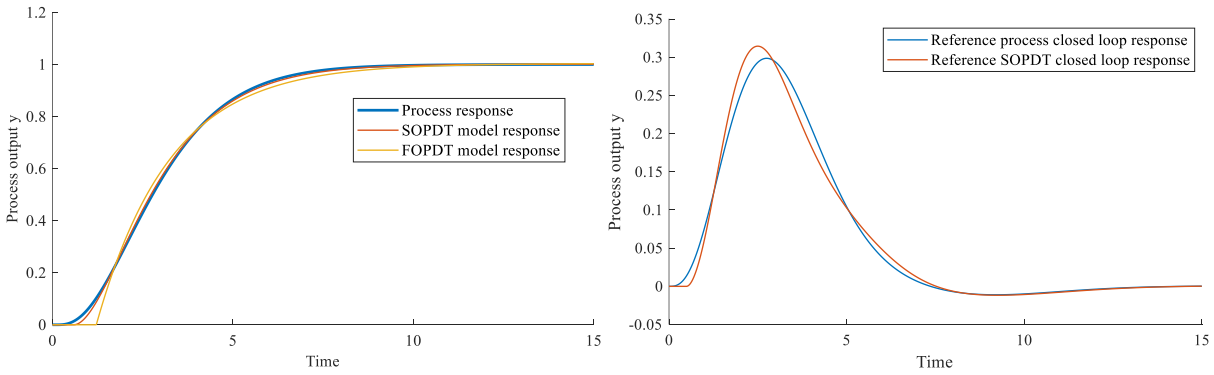

Fig. 2.53 A comparison of the responses of SOPDT and FOPDT processes (left). Reference closed loop responses of a higher order process and its SOPDT approximation (right).

Next, the closed loop system with an SOPDT approximation of the *P1* process was used to generate sets of PID tunings for the validation dataset. The results of assessment with the proposed CPA system

are presented in Fig. 2.54. The responses were assessed as OK (green) or NOK (red). The corresponding reference is also provided (black line).

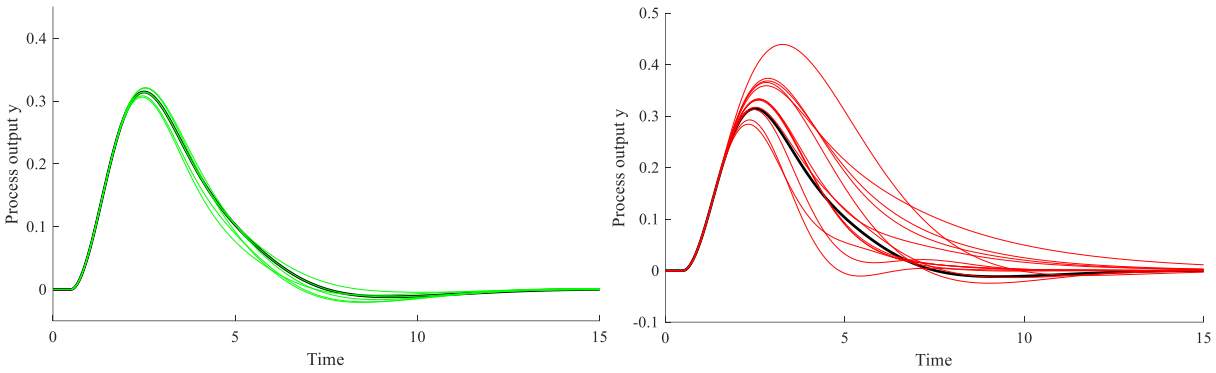

Fig. 2.54 The results of assessment with the proposed CPA system based on an SOPDT approximation of a higher order process. Both OK (left) and NOK (right) closed loop responses are shown, together with its corresponding reference (black line).

Finally, the PID tunings generated and labelled in this manner were transferred to a closed loop system with the higher order process *P1*. Again, closed loop responses were generated. Fig. 2.55 presents the results, with the responses grouped by the assessment made based on the SOPDT approximation. The results indicate that such an assessment also provides accurate results for the closed loop system with higher order process *P1*, given that those responses that are close to the reference response are assessed as OK.

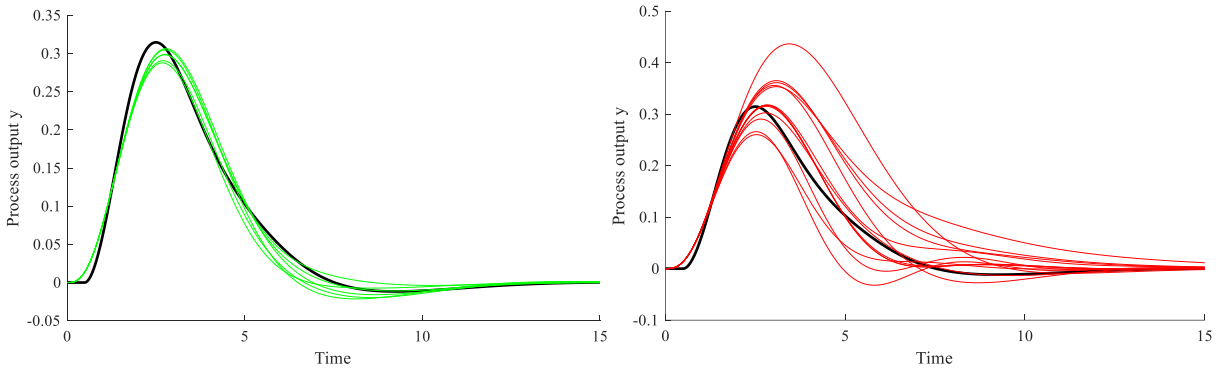

Fig. 2.55 The results of assessment with the proposed CPA system of a higher order process. Both OK (left) and NOK (right) closed loop responses are shown, together with its corresponding reference (black line).

Analysis was conducted of the assessment results for closed loop systems with processes *P2–P7*, as presented in Appendix 3. The analysis shows that for closed loop systems with processes described by the transfer function K_2 (*P3–P7*), the reference closed loop responses for both control systems with the higher order processes and those with the corresponding SOPDT approximations are similar. A more significant difference can be observed for the closed loop systems with processes described by the transfer function K_1 (*P1 – P2*). Despite the greater difference, the shapes and major properties of the responses are still preserved.

The proposed CPA provides a very accurate assessment of closed loop systems with the considered higher order processes. For all selected processes *P1–P7*, the closed loop responses that were classified as OK are very similar to the corresponding predefined references. Closed loop responses that deviated more strongly were classified as NOK. The SOPDT modelling was highly accurate for processes *P3–P7*, and thus the performance assessment based on the SOPDT approximations of these processes can be directly applied to the assessment of the corresponding higher order processes without a reduction in efficiency of the proposed CPA. However, the SOPDT modelling was less accurate for the *P1* and *P2* processes. As such, the difference between the higher order processes and the corresponding SOPDT approximations is greater.

To summarize, the proposed CPA system can be used to assess the responses of closed loop systems with higher order processes that have significantly different dynamics from the SOPDT model that was used to generate the training dataset for the ML methods. Closed loop performance can be successfully assessed using SOPDT approximations of higher order processes, while maintaining high accuracy. These findings are critical for the practical application of the proposed CPA system, given that in practice SOPDT approximations are not perfect, and differences exist between real processes and their SOPDT approximations.

2.7.4 Experimental validation

Following validation via simulation, the performance of the CPA system was verified using a laboratory control system. For this purpose, a cloud-based implementation of the CPA system was developed.

2.7.4.1 Cloud-based implementation of the proposed CPA system

The CPA system was initially implemented with a full vector of features, consisting of thirty selected CPIs. Due to the system's high demand for computational and memory resources, the initial implementation was given the form of a cloud-based application. The assessing control system is implemented on a Siemens S7-1500 PLC in the form of a *PID_Compact* library function block. Only the necessary functionalities are implemented directly within the PLC, in the form of a *ControlPerformanceAssessment* function block. The integration of the testing PID controller with the PLC implementation of the CPA system is shown in Fig. 2.56. The integration requires that the controller output is connected directly to the CPA dedicated function block. This is a software modification, and can be downloaded to the PLC without interrupting its normal operation. The main tasks of the dedicated function block are initializing the CPA procedure and gathering the closed loop response data for further analysis. The CPA procedure can be initialized by either the master system or the process operator, periodically or upon request. When the *InitializeCPA* input is set, the CPA procedure is initialized. If a steady state is detected by the ICM algorithm [50], a load disturbance is introduced to the control system via the *ControlSignal* output, the amplitude of which is set as 10% of

the range of the manipulating variable stored in the structure *PID_CompactConfig*. If a transient state is detected, the load disturbance is no longer applied, until a steady state is once again detected by the function block. Following the introduction of the load disturbance, the closed loop response data (*Setpoint* and *ProcessSignal*) is collected and stored within the internal PLC memory together with the constant *SamplingTime*, which is ensured by calling the CPA function block in the cyclic interrupt organization block OB30. Once a transient state is detected, data is collected until a new steady state is detected. After all the necessary closed loop response data have been gathered, they are sent to an OPC Unified Architecture (UA) server, together with the actual PID tunings (stored in *PID_CompactCtrlParams*) and a notification that the data is ready for further analysis. As an additional functionality, the CPA function block constantly monitors steady and transient states, the presence of which are indicated at the *SteadyState* and *TransientState* outputs, respectively.

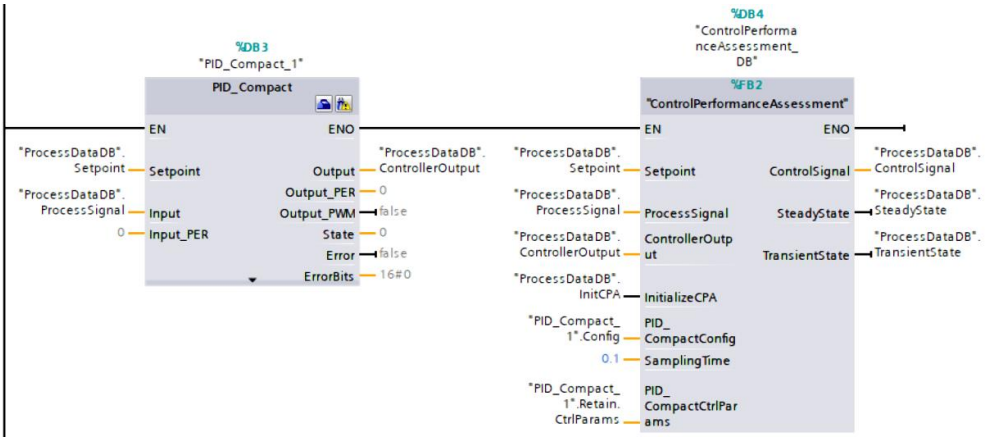

Fig. 2.56 The PLC implementation of the proposed CPA system on a Siemens S7-1200/1500, and its integration with a PID control system, implemented using TIA Portal software.

The architecture of the cloud-based CPA system is presented in Fig. 2.57. The collected data is stored in the OPC UA server, which is connected to the cloud-based application. Once the server receives a notification that the gathered data is ready for further analysis, the process model identification stage is initialized. The SOPDT process parameters are identified by minimizing modelling error with a nonlinear Nelder-Mead simplex algorithm. Then, based on the identified model and the actual PID parameters of the tested closed loop system, the process dynamics are normalized. Finally, the normalized closed loop response is simulated and used to calculate the CPIs. Once the feature vector is formed, the SVM classifier conducts the final assessment of control performance, as it provides the greatest accuracy of all selected algorithms. The final decision (OK or NOK) is sent to the OPC UA server, from where it can be read by the PLC or visualized using additional tools. OPC UA is an open communication standard, and does not restrict user specific client implementation. The OPC UA communication libraries are implemented in a wide range of programming languages, including Python, C#, and MATLAB. Fig. 2.58 shows an example client application for the proposed CPA system, developed using MATLAB. All necessary functionalities are implemented, including the ability to establish a connection with the OPC UA server, manual initialization of the CPA procedure,

and process data visualization. In addition, if poor control performance is detected, the application calculates suggested PID tunings based on its SOPDT approximation. Users can confirm the classification accuracy by a visual comparison of the reference and simulated responses.

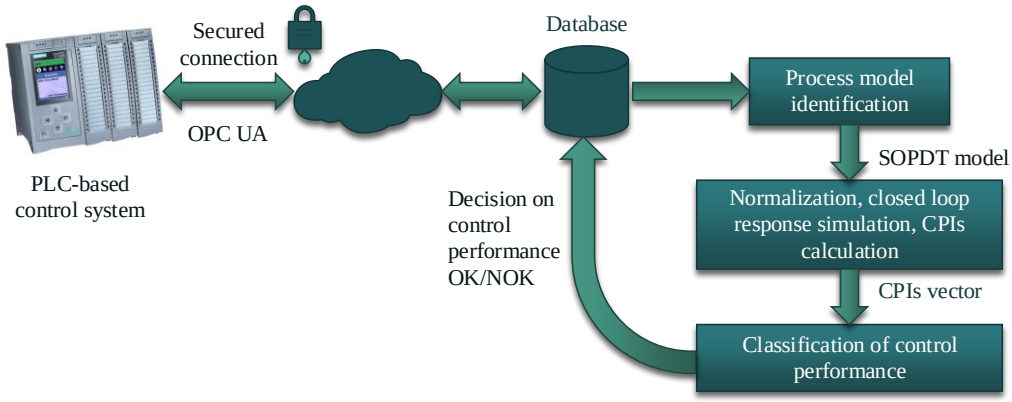

Fig. 2.57 The architecture of the cloud-based implementation of the proposed CPA system.

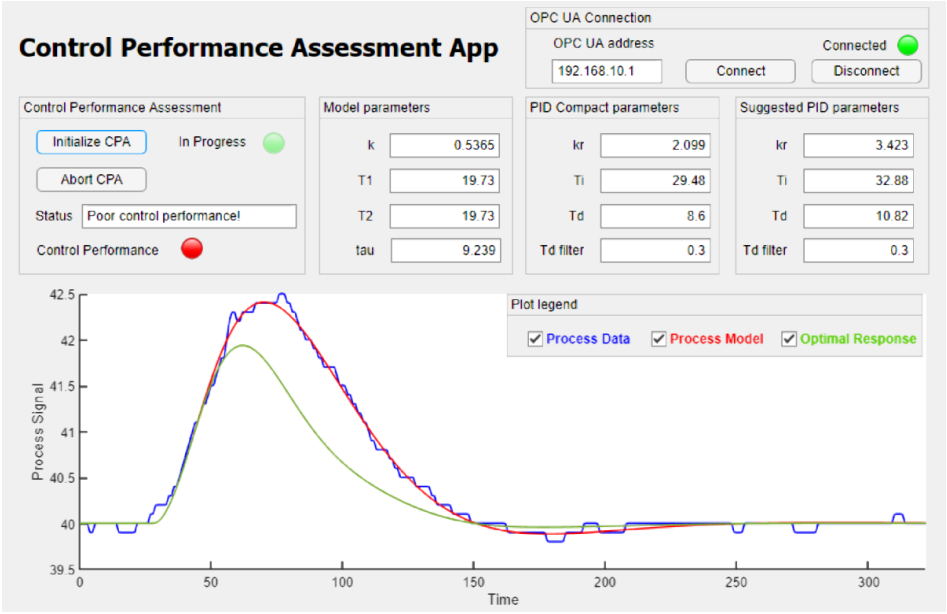

Fig. 2.58 An example implementation of a client application for the proposed CPA system, developed using MATLAB.

2.7.4.2 Verification of the system as applied to a heat exchange and distribution plant

For verification, the implemented system was applied to a heat exchange and distribution plant (described in Appendix 4. Such plants are characterized by complex, higher order dynamics with transportation delay time. Thus, this experimental validation investigates the overall performance of the CPA system, both in terms of its practical applicability and its capability to handle closed loop system with real higher order process.

For a constant flow rate of F = 3.5 L/min, 20 different sets of PID tunings were selected, representing 20 different control systems to be assessed. Each control system was assessed by initializing the CPA procedure with a load disturbance of ΔP_h = 10% applied to the closed loop system. The resultant

closed loop response was then obtained and used to create an SOPDT model of the real process. Finally, the SOPDT model parameters were normalized, and the new closed loop response was simulated to allow the CPIs to be calculated.

The classification results are presented in Fig. 2.59, with OK (green) and NOK (red) closed loop responses indicates. Both plots show the reference responses calculated based on the identified SOPDT model. Note that the identified parameters were slightly different for each of the obtained closed loop responses, and thus the resulting reference responses differ slightly from one another. Nevertheless, the results confirm the high assessment accuracy, with the closed loop responses that displayed similar dynamic behavior to the reference response classified as OK. Correspondingly, those closed loop responses that differed substantially from the reference were classified as NOK. The results show that such control systems are also unacceptable due to their sluggish or oscillatory behavior.

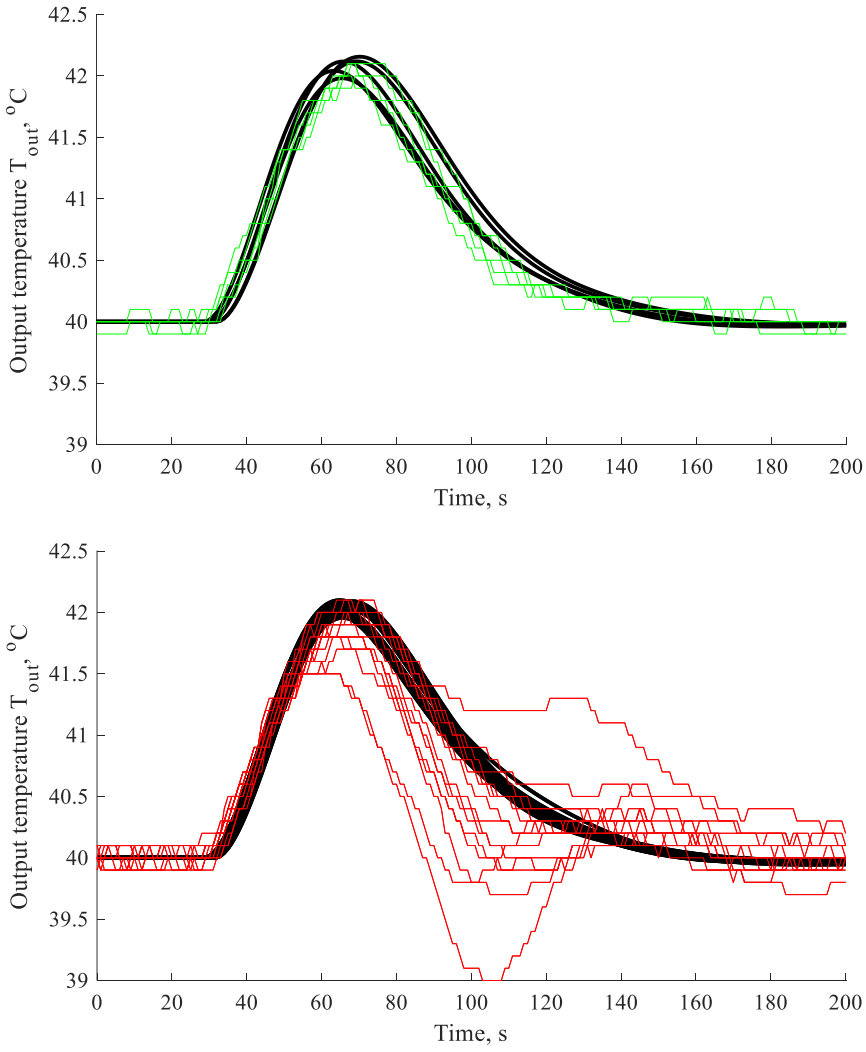

Fig. 2.59 The results of the assessment of real control systems by the proposed CPA system, with OK (top, green) and NOK (bottom, red) closed loop responses indicated, together with the corresponding reference responses (black lines).

2.8 Summary of the proposed CPA system

Section 2 described the synthesis of the proposed CPA system, in addition to validation via simulations and experiments. Note that each of the design stages described in the previous sections were conducted based on simulation studies, and only the rejection response of a closed loop system to a load disturbance is required for assessment. Validation via simulation demonstrated that the system can perform highly accurate assessments of PID-based closed loop systems with SOPDT or higher order processes. The latter is particularly important for the practical usability of the proposed CPA system. Given its high classification accuracy (greater than 96% when using the SVM classifier) and the possibility for parametrization, the system is an ideal tool for the assessment of industrial PID-based closed loop systems.

Fig 2.60 shows a complete block diagram describing the stages of development of the proposed CPA system. This process ultimately produced a functioning classifier. All of the predefined criteria used for the generation of reference tunings or the assessment of performance, among other tasks, should be considered as examples. These criteria can be modified to meet the technological requirements of the assessed closed loop systems. Thus, the presented developmental stages of the CPA system represent the general procedure, which can be freely parametrized. Potentially, this approach can also be applied for closed loop systems that are not based on PID controllers.

A cloud-based implementation of the proposed CPA system was developed for experimental validation. The integration of the CPA system with an existing PLC-based closed loop system is presented in Fig 2.61. The assessment procedure can be summarized as follows:

1. The CPA procedure is initialized with an external signal. Upon detection of a steady state, a step load disturbance is introduced to the control system. This load disturbance should not be introduced during a transient state, as external disturbances effecting the response will render the collected closed loop response data inappropriate for further analysis. Thus, for best possible accuracy a steady state must be detected first, which is done using ICM method proposed by author in [50] and described in details in Appendix 1.
2. The closed loop response data is collected during the transient state (load disturbance rejection). Upon detection of a steady state once again, data collection should cease, as the disturbance rejection is complete.
3. The obtained closed loop response data is used for SOPDT model identification, as the model parameters are necessary to normalize the closed loop response. Following the normalization of the process dynamics, the closed loop response is reproduced by simulations. The newly generated data, now without measurement noise, is used for further analysis.

4. Selected CPIs are calculated from the generated normalized closed loop response, forming a feature vector for the classification algorithm. Finally, the selected classifier assesses the control performance.

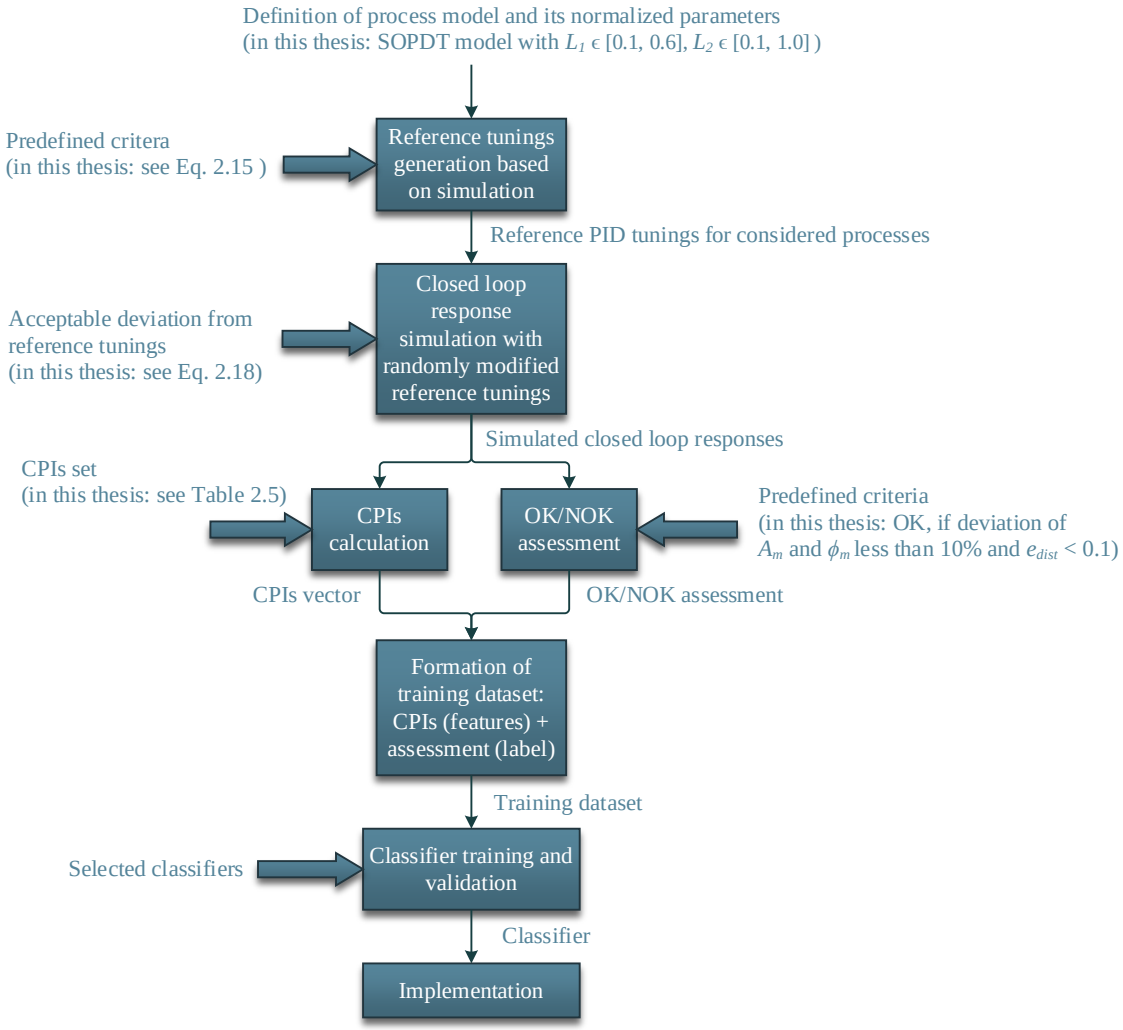

Fig. 2.60 A complete block diagram showing the development of the proposed CPA system.

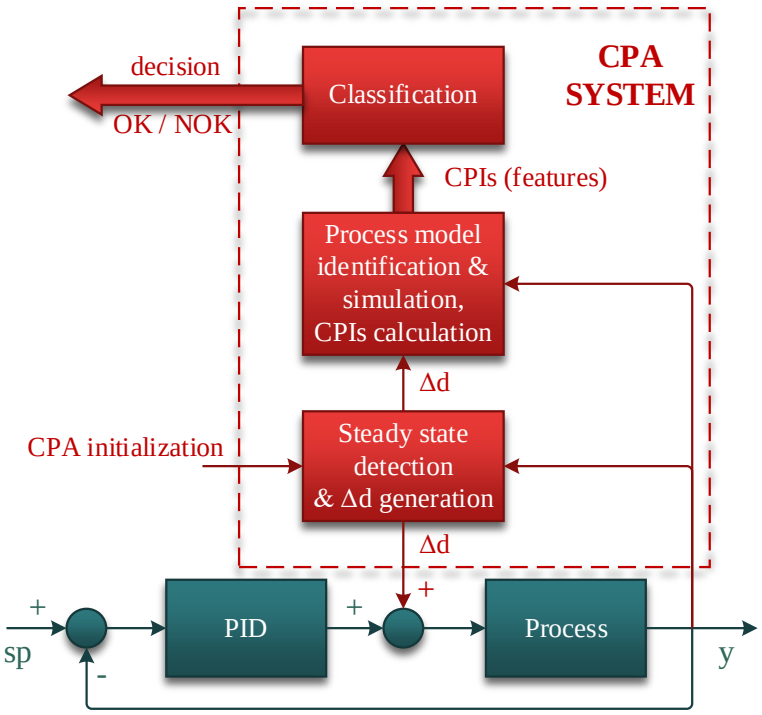

Fig. 2.61 A block diagram showing the CPA assessment procedure, and the integration of the CPA system with an existing PLC-based closed loop system.

The integration of the cloud-based implementation of the suggested CPA system with an existing PLC-based closed loop system requires additional hardware infrastructure. Such infrastructure is not always available in an industrial environment, given that its application requires relatively high additional costs. Existing commercial CPA systems for industrial environments are implemented in the form of external or master systems which also require integration with the existing hardware infrastructure. This strongly limits the practical usability of such CPA systems for industrial closed loop systems. These additional costs could potentially be reduced by implementing the CPA system directly within the control layer, that is, within a PLC. This approach would require no modifications of the existing hardware setup, and the only additional costs would be related to the licensing fees of the CPA system. Moreover, with a correctly prepared PLC-based implementation, the proposed CPA system could be implemented on-line without any interruption in the normal operation of the closed loop system. Note that such PLC-based implementations of CPA systems are not yet available. This represents an interesting research gap that is important from scientific and practical perspectives.

3 A PLC-based implementation of the proposed CPA system

This section investigates the possibility of implementing the proposed CPA system on a PLC.

3.1 General concept

Due to the very limited computational and memory resources of PLCs, the proposed CPA system must be simplified before being implemented on a PLC. Such a simplification can be achieved by reducing the dimensionality of the feature vector (the number of CPIs), which affects the structure of the final classifier. However, such a reduction must be accomplished without a significant reduction in classification accuracy. Thus, the generated dataset must first be analyzed to investigate any potential effects of reducing the number of CPIs. This analysis will take the form of a correlation analysis between the number of CPIs and the feature importance score of tree-based classifiers. Following this, the forward selection method will be used to derive a universal subset of features that are sufficient to ensure satisfactory classification accuracy for all of the selected classification algorithms. This approach will allow the user of the proposed CPA system to individually verify the implementation issues for all of the selected classifiers and chose the most convenient, without needing to determine a new subset of features. A successful reduction in the number of CPIs and the classifier itself could allow the proposed CPA system to be implemented on a PLC, thus avoided the need for either hardware modifications or a cloud-based implementation. This would significantly simplfy the pracitcal implementation of the CPA system.

3.2 Correlation analysis of the selected CPIs

The classifier structure can be simplified by reducing the number of CPIs that form the feature vector. This can be achieved without reducing accuracy by eliminating redundant CPIs that capture similar features of the closed loop response. To estimate the potential reduction of CPIs in this manner, an analysis of the distributions of and correlations between CPIs was conducted.

Initially, a dataset of 60000 samples was generated. A total of 30 CPIs were calculated based on the closed loop responses. Each feature (CPI) was normalized to have mean value of 0, with a unitary standard deviation. Fig. 3.1 presents a parallel coordinated plot of the OK and NOK classes, with the median values (continuous lines) and upper 0.75 and lower 0.25 quantiles (dashed lines) indicated. The feature names correspond to those shown in Table 2.5.

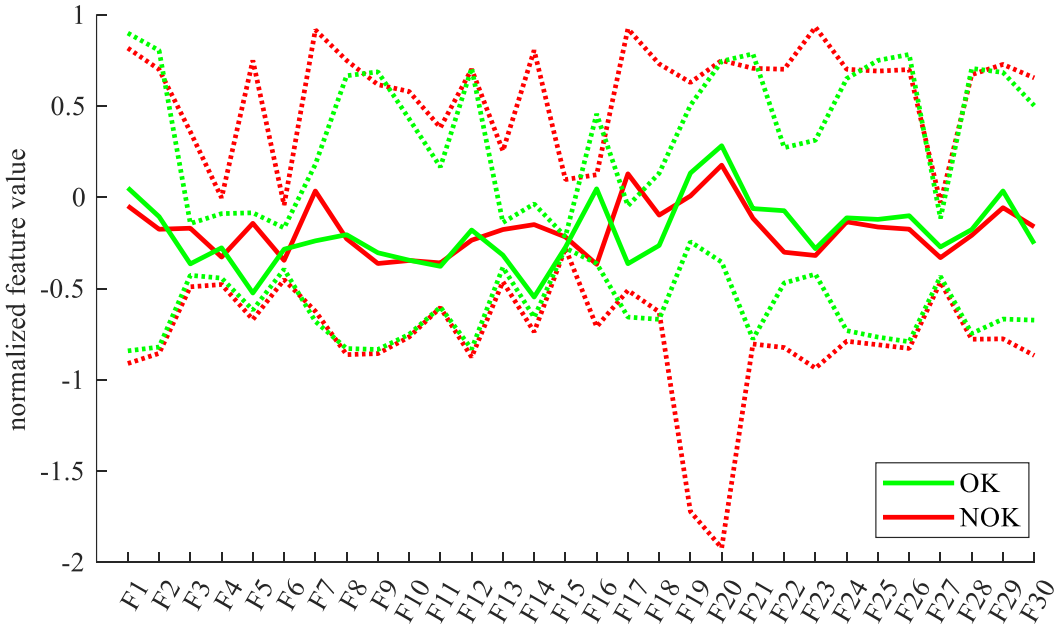

Fig. 3.1 A parallel coordinated plot of the generated training dataset, showing the median values (continuous lines) and upper (0.75) and lower quantiles (0.25) (dashed lines), for the OK (green) and NOK (red) classes.

The graph indicates that some CPIs produce similar distributions of OK and NOK labels. For example, the median, lower quartile, and upper quartile values for F8, F12, and F27 are similar. This is confirmed by Fig. 3.2, which compares the distributions of example features F8 and F12. Conversely, Fig. 3.3 presents example features F17 and F23, which have very different distributions. Such features are potentially important for the decision making of classification algorithms.

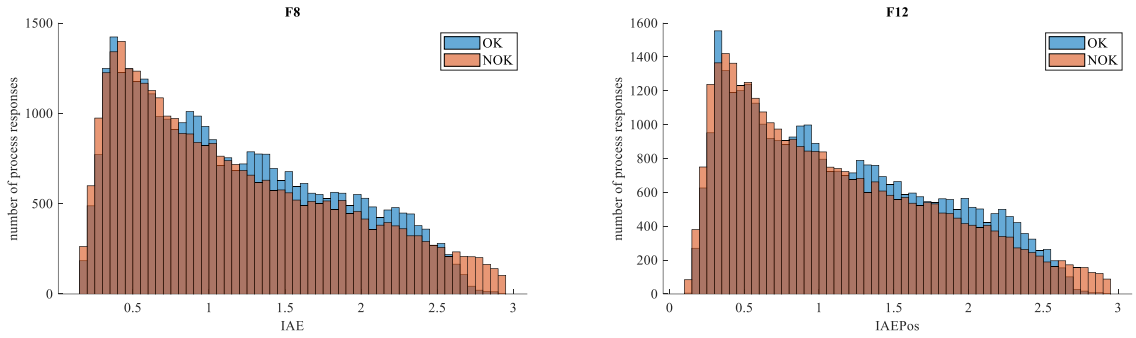

Fig. 3.2 The similar distributions of example CPIs F8 (left) and F12 (right) for both OK and NOK samples.

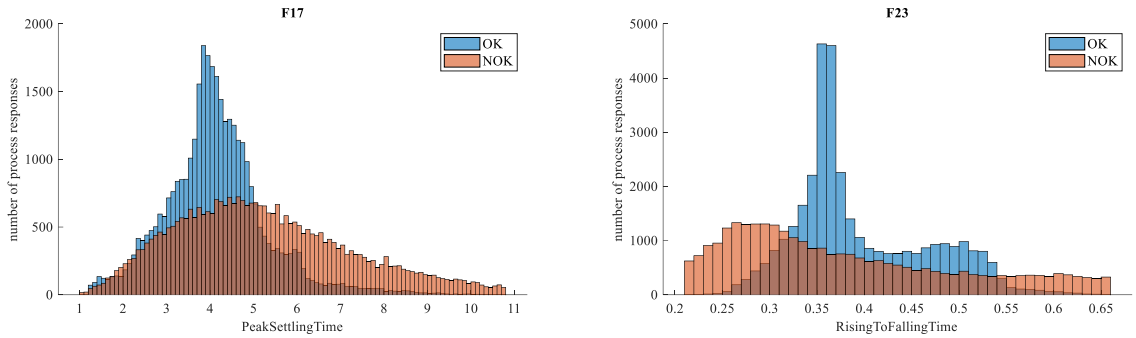

Fig. 3.3 The dissimilar distributions of example CPIs F17 (left) and F23 (right) for both OK and NOK samples.

Further analysis can be used to determine the correlations between features. The obtained dataset can then be divided into clusters of highly correlated CPIs. Highly correlated CPIs provide redundant capture of features of the closed loop responses, and thus represent an opportunity to simplify the training dataset. Table 3.1 presents the correlation matrix for the obtained dataset. From this matrix clusters of highly correlated features were extracted, for which the correlation between cluster members was greater than 0.8. This process produced four highly correlated groups of features.

This first group consisted of the features F1, F2, F8, F9, F10, F12, F21, F24, F25, and F26. From these CPIs, some subgroups can be extracted: F1 and F2 describe the value of the maximum peak and the time at which it occurs; F8, F9, F10 and F12 are integral indices that are calculated for the entire closed loop response; F24, F25 and F26 indicate the times at which 25%, 50%, and 75% of the maximum peak is damped, respectively. The final feature within the group is F21, which describes the rising time of the maximum peak. Notably, the large majority of these indices (F1, F2, F21, F24, F25, F26) refer to the first maximum peak. The integral indices (F8, F9, F10, F12) are also heavily dependent upon the shape of the maximum peak, given that the index values increase for higher values of control error, which correlate with the size of the maximum peak.

The second group contains the features F3, F5, F13, and F14. Features F3 and F5 describe the value of the minimum peak and its ratio when compared to the size of the maximum peak. The integral index F13 depends on the negative values of the control error. Given that the size of the minimum peak is correlated with the negative values of control error, features F13 and F3 are related. Feature F14 is a ratio-based index, and is calculated together with F13. However, F15 is similar to feature F5.

The third group contains the features F4, F6, and F27. Features F4 and F6 are time-based indices, calculated based upon the time at which the maximum and minimum peaks occur. Feature F27 describes the time at which the closed loop response crosses the zero value between the minimum and maximum peaks. Thus, a high value of F4 (the time at which the maximum peak occurs) correlates with a high value of F27 (the time at which zero crossing occurs between the maximum and minimum peaks) and a high value of F6 (the time at which the minimum peak occurs).

The fourth group contains the features F7, F11, F17, and F18. Features F7, F17, and F18 are time-based indices; they describe the settling time of the entire closed loop response, the difference between the settling time and the time at which the maximum peak occurs, and the time during which the control error has a positive sign, respectively. Feature F11 is an integral-based index, and hence also time-based, and is given by the integral of the control error multiplied by time squared. In this case, F11 increases substantially as settling time increases. As such, all of these indices are based on the settling time.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
F1	1.000	0.888	0.103	0.386	-0.224	0.326	0.651	0.920	0.927	0.822	0.667	0.920	0.145	-0.308	0.038	0.238	0.473	0.619	0.064	-0.388	0.802	0.480	0.343	0.803	0.844	0.869	0.373	0.417	0.669	-0.620
F2		1.000	0.122	0.383	-0.145	0.314	0.793	0.959	0.936	0.914	0.776	0.956	0.185	-0.219	0.059	0.237	0.610	0.728	0.211	-0.316	0.982	0.572	0.406	0.949	0.986	0.997	0.364	-0.020	0.350	-0.611
F3			1.000	-0.193	0.897	-0.208	0.404	0.134	0.108	0.201	0.301	0.041	0.967	0.839	0.539	-0.183	0.474	0.089	0.576	0.287	0.110	-0.323	0.684	0.129	0.086	0.102	-0.194	0.082	0.270	-0.280
F4				1.000	-0.297	0.997	0.507	0.487	0.425	0.488	0.462	0.509	-0.156	-0.316	-0.015	0.072	0.498	0.766	-0.498	-0.661	0.372	0.741	-0.295	0.458	0.417	0.398	0.996	0.082	0.070	-0.075
F5					1.000	-0.294	0.201	-0.149	-0.159	-0.050	0.080	-0.233	0.828	0.987	0.517	-0.287	0.323	-0.084	0.526	0.398	-0.129	0.604	-0.126	-0.164	-0.155	-0.294	0.995	-0.221	0.001	-0.072
F6						1.000	0.458	0.424	0.362	0.429	0.413	0.447	-0.175	-0.307	-0.020	0.055	0.464	0.729	-0.529	-0.654	0.305	0.716	-0.336	0.395	0.350	0.330	0.086	0.044	-0.029
F7							1.000	0.829	0.753	0.866	0.855	0.795	0.475	0.146	0.266	0.070	0.966	0.855	0.227	-0.289	0.799	0.615	0.320	0.857	0.820	0.805	0.495	-0.146	0.155	-0.383
F8								1.000	0.981	0.972	0.864	0.995	0.203	-0.222	0.068	0.212	0.674	0.799	0.112	-0.392	0.914	0.656	0.286	0.949	0.959	0.960	0.471	0.135	0.386	-0.491
F9									1.000	0.950	0.829	0.980	0.168	-0.228	0.045	0.190	0.586	0.713	0.115	-0.344	0.867	0.564	0.311	0.895	0.924	0.932	0.410	0.206	0.428	-0.469
F10										1.000	0.951	0.959	0.281	-0.115	0.127	0.142	0.740	0.828	0.141	-0.349	0.878	0.657	0.273	0.936	0.927	0.920	0.474	0.044	0.280	-0.413
F11											1.000	0.838	0.400	0.026	0.220	0.056	0.785	0.808	0.165	-0.290	0.749	0.610	0.233	0.845	0.804	0.785	0.452	-0.005	0.194	-0.311
F12												1.000	0.108	-0.301	0.018	0.228	0.631	0.797	0.053	-0.424	0.910	0.690	0.229	0.943	0.958	0.958	0.494	0.142	0.369	-0.472
F13													1.000	0.778	-0.128	0.524	0.540	0.149	0.619	0.269	0.176	-0.243	0.618	0.216	0.161	0.169	-0.160	-0.047	0.228	-0.265
F14														1.000	0.482	-0.298	0.283	-0.131	0.514	0.434	-0.197	-0.491	0.534	-0.189	-0.231	-0.226	-0.313	-0.261	-0.081	0.001
F15															1.000	-0.184	0.321	0.162	0.222	0.019	0.051	-0.130	0.350	0.086	0.037	0.041	-0.013	-0.034	0.194	-0.217
F16																1.000	-0.009	0.101	-0.070	-0.267	0.242	0.232	-0.126	0.254	0.250	0.243	0.049	0.038	0.048	-0.100
F17																	1.000	0.806	0.206	-0.242	0.626	0.559	0.245	0.715	0.651	0.627	0.490	-0.182	0.054	-0.241
F18																		1.000	-0.186	-0.646	0.733	0.828	0.024	0.814	0.766	0.745	0.757	-0.080	0.112	-0.290
F19																			1.000	0.745	0.239	-0.326	0.568	0.198	0.204	0.208	-0.511	-0.282	-0.022	-0.199
F20																				1.000	-0.296	-0.592	0.275	-0.344	-0.324	-0.320	-0.660	-0.217	-0.209	0.156
F21																					1.000	0.593	0.398	0.952	0.983	0.987	0.353	-0.180	0.219	-0.592
F22																						1.000	-0.414	0.709	0.648	0.607	0.733	-0.079	-0.093	-0.043
F23																							1.000	0.251	0.322	0.371	-0.302	-0.074	0.432	-0.642
F24																								1.000	0.977	0.963	0.439	-0.108	0.181	-0.439
F25																									1.000	0.995	0.398	-0.091	0.230	-0.499
F26																										1.000	0.379	-0.059	0.289	-0.559
F27																											1.000	0.090	0.071	-0.068
F28																												1.000	0.768	-0.116
F29																													1.000	-0.678
F30																														1.000

Table 3.1 The correlation matrix obtained for the features in the generated training dataset.

An agglomerative hierarchical clustering methodology was used to verify the obtained feature clusters. A measure of distance r_{dist} between features X and Y is defined as

$$r_{dist} = 1 - r(X, Y), \tag{3.1}$$

where $r(X, Y)$ is the correlation coefficient. As such, a distance of zero indicates perfect correlation. Three linkage methods were applied: single (shortest distance), complete (farthest distance), and average (average distance). Figs. 3.4–3.6 present the hierarchical clustering results in the form of dendrograms, for each of the three linkage methods. Distance thresholds of 0.1 and 0.2 were used to extract clusters from the obtained dendrograms. Figs. 3.4–3.6 present color-coded groups of features for the two distance threshold values.

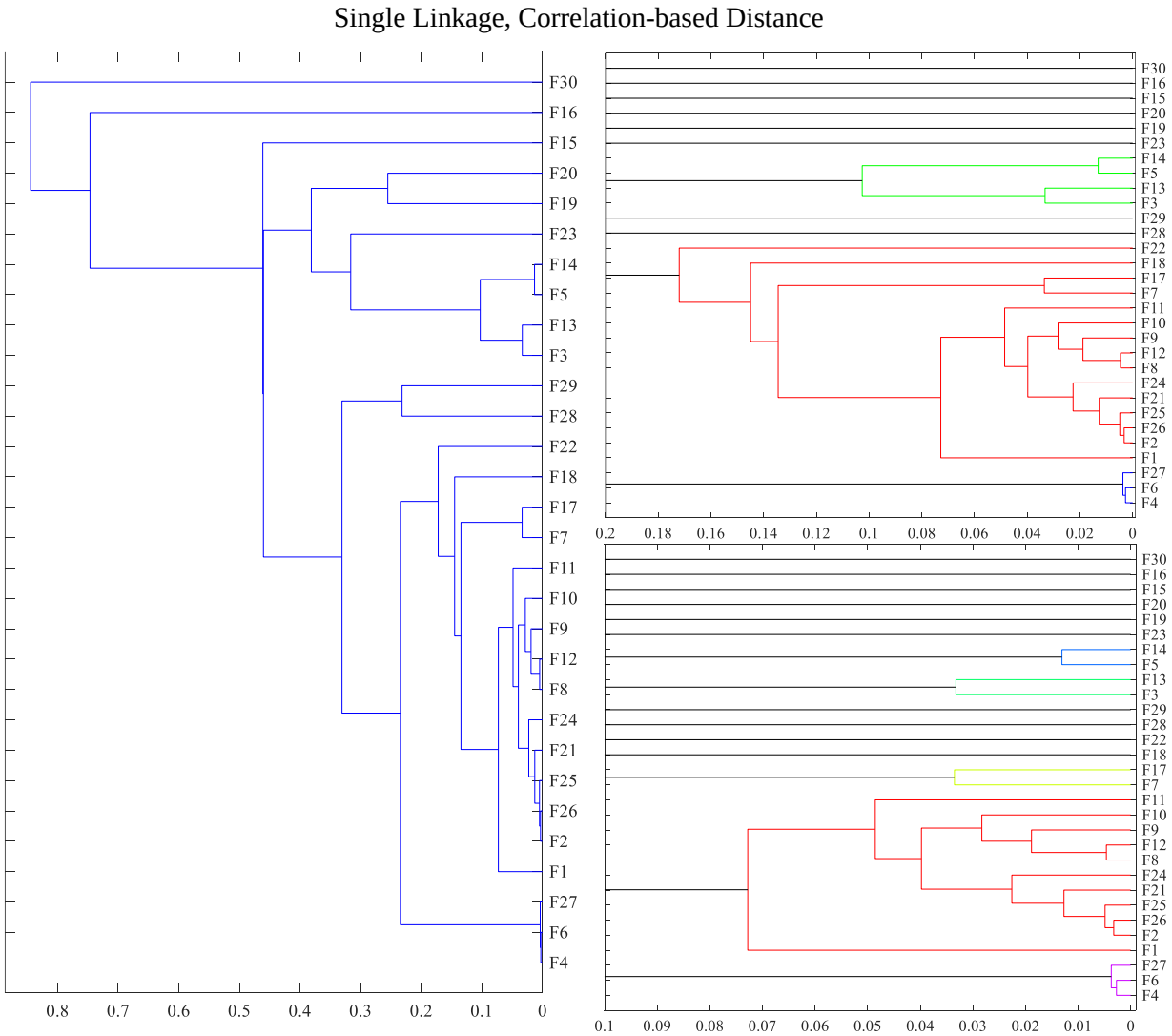

Fig. 3.4 The dendrogram obtained for the correlation-based distance metrics and single linkage method (left), with color-coded groups for threshold distance 0.2 (right, top) and 0.1 (right, bottom).

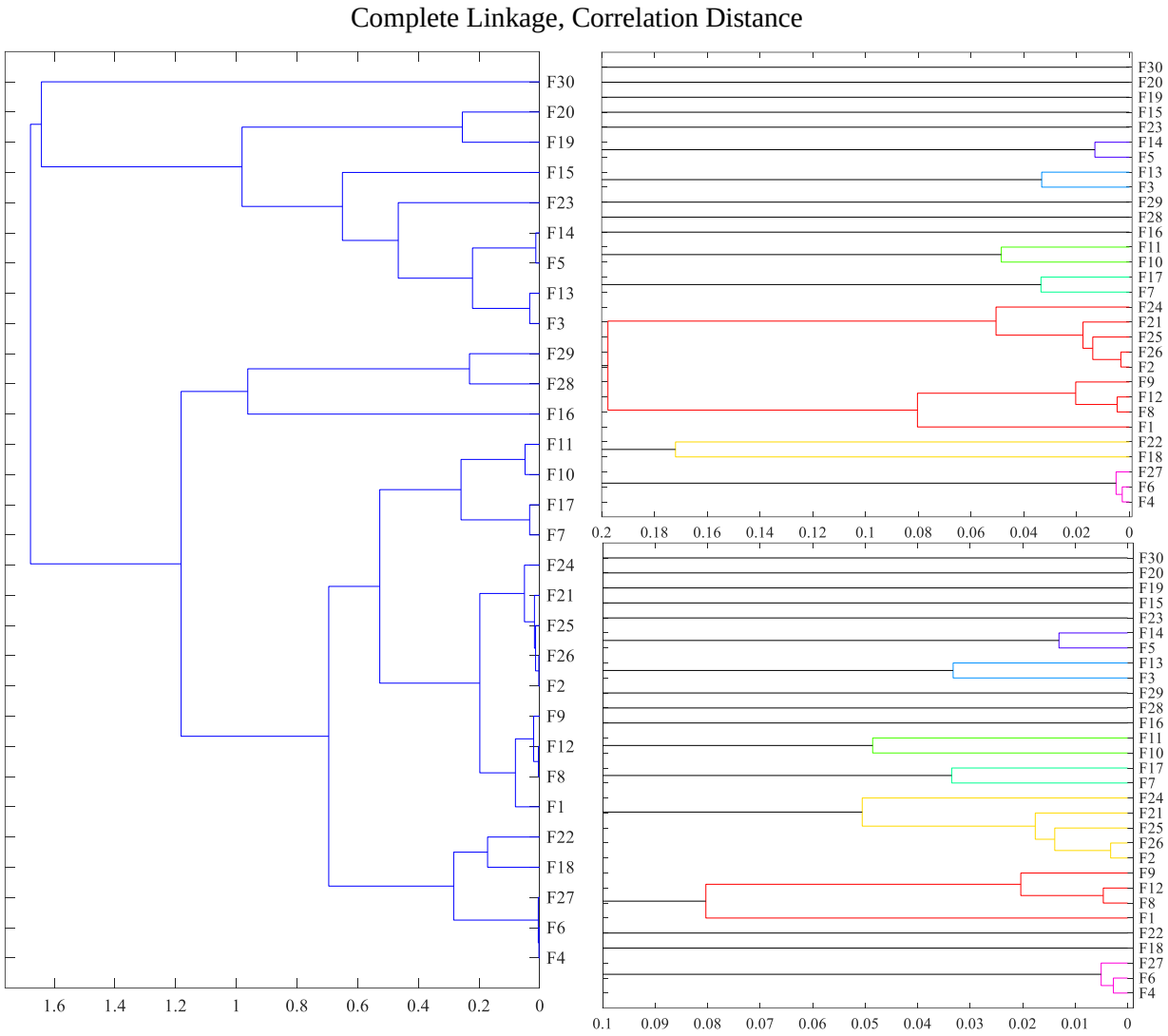

Fig. 3.5 The dendrogram obtained for the correlation-based distance metrics and complete linkage method (left) with color-coded groups for threshold distance 0.2 (right, top) and 0.1 (right, bottom).

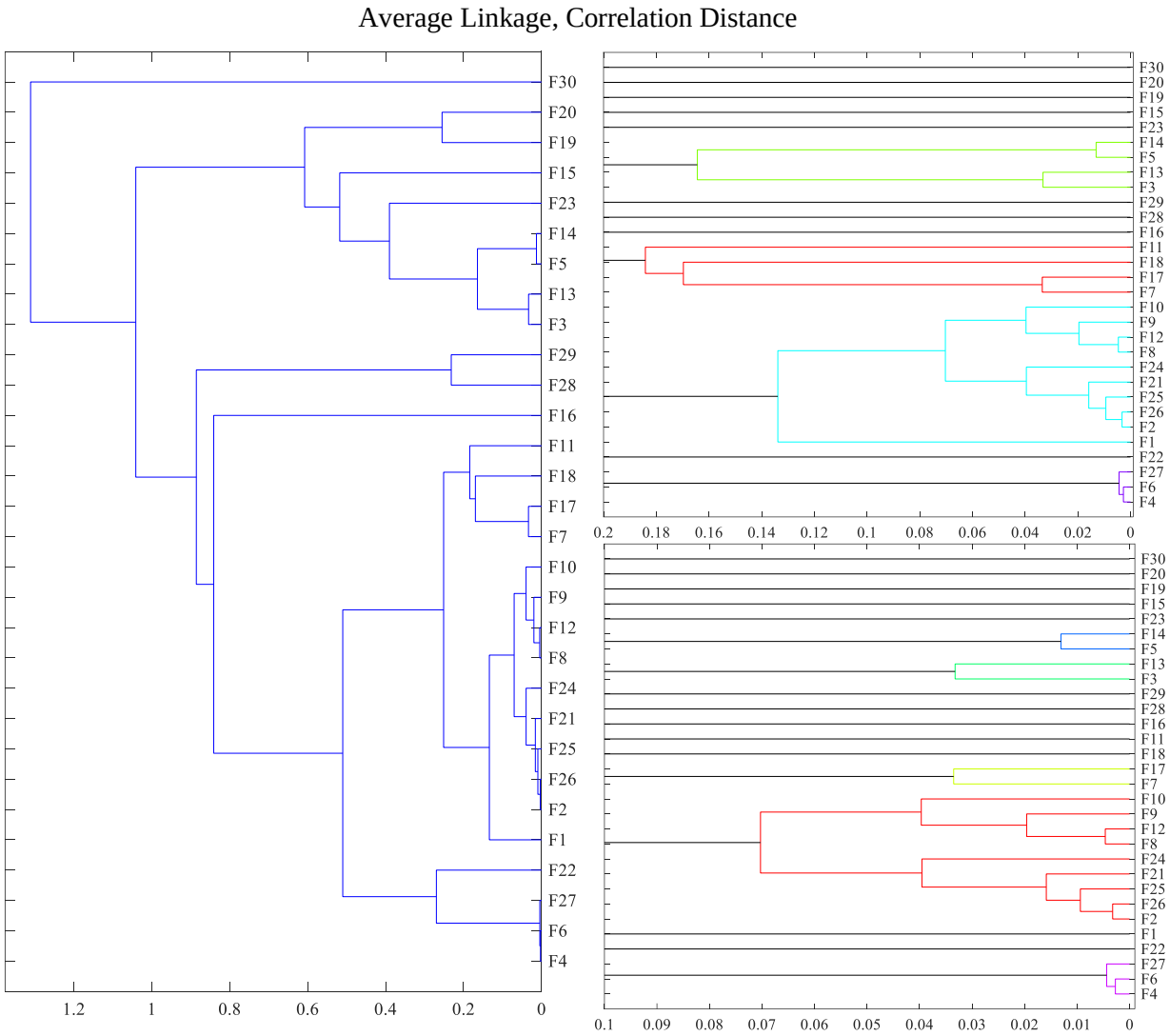

Fig. 3.6 The dendrogram obtained for the correlation-based distance metrics and average linkage method (left) with color-coded groups for threshold distance 0.2 (right, top) and 0.1 (right, bottom).

All of the obtained clusters are summarized in Table 3.2. The clusters are color-coded according to the results provided by the correlation matrix. The results display a high degree of consistency. The blue cluster (F3, F5, F13, and F14) was derived from all obtained dendrograms. However, with a distance threshold of 0.1, or for complete linkage with a threshold of 0.2, it can be divided into two, smaller clusters: F3 and F13, and F5 and F14. The green cluster (F4, F6, and F27) was obtained for all studied dendrograms. The yellow cluster (F1, F2, F8, F9, F10, F12, F21, F24, F25, and F26) was obtained by the correlation matrix, and for complete linkage and average linkage. In the case of single linkage, features from the yellow cluster are grouped with some features from the orange cluster (F7, F11, F17, and F18). However, the other linkage methods do not indicate such overlapping.

Method	Single Linkage (0.1)	Single Linkage (0.2)	Complete Linkage (0.1)	Complete Linkage (0.2)	Average Linkage (0.1)	Average Linkage (0.2)	Correlation Matrix
Clusters	F5, F14	F3, F5, F13, F14	F5, F14	F5, F14	F5, F14	F3, F5, F13, F14	F1, F2, F8, F9, F10, F12, F21, F24, F25, F26
	F3, F13	F1, F2, F7, F8, F9, F10, F11, F12, F17, F18, F21, F22, F24, F25, F26	F3, F13	F3, F13	F3, F13	F7, F11, F17, F18	F3, F5, F13, F14
	F7, F17	F4, F6, F27	F10, F11	F10, F11	F7, F17	F1, F2, F8, F9, F10, F12, F21, F24, F25, F26	F4, F6, F27
	F1, F2, F8, F9, F10, F11, F12, F21, F24, F25, F26		F7, F17	F7, F17	F2, F8, F9, F10, F12, F21, F24, F25, F26	F4, F6, F27	F7, F11, F17, F18
	F4, F6, F27		F2, F21, F24, F25, F26	F1, F2, F8, F9, F12, F21, F24, F25, F26	F4, F6, F27		
			F1, F8, F9, F12	F18, F22			
			F4, F6, F27	F4, F6, F27			

Table 3.2 The clusters obtained from the correlation matrix and the generated dendrograms with selected linkage methods and distance thresholds.

These results suggest that some of the suggested CPIs are redundant, given that the correlation between some pairs of CPIs is relatively high. Thus, this initial analysis confirms that the number of features can potentially be decreased, without decreasing the useful information that is available to the classification algorithms.

Tree-based models can be used to obtain the importance scores of individual CPIs. The importance score determines the usefulness of each CPI for the classifier prediction. The tree-based classifiers were trained using the K most important features, with K ranging across all features, from 1 to 30. The classification accuracies obtained with the K most important features, together with the feature rankings, are presented in Table 3.3. Particular features are color-coded according to the clusters obtained by the correlation matrix. The rankings indicate that the most important features for classification are the individual CPIs that are not associated with any cluster. CPIs that are associated with the blue cluster are also important. CPIs from the yellow and orange clusters are less important, but still relatively high in the feature ranking. The rankings indicate that CPIs from the green cluster are not important for tree-based classification. The individual features that are important for all tree-

based models are F1, F3, F13, F17, F22, F23, F29, and F30. For the XGBoost model, features F3, F23, and F30 are particularly important. Note that F23 and F30 are newly proposed indices.

Rank	Decision Tree		Random Forest		Extra Trees		Light GBM		XGBoost		AdaBoost	
	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %
1	F23	73.70	F23	74.36	F23	74.76	F30	64.28	F13	70.36	F30	63.61
2	F3	78.50	F30	81.06	F30	81.5	F23	77.99	F3	72.45	F23	78.82
3	F30	84.98	F3	88.19	F3	86.15	F29	86	F22	75.36	F29	84.95
4	F22	89.23	F13	89.72	F22	88.38	F1	88.78	F17	80.42	F1	93.18
5	F29	90.92	F17	90.93	F17	90.36	F28	93.31	F23	86.87	F20	95.14
6	F28	90.63	F22	91.65	F19	89.77	F9	93.35	F30	90.74	F9	94.92
7	F1	91.99	F15	92.28	F14	89.72	F20	94.2	F5	92.89	F28	94.88
8	F19	91.48	F29	93.26	F20	91.23	F22	94.35	F12	93.52	F3	95.55
9	F15	91.96	F19	92.98	F13	90.61	F3	95.04	F26	93.94	F14	95.69
10	F13	92.09	F28	93.14	F29	92.15	F14	95.17	F15	94.4	F17	95.72
11	F5	91.76	F5	93.24	F5	92.17	F5	95.17	F29	95.1	F19	95.58
12	F9	91.83	F20	93.25	F6	91.78	F19	95.11	F1	95.1	F5	95.68
13	F20	91.93	F1	93.76	F4	91.89	F15	95.3	F20	95.08	F15	95.64
14	F17	91.64	F16	93.74	F1	92.04	F16	95.39	F14	95.33	F13	95.66
15	F14	91.67	F14	93.55	F27	92.49	F6	95.35	F19	95.37	F12	95.56
16	F16	91.81	F9	93.79	F16	92.38	F2	95.46	F2	95.43	F18	95.51
17	F24	91.81	F8	93.65	F28	92.33	F17	95.2	F8	95.33	F22	95.71
18	F11	91.64	F12	93.69	F9	92.67	F18	95.17	F16	95.29	F16	95.82
19	F12	91.65	F6	93.75	F15	92.51	F12	95.34	F28	95.41	F8	95.53
20	F6	91.47	F7	93.71	F2	92.62	F13	95.18	F6	95.38	F6	95.69
21	F18	91.44	F2	93.65	F8	92.81	F8	95.06	F9	95.25	F27	95.65
22	F8	91.52	F26	93.64	F18	92.49	F26	95.06	F10	95.47	F4	95.63
23	F25	91.49	F10	93.66	F10	92.72	F7	95.23	F21	95.3	F7	95.58
24	F21	91.73	F18	93.59	F26	92.53	F21	94.84	F25	95.34	F11	95.55
25	F7	91.61	F24	93.6	F12	92.94	F27	95.43	F27	95.42	F24	95.41
26	F10	91.54	F25	93.6	F24	92.54	F11	95.23	F18	95.17	F25	95.55
27	F2	91.6	F27	93.47	F7	92.63	F24	95.48	F7	95.12	F10	95.43
28	F4	91.52	F21	93.61	F21	92.58	F25	95.24	F24	95.38	F2	95.52
29	F26	91.49	F11	93.51	F25	92.72	F4	95.13	F11	95.23	F21	95.69
30	F27	91.54	F4	93.7	F11	92.85	F10	95.23	F4	95.26	F26	95.48

Table 3.3 The classification accuracy of tree-based models with a training dataset consisting of the K most important features.

Fig 3.7 presents the classification accuracy of the selected tree-based classifiers obtained for the K of most important features. The graph indicates that classification accuracy plateaus after approximately ten features, with further increases in the number of features providing no corresponding increase in accuracy. Thus, the initial number of 30 CPIs can be reduced without decreasing classification accuracy. This will potentially reduce the computational and memory resources that are required for implementation of the CPA system. However, the optimal subset of features remains to be determined. This issue will be addressed in the following section.

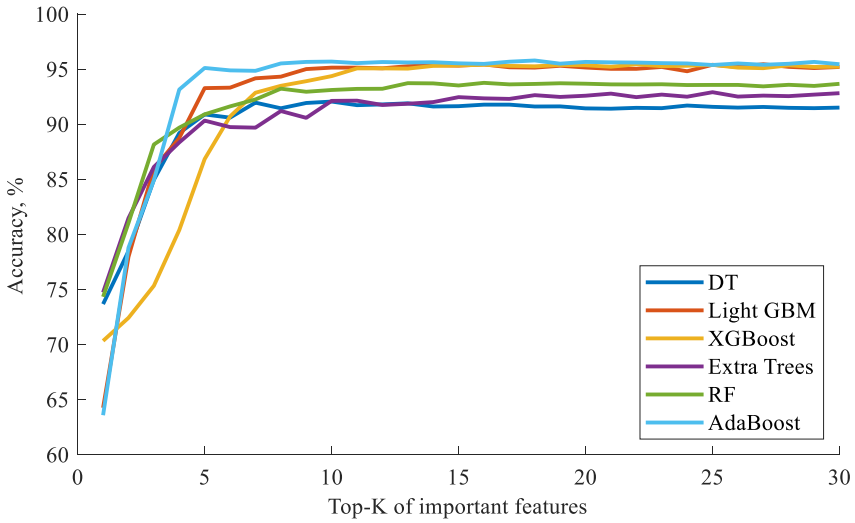

Fig. 3.7 A graphical representation of the classification accuracy of tree-based models with a training dataset consisting of the K most important features.

3.3 Reduction of the feature vector

The reduced CPI vector should be as universal as possible so as to provide a high classification accuracy for all of studied classification algorithms. With this approach, the most suitable classifier can be chosen for each implementation, without requiring the generation of a new feature vector.

For this purpose, the effect on classification accuracy of the addition of new features to the training dataset was determined. This iterative procedure was initialized with the single feature that provides the greatest accuracy. Following this, additional features were added, one by one. Each added feature was that which provided the greatest increase in accuracy. This process was repeated for the training and validation datasets, with an increasing number of CPIs. This forward feature selection procedure was repeated for each selected classifier, other than the fuzzy models and simple LDA and GNB. The obtained results are presented in Table 3.4, together with both feature rank and classification accuracy obtained for the validation dataset.

DT		Light GBM		XGBoost		Extra Trees		RF		AdaBoost		SVM		KNN	
Feature	Acc, %	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %	Feature	Acc, %
F23	73.7	F23	75.12	F23	73.56	F23	74.76	F23	74.36	F23	74.5	F23	74.29	F23	72.06
F25	80.8	F24	75.12	F3	80.88	F25	84.19	F25	83.19	F21	81.64	F24	84.01	F11	81.75
F30	86.68	F25	83.16	F29	81.68	F30	88.66	F30	89.34	F30	86.95	F30	88.23	F30	85.69
F1	91.06	F30	89.99	F9	88.01	F1	90.63	F1	92.87	F28	92.96	F28	94.14	F14	91.35
F15	92	F1	93.28	F30	94.56	F15	92.78	F15	93.76	F14	95.15	F14	95.97	F7	92.67
F6	92.3	F5	94.44	F20	94.69	F29	93.07	F29	94.39	F9	95.42	F6	96.4	F29	94.26
F4	92.47	F20	95.03	F5	95.43	F14	92.86	F16	94.43	F27	95.6	F9	96.55	F27	94.59
F29	92.44	F12	95.15	F14	95.23	F24	93.41	F26	94.27	F24	95.72	F11	96.56	F5	94.61
F16	92.65	F29	95.17	F1	95.4	F12	93.11	F9	94.19	F11	95.75	F3	96.57	F3	94.75
F26	92.63	F13	95.37	F15	95.77	F28	93.6	F11	94.19	F6	95.63	F19	96.61	F15	94.75
F21	92.64	F14	95.51	F24	95.58	F21	93.44	F6	94.22	F20	95.67	F29	96.72	F13	94.7
F9	92.67	F6	95.34	F7	95.6	F16	93.28	F12	94.18	F29	95.64	F10	96.78	F10	94.68
F2	92.6	F9	95.48	F16	95.59	F27	93.22	F4	94.11	F3	95.84	F13	96.79	F8	94.52
F11	92.63	F11	95.58	F19	95.56	F9	93.05	F14	94.1	F8	95.78	F22	96.76	F4	94.46
F8	92.61	F18	95.52	F26	95.62	F8	93.41	F18	94.17	F16	95.87	F8	96.75	F6	94.36
F22	92.67	F28	95.45	F18	95.56	F26	93.3	F13	94.06	F2	95.8	F15	96.74	F22	94.29
F10	92.64	F4	95.42	F27	95.58	F20	93.13	F22	94.06	F1	95.76	F12	96.73	F18	94.23
F24	92.5	F16	95.45	F2	95.45	F11	93.25	F8	94.03	F12	95.82	F5	96.73	F1	94.11
F27	92.31	F27	95.43	F12	95.6	F4	93.01	F10	93.97	F15	95.83	F18	96.71	F19	94.11
F12	92.35	F22	95.48	F8	95.55	F18	93.21	F28	94.04	F4	95.83	F25	96.72	F17	94.07
F5	92.25	F17	95.41	F10	95.48	F2	93.14	F21	94.04	F17	95.78	F4	96.68	F24	94.05
F17	92.17	F7	95.42	F13	95.57	F6	93.07	F17	94.11	F26	95.73	F26	96.61	F16	93.88
F13	92.32	F26	95.32	F22	95.46	F19	92.94	F7	93.99	F22	95.8	F2	96.63	F9	93.9
F20	92.33	F2	95.27	F4	95.36	F17	92.93	F20	93.86	F13	95.71	F27	96.67	F12	93.68
F28	92.46	F15	95.42	F17	95.51	F5	92.81	F2	93.78	F25	95.74	F21	96.59	F26	93.52
F19	92.36	F19	95.32	F25	95.32	F22	92.63	F24	93.84	F7	95.76	F7	96.56	F25	93.38
F18	92.43	F3	95.3	F6	95.49	F7	92.8	F27	94	F18	95.68	F17	96.49	F2	93.2
F7	92.35	F21	95.22	F21	95.32	F3	92.84	F19	93.86	F19	95.82	F1	96.47	F21	93.07
F14	92.22	F8	95.14	F11	95.41	F13	92.81	F5	93.68	F5	95.69	F16	96.44	F28	92.72
F3	91.54	F10	94.84	F28	95.32	F10	92.85	F3	93.7	F10	95.62	F20	96.17	F20	92.24

Table 3.4 The classification accuracy obtained during consecutive steps of the forward feature selection procedure for selected classification models.

The feature rankings provide some initial findings:

1. Feature F23 is clearly the most important for classification, as is provides the highest possible accuracy for all of the selected classifiers. A training dataset that consists only of feature F23 provides a classification accuracy of approximately 74–75%. Thus, three out of four closed loop responses can be accurately assessed using only feature F23. This feature describes the ratio of rising and falling time of the first maximum peak. Thus, it directly describes the shape of the first maximum peak.
2. For six of the eight selected classifiers, the second added feature is associated with the yellow cluster.
3. The derivative-based indices F28, F29, and F30 rank relatively highly. This suggests that information concerning the derivative of the closed loop response is also highly important for accruate performance assessment. The maximal and minimal values of the derivative describe the maximum rising and falling rates of the first maximum peak. As described above, these indices indirectly describe the maximum peak of closed loop response.
4. From the set of CPIs that form the top five most important for each classifier, 30% are associated with the yellow cluster, 10% with the blue cluster, 5% with the orange cluster, and

0% with the green cluster. The remaining 55% of features are individual CPIs that are not associated with any cluster.

5. The five highest ranked CPIs are similar for each of the tree-based models: F23, F25, F30, F1, and F15.
6. For the majority of classifiers, among the five highest ranked CPIs there are a maximum of two CPIs from any given cluster. The Light GBM classifier is an exception, with three out of the five highest ranked CPIs being associated with the yellow cluster.

Fig 3.8 presents the classification accuracy obtained for various number of features within the training dataset. Correspondingly, Fig. 3.9 presents the change in classification accuracy obtained during each step of the forward feature selection procedure.

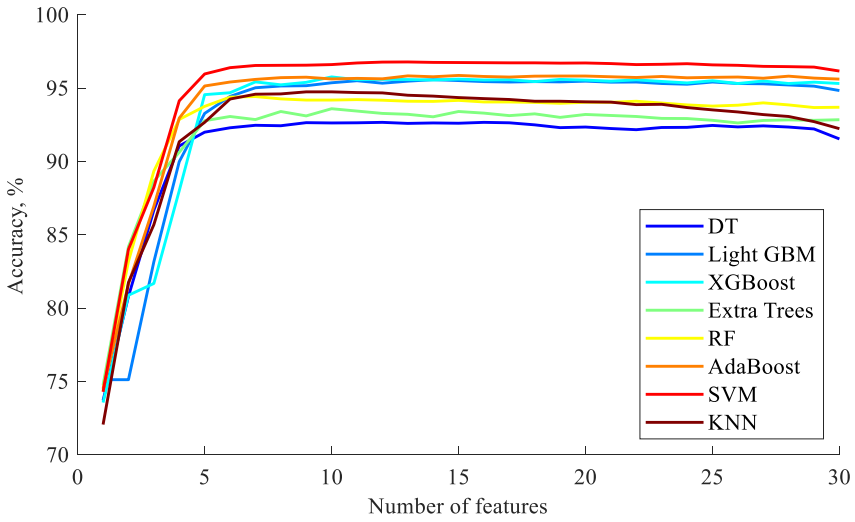

Fig. 3.8 A graphical representation of the classification accuracy obtained during consecutive steps of the forward feature selection procedure for selected classifiers.

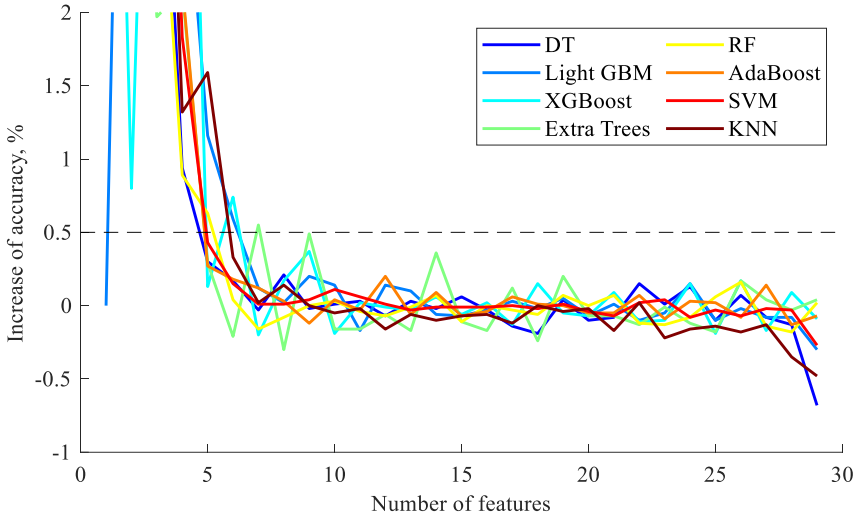

Fig. 3.9 A graphical representation of the change in classification accuracy obtained during consecutive steps of the forward feature selection procedure for selected classifiers.

The graphs indicate that for each selected classifier, the best possible classification accuracy was achieved after as few as five iterations of the forward feature selection procedure, and hence for five features within the training dataset. A further increase in the number of features does not produce a significant increase in assessment accuracy. These results alone are insufficient to derive a universal set of features that will ensure the highest possible classification accuracy, independent of the classifier used. The top five features are similar only for the tree-based models (DT, Extra trees, and RF); for the other models, the features are ranked differently. Hence, further analysis is required.

The detailed results of the iterative procedure for all selected classifiers are presented in Table 3.5–Table 3.12. Detailed results for the first six iterations of the forward feature selection procedure are presented together with the achieved accuracy and the difference between the accruacy achieved by that feature and the best accuracy achieved in that iteration. Table 3.5 shows the results for the DT model. During the first iteration, the best possible accuracy (73.7%) is displayed by feature F23. No other feature provides an accuracy within 1% of this. The addition of a second feature to F23 provides an increase in classification accuracy. The best possible accuracy (80.8%) is obtained by feature F25, although features F2, F4, F8, F11, F21, F22, F24, and F26 also provide a relatively high accuracy, within 1% of that obtained by F25. Moreover, the difference between the accuracies of F25 and F22 is only 0.02%. Thus, when selecting a subset of features during the second iteration, F23 can be supplemented with F25 or F22 while retaining a high classification accuracy. Hence, the subset of features for each classifier can be slightly modified from those presented in Table 3.4. Each selected classifier was analyzed in this manner, and common subsets of features were investigated.

DT

Iteration	Feature	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
1	Accuracy, %	51,45	50,8	72,54	60,73	70,3	66,38	61,44	54,08	52,54	53,68	54,25	53,26	70,43	69,72	71,44	66,83	66,09	58,43	64,56	61,56	51,39	70,09	73,7	53,79	51,4	51,6	62,67	51,17	55,98	63,02
	Difference from best	-22,25	-22,9	-1,16	-12,97	-3,4	-7,32	-12,26	-19,62	-21,16	-20,02	-19,45	-20,44	-3,27	-3,98	-2,26	-6,87	-7,61	-15,27	-9,14	-12,14	-22,31	-3,61	0	-19,91	-22,3	-22,1	-11,03	-22,53	-17,72	-10,68
2	Accuracy, %	78,95	80,36	78,5	80,25	78,45	79,71	78,45	79,95	79,45	79,49	80,05	79,43	78,35	79,25	79,25	79,69	76,97	76,95	75,84	75,76	80,19	80,78		80,63	80,8	80,47	79,65	77,19	77,08	78,06
	Difference from best	-1,85	-0,44	-2,3	-0,55	-2,35	-1,09	-2,35	-0,85	-1,35	-1,31	-0,75	-1,37	-2,45	-1,55	-1,55	-1,11	-3,83	-3,85	-4,96	-5,04	-0,61	-0,02		-0,17	0	-0,33	-1,15	-3,61	-3,72	-2,74
3	Accuracy, %	82,82	84,89	84,29	83,01	83,76	83,04	84,44	82,59	83,14	82,29	83,39	81,96	83,66	83,3	83,61	83,75	83,72	83,2	82,76	82,82	82,01	82,92		82,74		83,4	82,52	82,54	83,71	86,68
	Difference from best	-3,86	-1,79	-2,39	-3,67	-2,92	-3,64	-2,24	-4,09	-3,54	-4,39	-3,29	-4,72	-3,02	-3,38	-3,07	-2,93	-2,96	-3,48	-3,92	-3,86	-4,67	-3,76		-3,94		-3,28	-4,16	-4,14	-2,97	0
4	Accuracy, %	91,06	87,96	89,78	87,39	90,6	87,26	89,04	89,54	90,45	88,53	87,53	90,33	89,43	90,22	90,88	87,59	88,66	87,6	88,69	88,93	88,04	87,43		87,16		87,31	88,17	90,17	90,29	
	Difference from best	0	-3,1	-1,28	-3,67	-0,46	-3,8	-2,02	-1,52	-0,61	-2,53	-3,53	-0,73	-1,63	-0,84	-0,18	-3,47	-2,4	-3,46	-2,37	-2,13	-3,02	-3,63		-3,9		-3,75	-2,89	-0,89	-0,77	
5	Accuracy, %		90,91	91,54	91,04	91,89	90,8	90,82	91,02	91,18	91,06	91,02	91,16	91,09	91,79	92	90,61	91,02	91	91,38	91,52	90,75	90,48		91		90,89	91,07	91,51	91,43	
	Difference from best		-1,09	-0,46	-0,96	-0,11	-1,2	-1,18	-0,98	-0,82	-0,94	-0,98	-0,84	-0,91	-0,21	0	-1,39	-0,98	-1	-0,62	-0,48	-1,25	-1,52		-1		-1,11	-0,93	-0,49	-0,57	
6	Accuracy, %		91,91	92,07	92,18	91,32	92,3	91,84	92,09	92,04	91,91	91,86	91,79	91,64	91,78		92,05	91,56	91,76	91,73	91,77	91,74	91,9		92,3		91,78	92,16	92,21	92,29	
	Difference from best		-0,39	-0,23	-0,12	-0,98	0	-0,46	-0,21	-0,26	-0,39	-0,44	-0,51	-0,66	-0,52		-0,25	-0,74	-0,54	-0,57	-0,53	-0,56	-0,4		0		-0,52	-0,14	-0,09	-0,01	

Table 3.5 The detailed results of the first six consecutive steps of the forward feature selection procedure for the DT model.

Light GBM

Iteration	Feature	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
1	Accuracy, %	51,12	50,49	74,36	62,18	71,87	67,05	62,43	54,2	53,46	54,21	54,49	52,92	71,9	71,66	72,04	67,72	66,48	59,8	64,49	62,08	51,4	70,04	75,12	53,58	51,08	52,13	63	50,79	57,38	64,28
	Difference from best	-24	-24,63	-0,76	-12,94	-3,25	-8,07	-12,69	-20,92	-21,66	-20,91	-20,63	-22,2	-3,22	-3,46	-3,08	-7,4	-8,64	-15,32	-10,63	-13,04	-23,72	-5,08	0	-21,54	-24,04	-22,99	-12,12	-24,33	-17,74	-10,84
2	Accuracy, %	51,12	50,49	74,36	62,18	71,87	67,05	62,43	54,2	53,46	54,21	54,49	52,92	71,9	71,66	72,04	67,72	66,48	59,8	64,49	62,08	51,4	70,04		75,12	75,12	75,12	75,12	75,12	75,12	75,12
	Difference from best	-24	-24,63	-0,76	-12,94	-3,25	-8,07	-12,69	-20,92	-21,66	-20,91	-20,63	-22,2	-3,22	-3,46	-3,08	-7,4	-8,64	-15,32	-10,63	-13,04	-23,72	-5,08		0	0	0	0	0	0	0
3	Accuracy, %	81,51	82,04	82,33	81,52	81,56	81,57	80,31	81,75	81,83	82,11	81,91	81,62	81,23	81,47	81,96	81,59	78,37	79,26	78,58	77,21	82,91	82,54			83,16	83,16	83,16	83,16	83,16	83,16
	Difference from best	-1,65	-1,12	-0,83	-1,64	-1,6	-1,59	-2,85	-1,41	-1,33	-1,05	-1,25	-1,54	-1,93	-1,69	-1,2	-1,57	-4,79	-3,9	-4,58	-5,95	-0,25	-0,62			0	0	0	0	0	0
4	Accuracy, %	86,6	89,02	87,59	86,89	87,46	86,73	87,98	86,71	86,5	86,61	87,34	86,8	87,46	87,15	87,66	87,28	86,96	86,76	86,6	86,04	86,67	85,91				87,32	87,38	87,19	87,36	89,99
	Difference from best	-3,39	-0,97	-2,4	-3,1	-2,53	-3,26	-2,01	-3,28	-3,49	-3,38	-2,65	-3,19	-2,53	-2,84	-2,33	-2,71	-3,03	-3,23	-3,39	-3,95	-3,32	-4,08				-2,67	-2,61	-2,8	-2,63	0
5	Accuracy, %	93,28	91,37	92,04	90,84	92,74	90,83	91,12	92,49	92,97	91,93	91,32	92,64	91,84	92,61	92,43	90,91	91,08	91,33	91,75	91,68	91	90,35				90,79	91,03	93,15	93,14	
	Difference from best	0	-1,91	-1,24	-2,44	-0,54	-2,45	-2,16	-0,79	-0,31	-1,35	-1,96	-0,64	-1,44	-0,67	-0,85	-2,37	-2,2	-1,95	-1,53	-1,6	-2,28	-2,93				-2,49	-2,25	-0,13	-0,14	
6	Accuracy, %		93,08	94,38	93,69	94,44	93,74	93,77	93,26	93,11	93,29	93,67	93,13	94	94,08	94,09	93,81	93,51	93,85	93,58	93,91	93,33	92,91				93,12	93,79	93,6	93,36	
	Difference from best		-1,36	-0,06	-0,75	0	-0,7	-0,67	-1,18	-1,33	-1,15	-0,77	-1,31	-0,44	-0,36	-0,35	-0,63	-0,93	-0,59	-0,86	-0,53	-1,11	-1,53				-1,32	-0,65	-0,84	-1,08	

Table 3.6 The detailed results of the first six consecutive steps of the forward feature selection procedure for the Light GBM model.

XGBoost

Iteration	Feature	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
1	Accuracy, %	49,95	50,67	72,36	60,92	70	66,33	61,54	52,39	50,93	51,78	52,6	51,02	70,36	69,9	70,71	66,55	66,04	58,33	64,26	60,52	51,26	70,2	73,56	53,2	50,69	52,01	62,44	50,76	54,22	62,07
	Difference from best	-23,61	-22,89	-1,2	-12,64	-3,56	-7,23	-12,02	-21,17	-22,63	-21,78	-20,96	-22,54	-3,2	-3,66	-2,85	-7,01	-7,52	-15,23	-9,3	-13,04	-22,3	-3,36	0	-20,36	-22,87	-21,55	-11,12	-22,8	-19,34	-11,49
2	Accuracy, %	76,22	77,22	80,88	76,74	80,21	76,92	74,46	77	76,62	76,69	76,11	76,56	79,53	79,85	79,83	77,11	75,8	74,27	74,53	74,67	76,52	78,32		76,31	76,39	76,93	77,3	75,98	77,21	75,9
	Difference from best	-4,66	-3,66	0	-4,14	-0,67	-3,96	-6,42	-3,88	-4,26	-4,19	-4,77	-4,32	-1,35	-1,03	-1,05	-3,77	-5,08	-6,61	-6,35	-6,21	-4,36	-2,56		-4,57	-4,49	-3,95	-3,58	-4,9	-3,67	-4,98
3	Accuracy, %	80,98	81,22		80,6	80,2	80,82	81,05	81,11	81,29	80,96	80,84	80,96	80,58	80,34	80,94	81,02	80,52	80,82	81,02	80,27	81,18	81,38		80,91	80,93	81,36	80,58	80,73	81,68	81,68
	Difference from best	-0,7	-0,46		-1,08	-1,48	-0,86	-0,63	-0,57	-0,39	-0,72	-0,84	-0,72	-1,1	-1,34	-0,74	-0,66	-1,16	-0,86	-0,66	-1,41	-0,5	-0,3		-0,77	-0,75	-0,32	-1,1	-0,95	0	0
4	Accuracy, %	87,45	87,52		86,99	85,88	86,77	86,91	87,76	88,01	87,73	87,46	87,84	85,51	85,82	85,96	86,69	86,76	86,42	86,92	86,29	87,24	87,29		87,17	87,51	87,17	86,97	86,75		87,04
	Difference from best	-0,56	-0,49		-1,02	-2,13	-1,24	-1,1	-0,25	0	-0,28	-0,55	-0,17	-2,5	-2,19	-2,05	-1,32	-1,25	-1,59	-1,09	-1,72	-0,77	-0,72		-0,84	-0,5	-0,84	-1,04	-1,26		-0,97
5	Accuracy, %	90,15	91,07		89,84	89,02	89,55	89,81	89,48		90,39	90,42	89,4	89,34	89,54	89,61	89,99	89,71	89,96	90,05	90,79	90,67	89,96		90,62	90,49	90,96	89,82	92,58		94,56
	Difference from best	-4,41	-3,49		-4,72	-5,54	-5,01	-4,75	-5,08		-4,17	-4,14	-5,16	-5,22	-5,02	-4,95	-4,57	-4,85	-4,6	-4,51	-3,77	-3,89	-4,6		-3,94	-4,07	-3,6	-4,74	-1,98		0
6	Accuracy, %	94,35	94,08		94,37	94,53	94,38	94,37	94,31		94,45	94,24	94,42	94,46	94,48	94,37	94,44	94,2	94,49	94,57	94,69	94,54	94,21		94,31	94,22	94,32	94,39	94,45		
	Difference from best	-0,34	-0,61		-0,32	-0,16	-0,31	-0,32	-0,38		-0,24	-0,45	-0,27	-0,23	-0,21	-0,32	-0,25	-0,49	-0,2	-0,12	0	-0,15	-0,48		-0,38	-0,47	-0,37	-0,3	-0,24		

Table 3.7 The detailed results of the first six consecutive steps of the forward feature selection procedure for the XGBoost model.

Extra Trees

Iteration	Feature	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
1	Accuracy, %	50,71	51,09	74,11	61,68	71,77	66,83	61,75	53,72	52,34	52,89	53,84	52,79	71,48	71,6	71,78	67	66,04	58,5	64,57	62,26	51,41	70,15	74,76	53,21	51,01	51,37	62,72	51,16	56,44	63,2
	Difference from best	-24,05	-23,67	-0,65	-13,08	-2,99	-7,93	-13,01	-21,04	-22,42	-21,87	-20,92	-21,97	-3,28	-3,16	-2,98	-7,76	-8,72	-16,26	-10,19	-12,5	-23,35	-4,61	0	-21,55	-23,75	-23,39	-12,04	-23,6	-18,32	-11,56
2	Accuracy, %	82,45	83,61	83,25	83,14	82,99	83,01	81,51	83,12	82,92	83,31	83,61	83,12	82,59	83,25	83	83,09	80,14	80,65	79,38	78,61	83,97	83,85		84,16	84,19	83,89	83,07	80,64	81,11	81,5
	Difference from best	-1,74	-0,58	-0,94	-1,05	-1,2	-1,18	-2,68	-1,07	-1,27	-0,88	-0,58	-1,07	-1,6	-0,94	-1,19	-1,1	-4,05	-3,54	-4,81	-5,58	-0,22	-0,34		-0,03	0	-0,3	-1,12	-3,55	-3,08	-2,69
3	Accuracy, %	86,09	86,07	87,22	84,35	86,93	84,2	86	85,83	86	85,25	84,99	85,86	87,02	87	85,86	84,73	86	85,32	86,1	85,12	85,12	84,39		84,96		85,16	84,26	86,32	86,78	88,66
	Difference from best	-2,57	-2,59	-1,44	-4,31	-1,73	-4,46	-2,66	-2,83	-2,66	-3,41	-3,67	-2,8	-1,64	-1,66	-2,8	-3,93	-2,66	-3,34	-2,56	-3,54	-3,54	-4,27		-3,7		-3,5	-4,4	-2,34	-1,88	0
4	Accuracy, %	90,63	88,49	89,16	87,32	89,23	87,46	88,67	89,27	89,75	88,37	87,51	89,05	88,41	89,27	89,4	88,35	88,52	87,97	87,49	87,83	88,39	87,54		88,32		88	87,49	90,23	90,18	0
	Difference from best	0	-2,14	-1,47	-3,31	-1,4	-3,17	-1,96	-1,36	-0,88	-2,26	-3,12	-1,58	-2,22	-1,36	-1,23	-2,28	-2,11	-2,66	-3,14	-2,8	-2,24	-3,09		-2,31		-2,63	-3,14	-0,4	-0,45	
5	Accuracy, %		92	92,63	91,29	92,53	91,43	91,83	92,14	92,27	92,1	91,87	91,61	92,25	92,34	92,78	91,67	91,99	92,19	91,52	91,88	91,97	91,5		91,71		91,73	91,21	92,37	92,67	
	Difference from best		-0,78	-0,15	-1,49	-0,25	-1,35	-0,95	-0,64	-0,51	-0,68	-0,91	-1,17	-0,53	-0,44	0	-1,11	-0,79	-0,59	-1,26	-0,9	-0,81	-1,28		-1,07		-1,05	-1,57	-0,41	-0,11	
6	Accuracy, %		92,12	91,78	91,58	92,13	91,3	92,09	92,25	92,78	92,01	92,3	92,39	92,06	91,88		92,16	92,05	91,84	91,68	91,88	92,24	91,45		92,12		92,05	91,41	93,03	93,07	
	Difference from best		-0,95	-1,29	-1,49	-0,94	-1,77	-0,98	-0,82	-0,29	-1,06	-0,77	-0,68	-1,01	-1,19		-0,91	-1,02	-1,23	-1,39	-1,19	-0,83	-1,62		-0,95		-1,02	-1,66	-0,04	0	

Table 3.8 The detailed results of the first six consecutive steps of the forward feature selection procedure for the Extra trees model.

RF

Iteration	Feature	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
1	Accuracy, %	51,47	50,57	73,49	61,43	71,46	66,52	61,59	53,91	51,89	53,24	54,34	52,73	71,66	71,24	71,66	66,86	66,08	58,44	64,54	61,52	51,3	70,11	74,36	53,08	50,53	51,42	62,67	50,92	56,58	63,14
	Difference from best	-22,89	-23,79	-0,87	-12,93	-2,9	-7,84	-12,77	-20,45	-22,47	-21,12	-20,02	-21,63	-2,7	-3,12	-2,7	-7,5	-8,28	-15,92	-9,82	-12,84	-23,06	-4,25	0	-21,28	-23,83	-22,94	-11,69	-23,44	-17,78	-11,22
2	Accuracy, %	81,66	82,42	82,3	82,13	82,19	82,1	80,52	82,55	82,61	82,63	82,44	82,54	81,68	81,97	82,25	81,69	78,84	79,32	78,61	78,12	83,1	82,71		83,1	83,19	82,64	82,28	80,19	80,59	81,06
	Difference from best	-1,53	-0,77	-0,89	-1,06	-1	-1,09	-2,67	-0,64	-0,58	-0,56	-0,75	-0,65	-1,51	-1,22	-0,94	-1,5	-4,35	-3,87	-4,58	-5,07	-0,09	-0,48		-0,09	0	-0,55	-0,91	-3	-2,6	-2,13
3	Accuracy, %	86,25	87,29	87,1	85,93	87,44	85,55	86,85	85,94	85,97	86,05	85,84	85,91	87,17	86,94	87,41	86,35	86,41	85,94	86,27	85,67	85,99	84,57		85,16		85,78	86,2	86,57	87,15	89,34
	Difference from best	-3,09	-2,05	-2,24	-3,41	-1,9	-3,79	-2,49	-3,4	-3,37	-3,29	-3,5	-3,43	-2,17	-2,4	-1,93	-2,99	-2,93	-3,4	-3,07	-3,67	-3,35	-4,77		-4,18		-3,56	-3,14	-2,77	-2,19	0
4	Accuracy, %	92,87	90,08	91,65	89,56	91,97	89,6	90,56	91,31	92,22	90,44	89,94	91,41	91,64	91,79	92,17	89,86	90,72	90,33	91,04	91,12	90,58	89,78		90,06		89,53	89,79	92,6	92,46	
	Difference from best	0	-2,79	-1,22	-3,31	-0,9	-3,27	-2,31	-1,56	-0,65	-2,43	-2,93	-1,46	-1,23	-1,08	-0,7	-3,01	-2,15	-2,54	-1,83	-1,75	-2,29	-3,09		-2,81		-3,34	-3,08	-0,27	-0,41	
5	Accuracy, %		92,71	93,59	92,77	93,61	92,79	92,63	93,05	92,85	92,94	92,63	92,93	93,61	93,75	93,76	92,93	92,91	93,26	93,17	93,58	92,75	92,56		92,71		92,64	92,85	92,98	93,05	
	Difference from best		-1,05	-0,17	-0,99	-0,15	-0,97	-1,13	-0,71	-0,91	-0,82	-1,13	-0,83	-0,15	-0,01	0	-0,83	-0,85	-0,5	-0,59	-0,18	-1,01	-1,2		-1,05		-1,12	-0,91	-0,78	-0,71	
6	Accuracy, %		93,51	93,65	93,75	93,56	93,8	93,33	93,73	93,78	93,53	93,5	93,7	93,32	93,55		93,58	93,32	93,59	93,58	93,53	93,58	93,6		93,52		93,56	93,51	93,98	94,39	
	Difference from best		-0,88	-0,74	-0,64	-0,83	-0,59	-1,06	-0,66	-0,61	-0,86	-0,89	-0,69	-1,07	-0,84		-0,81	-1,07	-0,8	-0,81	-0,86	-0,81	-0,79		-0,87		-0,83	-0,88	-0,41	0	

Table 3.9 The detailed results of the first six consecutive steps of the forward feature selection procedure for the RF model.

AdaBoost

Iteration	Feature	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
1	Accuracy, %	51,19	50,75	73,75	61,86	71,5	66,81	61,95	54,17	53,06	53,84	54,42	53,07	71,75	71,32	71,77	67,16	66,04	58,77	64,63	62,05	51,36	70,09	74,5	53,07	51,1	51,86	62,66	50,9	56,82	63,61
	Difference from best	-23,31	-23,75	-0,75	-12,64	-3	-7,69	-12,55	-20,33	-21,44	-20,66	-20,08	-21,43	-2,75	-3,18	-2,73	-7,34	-8,46	-15,73	-9,87	-12,45	-23,14	-4,41	0	-21,43	-23,4	-22,64	-11,84	-23,6	-17,68	-10,89
2	Accuracy, %	79,57	81,23	79,99	80,35	80,21	80,12	79,09	80,67	80,4	80,39	80,73	80,29	80,16	80,27	79,76	79,9	77,72	77,57	77,41	76,93	81,64	80,99		81,39	81,3	80,79	80,05	77,83	78,94	78,51
	Difference from best	-2,07	-0,41	-1,65	-1,29	-1,43	-1,52	-2,55	-0,97	-1,24	-1,25	-0,91	-1,35	-1,48	-1,37	-1,88	-1,74	-3,92	-4,07	-4,23	-4,71	0	-0,65		-0,25	-0,34	-0,85	-1,59	-3,81	-2,7	-3,13
3	Accuracy, %	83,29	84,13	84,32	83,81	84,85	83,94	84,69	83,51	83,35	83,29	83,79	83,52	84,68	84,22	84,37	84,41	84,26	82,7	83,17	83,09		81,37		83,5	84,09	83,45	84,28	83,34	84,91	86,95
	Difference from best	-3,66	-2,82	-2,63	-3,14	-2,1	-3,01	-2,26	-3,44	-3,6	-3,66	-3,16	-3,43	-2,27	-2,73	-2,58	-2,54	-2,69	-4,25	-3,78	-3,86		-5,58		-3,45	-2,86	-3,5	-2,67	-3,61	-2,04	0
4	Accuracy, %	92,83	90,55	91,56	89,62	92,65	88,95	90,7	92,18	92,59	90,99	89,88	92,63	91,28	92,37	92,22	89,87	90,39	89,99	90,22	90,65		86,7		88,49	89,26	90,69	90,73	92,96	92,62	
	Difference from best	-0,13	-2,41	-1,4	-3,34	-0,31	-4,01	-2,26	-0,78	-0,37	-1,97	-3,08	-0,33	-1,68	-0,59	-0,74	-3,09	-2,57	-2,97	-2,74	-2,31		-6,26		-4,47	-3,7	-2,27	-2,23	0	-0,34	
5	Accuracy, %	93,44	92,96	94,76	93,55	95,07	93,95	93,8	93,16	93,5	93,1	93,41	93,1	94,68	95,15	95,15	93,8	94,01	94,31	94,68	94,47		92,69		92,74	92,82	93,02	93,83		93,19	
	Difference from best	-1,71	-2,19	-0,39	-1,6	-0,08	-1,2	-1,35	-1,99	-1,65	-2,05	-1,74	-2,05	-0,47	0	0	-1,35	-1,14	-0,84	-0,47	-0,68		-2,46		-2,41	-2,33	-2,13	-1,32		-1,96	
6	Accuracy, %	95,3	94,97	95,34	95,06	95,3	95,29	95,12	95,16	95,42	95,02	94,97	95,06	94,97		94,98	95,32	95,03	94,91	94,9	95,23		94,9		95,09	95,07	94,91	95,16		95,12	
	Difference from best	-0,12	-0,45	-0,08	-0,36	-0,12	-0,13	-0,3	-0,26	0	-0,4	-0,45	-0,36	-0,45		-0,44	-0,1	-0,39	-0,51	-0,52	-0,19		-0,52		-0,33	-0,35	-0,51	-0,26		-0,3	

Table 3.10 The detailed results of the first six consecutive steps of the forward feature selection procedure for the AdaBoost model.

SVM

Iteration	Feature	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
1	Accuracy, %	52,96	52,51	64,46	62,76	61,87	67,52	62,46	55,58	53,92	55,28	53,52	55,61	61,47	61,38	62,42	67,44	66,03	59,37	64,12	61,03	51,71	70,48	74,29	54,95	53,44	52,64	63,54	51,8	57,03	62,44
	Difference from best	-21,33	-21,78	-9,83	-11,53	-12,42	-6,77	-11,83	-18,71	-20,37	-19,01	-20,77	-18,68	-12,82	-12,91	-11,87	-6,85	-8,26	-14,92	-10,17	-13,26	-22,58	-3,81	0	-19,34	-20,85	-21,65	-10,75	-22,49	-17,26	-11,85
2	Accuracy, %	82,12	83,31	81,51	83,19	81,76	82,72	80,97	83,01	82,69	82,67	81,46	83,14	78,61	81,9	78,82	81,81	79,06	79,91	77,57	77,24	83,5	83,44		84,01	83,89	83,39	82,76	79,94	80,34	80,68
	Difference from best	-1,89	-0,7	-2,5	-0,82	-2,25	-1,29	-3,04	-1	-1,32	-1,34	-2,55	-0,87	-5,4	-2,11	-5,19	-2,2	-4,95	-4,1	-6,44	-6,77	-0,51	-0,57		0	-0,12	-0,62	-1,25	-4,07	-3,67	-3,33
3	Accuracy, %	86,26	87,08	86,37	84,57	86,34	84,46	85,74	85,5	85,56	84,44	84,24	85,62	85,48	86,29	84,5	84,62	85,53	85,32	85,01	85,6	84,88	84,72			84,9	85,86	84,9	86,32	86,8	88,23
	Difference from best	-1,97	-1,15	-1,86	-3,66	-1,89	-3,77	-2,49	-2,73	-2,67	-3,79	-3,99	-2,61	-2,75	-1,94	-3,73	-3,61	-2,7	-2,91	-3,22	-2,63	-3,35	-3,51			-3,33	-2,37	-3,33	-1,91	-1,43	0
4	Accuracy, %	93,93	91,57	90,35	89,37	91,62	89,08	89,57	91,84	92,76	89,19	88,34	92,62	89,59	91,67	89,24	89,3	89,87	89,49	90,47	90,06	89,64	89,17			90,63	91,51	89,84	94,14	93,9	
	Difference from best	-0,21	-2,57	-3,79	-4,77	-2,52	-5,06	-4,57	-2,3	-1,38	-4,95	-5,8	-1,52	-4,55	-2,47	-4,9	-4,84	-4,27	-4,65	-3,67	-4,08	-4,5	-4,97			-3,51	-2,63	-4,3	0	-0,24	
5	Accuracy, %	94,92	94,81	95,25	94,81	95,94	94,82	94,51	94,59	94,65	94,15	94,21	94,77	94,75	95,97	94,3	94,62	94,64	94,79	94,92	95,23	94,86	94,58			94,69	94,88	94,92		94,16	
	Difference from best	-1,05	-1,16	-0,72	-1,16	-0,03	-1,15	-1,46	-1,38	-1,32	-1,82	-1,76	-1,2	-1,22	0	-1,67	-1,35	-1,33	-1,18	-1,05	-0,74	-1,11	-1,39			-1,28	-1,09	-1,05		-1,81	
6	Accuracy, %	96,07	96,04	96,04	96,36	95,98	96,4	96,04	96,2	96,28	95,98	95,93	96,23	95,98		96	96,19	96,01	96,26	96,15	96,09	96,16	96,14			96,08	96,03	96,32		96,08	
	Difference from best	-0,33	-0,36	-0,36	-0,04	-0,42	0	-0,36	-0,2	-0,12	-0,42	-0,47	-0,17	-0,42		-0,4	-0,21	-0,39	-0,14	-0,25	-0,31	-0,24	-0,26			-0,32	-0,37	-0,08		-0,32	

Table 3.11 The detailed results of the first six consecutive steps of the forward feature selection procedure for the SVM model.

KNN

Iteration	Feature	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30
1	Accuracy, %	49,96	50,38	71,47	58,46	69,11	62,58	57,02	51,74	51,17	51,98	51,77	50,51	69,73	69,14	69,27	63,07	60,91	56,08	60,31	59,58	50,49	64,06	72,06	52,6	51,51	51,71	58,91	50,26	53,57	61,1
	Difference from best	-22,1	-21,68	-0,59	-13,6	-2,95	-9,48	-15,04	-20,32	-20,89	-20,08	-20,29	-21,55	-2,33	-2,92	-2,79	-8,99	-11,15	-15,98	-11,75	-12,48	-21,57	-8	0	-19,46	-20,55	-20,35	-13,15	-21,8	-18,49	-10,96
2	Accuracy, %	79,26	81,05	80,63	80,89	80,11	80,3	78,57	80,12	80,45	81,64	81,75	80,73	80,03	80,06	80,95	80,34	76,7	77,87	76,42	76,25	81,26	81,21		81,55	81,46	81	80,76	77,1	77,76	79,2
	Difference from best	-2,49	-0,7	-1,12	-0,86	-1,64	-1,45	-3,18	-1,63	-1,3	-0,11	0	-1,02	-1,72	-1,69	-0,8	-1,41	-5,05	-3,88	-5,33	-5,5	-0,49	-0,54		-0,2	-0,29	-0,75	-0,99	-4,65	-3,99	-2,55
3	Accuracy, %	81,57	81,65	85,4	81,7	84,52	81,94	83,47	81,15	82,12	82,24		81,45	85,44	83,98	84,84	82,55	83,31	81,55	83,44	81,65	82,58	82,58		82,13	82,2	81,53	82,24	83,44	84,03	85,69
	Difference from best	-4,12	-4,04	-0,29	-3,99	-1,17	-3,75	-2,22	-4,54	-3,57	-3,45		-4,24	-0,25	-1,71	-0,85	-3,14	-2,38	-4,14	-2,25	-4,04	-3,11	-3,11		-3,56	-3,49	-4,16	-3,45	-2,25	-1,66	0
4	Accuracy, %	90,32	87,87	90,3	87,89	91,28	87,89	87,19	88,86	90,15	88,25		89,45	88,39	91,35	88,55	88,08	88,16	87,87	87,11	88,33	87,82	87,39		87,63	87,77	88,01	88,47	87,45	87,83	
	Difference from best	-1,03	-3,48	-1,05	-3,46	-0,07	-3,46	-4,16	-2,49	-1,2	-3,1		-1,9	-2,96	0	-2,8	-3,27	-3,19	-3,48	-4,24	-3,02	-3,53	-3,96		-3,72	-3,58	-3,34	-2,88	-3,9	-3,52	
5	Accuracy, %	92,19	92	91,83	92,4	91,15	92,48	92,67	92,41	92,3	92,66		92,4	91,9		91,44	92,33	92,46	92,29	92,19	90,92	91,93	92,53		92,49	92,44	92,14	92,49	91,32	92,03	
	Difference from best	-0,48	-0,67	-0,84	-0,27	-1,52	-0,19	0	-0,26	-0,37	-0,01		-0,27	-0,77		-1,23	-0,34	-0,21	-0,38	-0,48	-1,75	-0,74	-0,14		-0,18	-0,23	-0,53	-0,18	-1,35	-0,64	
6	Accuracy, %	93,66	92,33	92,79	92,72	92,85	92,74		93,29	93,43	92,98		93,25	92,78		92,7	92,68	92,58	92,69	92,51	92,57	92,51	93,07		93,01	92,67	92,52	92,58	93,42	94,26	
	Difference from best	-0,6	-1,93	-1,47	-1,54	-1,41	-1,52		-0,97	-0,83	-1,28		-1,01	-1,48		-1,56	-1,58	-1,68	-1,57	-1,75	-1,69	-1,75	-1,19		-1,25	-1,59	-1,74	-1,68	-0,84	0	

Table 3.12 The detailed results of the first six consecutive steps of the forward feature selection procedure for the KNN model.

Following an in-depth analysis, four subsets of features were selected, with each associated with a set of classifiers:

1. DT, Light GBM, Extra Trees, RF: F23, F25, F30, and F1.
2. AdaBoost, SVM: F23, F25, F30, F29, and F5.
3. KNN: F23, F11, F30, and F5.
4. XGBoost: F23, F5, F29, and F25.

Each subset was selected to minimize the diversity of features without a significant decrease in classification accuracy. In addition to the dedicated features for each selected classifier, a universal subset of features was extracted for use with all selected classifiers. This universal subset is the sum of the four dedicated subsets, and thus consists of seven CPIs: F1, F5, F11, F23, F25, F29, and F30. Fig. 3.10 presents the universal subset together with the dedicated subsets for selected classifiers, in the form of decision paths.

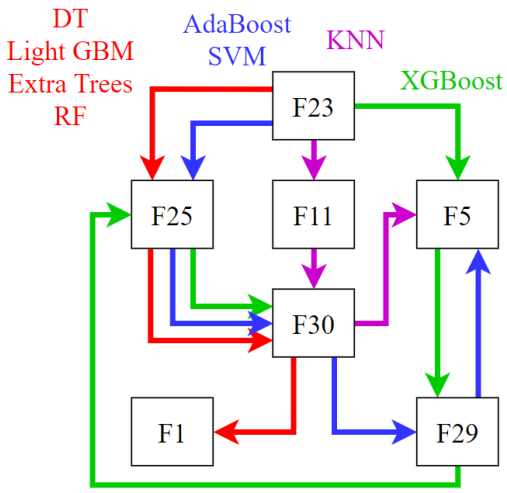

Fig. 3.10 A graphical representation of the universal and dedicated subsets of features for selected classifiers, obtained by the feature reduction procedure.

Table 3.13 presents the classification accuracy of selected algorithms for the full set of features (30 CPIs), the universal subset of features (seven CPIs), and the dedicated subset of features for each selected classifier. The results show that the universal subset of features ensures high classification accuracy, independent of the algorithm used. For some classifiers, the classification accuracy is lower when using the universal subset than when using the full set of features. However, the maximum difference of 0.45% is negligible from a practical perspective. Conversely, for some classifiers the accuracy increases slightly when using the universal subset, which confirms the usefulness of the universal subset. The dedicated subsets of features also provide relatively high classification accuracy. However, the decrease in accuracy is somewhat significant when compared to the full set of features. Note that the subset of features is even smaller than the universal subset, and thus the number of

captured key features decreases. The best possible classification accuracy is again obtained by the SVM classifier, including when using the appropriate dedicated subset.

Number of features	Full set of features	Universal subset of features	Dedicated subset of features
DT	91.54	91.97	90.6
Light GBM	95.23	95.00	91.53
XGBoost	95.26	95.01	94.0
Extra Trees	92.85	92.40	89.23
RF	93.70	93.87	91.97
AdaBoost	95.48	95.57	95.3
SVM	**96.17**	**96.18**	**95.8**
KNN	92.24	93.62	92.21

Table 3.13 The classification accuracy of selected classification models trained with a full set, universal subset, and dedicated subset of features.

The high classification accuracy obtained when using the universal subset confirms that the features captured by the CPIs in this subset are sufficient for precise classification. In the following sections, the universal subset of features is analyzed further and then used to implement the proposed CPA system on a PID.

3.3.1 In-depth analysis of the universal subset features

As described in the previous section, the most important feature for classification is F23. Use of this feature alone provides a classification accuracy of approximately 74–75%. The feature is defined as the ratio of rising and falling time of the first maximum peak of the closed loop response, and thus directly describes the shape of maximum peak. This is a key aspect of the entire response. As the value of F23 goes to zero (F23 → 0), the falling time dominates and the closed loop response is sluggish. This can be caused by an integral constant that is too high, for example. As the value of F23 goes to infinity (F23 → ∞), the rising time dominates and the closed loop response is aggressive, with oscillatory behavior. Thus, a good ratio is critical for satisfactory control performance. However, even if F23 indicates satisfactory performance, the overall closed loop response can be still poor due to conservative or sluggish dynamic behavior. Such cases are presented in Fig. 3.11, which presents three closed loop responses with the same value of the F23 index.

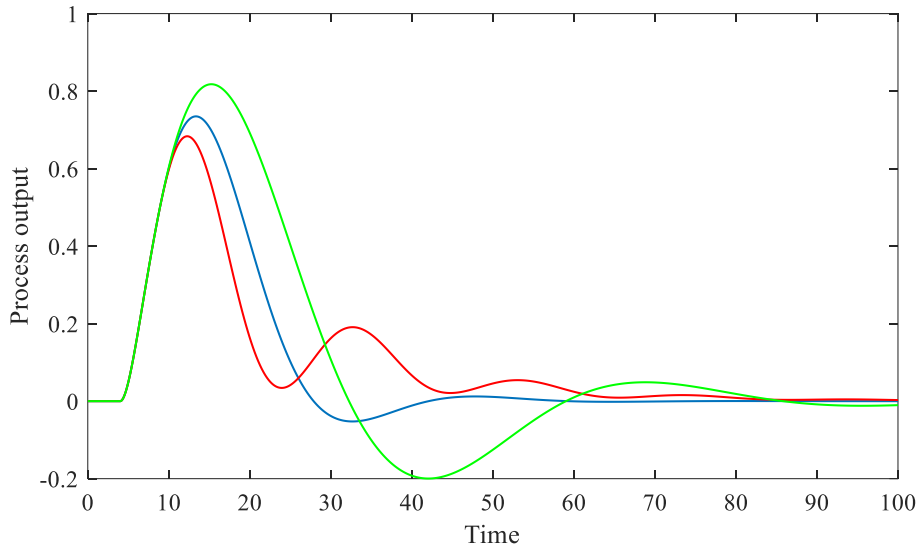

Fig. 3.11 The first comparison between three example closed loop responses, each with the same value of F23.

The shape of the first maximum peak is similar for each closed loop response. However, further stabilization about the setpoint differs substantially. This explains why additional indices must be included to provide a more accurate assessment. As indicated by Fig. 3.10, the F23 index should be further supplemented by the F5 index, which describes the ratio of undershoot and overshoot of the process response, the F11 integral index, which considers the product of the control error and the squared time vector, or the F25 index, which describes the time at which the response reaches 50% of the maximum peak value. Notably, indices F11 and F25 directly capture features of the entire response, while F5 describes the closed loop response once the maximum peak has been damped. Thus, when combined with F23, each index increases the effectiveness with which the dynamic behavior of the system is described. Cases in which the F23 index indicates acceptable behavior but the response is too sluggish or too aggressive can be caused by the following:

- the undershoot, as described by F5, is too large or too small, producing aggressive or sluggish dynamic behavior, respectively;
- a high control error, as described by the integral index F11, causes either aggressive and conservative behavior, to persist for a relatively long period of time; or
- as described by F25, the closed loop response reaches 50% of the maximum peak value too slowly for sluggish or aggressive behavior, or too quickly for aggressive behavior.

If two selected indices are similar, the remaining CPIs will correctly assess performance. Fig. 3.12 presents a further three closed loop responses. The first maximum peak is similar for both the blue and red closed loop responses. This indicates similar values of F23 and F25. However, the dynamic behavior differs once the first maximum peak is damped. Despite this, the overall behavior of these closed loop responses is more similar than that of the three responses depicted in Fig. 3.11, for which

only the F23 index takes similar values. The difference between the blue and red responses shown in Fig. 3.12 is successfully captured by the F5 and F11 indices.

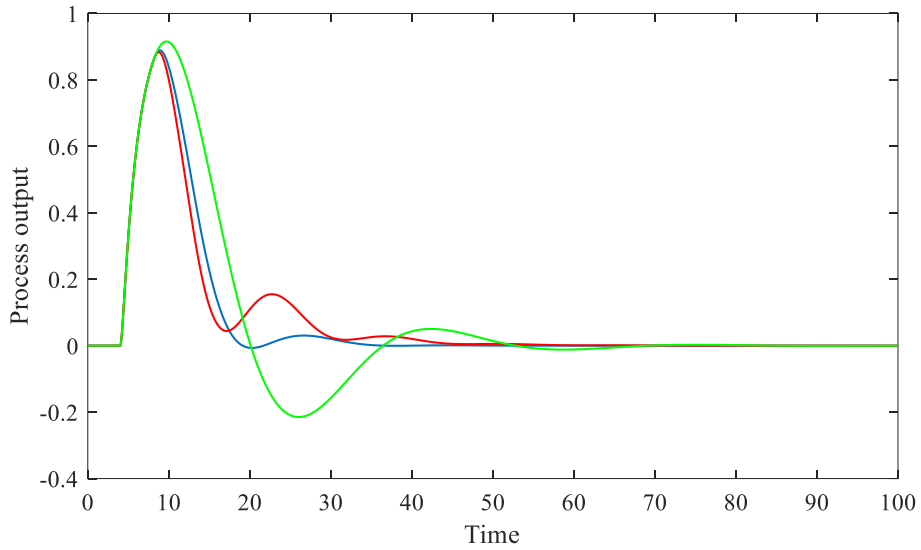

Fig. 3.12 The second comparison between three example closed loop responses. The red and blue responses have similar values of the F23 and F25 indices.

A third comparison is presented in Fig. 3.13. The blue and green closed loop responses have similar shapes, with the same F23 index, but the green response is too conservative. This behavior is captured by the F5, F11, or F25 indices. The value of the F23 index would also be similar for the red closed loop response that is dominated by oscillatory behavior. This is also be captured by additional CPIs.

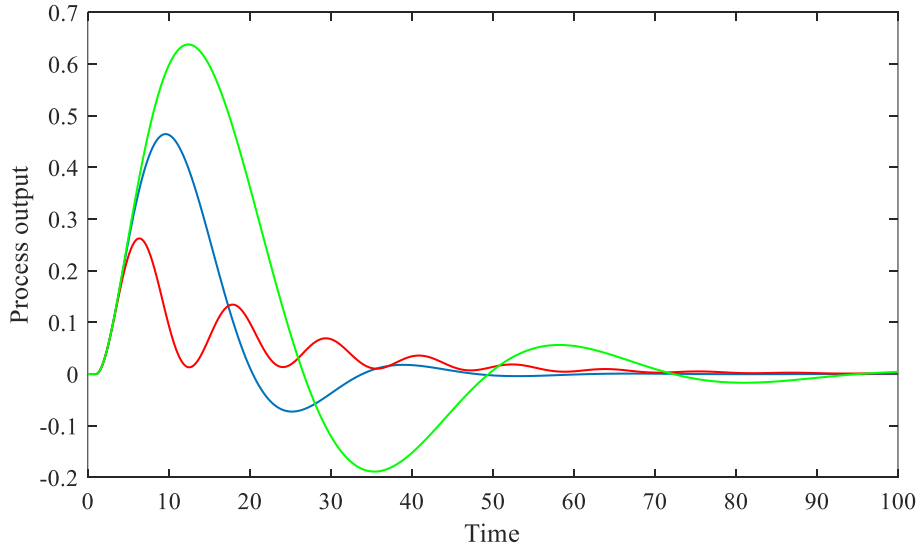

Fig. 3.13 The third comparison between three example closed loop responses.

Notably, the subsets of features are supplemented by the derivative-based indices F29 and F30. These indices indirectly describe the shape of the maximum peak, and thus provide a more precise assessment of that section of the response, in conjunction with F23.

An analysis of all seven features within the universal subset reveals the presence of two ratio-based indices, F5 and F30, together with their components, F1 and F29, respectively. This suggests that both ratio-based features and value-based features are important for classification.

3.4 Technical aspects of the PLC implementation

The universal subset of features allows the classifier structure to be simplified substantially, with the number of parameters decreasing greatly. This simplified structure allows the system to be implemented directly on a PLC as a general-purpose function block. The SVM algorithm was selected for practical implementation, as it provides the highest possible accuracy. This algorithm is also the most suitable for implementation on a PLC, because its output can be calculated directly from an analytical formula. For this particular classifier, the use of a universal subset of features provides a 63.2% reduction in the memory required to store the SVM parameters, from 475.1 kB for the full set of features to 174.8 kB for the universal subset. The CPA system was implemented in the form of a *ControlPerformanceAssessment* function block for Siemens S7-1200/1500 PLCs using TIA Portal v15.1. The function block incorporates several functionalities, including steady state detection, closed loop response data collection, SOPDT model identification, process dynamics normalization, CPI calculation, and SVM-based classification. The implemented function block can be integrated with *PID_Compact*, as presented in Fig. 3.14. Note that such an integration does not require normal operation of the closed loop system to be halted or otherwise interrupted.

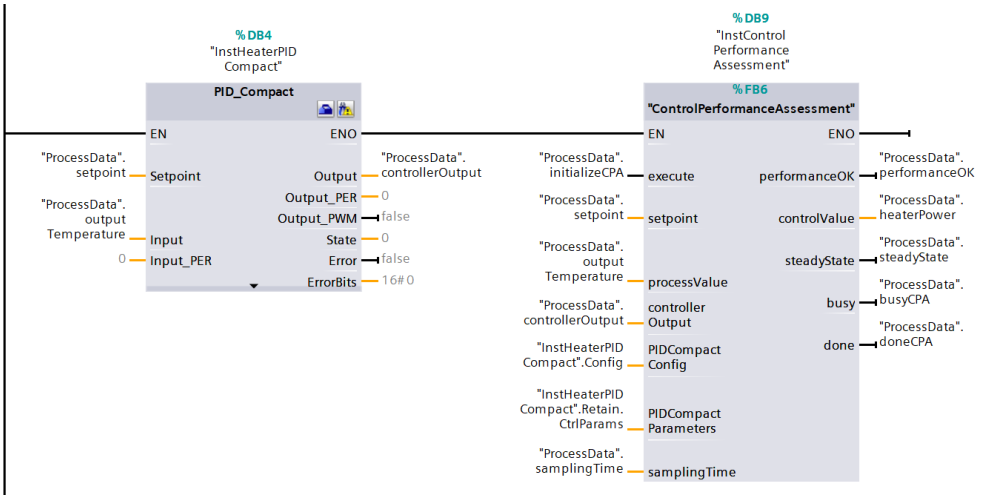

Fig. 3.14 The integration of a PLC-based implementation of the proposed CPA system with an existing PID control system for Siemens S7-1200/1500 PLCs, using the TIA Portal.

Fig 3.15 presents the sequence of operations implemented within the function block. The CPA procedure is initialized with the *execute* input. Then, once the ICM method [50] has detected a steady-state, the step load disturbance is introduced to the control system, with an amplitude given by 10% of the total range of the control signal (stored in the *PIDCompactConfig* structure). During the subsequent transient behavior, the closed loop response data (*setpoint* and *processValue*) are collected according to the constant *samplingTime*. This sampling rate is ensured by the cyclic interrupt OB30

organization block. Once the steady state is detected again, the gathered data is used for SOPDT model identification. For this purpose, the leapfrogging optimization method is used [101], [102] to minimize the modelling error of the simulated closed loop response. This simple yet powerful optimization method does not require complex calculations, and so is suitable for the PLC-based implementation. The obtained SOPDT parameters are then normalized, and new closed loop response data is simulated, allowing the seven CPIs to be calculated. Once the feature vector is generated, the output of SVM classifier is calculated, based on the sign of the expression

$$\sum_{i \in SV} y_i \beta_i K(x_i, x) + b, \tag{3.2}$$

where x is a testing sample vector consisting of the seven calculated CPIs, the x_i are training support vectors (SV) with class y_i, the β_i are dual coefficients, K is the Gaussian kernel, and b is an independent term. The SVM parameters were obtained during offline simulation studies with the Python scikit-learn library, and transferred to the internal PLC memory in the form of an array, stored within a data block. Satisfactory control performance (OK) is indicated by a high *performanceOK* output.

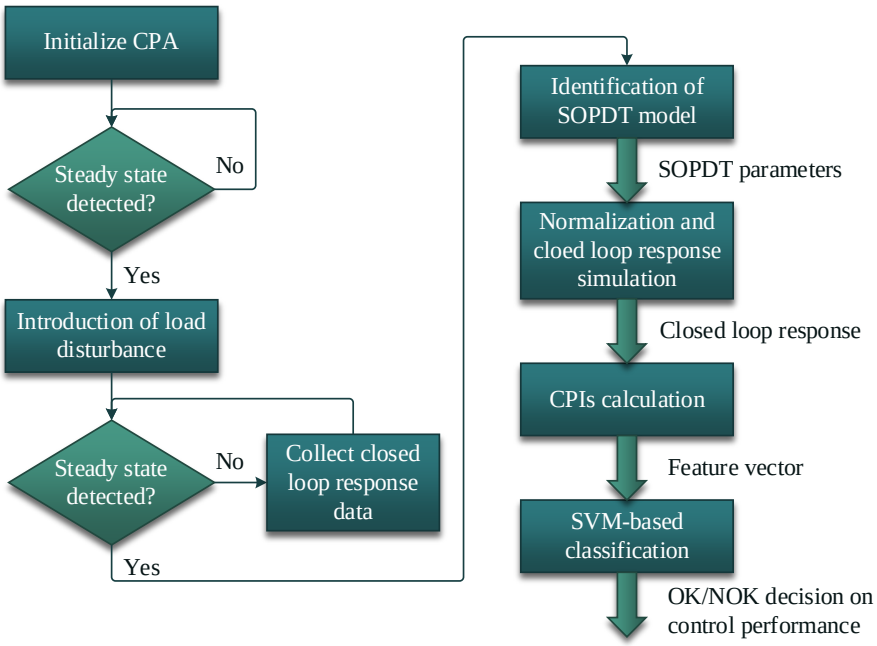

Fig. 3.15 A block diagram showing the PLC-based implementation of the proposed CPA system.

The *ControlPerformanceAssessment* function block was implemented using several sub-functions. Fig. 3.16 presents the block diagram of the *ControlPerformanceAssessment* function block, including a description of the implemented sub-functions and the flow of key signals.

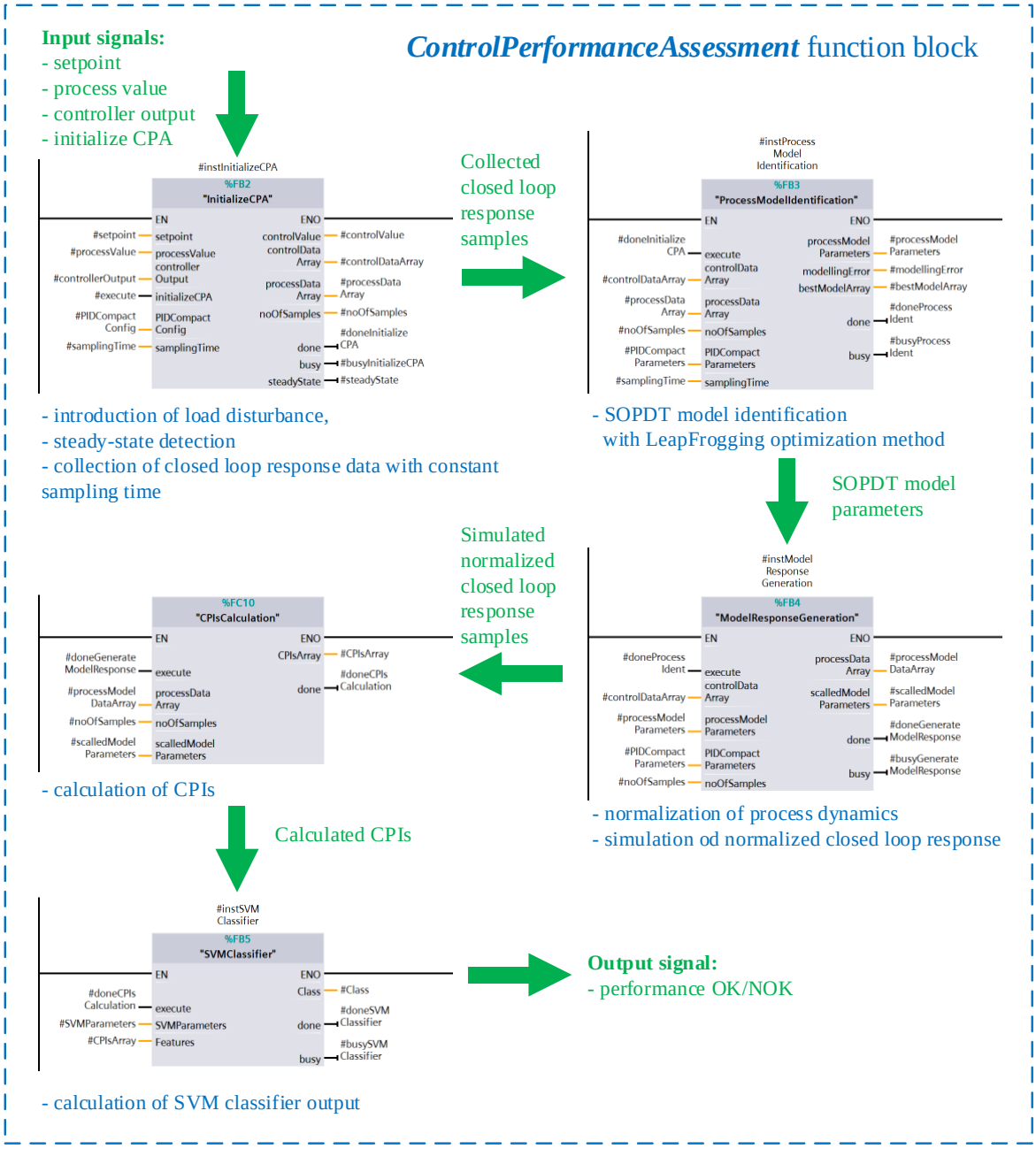

Fig. 3.16 A block diagram for the *ControlPerformanceAssessment* function block, including a description of the implemented sub-functions and the flow of key signals.

Note that the operating cycle time of a PLC should be as small as possible. Typically, the maximum cycle time cannot exceed 150 ms—the default watchdog value. During this interval all logical and arithmetical calculations must be performed. In practice, the typical cycle time is much shorter. All functionalities that are executed followed data collection still consume resources, and are thus implemented in an asynchronous manner, to ensure that the maximum cycle time is not violated.

3.5 Experimental verification

The implemented CPA system was used to assess a temperature controller *TC* within the heat exchange and distribution system (Appendix 4). For clarity, 10 sets of PID tunings were selected. Half

of them were calculated by introducing small deviations into the reference tunings to ensure satisfactory (OK) performance, based on predefined criteria. The other half were derived from Appendix 5. These tunings ensured stable behavior, but were significantly different from the predefined reference. This approach produced 10 different closed loop systems for assessment with the implemented CPA system. Fig. 3.16 presents the closed loop responses that were assessed as either OK or NOK (continuous lines), compared with the predefined reference responses (dashed lines). Each closed loop response was used independently for SOPDT model identification, resulting in small variations in the obtained model parameters. Thus, the calculated reference response differs slightly for each closed loop system. Despite this, the classification accuracy is high, given that the closed loop responses that were assessed as OK are similar to the predefined reference responses. Conversely, the dynamic behavior of the NOK closed loop responses is significantly different from the predefined reference behavior.

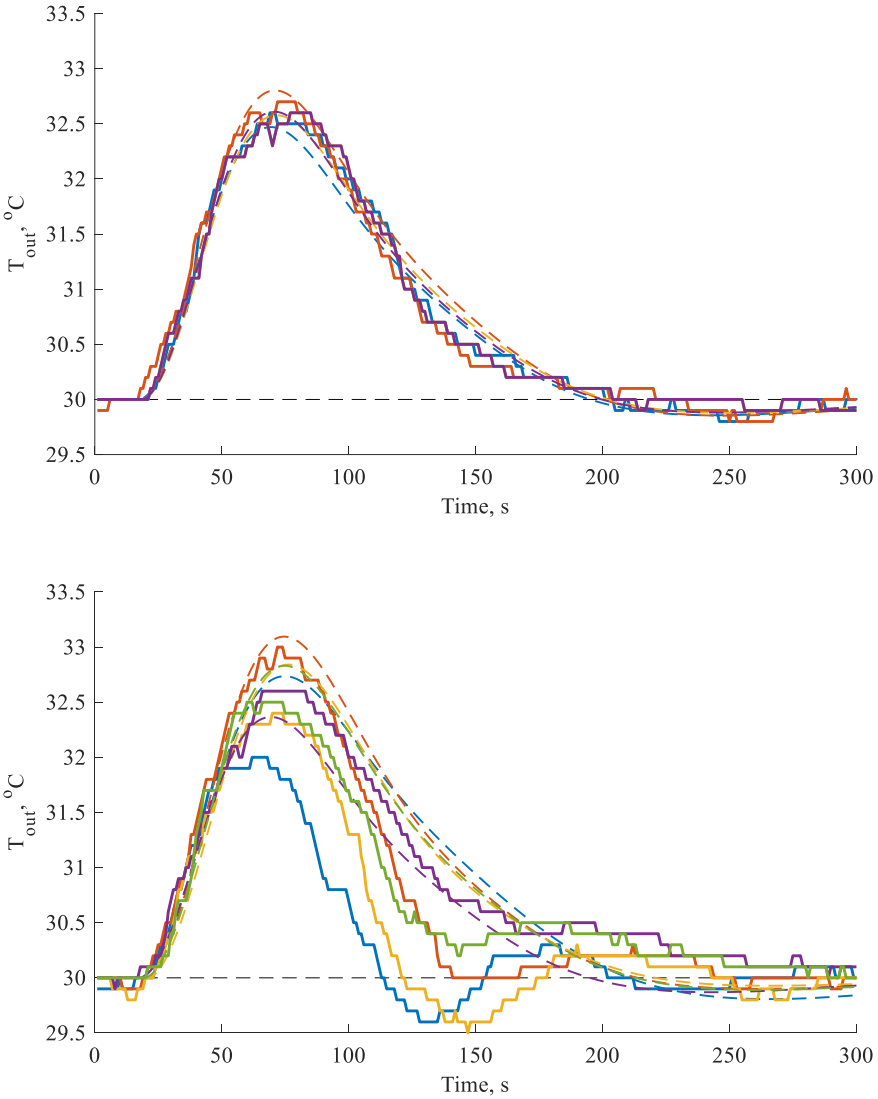

Fig. 3.17 The results of the assessment of real control systems using the PLC implementation of the proposed CPA system, showing OK (top) and NOK (bottom) closed loop responses (continuous lines), compared with the corresponding reference responses (dashed lines).

4 Conclusion

This dissertation described the development of an automatic system for the performance assessment of PID-based control systems. To increase the practical usability of the system, key requirements were first defined. The most important requirements were generality, so that the proposed system could be used with a wide range of self-stabilizing processes, the possibility of parametrization, so that the system could be adapted to predefined technological requirements, and the possibility of implementation on a PLC. The proposed system assesses control performance based on the closed loop response to a step load disturbance. This assessment procedure is convenient from a practical perspective, as it can be manually introduced to any considered control system, without significant disturbance from its operating point. The central principle of the proposed system is similar to that of other existing CPA algorithms: to compare actual control performance with a predefined reference performance. To generate this reference closed loop response, reference PID tunings were derived following the minimization of the *IAE* index for a closed loop response with constraints on the gain and phase margins to ensure proper robustness.

However, different tunings can be used as a reference without loss of generality. The training dataset for the ML algorithms was generated for a wide range of SOPDT processes by slightly modifying the corresponding reference tunings. A method to automatically label the closed loop responses was developed. The method is based on the acceptable deviation of gain and phase margins, with an appropriate normalized distance from the predefined reference. Thirty simple CPIs were used to create a feature vector describing the closed loop response. The CPIs require neither high computational or memory resources. This approach was based on the assumption that the combined application of a large number of the suggested CPIs could capture key features of the closed loop response and provide potentially useful information for performance classification. Various classification algorithms were verified. For the majority of classifiers, classification accuracy was higher than 90%. The SVM classifier achieved the highest accuracy, of greater than 96%.

From a practical perspective, the obtained level of accuracy is sufficient for the application of the proposed CPA system to industrial control systems. The proposed approach was compared with other existing methods, and the obtained results show that it provides additional, useful information concerning control performance. A Cloud-based implementation of the system was developed, and the entire system was verified by applying it to a laboratory heat distribution and exchange system. Finally, the structure of the classifier was minimized to allow the system to be implemented on a PLC. Following deep analysis of index correlations and feature importance, a forward feature selection algorithm was used to determine a universal subset of features that can be used by all selected classifiers. The dimensionality of the feature vector was thus reduced from 30 CPIs to seven, without a significant decrease in classification accuracy. This reduction was sufficient to allow the proposed

CPA system to be directly implemented on a PLC as a ready-to-use, general purpose function block. This block can be easily integrated with existing PID algorithms, without negatively effecting normal operation. Operation of the PLC-based CPA system was again verified using a laboratory control system. The obtained results confirmed the practical usability of the implementation.

In consideration of the above conclusions, the following thesis statement was confirmed: *the developed CPA system can explicitly assess control performance based on load disturbance rejection response data for closed loop systems in process automation, while clearly and objectively considering the predefined assumptions and limitations*.

When conducting the studies described in this thesis, many interesting directions of development of the proposed CPA system emerged. One of the most promising is the use of fuzzy models for more precise assessment of control performance. The use of fuzzy model membership functions could potentially allow the gradual degradation of control performance to be monitored. Such degradation could be caused by slow fouling, for example.

At the same time, this PhD thesis describes complete results of the successful generation and implementation of CPA system. However it should be considered as important but still an initial stage of the studies. Based on the obtained results, another interesting direction of development is to derive CPA system that does not require any identification of process parameters. It probably requires determining a different set of CPIs (features), which are dimensionless and insensitive to process dynamical parameters and measurement noise. Very initial studies on this subject have been already conducted and results show that it is possible to suggest set of dimensionless CPIs, which do not require normalization of closed loop rejection response and ensure high classification accuracy of control performance. Thus, this new approach will significantly simplify the assessing procedure by removing the stages of necessity of process parameters identification, normalization and simulation of the closed loop response.

The results of this thesis have been already published in:

- [50], which describes ICM method for determining transportation delay time with high accuracy, which is essential for many CPA algorithms, published in IEEE Transactions on Industrial Electronics (IF = 8.236, TOP10 based on Scopus)
- [84], which describes correlation between stochastic and deterministic-based CPA indices, published in conference proceedings of the 22nd International Conference on Process Control.

All of the results presented in this thesis were developed with the participation of a research team from the Silesian University of Technology and the University of Technology Sydney. However, the author

of this dissertation made the most significant contribution to the development of the presented CPA system, including suggesting and developing:

- the general principles of the CPA system,
- the normalization of the closed loop responses,
- the generation of reference responses,
- the use and deep analysis of novel CPIs,
- the generation of closed loop responses to create a training dataset,
- the automatic labelling method,
- the comparison of simulation results with other existing methods,
- the cloud- and PLC-based implementations of the CPA system, together with their experimental verification.

The dissertation author also played an active role in the development of the classifiers and the feature reduction process.

Appendix 1 – Use of the ICM for steady-state detection and time delay estimation

For the accurate assessment of control performance, many studied algorithms require that information concerning the transportation delay time (TDT) is estimated with the highest possible accuracy. The literature suggests many different approaches for TDT estimation [83], [103]–[105], with reference to a cross-correlation method (CCM) [13]. A complete review is provided in [106]. To improve practical usability, the TDT estimation procedure should be implemented directly on the PLC, and thus should not require a high level of computational or memory resources. To fulfil this requirement, use of the novel ICM was suggested for the estimation of TDT [50]. This method is used for the detection of both steady and transient states, and can detect the precise moment of transition into the transient state. Thus, the TDT can be calculated as the time between the introduction of an external excitation into the dynamical system, and the time at which the system transitions into a transient state. The underlying concept of the ICM is based on the consecutive counting of positive and negative increments of the process output

$$\Delta y_i = y_i - y_{i-1} \tag{Ap1.1}$$

for each sample i, within a predefined discretized period of time L. Detection of the steady and transient states is based upon the assumption that during steady state, the dynamical system is only affected by measurement noise. Thus, the number of positive and negative increments Δy should be equal, during a given period of time. During transient-state, the number of positive or negative increments should be unequal.

Consider an L element vector of counters inc_j, with $j = 1, ..., L$, where inc_1 is a counter calculated at the current sampling instant i, and inc_2, …, inc_L are calculated at previous instants (i-1), …, (i-L+1). At each sampling instant i, elements of inc vector are moving backwards, as

$$for\ j = 2,..,L, \quad inc_{L-j+2} = inc_{L-j+1}, \tag{Ap1.2a}$$

and the inc_1 counter is updated according to the sign of the increment of the process output:

$$inc_1 = inc_2 + sign(\Delta y_i). \tag{Ap1.2b}$$

Thus, the vector inc stores the last L counters calculated at sampling instants i to (i-L+1). Linear regression can then be applied by computing the least squares method for the pairs (j, inc_j), where j = 1, …, L. This analysis will detect potential trends in the vector inc. The slope of the regression $a_{inc} \in$ [-1, 1] provides information concerning the state of the process: $|a_{inc}| < \gamma_{ss}$ indicates steady state and $|a_{inc}| > \gamma_{ts}$ indicates transient state. For the case $\gamma_{ts} > |a_{inc}| > \gamma_{ss}$ it is assumed that the previously detected state did not change (hysteresis). For the accurate detection of steady and transient states, and

hence accurate TDT estimation, the following general rules are suggested: $\gamma_{ss} = 1/L$ and $\gamma_{ts} = (1 - 1/L)$. The presented approach can be implemented directly on a PLC, as it has low computational and memory resource requirements. At each iteration (sampling instant), $5(L+1)$ floating point operations are performed, which results in $O(L)$ complexity. Moreover, $4(L+10)$ bytes are required for data storage. Note that the detection of steady and transient states is delayed by L samples, as only increments of the process output are considered. Despite its disadvantages, this property makes the ICM general and insensitive to process gain.

The accuracy with which the ICM can detect steady states was compared with the commonly used R-statistics method (RSM) [107], which calculates the ratio between two variances calculated for the measurement data using two different methods. The RSM requires the calculation of three dynamical filters, and hence is suitable for a PLC-based implementation. However, detection of the precise moment of transition into transient-state is difficult, due to the additional dynamics of the filters. A heat distribution system (described in Appendix 4) was used for the steady-state detection comparison. Several step changes in heater power were introduced to the real laboratory system, and the outlet temperature samples, gathered with constant sampling time 2 s, were used as input data for steady state detection. Fig. Ap1.1 presents the accuracy of steady-state detection when using ICM with $L = 5$, ICM with $L = 10$, and RSM. The results show that both RSM and ICM achieve high accuracy.

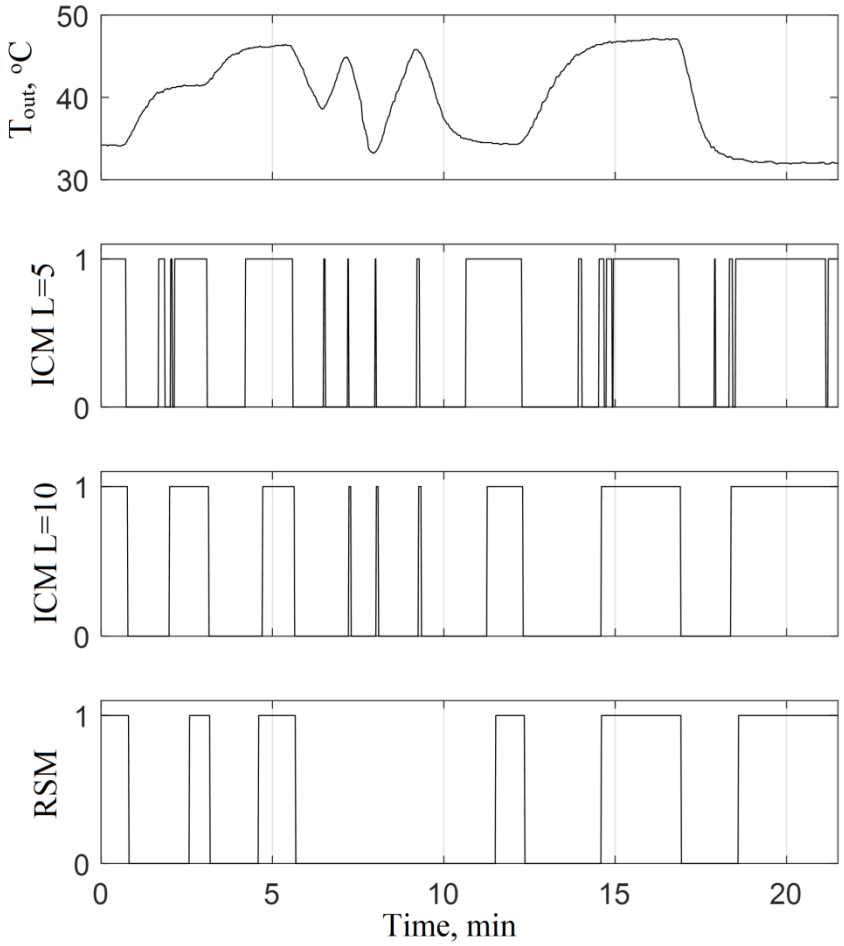

Fig. Ap1.1 A comparison of steady-state detection when using ICM and RSM, showing outlet temperature (top), steady-state detection using ICM with *L*=5 (second from top) and *L*=10 (second from bottom), and steady-state detection using RSM (bottom).

Following this, the output temperature response for the first three step changes in heater power was used to verify the accuracy of the TDT estimation. The results of steady-state detection using the ICM is presented in Fig. Ap1.2.

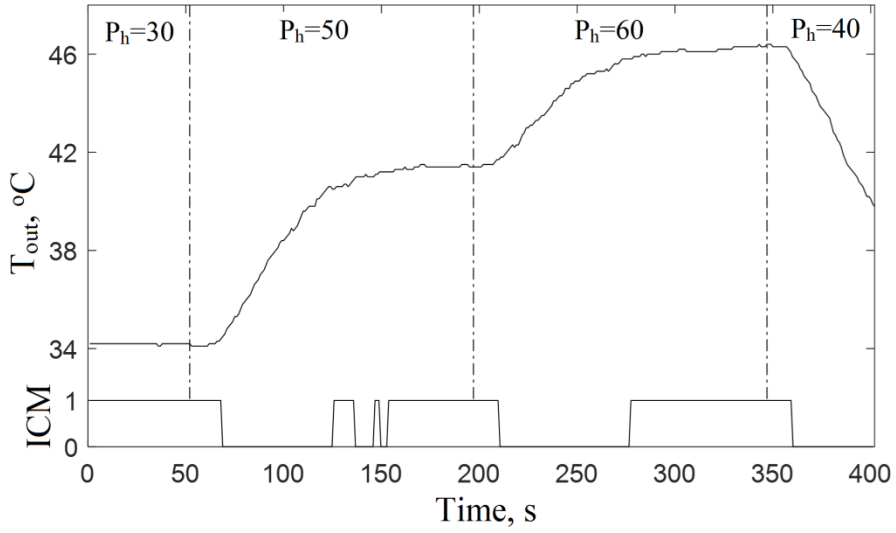

Fig. Ap1.2 Response of the outlet temperature to step changes in heater power, with TDT estimation based on steady-state detection by the ICM.

Fig. Ap1.3 presents the accuracy of TDT estimation based on the ICM, compared to the reference CCM.

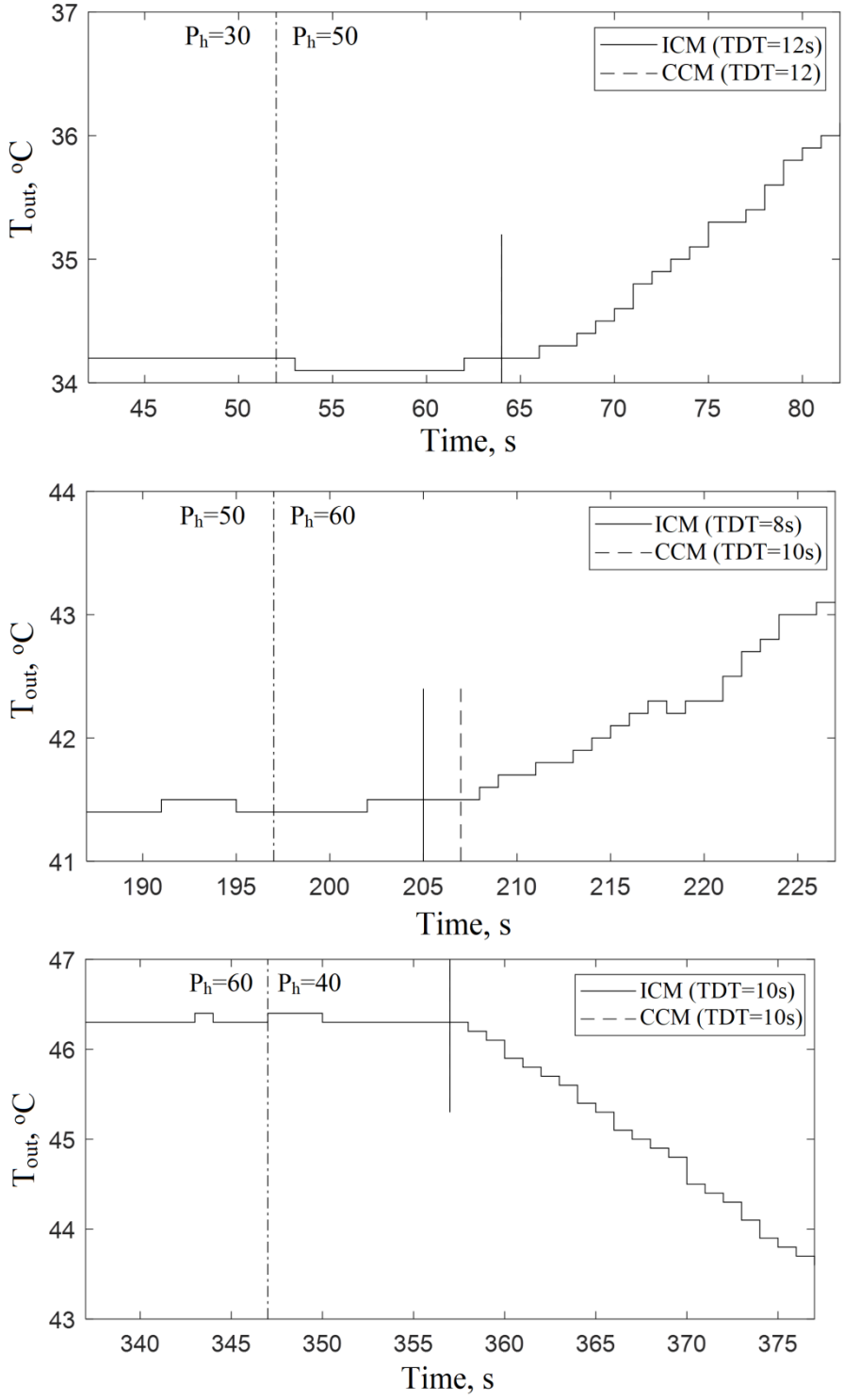

Fig. Ap1.3 The results of TDT estimation using the ICM and CCM for three outlet temperature responses to step changes in heater power of 30% to 50% (top), 50% to 60% (middle) and 60% to 40% (bottom).

The accuracy of TDT estimation using the ICM is relatively high, and comparable with that provided by the reference CCM. A deeper comparison of the ICM and CCM was conducted via simulation for the FOPDT $K_1(s)$ and SOPDT $K_2(s)$ processes (see Table Ap1.1). For five different levels of

measurement noise, denoted here by the signal to noise ratio (SNR), 100 noisy step responses were generated with a constant sampling time of 0.5. For each response, the TDT was estimated by the ICM and CCM. These results were used to calculate the distribution features of the TDT: the median, and the first and third quartiles. Table Ap1.1 presents the results for both K_1 and K_2 processes, and the different SNR values. Within each table cell, the TDT median is presented on the first row, with the first and third quartiles given inside parentheses on the second row. Expected values of the TDT are highlighted in bold. The results again confirm the high accuracy of the ICM, even in the presence of second order dynamics.

	$\mathbf{K_1}(s) = \frac{1}{(1+5s)}e^{-3s}$		$\mathbf{K_2}(s) = \frac{1}{(1+5s)^2}e^{-3s}$	
SNR (dB)	**TDT by ICM**	**TDT by CCM**	**TDT by ICM**	**TDT by CCM**
79.52	2.5 (2.25, **3.0**)	**3.0** **(3.0, 3.0)**	2.5 (2.5, **3.0**)	3.5 (3.5, 3.5)
65.53	2.5 (2.5, **3.0**)	**3.0** **(3.0, 3.0)**	2.5 (2.25, **3.0**)	3.5 (3.5, 3.5)
59.47	2.5 (2.0, **3.0**)	**3.0** **(3.0, 3.0)**	2.5 (2.25, **3.0**)	3.5 (3.5, 3.5)
45.58	2.5 (2.5, **3.0**)	**3.0** **(3.0, 3.0)**	**3.0** (2.5, 3.5)	3.5 (3.5, 4.0)
39.50	**3.0** (2.5, **3.0**)	**3.0** **(3.0, 3.0)**	3.5 (**3.0**, 4.0)	3.5 (3.5, 4.0)

Table Ap1.1 The accuracy of TDT estimation by the ICM and CCM for two processes with various noise levels. In each cell, the first row presents the median TDT and the second row the first and third quartiles (in parentheses). Expected values are highlighted in bold.

Further details concerning the ICM, together with an example of its application for the autotuning of the Reduced Order ADRC algorithm, can be found in [50].

Appendix 2 – Graphical representation of the selected CPIs

This appendix presents graphical representations of all of the selected CPIs in the form of an example process response.

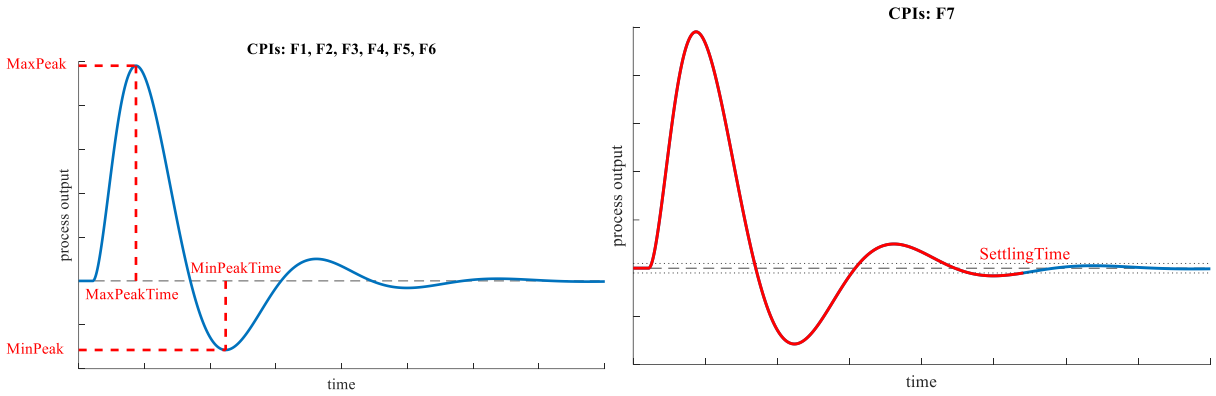

Fig. Ap1.1 Graphical representations of the F1, F2, F3, F4, F5, F6 (left), and F7 (right) CPIs.

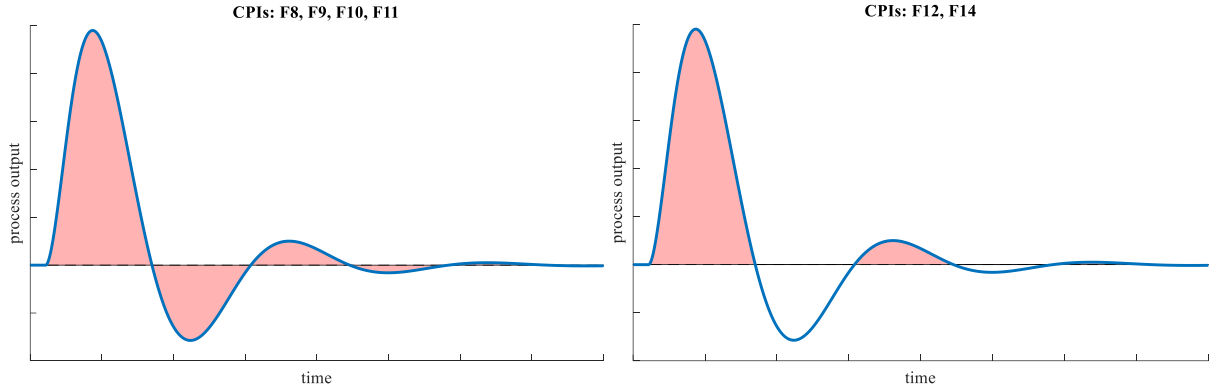

Fig. Ap1.2 Graphical representations of the F8, F9, F10, F11 (left), F12, and F14 (right) CPIs.

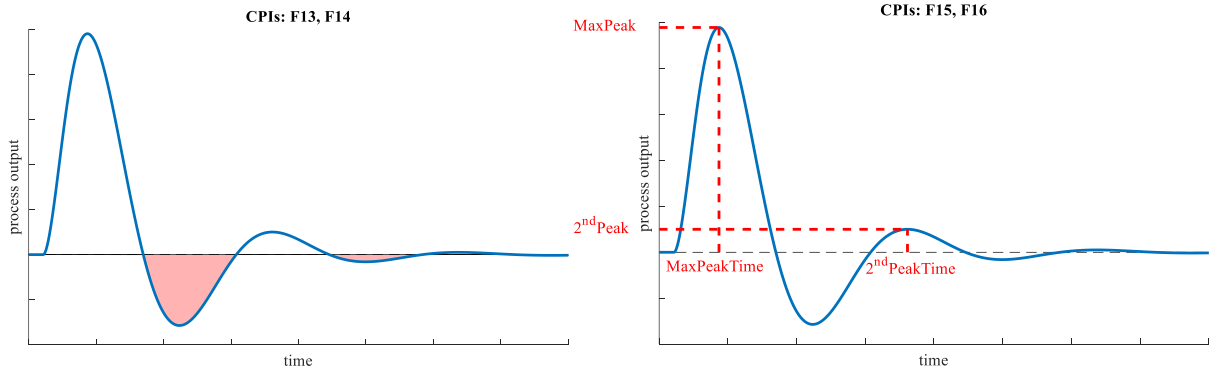

Fig. Ap1.3 Graphical representations of the F13, F14 (left), F15, and F16 (right) CPIs.

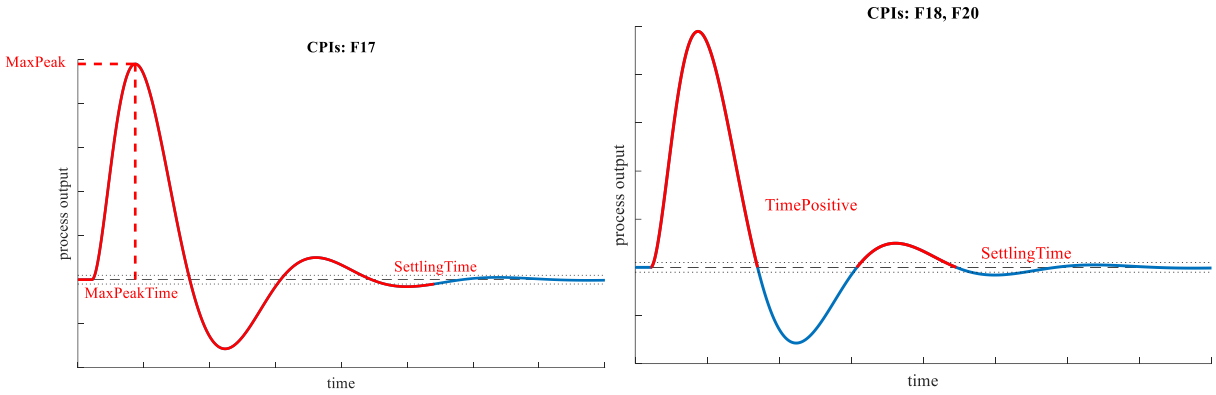

Fig. Ap1.4 Graphical representations of the F17 (left), F18, and F20 (right) CPIs.

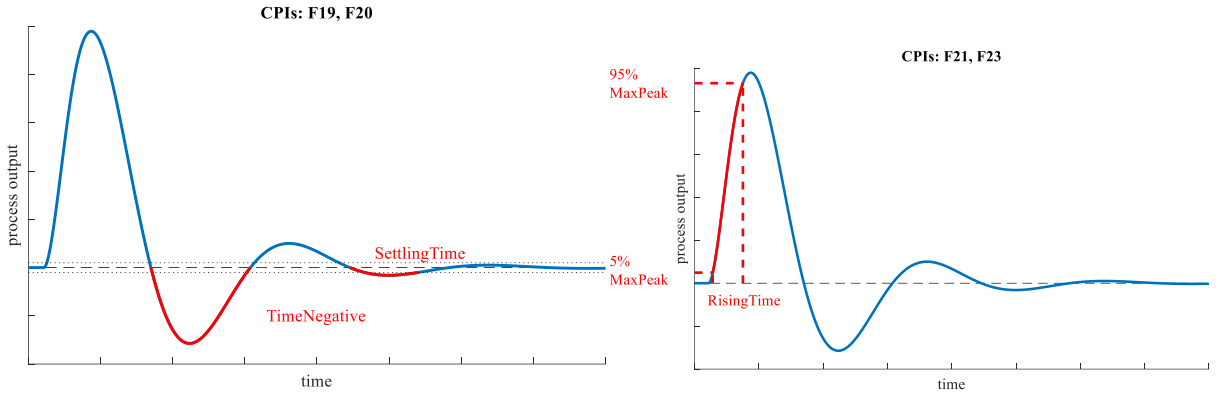

Fig. Ap1.5 Graphical representations of the F19, F20 (left), F21, and F23 (right) CPIs.

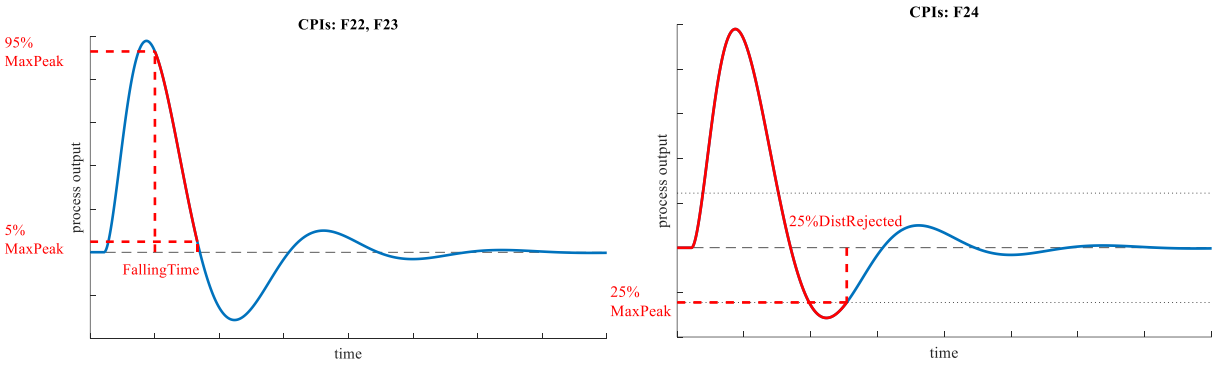

Fig. Ap1.6 Graphical representations of the F22, F23 (left), and F24 (right) CPIs.

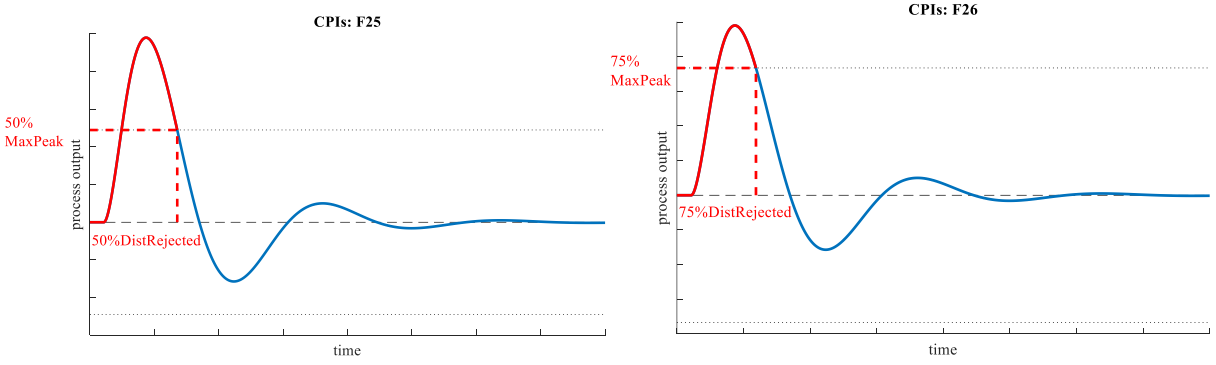

Fig. Ap1.7 Graphical representations of the F25 (left), and F26 (right) CPIs.

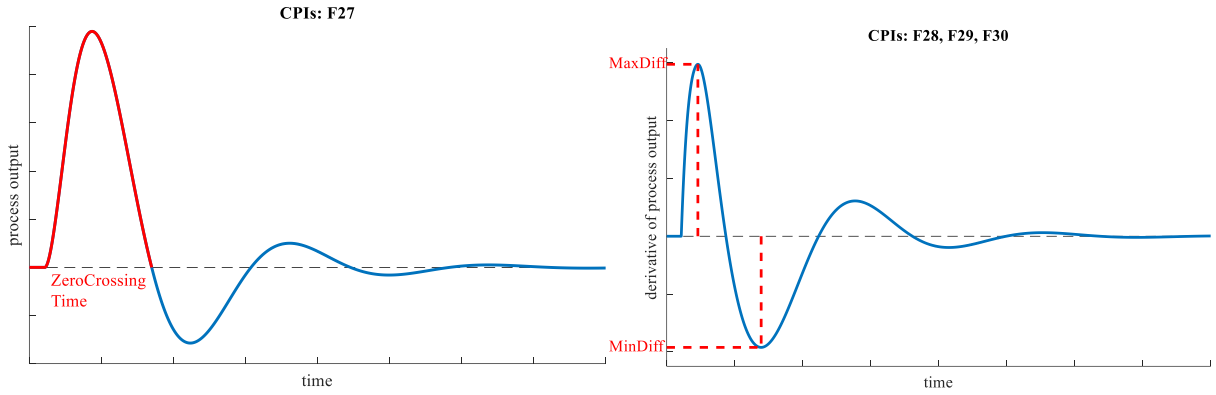

Fig. Ap1.8 Graphical representations of the F27 (left), F28, F29, and F30 (right) CPIs.

Appendix 3 – Validation of the CPA system for higher order dynamics

This appendix presents the results of applying the CPA system to the higher order systems *P2–P7*. The results show the accuracy of modelling higher order systems with FOPDT and SOPDT approximations, a comparison between reference responses generated for a higher order system and its SOPDT approximation, the classification results of various process responses based on a SOPDT approximation, and the process responses generated for a higher order system and labelled according to classification results based on a SOPDT approximation.

- Process *P2*:

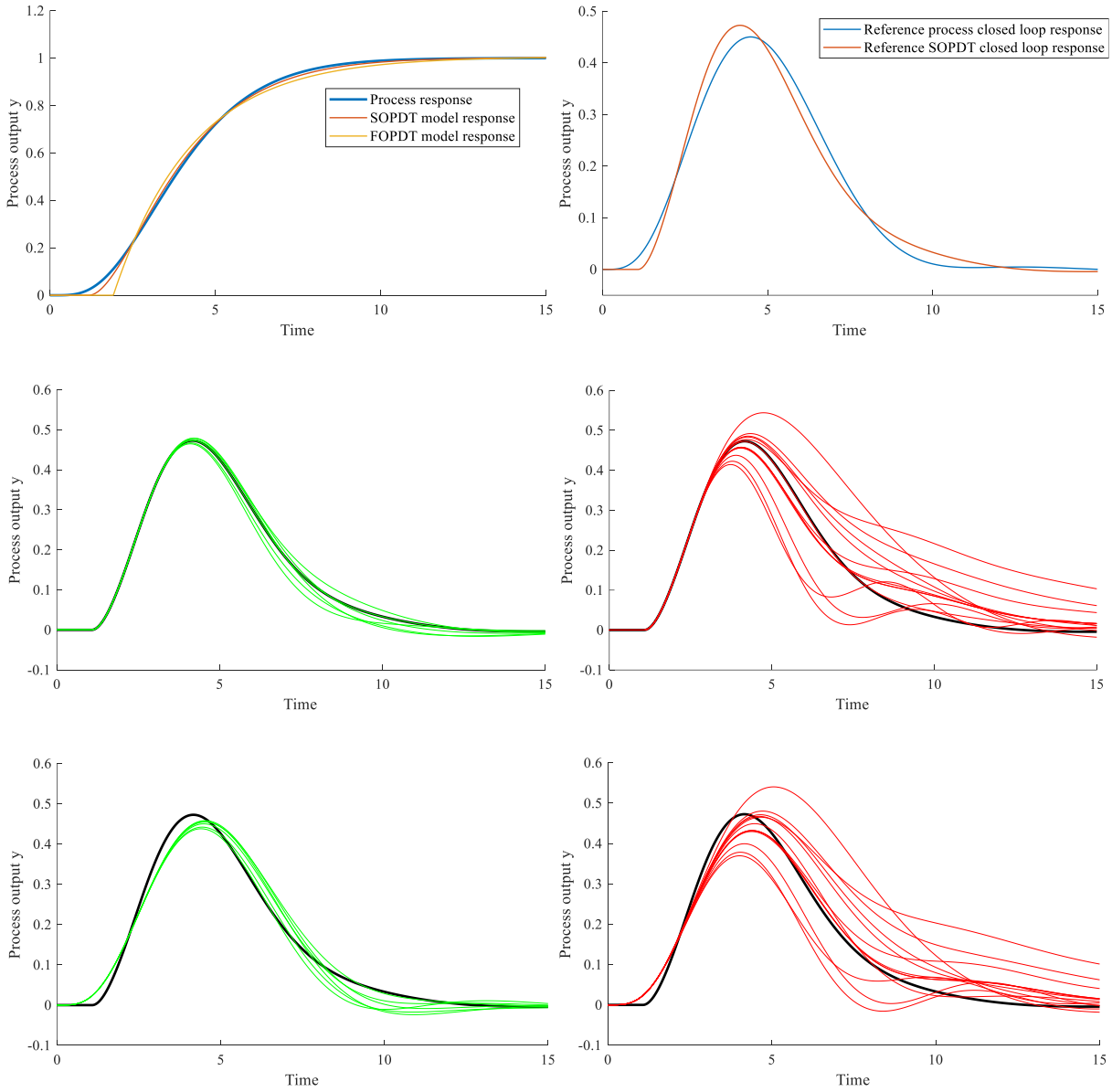

Fig. Ap2.1 The modelling accuracy of FOPDT and SOPT approximations (top left), the reference closed loop responses for reference PID tunings computed based on a SOPDT approximation (top right), the classification results of closed loop responses based on a SOPDT approximation, simulated based on a SOPDT approximation (middle left and middle right), and the classification results of closed loop responses based on an SOPDT approximation, simulated based on a higher order process (lower left and lower right). All results for process *P*2. The green and red responses denote OK and NOK control performance, respectively.

- Process *P3*:

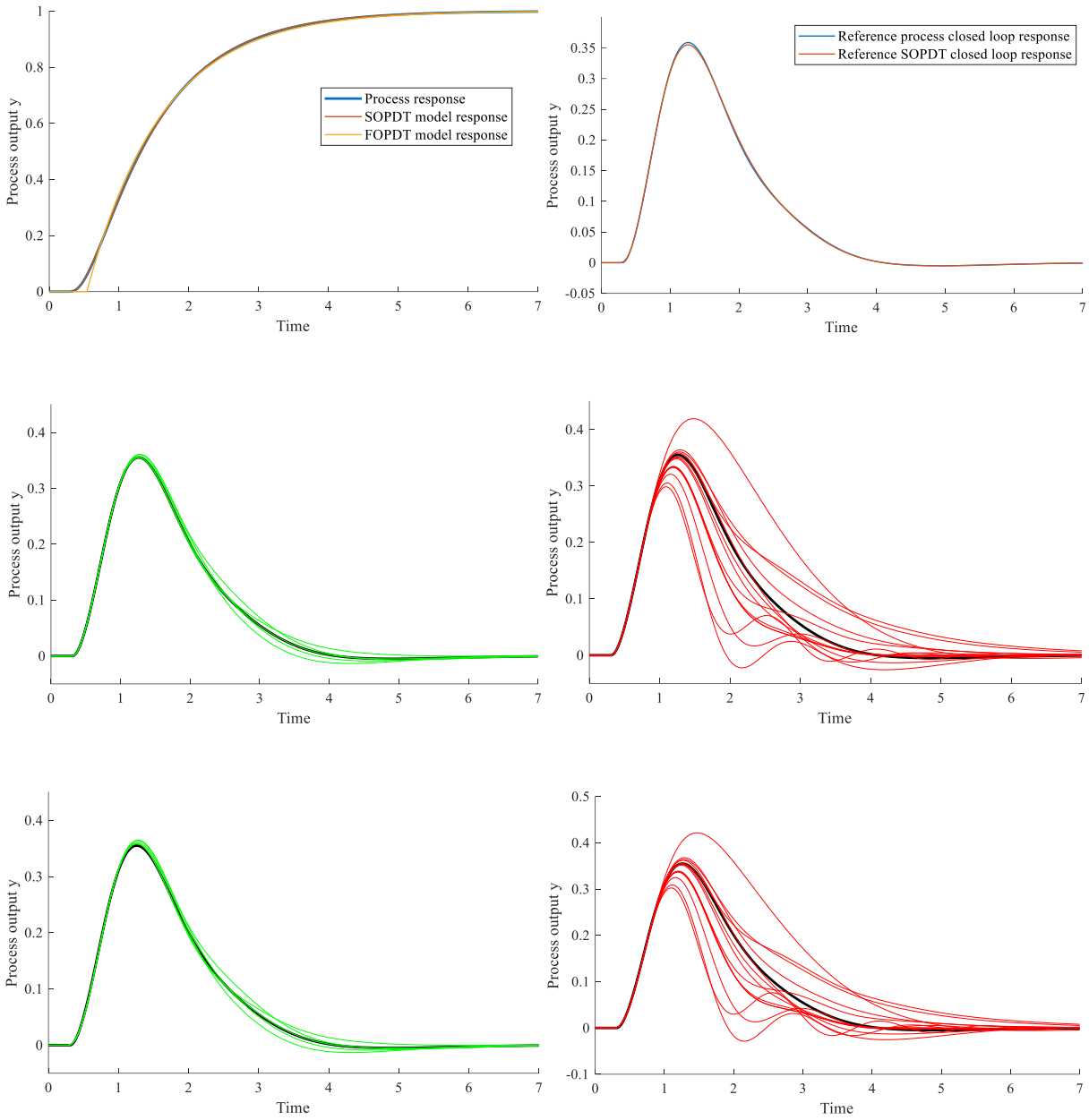

Fig. Ap2.2 The modelling accuracy of FOPDT and SOPT approximations (top left), the reference closed loop responses for reference PID tunings computed based on a SOPDT approximation (top right), the classification results of closed loop responses based on a SOPDT approximation, simulated based on a SOPDT approximation (middle left and middle right), and the classification results of closed loop responses based on an SOPDT approximation, simulated based on a higher order process (lower left and lower right). All results for process *P*3. The green and red responses denote OK and NOK control performance, respectively.

- Process *P4*:

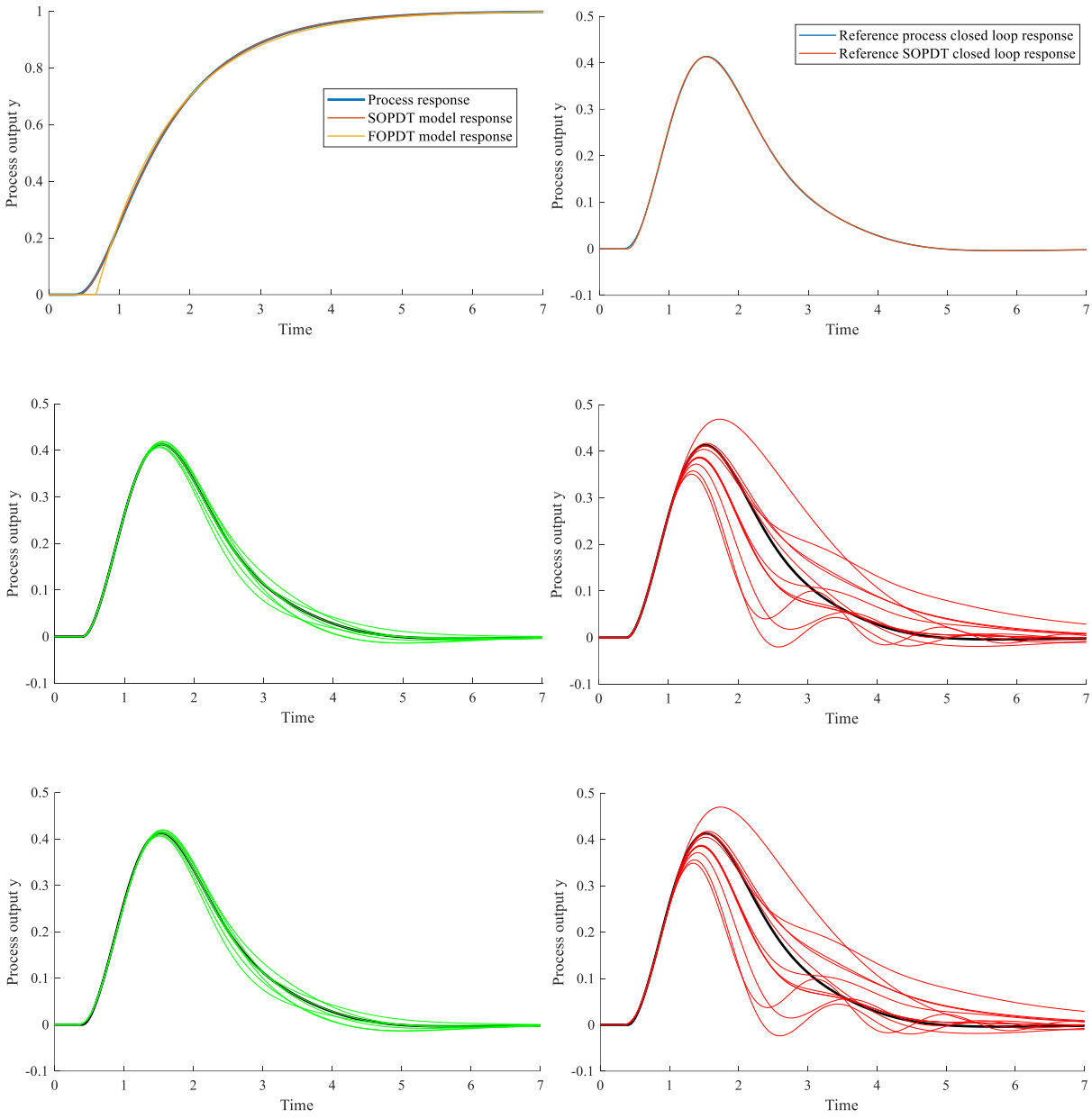

Fig. Ap2.3 The modelling accuracy of FOPDT and SOPT approximations (top left), the reference closed loop responses for reference PID tunings computed based on a SOPDT approximation (top right), the classification results of closed loop responses based on a SOPDT approximation, simulated based on a SOPDT approximation (middle left and middle right), and the classification results of closed loop responses based on an SOPDT approximation, simulated based on a higher order process (lower left and lower right). All results for process *P*4. The green and red responses denote OK and NOK control performance, respectively.

- Process *P5*:

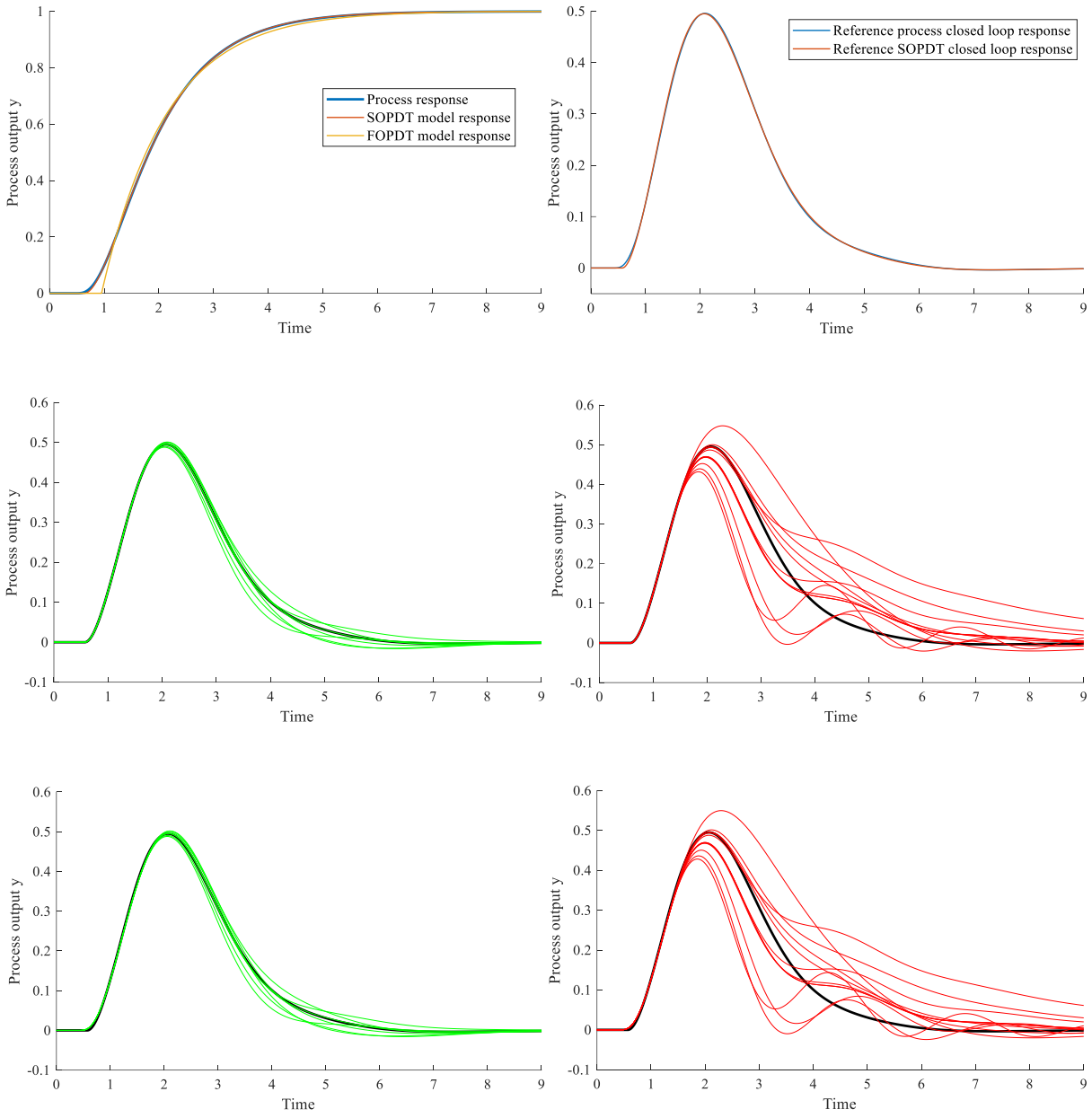

Fig. Ap2.4 The modelling accuracy of FOPDT and SOPT approximations (top left), the reference closed loop responses for reference PID tunings computed based on a SOPDT approximation (top right), the classification results of closed loop responses based on a SOPDT approximation, simulated based on a SOPDT approximation (middle left and middle right), and the classification results of closed loop responses based on an SOPDT approximation, simulated based on a higher order process (lower left and lower right). All results for process *P*5. The green and red responses denote OK and NOK control performance, respectively.

- Process *P6*:

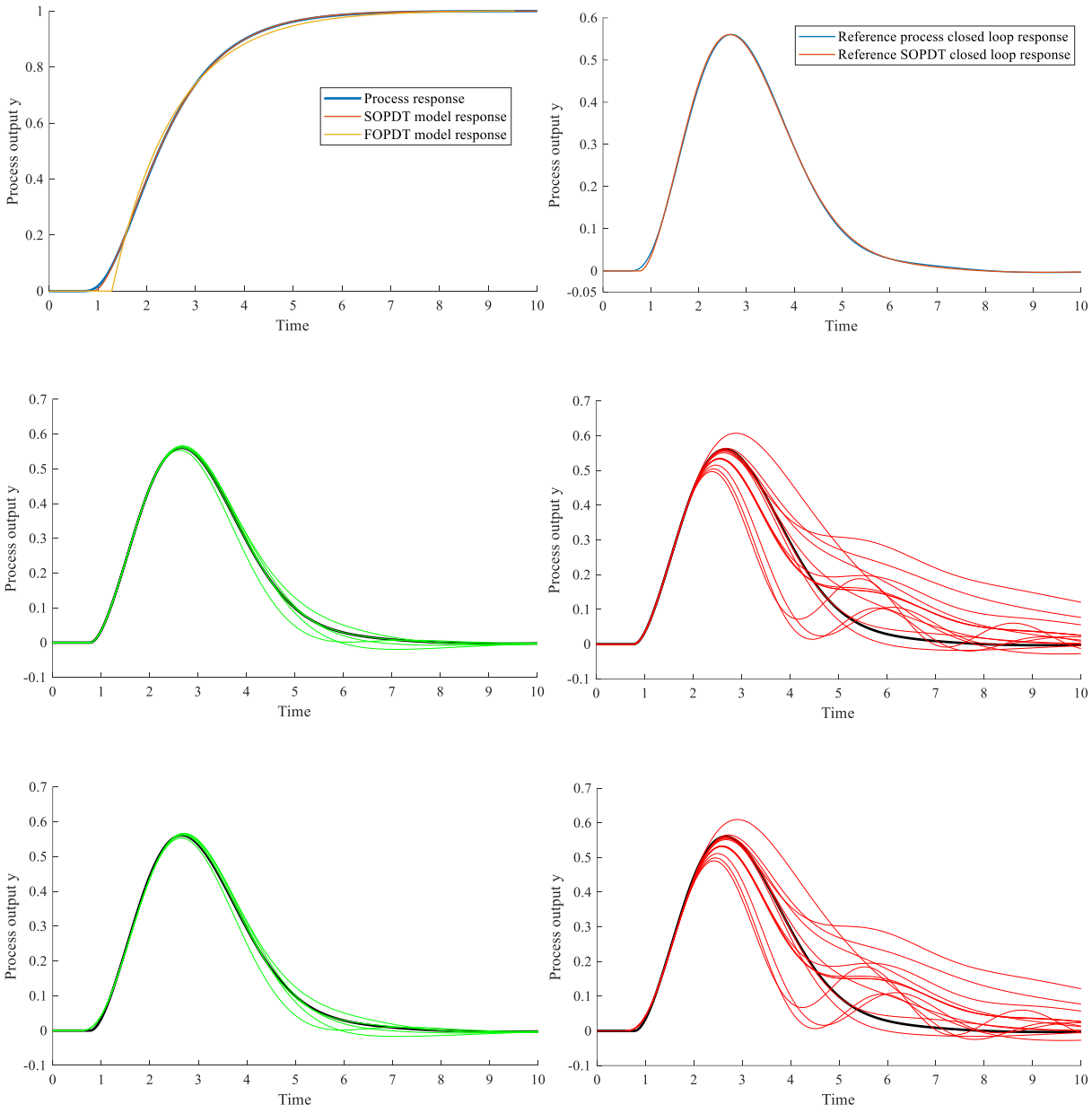

Fig. Ap2.5 The modelling accuracy of FOPDT and SOPT approximations (top left), the reference closed loop responses for reference PID tunings computed based on a SOPDT approximation (top right), the classification results of closed loop responses based on a SOPDT approximation, simulated based on a SOPDT approximation (middle left and middle right), and the classification results of closed loop responses based on an SOPDT approximation, simulated based on a higher order process (lower left and lower right). All results for process *P*6. The green and red responses denote OK and NOK control performance, respectively.

- Process *P7*:

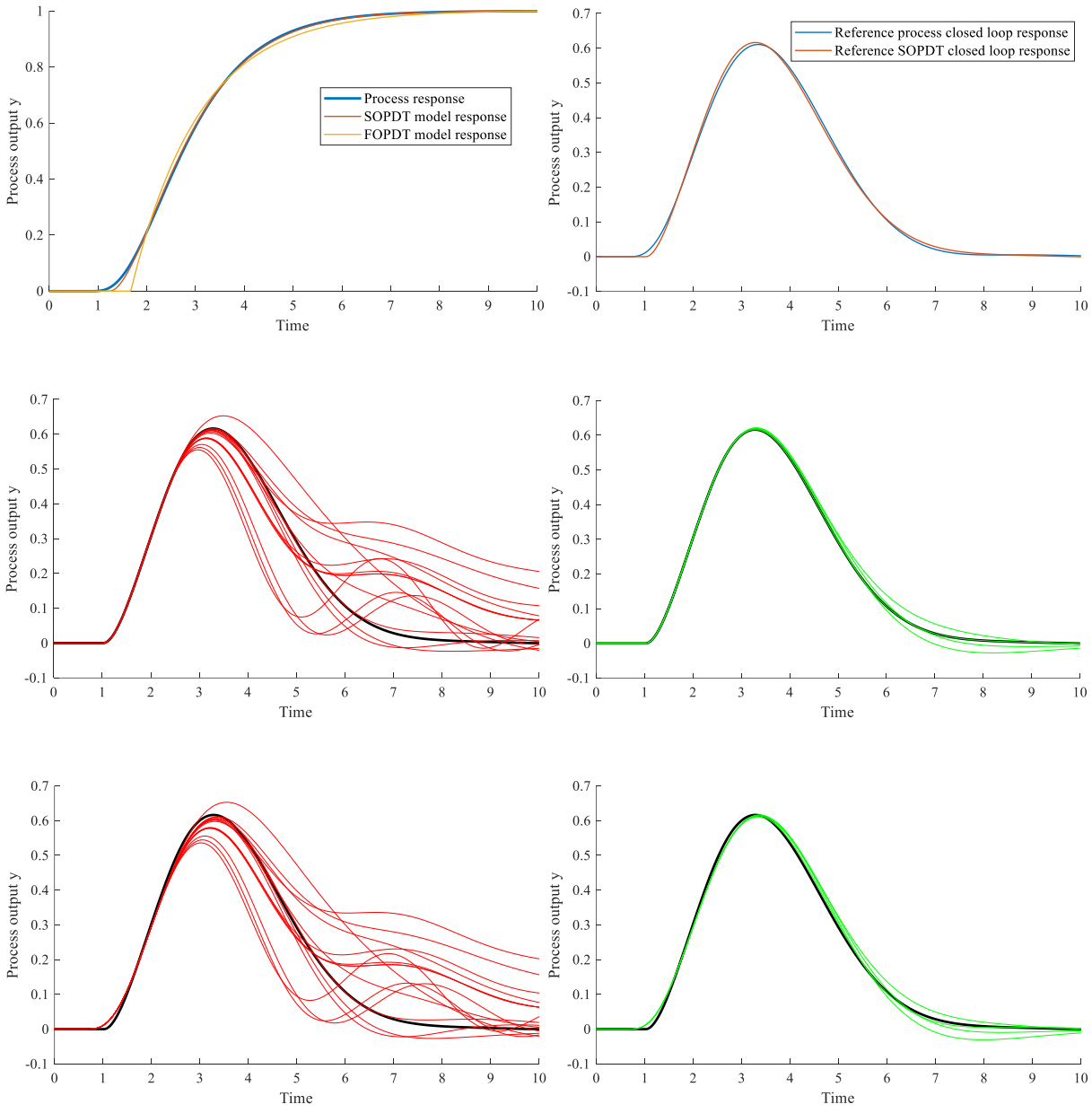

Fig. Ap2.6 The modelling accuracy of FOPDT and SOPT approximations (top left), the reference closed loop responses for reference PID tunings computed based on a SOPDT approximation (top right), the classification results of closed loop responses based on a SOPDT approximation, simulated based on a SOPDT approximation (middle left and middle right), and the classification results of closed loop responses based on an SOPDT approximation, simulated based on a higher order process (lower left and lower right). All results for process *P7*. The green and red responses denote OK and NOK control performance, respectively.

Appendix 4 – Description of the heat distribution and exchange system

The central component of the heat distribution and exchange system is the electric flow heater. A block diagram of the laboratory system is presented in Fig. Ap4.1.

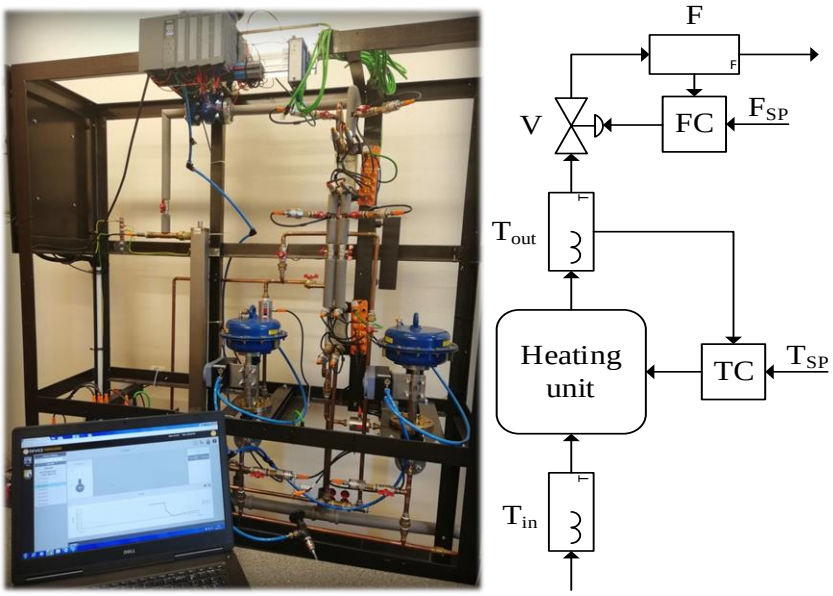

Fig. Ap4.1 A picture of the heat distribution and exchange system (left) and the corresponding block diagram (right).

Water flows through the electric flow heater of constant volume V = 1.2 L and of maximum power P_{max} = 12 kW. The mean power of the heater P_h is adjusted from 0% to 100% of its maximal power using a pulse width modulation (PWM) signal. The inlet T_{in} and outlet T_{out} temperatures of the water are measured with PT100 sensors. The flow rate F of the water is controlled with a local PI controller. Both the temperature and flow sensors are of IO-Link type. The primary purpose of the control system is to keep the output temperature T_{out} of the heater at the desired temperature T_{SP} by manipulating the mean power P_h. The control algorithm is implemented in a Siemens S7-1516-3 PN/DP PLC unit, using 5 MB of data memory and 1 MB of code memory. All of these signals are directly connected to distributed input and output modules. The complete topology of the Profinet network is presented in Fig. Ap4.2.

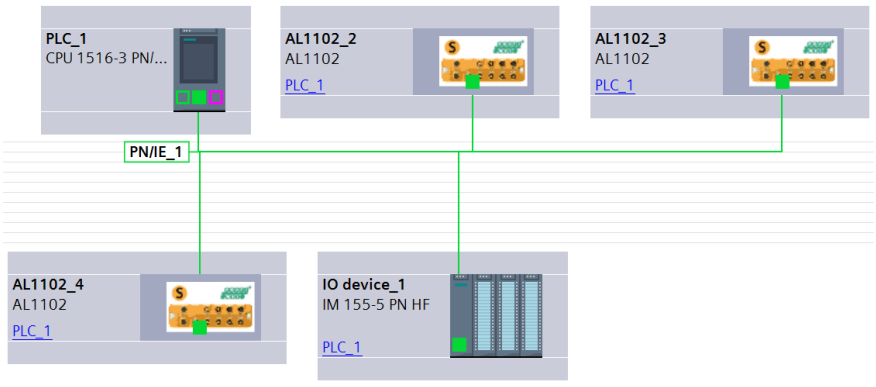

Fig. Ap4.2 The topology of the Profinet network.

As part of the research into virtual commissioning methodologies, a model of the studied section of the heat distribution and exchange system was identified. The structure of the suggested model is presented in Fig. Ap4.3.

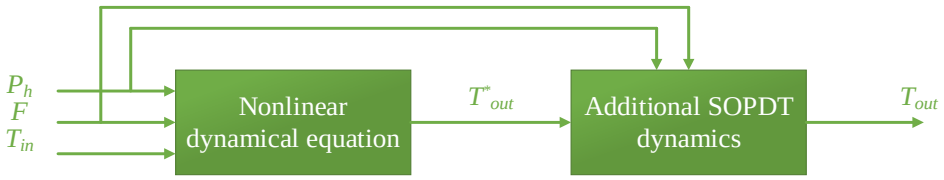

Fig. Ap4.3 The complete structure of the suggested model of the studied section of the heat distribution and exchange system.

The model primarily consists of the nonlinear dynamical equation

$$\frac{dT^*_{out}}{dt} = \frac{F}{60V}(T_{in} - T^*_{out}) + \frac{P_h P_{max}}{100 c_s \rho_s V}, \tag{Ap4.1}$$

where c_s = 4200 J/(kg °C) is a water specific heat capacity, ρ_s = 1 kg/L is water density, and T_{out}^* in °C is the unmeasurable output temperature. However, to increase the accuracy with which transients are modelled, additional SOPDT dynamics were included:

$$K(s) = \frac{T_{out}(s)}{T^*_{out}} = \frac{k(P_h)}{(1+s\tau_1(F))(1+s\tau_2(F))} e^{-s\tau_0(F)} \tag{Ap4.2}$$

where gain k is dependent upon the mean power P_h, and the time constants τ_1 and τ_2 and the transportation time τ_0 are dependent upon the flow rate F. These functions were identified using real process data, and take the following form:

$$\begin{gathered} k(P_h) = -0.0002347 P_h + 1.012, \\ \tau_1(F) = \tau_2(F) = 19.08 F^{-0.4293} - 4.042, \\ \tau_0(F) = 11.93 F^{-0.7838} + 2.365. \end{gathered} \tag{Ap4.3}$$

The modelling accuracy is presented in Fig. Ap4.4, which presents the responses of the real system and the identified model for several step changes in mean power P_h and flow rate F.

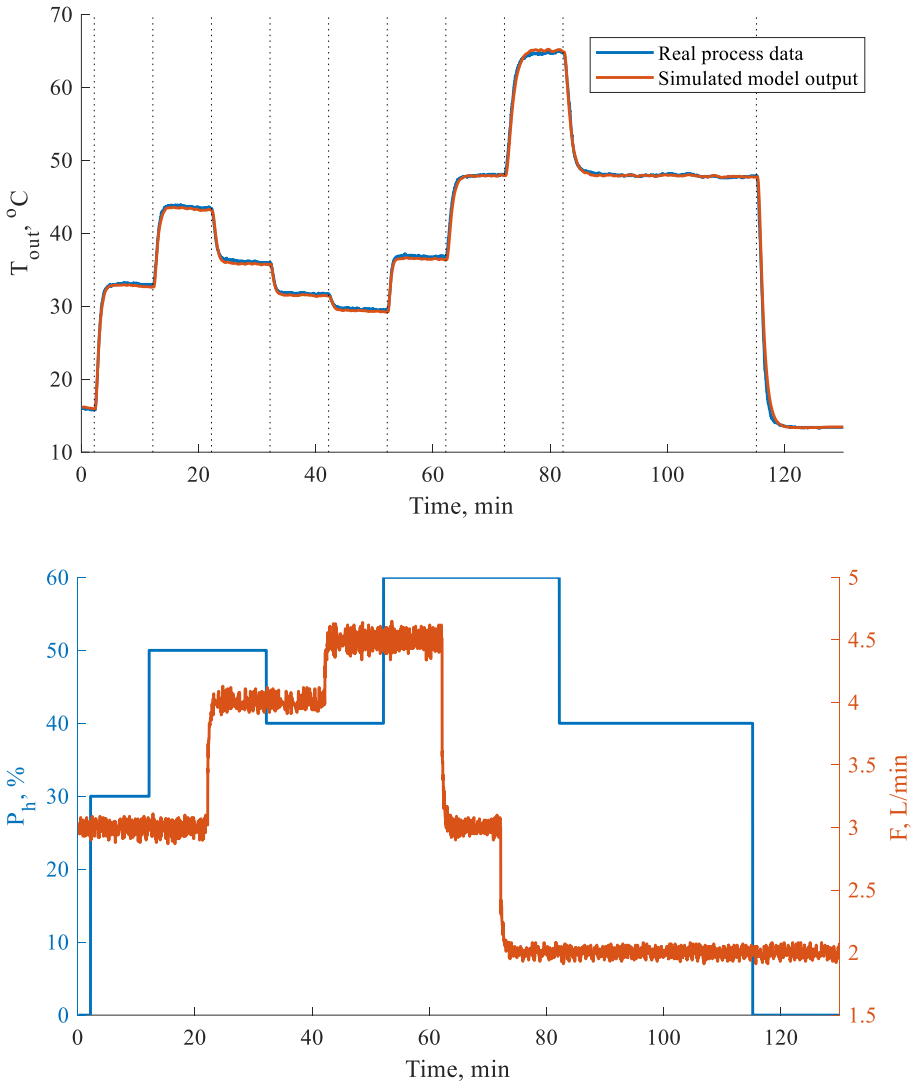

Fig. Ap4.4 A comparison of real process data and the output of the identified model.

Appendix 5 - PID tuning rules for CPA

Based on the FOPDT approximation, countless methods have been developed in recent decades to determine PID tuning parameters. To validate the proposed CPA system, various tuning rules were selected. All of them are summarized in Table Ap5.1. The majority of them were sourced from scientific papers, but industrial examples are also included.

Tuning rule	k_r	T_i	T_d
Ziegler-Nichols [108]	$\frac{1.6\tau}{k\tau_0}$	$2\tau_0$	$0.5\tau_0$
SIMC PI [28]	$\frac{\tau}{2k\tau_0}$	$\min(\tau, 8\tau_0)$	0
Chien-Hrones-Reswick Regulator mode; 0% overshoot [25]	$\frac{0.95\tau}{k\tau_0}$	$2.38\tau_0$	$0.42\tau_0$
Chien-Hrones-Reswick Regulator mode; 20% overshoot [25]	$\frac{1.2\tau}{k\tau_0}$	$2\tau_0$	$0.42\tau_0$
Chien-Hrones-Reswick Servo mode; 0% overshoot [25]	$\frac{0.6\tau}{k\tau_0}$	τ	$0.5\tau_0$
Chien-Hrones-Reswick Servo mode; 20% overshoot [25]	$\frac{0.95\tau}{k\tau_0}$	1.36τ	$0.47\tau_0$
AMIGO PI [109]	$\frac{0.15}{k} + \left(0.35 - \frac{\tau_0\tau}{(\tau_0+\tau)^2}\right)\frac{\tau}{k\tau_0}$	$0.35\tau_0 + \frac{6.7\tau_0\tau^2}{\tau^2 + 2\tau_0\tau + 10\tau_0}$	0
AMIGO PID [109]	$\frac{1}{k}\left(0.2 + 0.45\frac{\tau}{\tau_0}\right)$	$\frac{0.4\tau_0 + 0.8\tau}{\tau_0 + 0.1\tau}\tau_0$	$\frac{0.5\tau_0\tau}{0.3\tau_0 + \tau}$
Cohen-Coon [110]	$\frac{1}{k}\left(1.35\frac{\tau}{\tau_0} + 0.25\right)$	$\tau\left(\frac{2.5(\tau_0/\tau) + 0.46(\tau_0/\tau)^2}{1 + 0.61(\tau_0/\tau)}\right)$	$\frac{0.37\tau_0}{1 + 0.19(\tau_0/\tau)}$

Parr [111]	$\frac{1.25\tau}{k\tau_0}$	$2.5\tau_0$	$0.4\tau_0$
Sain-Özgen	$\frac{1}{k}\left(0.6939\frac{\tau}{\tau_0}+0.1814\right)$	$\frac{0.8647\tau+0.226\tau_0}{\frac{\tau}{\tau_0}+0.8647}$	$\frac{0.0565\tau}{0.8647\frac{\tau}{\tau_0}+0.226}$
Connell [112]	$\frac{1.6\tau}{k\tau_0}$	$1.6667\tau_0$	$0.4\tau_0$
Moros (Attributed to Oppelt)	$\frac{1.2\tau}{k\tau_0}$	$T_i=2\tau_0$	$0.42\tau_0$
Moros (Attributed to Rosenberg)	$\frac{1.2\tau}{k\tau_0}$	$2\tau_0$	$0.44\tau_0$
Liptak	$\frac{0.85\tau}{k\tau_0}$	$1.6\tau_0$	$0.6\tau_0$
Chidambaram [113]	$\frac{1}{k}\left(1.8\frac{\tau}{\tau_0}+0.45\right)$	$2.4\tau_0$	$0.4\tau_0$
ControlSoft Inc. (Slow loop)	$\frac{2}{k}$	$\tau+\tau_0$	$\max\left(\frac{\tau_0}{3},\frac{\tau}{6}\right)$
ControlSoft Inc. (Fast loop)	$\frac{2}{k}$	$\tau+\tau_0$	$\min\left(\frac{\tau_0}{3},\frac{\tau}{6}\right)$
PMA [114]	$\frac{0.59\tau}{k\tau_0}$	$2\tau_0$	$2\tau_0$
Minimum IAE Murrill [115]	$\frac{1.435}{k}\left(\frac{\tau}{\tau_0}\right)^{0.921}$	$\frac{\tau}{0.878}\left(\frac{\tau_0}{\tau}\right)^{0.749}$	$0.482\tau\left(\frac{\tau_0}{\tau}\right)^{1.137}$
Minimum IAE Arrieta Orozco [116]	$\frac{1}{k}(0.2068+1.1597\left(\frac{\tau}{\tau_0}\right)^{1.0158})$	$\tau(-0.2228+1.3009\left(\frac{\tau_0}{\tau}\right)^{0.5022})$	$0.3953\tau\left(\frac{\tau_0}{\tau}\right)^{0.8469}$
Minimum ISE Murrill [115]	$\frac{1.473}{k}\left(\frac{\tau}{\tau_0}\right)^{0.97}$	$\frac{\tau}{1.101}\left(\frac{\tau_0}{\tau}\right)^{0.771}$	$0.56\tau\left(\frac{\tau_0}{\tau}\right)^{1.006}$
Minimum ISE Zhuang-Atherton [117]	$\frac{1.473}{k}\left(\frac{\tau}{\tau_0}\right)^{0.97}$	$\frac{\tau}{1.115}\left(\frac{\tau_0}{\tau}\right)^{0.753}$	$0.55\tau\left(\frac{\tau_0}{\tau}\right)^{0.948}$
Minimum IAE Rovira [118]	$\frac{1.086}{k}\left(\frac{\tau}{\tau_0}\right)^{0.869}$	$\frac{\tau}{0.74-0.13\frac{\tau_0}{\tau}}$	$0.348\tau\left(\frac{\tau_0}{\tau}\right)^{0.914}$
Minimum IAE Sadeghi-Tych [119]	$\frac{1}{k}(0.26266+0.82714\frac{\tau}{\tau_0})$	$\tau(0.28743\frac{\tau_0}{\tau}+1.35955)$	$\frac{1.45933\tau_0}{\frac{\tau_0}{\tau}+3.27873}$

Minimum IAE Arrieta Orozco [116]	$\frac{1}{k}(0.3295 + 0.7182\left(\frac{\tau}{\tau_0}\right)^{0.9971})$	$\tau(0.9781 + 0.3723\left(\frac{\tau_0}{\tau}\right)^{0.8456})$	$0.3416\tau(\frac{\tau_0}{\tau})^{0.9414}$
Minimum IAE Wang [120]	$\frac{(0.7645+\frac{0.6032}{\tau_0/\tau})(\tau+0.5\tau_0)}{k(\tau+\tau_0)}$	$\tau+0.5\tau_0$	$\frac{0.5\tau\tau_0}{\tau+0.5\tau_0}$
Minimum ISE Wang [120]	$\frac{(0.9155+\frac{0.7524}{\tau_0/\tau})(\tau+0.5\tau_0)}{k(\tau+\tau_0)}$	$\tau+0.5\tau_0$	$\frac{0.5\tau\tau_0}{\tau+0.5\tau_0}$
Minimum ISE Zhuang-Atherton [117]	$\frac{1.048}{k}(\frac{\tau}{\tau_0})^{0.897}$	$\frac{\tau}{1.195-0.368\frac{\tau_0}{\tau}}$	$0.489\tau(\frac{\tau_0}{\tau})^{0.888}$
Minimum ISE Sadeghi-Tych [119]	$\frac{1}{k}(0.40455+0.96441\frac{\tau}{\tau_0})$	$\tau(0.4377\frac{\tau_0}{\tau}+1.39588)$	$\frac{1.93576\tau_0}{\frac{\tau_0}{\tau}+3.83528}$
Van der Grinten for step disturbance	$\frac{1}{k}(0.5+\frac{\tau}{\tau_0})$	$\tau+0.5\tau_0$	$\frac{\tau\tau_0}{2\tau+\tau_0}$
Mann [121]	$\frac{\tau(0.77+0.245(\tau_0/\tau)^{0.854})}{k\tau_0}$	$\tau(1.262+0.147\left(\frac{\tau_0}{\tau}\right)^{0.854})$	$\tau_0\frac{0.262+0.147(\tau_0/\tau)^{0.854}}{0.77+0.245(\tau_0/\tau)^{0.854}}$
Sree [122]	$\frac{1.377}{k}\left(\frac{\tau}{\tau_0}\right)^{0.8422}$	$1.085\tau\left(\frac{\tau_0}{\tau}\right)^{0.4777}$	$\tau(0.3899\frac{\tau_0}{\tau}0.0195)$
Brambilla [123]	$\frac{\tau+0.5\tau_0}{0.85k\tau_0}$	$\tau+0.5\tau_0$	$\frac{\tau\tau_0}{2\tau+\tau_0}$
Gerry [124]	$\frac{\tau}{2.5k\tau_0}$	τ	$0.5\tau_0$
Gong [125]	$\frac{0.7556\tau}{k\tau_0}(1+0.3866\frac{\tau_0}{\tau})$	$\tau+0.3866\tau_0$	$\frac{0.3866\tau_0}{1+0.3866(\tau_0/\tau)}$

Table Ap5.1 Selected PID tuning rules.

References

[1] H. Kagermann, “Change Through Digitization—Value Creation in the Age of Industry 4.0 BT - Management of Permanent Change,” H. Albach, H. Meffert, A. Pinkwart, and R. Reichwald, Eds. Wiesbaden: Springer Fachmedien Wiesbaden, 2015, pp. 23–45.

[2] T. Zheng, M. Ardolino, A. Bacchetti, and M. Perona, “The applications of Industry 4.0 technologies in manufacturing context: a systematic literature review,” *Int. J. Prod. Res.*, vol. 59, no. 6, pp. 1922–1954, 2021, doi: 10.1080/00207543.2020.1824085.

[3] H. Kagermann, W. Wahlster, and J. Helbig, “Securing the future of German manufacturing industry: Recommendations for implementing the strategic initiative INDUSTRIE 4.0,” 2013.

[4] J. Malinauskaite, H. Jouhara, B. Egilegor, F. Al-Mansour, L. Ahmad, and M. Pusnik, “Energy efficiency in the industrial sector in the EU, Slovenia, and Spain,” *Energy*, vol. 208, p. 118398, 2020, doi: 10.1016/j.energy.2020.118398.

[5] I. Grossmann, “Enterprise-wide optimization: A new frontier in process systems engineering,” *AIChE J.*, vol. 51, no. 7, pp. 1846–1857, 2005, doi: 10.1002/aic.10617.

[6] J. M. Simkoff, F. Lejarza, M. T. Kelley, C. Tsay, and M. Baldea, “Process Control and Energy Efficiency,” *Annu. Rev. Chem. Biomol. Eng.*, vol. 11, no. 1, pp. 423–445, Jun. 2020, doi: 10.1146/annurev-chembioeng-092319-083227.

[7] F. Pask, P. Lake, A. Yang, H. Tokos, and J. Sadhukhan, “Industrial oven improvement for energy reduction and enhanced process performance,” *Clean Technol. Environ. Policy*, vol. 19, no. 1, pp. 215–224, 2017, doi: 10.1007/s10098-016-1206-z.

[8] T. T. Nguyen and M. T. Le, “Optimization of the internal roller burnishing process for energy reduction and surface properties,” *Stroj. Vestnik/Journal Mech. Eng.*, vol. 67, no. 4, pp. 167–179, 2021, doi: 10.5545/sv-jme.2021.7106.

[9] M. Trafczynski, M. Markowski, and K. Urbaniec, “Energy saving potential of a simple control strategy for heat exchanger network operation under fouling conditions,” *Renew. Sustain. Energy Rev.*, vol. 111, no. December 2018, pp. 355–364, 2019, doi: 10.1016/j.rser.2019.05.046.

[10] J. J. Siirola and T. F. Edgar, “Process energy systems: Control, economic, and sustainability objectives,” *Comput. Chem. Eng.*, vol. 47, pp. 134–144, 2012, doi: 10.1016/j.compchemeng.2012.06.019.

[11] K. Stebel, M. Fratczak, P. Grelewicz, J. Czeczot, and T. Klopot, “Adaptive predictive

controller for energy-efficient batch heating process," *Appl. Therm. Eng.*, vol. 192, no. April, p. 116954, 2021, doi: 10.1016/j.applthermaleng.2021.116954.

[12] T. Klopot, P. Skupin, P. Grelewicz, and J. Czeczot, "Practical PLC-based Implementation of Adaptive Dynamic Matrix Controller for Energy-Efficient Control of Heat Sources," *IEEE Trans. Ind. Electron.*, vol. 0046, no. c, pp. 1–1, 2020, doi: 10.1109/tie.2020.2987272.

[13] M. Jelali, "An overview of control performance assessment technology and industrial applications," *Control Eng. Pract.*, vol. 14, no. 5, pp. 441–466, 2006, doi: 10.1016/j.conengprac.2005.11.005.

[14] W. L. Bialkowski, "Dreams versus reality: a view from both sides of the gap: manufacturing excellence with come only through engineering excellence," 1993.

[15] M. Jelali, *Control Performance Management in Industrial Automation: Assessment, Diagnosis and Improvement of Control Loop Performance*. Springer-Verlag London, 2013.

[16] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques-part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3757–3767, 2015, doi: 10.1109/TIE.2015.2417501.

[17] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques-part II: Fault diagnosis with knowledge-based and hybrid/active approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3768–3774, 2015, doi: 10.1109/TIE.2015.2419013.

[18] L. Desborough and R. Miller, "Increasing Customer Value of Industrial Control Performance Monitoring -Honeywell's Experience," *AIChE Symp. Ser.*, vol. 98, Jan. 2002.

[19] P. Ghorai, S. Majhi, V. R. Kasi, and S. Pandey, "Parameter Identification of Delayed Under-Damped Systems Using On-Line Relay Autotuning," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 66, no. 7, pp. 1197–1201, 2019, doi: 10.1109/TCSII.2018.2876476.

[20] R. D. Utomo and L. Chen, "A Relay-Feedback Autotuning PD Controller for Automatic Train Operation System," *2018 Int. Conf. Intell. Rail Transp. ICIRT 2018*, pp. 1–5, 2019, doi: 10.1109/ICIRT.2018.8641598.

[21] G. H. Nguyen, J. H. Shin, and W. H. Kim, "Autotuning controller for motion control system based on intelligent neural network and relay feedback approach," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 3, pp. 1138–1148, 2015, doi: 10.1109/TMECH.2014.2344692.

[22] P. Grelewicz, P. Nowak, M. Frątczak, and T. Kłopot, "Practical Veryfication of the Advanced Control Algorithms Based on the Virtual Commissioning Methodology – case study," 2018.

[23] P. Van Overschee and B. De Moor, “RAPID: The End of Heuristic PID Tuning,” *IFAC Proc. Vol.*, vol. 33, no. 4, pp. 595–600, 2000, doi: 10.1016/s1474-6670(17)38308-8.

[24] A. O’Dwyer, *Handbook of PI and PID Controller Tuning Rules*. Imperial College Press, 2009.

[25] J. B. R. K. L. Chien, J. A. Hrones, “On the Automatic Control of Generalized Passive Systems,” *Trans. Am. Soc. Mech. Engineeing*, vol. 74, pp. 175–185, 1972.

[26] W. K. Ho, K. W. Lim, and W. Xu, “Optimal gain and phase margin tuning for PID controllers,” *Automatica*, vol. 34, no. 8, pp. 1009–1014, 1998, doi: 10.1016/S0005-1098(98)00032-6.

[27] C. Grimholt and S. Skogestad, “Optimal PI and PID control of first-order plus delay processes and evaluation of the original and improved SIMC rules,” *J. Process Control*, vol. 70, pp. 36–46, 2018, doi: 10.1016/j.jprocont.2018.06.011.

[28] S. Skogestad and C. Grimholt, “The SIMC method for smooth PID controller tuning,” *Adv. Ind. Control*, pp. 147–175, Jan. 2012, doi: 10.1007/978-1-4471-2425-2_5.

[29] K. J. Åström and T. Hägglund, *PID Controllers: Theory, Design, and Tuning*. Research Triangle Park, North Carolina: ISA - The Instrumentation, Systems and Automation Society, 1995.

[30] Q. Bi, W. J. Cai, E. L. Lee, Q. G. Wang, C. C. Hang, and Y. Zhang, “Robust identification of first-order plus dead-time model from step response,” *Control Eng. Pract.*, vol. 7, no. 1, pp. 71–77, 1999, doi: 10.1016/S0967-0661(98)00166-X.

[31] F. S. Coelho and P. R. Barros, “Continuous-time identification of first-order plus dead-time models from step response in closed loop,” *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 36, no. 16. pp. 393–398, 2003, doi: 10.1016/S1474-6670(17)34793-6.

[32] G. A. Júnior, J. B. M. dos Santos, and P. R. Barros, “On Simple Identification Techniques for First-Order plus Time-Delay Systems,” *IFAC Proc. Vol.*, vol. 42, no. 10, pp. 605–610, 2009, doi: 10.3182/20090706-3-fr-2004.00100.

[33] A. A. M. Hakim and I. M. H. Sanhoury, “Adaptive Control for x Inverted Pendulum Utilizing Gain Scheduling Approach,” *2018 Int. Conf. Comput. Control. Electr. Electron. Eng. ICCCEEE 2018*, pp. 1–6, 2018, doi: 10.1109/ICCCEEE.2018.8515839.

[34] R. Divya, N. Pappa, and V. Govindan, “An Adaptive Gain Scheduled PID Controller for PWR type of Nuclear Reactor,” *Int. Conf. Innov. Control. Commun. Inf. Syst. ICICCI 2017*, no. 4, pp. 1–6, 2019, doi: 10.1109/ICICCIS.2017.8660905.

[35] Y. Xiujian, W. Zengcai, and L. Yufeng, “Gain-scheduling control of vehicle roll stiffness distribution,” *Proc. 2007 IEEE Int. Conf. Mechatronics Autom. ICMA 2007*, no. 3, pp. 399–404, 2007, doi: 10.1109/ICMA.2007.4303576.

[36] A. Maddi, A. Guessoum, and D. Berkani, “Design of nonlinear PID-smith predictor controllers with large time delays,” *Proc. 2015 IEEE World Conf. Complex Syst. WCCS 2015*, no. 2, pp. 2–5, 2016, doi: 10.1109/ICoCS.2015.7483219.

[37] V. J. Poorani and L. D. V. Anand, “Comparison of PID controller and Smith predictor controller for heat exchanger,” *2013 IEEE Int. Conf. Emerg. Trends Comput. Commun. Nanotechnology, ICE-CCN 2013*, no. Iceccn, pp. 217–221, 2013, doi: 10.1109/ICE-CCN.2013.6528496.

[38] S. M. Sundaram and P. R. Venkateswaran, “Smith Predictor Implementation of a High Dead Time Interacting Tank Process,” *2018 Int. Conf. Recent Innov. Electr. Electron. Commun. Eng. ICRIEECE 2018*, pp. 1386–1390, 2018, doi: 10.1109/ICRIEECE44171.2018.9009146.

[39] P. Tatjewski, *Advanced Control of Industrial Processes. Structures and Algorithms*. Springer-Verlag London, 2007.

[40] J. Oravec, M. Bakošová, M. Trafczynski, A. Vasičkaninová, A. Mészáros, and M. Markowski, “Robust model predictive control and PID control of shell-and-tube heat exchangers,” *Energy*, vol. 159, pp. 1–10, 2018, doi: 10.1016/j.energy.2018.06.106.

[41] P. Grelewicz, P. Nowak, M. Fratczak, and T. Klopot, “Practical Verification of the Advanced Control Algorithms Based on the Virtual Commissioning Methodology - A Case Study,” *2018 23rd Int. Conf. Methods Model. Autom. Robot. MMAR 2018*, pp. 217–222, 2018, doi: 10.1109/MMAR.2018.8485990.

[42] J. Han, “From PID to active disturbance rejection control,” in *IEEE Transactions on Industrial Electronics*, 2009, doi: 10.1109/TIE.2008.2011621.

[43] R. Madoński and P. Herman, “Survey on methods of increasing the efficiency of extended state disturbance observers,” *ISA Trans.*, vol. 56, pp. 18–27, 2015, doi: https://doi.org/10.1016/j.isatra.2014.11.008.

[44] S. Xingling and W. Honglun, “Back-stepping active disturbance rejection control design for integrated missile guidance and control system via reduced-order ESO,” *ISA Trans.*, vol. 57, pp. 10–22, 2015, doi: https://doi.org/10.1016/j.isatra.2015.02.013.

[45] T. Liu, S. Hao, D. Li, W. Chen, and Q. Wang, “Predictor-Based Disturbance Rejection Control

for Sampled Systems With Input Delay," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 2, pp. 772–780, 2019, doi: 10.1109/TCST.2017.2781651.

[46] T. Klopot, P. Skupin, M. Metzger, and P. Grelewicz, "Tuning strategy for dynamic matrix control with reduced horizons," *ISA Trans.*, Mar. 2018, doi: 10.1016/j.isatra.2018.03.003.

[47] G. Herbst, "Practical Active Disturbance Rejection Control: Bumpless Transfer, Rate Limitation, and Incremental Algorithm," *IEEE Trans. Ind. Electron.*, 2016, doi: 10.1109/TIE.2015.2499168.

[48] P. Nowak, J. Czeczot, and T. Klopot, "Robust tuning of a first order reduced Active Disturbance Rejection Controller," *Control Eng. Pract.*, 2018, doi: 10.1016/j.conengprac.2018.02.001.

[49] P. Nowak, K. Stebel, T. Klopot, J. Czeczot, M. Fratczak, and P. Laszczyk, "Flexible function block for industrial applications of active disturbance rejection controller," *Arch. Control Sci.*, 2018, doi: 10.24425/acs.2018.124708.

[50] P. Grelewicz, P. Nowak, J. Czeczot, and J. Musial, "Increment Count Method and Its PLC-Based Implementation for Autotuning of Reduced-Order ADRC with Smith Predictor," *IEEE Trans. Ind. Electron.*, vol. 68, no. 12, pp. 12554–12564, 2021, doi: 10.1109/TIE.2020.3045696.

[51] T. J. Harris, "Assessment of control loop performance," *Can. J. Chem. Eng.*, vol. 67, no. 5, pp. 856–861, Oct. 1989, doi: 10.1002/cjce.5450670519.

[52] F. G. Shinskey, *Process Control Systems : Application, Design, and Tuning*. McGraw-Hill Professional Publishing, 1996.

[53] K. J. Åström, *Introduction to stochastic control theory*. 1970.

[54] I. Filip, C. Vasar, I. Szeidert, and O. Prostean, "Consideration about the stability and performance of a minimum variance control system," *SACI 2020 - IEEE 14th Int. Symp. Appl. Comput. Intell. Informatics, Proc.*, pp. 83–88, 2020, doi: 10.1109/SACI49304.2020.9118827.

[55] M. Chen, L. Xie, and H. Su, "Impact of Model-Plant Mismatch to Minimum Variance Benchmark in Control Performance Assessment," *Chinese Control Conf. CCC*, vol. 2020-July, pp. 2252–2257, 2020, doi: 10.23919/CCC50068.2020.9188456.

[56] A. W. Ordys, D. Uduehi, M. A. Johnson, and M. A. Johnson, *Process Control Performance Assessment: From Theory to Implementation*. Springer-Verlag London, 2007.

[57] M. J. Grimble, "Generalized minimum variance control law revisited," *Optim. Control Appl.*

Methods, vol. 9, no. 1, pp. 63–77, 1988, doi: 10.1002/oca.4660090106.

[58] M. J. Grimble, "Controller performance benchmarking and tuning using generalised minimum variance control," *Automatica*, vol. 38, no. 12, pp. 2111–2119, 2002, doi: 10.1016/S0005-1098(02)00141-3.

[59] M. Grimble and P. Majecki, "Weighting Selection for Controller Benchmarking and Tuning," Jan. 2004.

[60] R. Kadali and B. Huang, "Controller performance analysis with LQG benchmark obtained under closed loop conditions," *ISA Trans.*, vol. 41, no. 4, pp. 521–537, 2007, doi: 10.1016/s0019-0578(07)60107-4.

[61] Z. Liu, H. Y. Su, L. Xie, and Y. Gu, *Improved LQG benchmark for control performance assessment on ARMAX model process*, vol. 8, no. PART 1. IFAC, 2012.

[62] Z. Liu, Y. Gu, and L. Xie, "MPC economic performance assessment based on equal-grid LQG benchmark.".

[63] B. S. Ko and T. F. Edgar, "Assessment of achievable pi control performance for linear processes with dead time," *Proc. Am. Control Conf.*, vol. 3, no. June, pp. 1548–1552, 1998, doi: 10.1109/ACC.1998.707239.

[64] B. S. Ko and T. F. Edgar, "PID control performance assessment: The single-loop case," *AIChE J.*, vol. 50, no. 6, pp. 1211–1218, 2004, doi: 10.1002/aic.10104.

[65] T. Hägglund and K. J. Åström, "Supervision of adaptive control algorithms," *Automatica*, vol. 36, no. 8, pp. 1171–1180, 2000, doi: 10.1016/S0005-1098(00)00026-1.

[66] F. G. Shinskey, *Process Control Systems: Application, Design, and Tuning*. 1996.

[67] J. Oravec, M. Horváthová, and M. Bakošová, "Energy efficient convex-lifting-based robust control of a heat exchanger," *Energy*, vol. 201, 2020, doi: 10.1016/j.energy.2020.117566.

[68] L. A. Sanchez, F. G. Arroyo, and R. A. Villavicencio, "Dynamic Matrix Control of Steam Temperature in Fossil Power Plants," *IFAC Proc. Vol.*, vol. 28, no. 26, pp. 275–280, 1995, doi: 10.1016/S1474-6670(17)44770-7.

[69] A. Sanchez-Lopez, G. Arroyo-Figueroa, and A. Villavicencio-Ramirez, "Advanced control algorithms for steam temperature regulation of thermal power plants," *Int. J. Electr. Power Energy Syst.*, vol. 26, no. 10, pp. 779–785, 2004, doi: 10.1016/j.ijepes.2004.08.003.

[70] T. Hägglund, "Automatic detection of sluggish control loops," *Control Eng. Pract.*, vol. 7, no.

12, pp. 1505–1511, 1999, doi: https://doi.org/10.1016/S0967-0661(99)00116-1.

[71] M. Veronesi and A. Visioli, “Performance assessment and retuning of PID controllers,” *Ind. Eng. Chem. Res.*, vol. 48, no. 5, pp. 2616–2623, 2009, doi: 10.1021/ie800812b.

[72] M. Veronesi and A. Visioli, *Performance assessment and retuning of PID controllers for load disturbance rejection*, vol. 2. IFAC, 2012.

[73] X. Gao, F. Yang, C. Shang, and D. Huang, “A Novel Data-Driven Method for Simultaneous Performance Assessment and Retuning of PID Controllers,” *Ind. Eng. Chem. Res.*, vol. 56, no. 8, pp. 2127–2139, 2017, doi: 10.1021/acs.iecr.6b03893.

[74] X. Gao, C. Shang, D. Huang, and F. Yang, “A novel approach to monitoring and maintenance of industrial PID controllers,” *Control Eng. Pract.*, vol. 64, no. May, pp. 111–126, 2017, doi: 10.1016/j.conengprac.2017.04.008.

[75] D. E. Rivera, M. Morari, and S. Skogestad, “Internal model control: PID controller design,” *Ind. Eng. Chem. Process Des. Dev.*, vol. 25, no. 1, pp. 252–265, Jan. 1986, doi: 10.1021/i200032a041.

[76] L. J. da Silva Moreira, G. A. Júnior, and P. R. Barros, “IMC PI Control Loops Frequency and Time Domains Performance Assessment and Retuning ∗,” *IFAC-PapersOnLine*, vol. 51, no. 4, pp. 148–153, 2018, doi: 10.1016/j.ifacol.2018.06.117.

[77] M. Bauer and I. K. Craig, “Economic assessment of advanced process control - A survey and framework,” *J. Process Control*, vol. 18, no. 1, pp. 2–18, 2008, doi: 10.1016/j.jprocont.2007.05.007.

[78] K. D. Starr, H. Petersen, and M. Bauer, “Control loop performance monitoring – ABB’s experience over two decades,” *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 526–532, 2016, doi: 10.1016/j.ifacol.2016.07.396.

[79] K. Bu, Y. Liu, and F. Wang, “Operating performance assessment based on multi-source heterogeneous information with deep learning for smelting process of electro-fused magnesium furnace,” *ISA Trans.*, 2021, doi: https://doi.org/10.1016/j.isatra.2021.10.024.

[80] M. Xu and P. Wang, “Evidential KNN-based Performance Monitoring Method for PID Control System,” in *2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, 2020, pp. 597–601, doi: 10.1109/ICMCCE51767.2020.00134.

[81] N. Pillay and P. Govender, “Multi-Class SVMs for automatic performance classification of closed loop controllers,” *Control Eng. Appl. Informatics*, vol. 19, no. 3, pp. 3–12, 2017.

[82] Z. Gao, “Scaling and Bandwidth-Parameterization based Controller Tuning,” in *Proceedings of the American Control Conference*, 2003, vol. 6, pp. 4989–4996, doi: 10.1109/acc.2003.1242516.

[83] G. Duffy, P. Mills, Q. Li, and L. Vlacic, “An on-line process dead-time estimation algorithm,” in *2017 Asian Control Conference, ASCC 2017*, 2017, pp. 339–344, doi: 10.1109/ASCC.2017.8287191.

[84] P. Grelewicz, P. Nowak, J. Czeczot, and M. Fratczak, “Corelation between Conventional and Data-Driven Control Performance Assessment Indices for Heating Process,” in *Proceedings of the 2019 22nd International Conference on Process Control, PC 2019*, 2019, pp. 86–90, doi: 10.1109/PC.2019.8815041.

[85] R. H. Byrd, J. C. Gilbert, and J. Nocedal, “A trust region method based on interior point techniques for nonlinear programming,” *Math. Program.*, vol. 89, no. 1, pp. 149–185, 2000, doi: 10.1007/PL00011391.

[86] H. Zhang, *The Optimality of Naive Bayes*. 2004.

[87] J. Ye, *Least squares linear discriminant analysis*. 2007.

[88] N. S. Altman, “An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression,” *Am. Stat.*, vol. 46, no. 3, pp. 175–185, Jan. 1992, doi: 10.2307/2685209.

[89] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification And Regression Trees*. Routledge, 1984.

[90] B. Gabrys and A. Bargiela, “General fuzzy min-max neural network for clustering and classification,” *IEEE Trans. Neural Networks*, vol. 11, pp. 769–783, Jun. 2000, doi: 10.1109/72.846747.

[91] B. Gabrys, “Agglomerative Learning Algorithms for General Fuzzy Min-Max Neural Network,” *J. VLSI signal Process. Syst. signal, image video Technol.*, vol. 32, pp. 67–82, Aug. 2002, doi: 10.1023/A:1016315401940.

[92] J. Suykens and J. Vandewalle, “Least Squares Support Vector Machine Classifiers,” *Neural Process. Lett.*, vol. 9, pp. 293–300, Jun. 1999, doi: 10.1023/A:1018628609742.

[93] G. Ke *et al.*, “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 3149–3157.

[94] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[95] Y. Freund and R. E. Schapire, “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting,” *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997, doi: https://doi.org/10.1006/jcss.1997.1504.

[96] P. Geurts, D. Ernst, and L. Wehenkel, “Extremely Randomized Trees,” *Mach. Learn.*, vol. 63, pp. 3–42, Apr. 2006, doi: 10.1007/s10994-006-6226-1.

[97] L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, pp. 5–32, 2001.

[98] T. I. Salsbury, “A practical method for assessing the performance of control loops subject to random load changes,” *J. Process Control*, vol. 15, no. 4, pp. 393–405, 2005, doi: 10.1016/j.jprocont.2004.08.004.

[99] A. Visioli, “Method for proportional-integral controller tuning assessment,” in *Industrial and Engineering Chemistry Research*, 2006, vol. 45, no. 8, pp. 2741–2747, doi: 10.1021/ie0508482.

[100] K. J. Åström and T. Hägglund, “Benchmark Systems for PID Control,” *IFAC Proc. Vol.*, vol. 33, no. 4, pp. 165–166, 2000, doi: https://doi.org/10.1016/S1474-6670(17)38238-1.

[101] U. Manimegalai-Sridhar, A. Govindarajan, and R. Russell Rhinehart, “Improved initialization of players in leapfrogging optimization,” *Comput. Chem. Eng.*, vol. 60, pp. 426–429, 2014, doi: 10.1016/j.compchemeng.2013.08.009.

[102] R. R. Rhinehart, M. Su, and U. Manimegalai-Sridhar, “Leapfrogging and synoptic Leapfrogging: A new optimization approach,” *Comput. Chem. Eng.*, vol. 40, pp. 67–81, 2012, doi: 10.1016/j.compchemeng.2012.02.011.

[103] X. Yang, X. Liu, and J. Shen, “The research of the explicit time delay and gain estimation algorithm based on fourth-order cumulants in acoustic pyrometry in the power plant boiler,” *Proc. - 2017 Chinese Autom. Congr. CAC 2017*, vol. 2017-Janua, no. 1, pp. 6091–6097, 2017, doi: 10.1109/CAC.2017.8243874.

[104] C. P. Vo, X. D. To, and K. K. Ahn, “A Novel Adaptive Gain Integral Terminal Sliding Mode Control Scheme of a Pneumatic Artificial Muscle System With Time-Delay Estimation,” *IEEE Access*, vol. 7, pp. 141133–141143, 2019, doi: 10.1109/access.2019.2944197.

[105] S. Mehrkanoon, Y. A. W. Shardt, J. A. K. Suykens, and S. X. Ding, “Estimating the unknown

time delay in chemical processes," *Eng. Appl. Artif. Intell.*, vol. 55, no. November 2017, pp. 219–230, 2016, doi: 10.1016/j.engappai.2016.06.014.

[106] S. Björklund, "A Survey and Comparison of Time-Delay Estimation Methods in Linear Systems," Linkopings Universitet, 2003.

[107] Songling Cao and R. Russell Rhinehart, "An efficient method for on-line identification of steady state," *J. Process Control*, vol. 5, no. 6, pp. 363–374, 1995.

[108] J. G. Ziegler and N. B. Nichols, "Optimum settings for automatic controllers," *InTech*, 1995.

[109] K. J. Åström and T. Hägglund, "Revisiting the Ziegler-Nichols step response method for PID control," *J. Process Control*, vol. 14, no. 6, pp. 635–650, 2004, doi: 10.1016/j.jprocont.2004.01.002.

[110] G. H. Cohen and G. A. Coon, "Theoretical Consideration of Retarded Control," *Trans. ASME*, 1953.

[111] E. A. Parr, *Industrial Control Handbook: Theory and Applications*. Industrial Press, 1989.

[112] B. Connell, *Process Instrumentation Applications Manual*. McGraw-Hill, 1996.

[113] M. Chidambaram, *Computer Control of Processes*. CRC Press, 2002.

[114] PMA (Prozeβ-und Maschinen-Automation) GmbH, "Industrieregler KS40-1, KS41-1, KS42- 1 Manuelle Optimierung, p. 18." [Online]. Available: https://www.west-cs.de/assets/DE-Manuals/PMA-KS40-1-KS41-1-KS42-1-Bedienungsanleitung-deutsch-9499-040-62718.pdf. [Accessed: 07-Feb-2020].

[115] P. W. Murrill, *Automatic Control of Processes*. International Textbook Company, 1967.

[116] A. O. Orlando, *Sintonización de controladores PI y PID empleando un índice de desempeño de criterio múltiple.* [San José, Costa Rica], 2006.

[117] M. Zhuang and D. P. Atherton, "Automatic tuning of optimum PID controllers," *IEE Proc. D Control Theory Appl.*, 1993, doi: 10.1049/ip-d.1993.0030.

[118] A. A. Rovira, P. W. Murrill, and C. L. Smith, "Tuning controllers for setpoint changes," *Instruments Control Syst.*, vol. 42, pp. 67–69, 1969.

[119] J. Sadeghi and W. Tych, "Deriving new robust adjustment parameters for PID controllers using scale-down and scale-up techniques with a new optimization method," in *Proceedings of ICSE: 16th Conference on Systems Engineering*, 2003, pp. 608–613.

[120] F.-S. Wang, W.-S. Juang, and C.-T. Chan, “Optimal tuning of PID controllers for single and cascade control loops,” *Chem. Eng. Commun.*, vol. 132, no. 1, pp. 15–34, Feb. 1995, doi: 10.1080/00986449508936294.

[121] G. K. I. Mann, B. G. Hu, and R. G. Gosine, “Time-domain based design and analysis of new PID tuning rules,” *IEE Proc. Control Theory Appl.*, 2001, doi: 10.1049/ip-cta:20010464.

[122] R. Padma Sree and C. Manickam, “Control of Unstable Bioreactor with Dominant Unstable Zero,” *Chem. Biochem. Eng. Q.*, vol. 17, Jun. 2003.

[123] A. Brambilla, C. Scali, and S. (Shanghai C. of P. T. Chen Shanghai (CN)), “Robust tuning of conventional controllers,” 1990.

[124] J. Gerry, “Tuning process controllers starts in manual,” 1999. [Online]. Available: http://www.expertune.com/artinTechMay99.aspx. [Accessed: 07-Feb-2020].

[125] X. Gong, “Normalised tuning method of PID controller parameters,” 2000.