

Machine learning for industrial sensing and control: A survey and practical perspective

Nathan P. Lawrence^a, Seshu Kumar Damarla^c, Jong Woo Kim^f, Aditya Tulsyan^c, Faraz Amjad^c, Kai Wang^g, Benoit Chachuat^d, Jong Min Lee^e, Biao Huang^c, R. Bhushan Gopaluni^b

^a*Department of Mathematics, University of British Columbia, Canada*

^b*Department of Chemical and Biological Engineering, University of British Columbia, Canada*

^c*Department of Chemical and Materials Engineering, University of Alberta, Canada*

^d*The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London SW7 2AZ, UK*

^e*School of Chemical and Biological Engineering, Institute of Chemical Processes, Seoul National University, 1, Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea*

^f*Department of Energy and Chemical Engineering, Incheon National University, Incheon 22012, Republic of Korea*

^g*School of Automation, Central South University, Changsha, 410083, China*

Abstract

With the rise of deep learning, there has been renewed interest within the process industries to utilize data on large-scale nonlinear sensing and control problems. We identify key statistical and machine learning techniques that have seen practical success in the process industries. To do so, we start with hybrid modeling to provide a methodological framework underlying several core applications. Specifically, we traverse soft sensing and process optimization and control. Soft sensing contains a wealth of industrial applications of statistical and machine learning methods. We quantitatively identify research trends, allowing insight into the most successful techniques in practice. We consider two distinct flavors for data-driven control: hybrid modeling in conjunction with mathematical programming techniques and reinforcement

learning. Throughout these application areas, we discuss their respective industrial requirements and challenges. A common challenge is the interpretability and efficiency of purely data-driven methods. This suggests a need to carefully balance deep learning techniques with domain knowledge. As a result, we highlight ways prior knowledge may be integrated into industrial machine learning applications. The treatment of methods, problems, and applications presented here is poised to inform and inspire practitioners and researchers to develop impactful data-driven sensing, optimization, and control solutions in the process industries.

Keywords: statistical machine learning, deep learning, hybrid modeling, soft sensing, reinforcement learning, control

1. Motivation

Data analytics and machine learning (ML) ideas are not new to the process industries¹. The review paper by Venkatasubramanian [1] provides an excellent overview of the history, successes, and failures of various attempts
5 over more than three decades to use ideas from artificial intelligence (AI) in the industry. In particular, statistical techniques such as principal component analysis, partial least squares, canonical correlation analysis, and time series methods for modeling, such as maximum likelihood estimation and prediction error methods, have been extensively used in industry [2]. Sev-
10 eral classification and clustering algorithms, such as k -means, support vector machines, and Fisher discriminant analysis, are also widely used in industry

¹For simplicity, we refer to chemical and biological industries as *process industries* or just *industries*.

[3, 4]. And several nonlinear approaches, such as kernel methods, Gaussian processes, and adaptive control algorithms, such as reinforcement learning, have been applied in some niche applications [5, 6, 7].

15 Despite the longstanding success of many statistical techniques in industry, there is also considerable interest in developing sensing and control technologies based on more recent ML architectures [1, 8, 9]. Broadly speaking, these aspirations are driven by the promises of increased autonomy: increased operational efficiency, consistency, and safety; improved scalability
20 beyond linear methods; upskilling of plant personnel [10]. Consequently, this paper addresses the need to dissect and organize the general use of modern ML techniques in industrial applications. In doing so, such a treatment will inform practitioners of the latest research trends and their potential practical impact. Conversely, researchers in core areas will benefit from a holistic view
25 of successful ML techniques and the industrial requirements they satisfy.

1.1. Overview and scope

This paper is a significant extension of Gopaluni et al. [11]: in addition to a more detailed and expansive treatment of the literature, we discuss the practical success of various methods. Note that this is primarily a problem-
30 driven survey; however, we have provided sufficient references for interested readers on the underlying methods discussed here. Moreover, we have included additional exposition on some of these methods in the supplementary material.

Hybrid modeling is first introduced to provide a conceptual framework

35 underlying core application areas, namely:²

1. Soft sensing
2. Process control

Process control also includes process optimization. In our survey, we identify several methodological areas of research: statistical learning and machine
40 learning, deep learning and its variants, and reinforcement learning. Algorithms from each of these methodological areas are used to varying degrees among the core applications. Soft sensing encompasses more statistical and machine learning methods, with some discussion of deep learning. On the optimization and control side, we discuss hybrid modeling in tandem with
45 mathematical programming and reinforcement learning.

This is by no means an exhaustive survey of the recent research on these topics. However, we have tried our best to include some of the most critical developments of ML tools in the process industries. In that vein, we only discuss methods that have seen industrial use or have received consid-
50 erable research attention within process systems engineering, either in real life or in simulations. Therefore, speculation about the potential use of very recent developments in the broader ML community, such as ChatGPT or other large language models, is beyond the scope of this paper. However, we provide insight into the practical deployment of ML techniques in the process
55 industries.

²Tackling the trio of sensing, monitoring, and control in detail is beyond the scope of this paper. Readers are referred to Ge et al. [4], Bi et al. [12], Sansana et al. [13] for more details on ML methods for process monitoring.

This paper surveys a large number of algorithms. Table 1 gives a convenient list to reference across all sections. Throughout this paper, *artificial intelligence* is the broadest term for classifying machines that aim to mimic human intelligence. It is intended to predict, automate, and optimize the tasks humans have traditionally performed, such as speech recognition, image
60 recognition, decision-making, and translation. *Machine learning* is an area of artificial intelligence and computer science where algorithms are developed to extract patterns from data and make predictions. *Supervised learning* is a
65 branch of machine learning comprised of algorithms for determining a predictive model based on labeled data with known outcomes. On the other hand, *unsupervised learning* is a branch of machine learning devoted to learning patterns from unlabeled data.

2. Mathematical modeling approaches

The core applications of this paper are soft sensing and process optimization and control. These areas rely on dynamic mathematical models to infer
70 measurements, make decisions, and synthesize controllers. Therefore, before describing the prominent machine learning (ML) techniques in these areas, it is useful to introduce the foundational assumptions and architectures underlying such models.

75 2.1. Knowledge-driven, data-driven, and hybrid modeling

Knowledge-driven (mechanistic or white box) modeling based on first principles and *data-driven* (or black box) modeling constitute two opposite strategies. Developing mechanistic models requires a deep understanding of

Table 1: Full forms for acronyms. Divided into three sections, top to bottom: 1) statistical learning, 2) machine learning & deep learning, and 3) reinforcement learning & control methods.

CCA	Canonical correlation analysis	LASSO	Least absolute shrinkage and selection operator
FA	Factorial analysis	LR	Logistic regression
GMM	Gaussian mixture model	PCA	Principal component analysis
ICA	Independent component analysis	PLS	Partial least squares
LARS	Least-angle regression	RBC	Reconstruction-based contribution
ANFIS	Adaptive network fuzzy inference system	GRNN	General regression neural network
ANN	Artificial neural network	MLP	Multilayer perceptron
BN	Bayesian network	RBFNN	Radial basis function neural network
CNN	Convolutional neural network	RNN	Recurrent neural network
DNNE	Decorrelated neural network ensemble	RT	Regression tree
DNN	Deep neural network	RVM	Relevance vector machine
ELM	Extreme learning machine	SFA	Slow feature analysis
ESN	Echo state network	SVM	Support vector machine
ENN	Elman neural network	TL	Transfer learning
GPR	Gaussian process regression	WNN	Wavelet neural network
A3C	Asynchronous advantage actor-critic	PI ²	Policy improvement with path integrals
ADP	Approximate dynamic programming	PID	Proportional-integral-derivative
DDPG	Deep deterministic policy gradient	PPO	Proximal policy optimization
DQN	Deep Q -network	RTO	Real-time optimization
HJB	Hamilton-Jacobi-Bellman	SAC	Soft actor-critic
MPC	Model predictive control	TD3	Twin-delayed DDPG

the processes at play. It is often labor-intensive, but embodying first principles may enable extrapolation beyond the conditions under which these models are trained. By construction, mechanistic models have a fixed structure and comprise a fixed number of parameters, often with a physical or empirical interpretation. For this reason, they may also be classified as *parametric* models.

By contrast, data-driven models require little physical knowledge and are fast to deploy or maintain. But a larger dataset is also typically needed for their construction, and their validity may not extend far beyond the conditions under which they are trained. The structure of a data-driven model does not need to be dictated by *a priori* knowledge but may be tailored to the training data at hand instead. A further distinction is whether a data-driven model tries to describe data with a set of parameters of fixed size, regardless of the size of the training dataset, in which case it is categorized as *parametric*, or whether its structure and number of parameters may evolve with the size of the dataset, commonly referred to as *nonparametric* [14, 15]. The so-called nonparametric regression models fall in the second category, whereby the predictor does not take a predetermined form, using techniques such as nearest-neighbor interpolation, local regression, and Gaussian process (GP) regression. However, the distinction between parametric and nonparametric models in statistical and ML is not without intricacies. For instance, a linear SVM is a typical example of a parametric model, having a fixed number of parameters—a weight for each input dimension. In contrast, RBF-kernel SVM may be considered nonparametric since the number of parameters grows with the size of the training set—a weight for each training point.

The basic idea behind *hybrid* models is to combine knowledge-driven and
105 data-driven models in such a way as to overcome their respective limitations.
This strategy is also frequently referred to as *gray box* or *block-oriented* mod-
eling in the literature. At the same time, the term hybrid *semi-parametric*
modeling is coined to describe those hybrid models where the data-driven
component is nonparametric [16]. *Multi-fidelity modeling* has also developed
110 fast in recent years and is akin to hybrid modeling. The idea is to use a (pos-
sibly inaccurate) knowledge-driven model as low-fidelity and correct it with
(noisy) process data, considered to be higher fidelity [17]. In particular, this
strategy has been applied in uncertainty propagation, inference, and opti-
mization and is also instrumental in small data problems (see supplementary
115 material).

It is worth noting that hybrid modeling has been investigated for over
25 years in chemical and biological process engineering [18, 19, 20, 21]. The
claimed benefits of hybrid modeling in these application domains include
faster prediction capability, better extrapolation capability, better calibration
120 properties, easier model life-cycle management, and higher benefit/cost ratio
to solve complex problems; see recent survey papers on the development
and applications of hybrid models by von Stosch et al. [16], Solle et al. [22],
Schuppert and Mrziglod [23], Zendehboudi et al. [24], Ahmad et al. [25],
Bradley et al. [26]. Hybrid models may be used to enable soft sensors (see
125 Section 3) or model-based optimization and control (see Section 4) in a first
principles approach.



Figure 1: Typology of hybrid models (see von Stosch et al. [16]). A and C represent serial structures: under A, a data-driven model is used as input to a knowledge-driven model; C is the reverse. B represents a parallel structure in which knowledge-driven predictions are corrected by data-driven predictions.

2.2. Hybrid modeling paradigms

2.2.1. Traditional serial and parallel hybrid models

The usual classification of hybrid model structures is either as *serial* or
 130 *parallel* [27]. In the serial approach, the data-driven model is most commonly
 used as an input to the mechanistic model (see Figure 1A), for instance, a
 material balance equation with a kinetic rate expressed using a data-driven
 model. This structure is especially suited to situations where precise knowl-
 edge about specific underlying mechanisms is lacking, yet sufficient process
 135 data exists to infer the corresponding relationship [18, 21]. However, when
 the mechanistic part of the model presents a structural mismatch, one should
 not expect the serial approach to perform better than a purely mechanistic
 approach. In the parallel approach, by contrast, the output of the data-driven
 model is used to correct the predictions of the mechanistic model [19, 20],
 140 most often in the form of an additive correction (see Figure 1B). This struc-
 ture can significantly improve the prediction accuracy of a mechanistic model
 when the data-driven component is trained on the residuals between process
 observations and mechanistic model predictions. However, this accuracy may
 not be better than the sole mechanistic model when the process conditions
 145 differ drastically from those in the training set.

Historically, the most common data-driven modeling techniques embedded in hybrid models have been multilayer perceptron (MLP) and RBF-based regression [16]. Recent representative applications include the development of a serial hybrid model to predict hydraulic fractures created by injecting fluid into a reservoir that accounts for the leak-off rate of the fracturing fluid using an MLP [28] and the development of a serial hybrid model of the thin film growth process coupling a macroscopic gas phase model described by partial differential equations to a microscopic thin-film model described by stochastic partial differential equations via an MLP [29]. Naturally, many other statistical and ML techniques have also been investigated in this context. For instance, Ghosh et al. [30] used subspace identification to construct the data-driven component in a parallel hybrid model and demonstrated the approach on a batch polymerization reactor. Lopez et al. [31] developed a serial hybrid model of a lignocellulosic fermentation process, whereby the glucose concentration is estimated from spectroscopic data using a partial least squares regression model. GP regression has also attracted attention due to its ability to estimate the predictor’s variance, for example, in bioprocess engineering applications [32].

Parallel hybrid models can significantly alleviate the issue of maintaining a complex mechanistic model since the data-driven component is trained to capture model mismatch in the first place, possibly in a nonparametric manner. For dynamic systems in particular, a popular approach entails training the data-driven model on the residuals between the predicted and observed states at given time instants [33]. Notice that such a data-driven model could either comprise algebraic or differential equations. By contrast, serial hybrid

models can prove more challenging to design, especially when the outputs of the data-driven component cannot be observed directly [34]. In such a case, training and assessing the performance of the data-driven component requires one to simulate the full serial hybrid model and compare its outputs
175 to the available observations. Identifying the unknown model parameters within such hybrid models has relied on regularized regression techniques, such as LASSO and LARS [35].

Another challenge shared by serial and parallel hybrid modeling paradigms is automatically detecting the best structure for the data-driven component.
180 Generally speaking, minimizing the number of parameters needed to capture the underlying mechanisms is desirable, that is, to neither underfit nor overfit the data. Classical approaches to help discriminate among multiple nonparametric model structures include the Akaike Information Criterion and Bayesian Information Criteria. Willis and von Stosch [36] proposed an
185 approach based on sparse regression and mixed-integer programming to simultaneously decide the structure and identify the parameters for a class of rational functions embedded into a serial hybrid model. Recently, Zhang et al. [37] applied hybrid modeling in combination with sparse identification of nonlinear dynamics [SINDy; 38] to a photo-production bioprocess,
190 whereby a sparse quadratic correction of the kinetic model is identified using mixed-integer nonlinear programming techniques. More generally, there is significant scope for extending sparse and symbolic regression techniques to enable the construction of hybrid models. Notably, the platform ALAMO [39] can enforce constraints on the response variables to incorporate first
195 principles knowledge, thereby revealing hidden relationships between regres-

sion parameters that may not be directly available to the modeler. One approach to incorporating such constraints is via semi-infinite programming [40]. Another promising direction entails using sum-of-squares optimization techniques to tackle this problem [41, 42].

200 *2.2.2. Emerging trends*

The traditional hybrid modeling approach has put a mechanistic model at its core. It uses data-driven elements to either describe specific unknown or poorly understood mechanisms or correct the predictions of the mechanistic model. Another way of incorporating domain knowledge and mechanistic
205 models is feature engineering, where the inputs to the data-driven elements are augmented by terms that would also appear in mechanistic models; for instance, think of enthalpy, which is not a measurement but a useful term in energy balances. Hybrid models whereby the mechanistic model is now used as an input to the data-driven component have become increasingly popular
210 in recent years (see Figure 1C). This approach includes physics-informed neural networks where the underlying conservation equations are imposed as extra constraints on the MLP’s parameters [43], like the classical orthogonal collocation theory on finite elements using piecewise polynomials [44, 45]. Co-Kriging techniques have also been developed where a GP trained using
215 data from a mechanistic model is combined with a second GP trained using process data (or a high-fidelity model) [46]. Such an approach also enables multi-fidelity modeling using linear or nonlinear autoregressive techniques [47, 48] and deep GPs [49], and finding applications, for instance, in the optimization of complex black box simulators and legacy codes. Another
220 body of research has been concerned with learning a dynamic system by

accounting for prior information, for instance, the regression of polynomial dynamic systems with prior information using sum-of-squares optimization methods [50].

Since there is no universal framework, a recurring challenge with hybrid
225 modeling is selecting the appropriate paradigm—for example, physics-driven
against data-driven backbone, or serial against parallel structure—for a par-
ticular application, such as small vs. large datasets or noisy vs. high-quality
data. This selection process still lacks a solid theoretical basis, although sys-
tematic computational comparisons of various hybridization techniques have
230 emerged in recent years [26]. Finally, looking beyond current hybrid models,
Venkatasubramanian [1] argued for the development of hybrid artificial intel-
ligence systems that would combine not only mechanistic with data-driven
models but also causal models-based explanatory systems or domain-specific
knowledge engines. Likewise, the mechanistic model could be replaced by
235 a graph-theoretical model, such as signed digraphs, or a production system
model, creating entirely new research fields.

3. Soft sensors in process industries

Soft sensing represents the most fundamental application of machine
learning (ML) techniques in the process industries. By extension, optimiza-
240 tion and control add complexity to a soft sensing core. As a result, based on
our analysis and own experience, soft sensing contains the most industrial
penetration of ML applications. We quantitatively analyze which ML meth-
ods have seen practical success and which are currently being researched.
We offer practical considerations and insights for implementing soft sensors

245 in practice to balance the apparent industrial-academic disconnect.

3.1. Motivation for soft sensing

In the process industries, some variables are difficult to measure online due to technological limitations or the high cost of sensors. These variables indicate a product's intermediate or final quality and must be continuously
250 monitored and controlled. In such circumstances, mathematical models are developed using easy-to-measure variables. These models provide a continuous estimate for quality variables in real time. The mathematical models devoted to the estimation of plant variables are called *soft sensors* [51, 52]. The process industries, such as refineries, steel plants, polymer industries, or
255 cement industries, remain the dominant users of soft sensors (see Figure 2).

Similar to hybrid modeling, soft sensors can be categorized as *knowledge-driven* and *data-driven*. Knowledge-driven soft sensors (or white box models), such as Kalman filters, are based on first principles models that describe the physical and chemical laws that govern the process, such as mass and
260 energy balance equations. In contrast, data-driven soft sensors (or black box models) have no information about the process and are based on empirical observations (historical process data). A third type of soft sensor, called *hybrid models* (or gray box models), uses a data-driven method to estimate the parameters of a knowledge-driven model. This special combination is closely
265 related to the general concept of *hybrid modeling*, as discussed in Section 2. For instance, a model may incorporate physics-based simulations and process measurements.

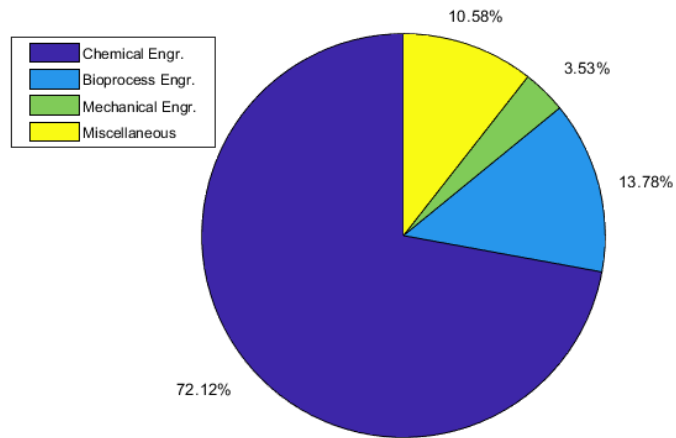


Figure 2: Distribution of soft sensor applications.

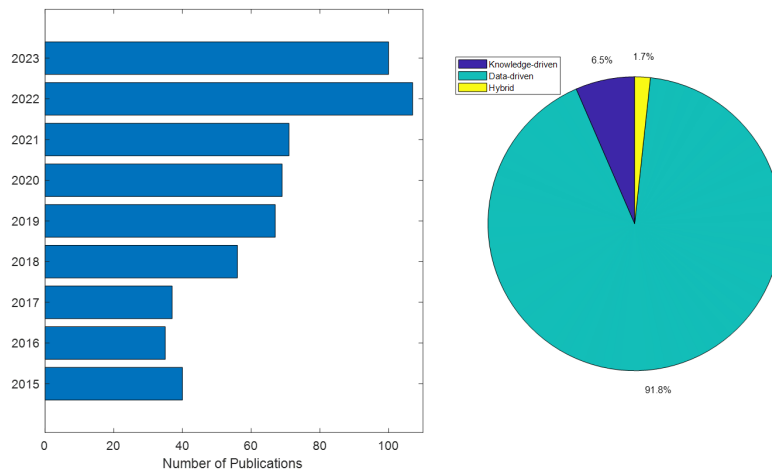


Figure 3: Research publication in soft sensors from 2015 to 2023.

3.2. A quantitative overview of soft sensing

Literature was collected by gathering articles published between 2015
270 and 2023 in relevant journals from publishing houses like Elsevier, Springer,
Wiley, Taylor and Francis, MDPI, World Scientific, Hindawi, De Gruyter,
AMSE, and IEEE. For the publication search, keywords such as “soft sensor”,
“virtual sensor” or “inferential model” were used. The statistics shown in
Figure 3 were computed based on the collected literature.

275 These statistics indicate that the research conducted in soft sensing be-
tween 2015 and 2023 was primarily focused on data-driven models. This is
unsurprising, as data-driven soft sensors can often capture complex and unex-
plained process dynamics more succinctly. In contrast, knowledge-driven soft
sensors require much expert process knowledge, which is not always available.
280 In addition, knowledge-driven soft sensors are difficult to calibrate, especially
for complex nonlinear processes. Note that hybrid model-based soft sensors
received the least research attention. Data-driven soft sensors can be further
categorized based on the learning technique used for modeling.

Tables 2 to 3 show the current trends in the data-driven soft sensing.
285 Table 1 contains the full forms for the acronyms used in Tables 2 to 3.
The research in soft sensing has dramatically shifted from statistical to ML
methods. Artificial neural networks (ANNs) received the greatest attention
among ML methods. The class of feedforward single hidden layer neural
networks (shallow networks)—encompassing multilayer perceptron (MLP),
290 GRNN, ELM, radial basis function neural network (RBFNN), wavelet neural
network (WNN) in Table 1—have more applications in soft sensing than
recurrent neural networks (RNNs) and deep learning. Aside from ANNs,

support vector machine (SVM) is the second most widely used ML method for developing inferential models.

295 Transfer learning is slowly gaining applications in inferential measurements. Transfer learning alludes to the scenario where knowledge gained while performing one specific task is exploited to carry out a different but related task. Especially when data collection becomes difficult in the task of interest, transfer learning still works by sharing information on relevant data
300 in other domains [53]. Transfer learning has yet to be applied to the online prediction of process variables.

Static (time-invariant) soft sensors are developed using data from a single operating mode. However, their prediction accuracy degrades over time as the process shifts to a new operating region. Adaptive soft sensors tackle this
305 issue by updating their parameters based on new samples.³ Less than one-third of soft sensors are adaptive, most of which use a just-in-time strategy to update model parameters in response to samples arriving in real time (see Figure 4 and Table 4). Therefore, computationally feasible methods are required. In particular, partial least squares (PLS) is the preferred algorithm
310 for local modeling. The training of global soft sensor models is performed offline. Then, trained soft sensor models are deployed online to obtain real-time estimates for key process or quality variables. Although training time is comparatively very high, most global soft sensors produce estimates quickly when used online [52, 54]

³One may further distinguish between unimodal and multimodal soft sensors. For instance, a static soft sensor may be developed for multimodal distributions. In any case, we are referring to the case of a static soft sensor degrading over time.

Table 2: Distribution of data-driven methods for soft sensors, split between statistical and ML methods.

Statistical methods	% of publications	ML methods	% of publications
PLS	11.38	ANN	47.72
PCA	4.54	TL	2.02
FA	0.95	RT	4.59
ICA	1.70	SFA	2.75
LASSO	1.51	RVM	2.57
GMM	2.64	SVM	7.53
-	-	ANFIS	1.65
-	-	GPR	5.87
-	-	BN	2.57
Total	22.73	Total	77.27

Table 3: Distribution of various types of ANNs for soft sensors.

Method	% of publications
MLP	14.23
DNNE	0.75
DNN	34.64
ELM	11.92
GRNN	6.92
WNN	2.31
RBFNN	2.68
RNN	21.94
ENN	0.38
ESN	4.23

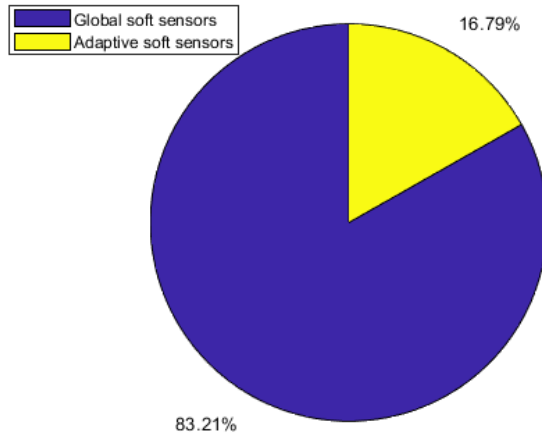


Figure 4: Distribution of global and adaptive soft sensors.

315 In Table 5, publications on each data-driven technique have been grouped into three categories: publications based on simulation data, publications based on industrial data, and publications that reported industrial implementation. Notice that most soft sensors have been developed and tested on industrial data. Still, only some of them—PLS, MLP, WNN, SVM, relevance
 320 vector machine (RVM), Gaussian process regression (GPR) and regression tree (RT)—have made it into actual industrial implementation. Of course, there may be a publication bias for academic examples and not all real-world industrial applications may be reported on.

3.3. Computational cost of soft sensors

325 The training time refers to the time taken to determine optimal values for the parameters of a soft sensor model. Once the developed soft sensor is implemented online in a distributed control system, it is used to estimate key

Table 4: Distribution of statistical and ML methods in local modeling of adaptive soft sensors.

Method	% of publications
PCA	7.05
MLP	3.52
SVR	11.77
GMM	2.35
GPR	15.30
BN	4.70
RVM	2.35
ELM	9.42
FA	3.52
PLS	31.78
LASSO	8.24

process or quality variables at regular sampling intervals. The time required to get the estimates is called *soft sensing time*.

330 Different techniques have various levels of computational complexity, that is, model training time. Since principal component analysis (PCA) [55], slow feature analysis [56], independent component analysis [57], and factorial analysis [58] can be developed in a single iteration, they require relatively low computational time compared to LASSO [59], and GMM [60] techniques, 335 which involve using iterative optimization algorithms to determine the model parameters. In general, ML methods need more computational time than statistical methods [54]. Further, the computational complexity of ML methods is influenced by the factors listed below [61] :

- Amount of training data.

Table 5: Breakdown of methods for soft sensors according to the level of industrial applications.

Method	Simulation data	Industrial data	Industrial use	Number of publications
PCA	8	16	0	24
PLS	12	45	3	60
SFA	4	11	0	15
ICA	7	2	0	9
LR	0	1	0	1
LASSO	1	6	0	7
FA	0	5	0	5
GMM	3	11	0	14
MLP	4	32	1	37
RBFNN	2	5	0	7
WNN	1	4	1	6
RNN	4	53	0	57
GRNN	5	13	0	18
ELM	6	25	0	31
ANFIS	3	6	0	9
DNNE	1	1	0	2
DNN	7	80	0	87
SVM	12	21	4	37
RVM	4	9	1	14
GPR	12	18	2	32
RT	5	19	1	25
BN	0	14	0	14
TL	4	7	0	11
ESN	2	9	0	11
ENN	0	1	0	1
Total	107	414	13	534

- 340 • Number of features or input variables.
- Type of training algorithm employed.
- Number of layers.
- Number of neurons (size) in layers.
- Type of device used (such as CPU or GPU).

345 The ELM is considered the fastest ML algorithm because it does not have parameters that need to be learned. The second fastest ML method is the GRNN, which has a single learnable parameter (spread or width of a radial basis function).⁴ Then come RTs and decorrelated neural network ensembles, which can be constructed more easily than shallow neural
350 networks like RBFNN, MLP, WNN, and adaptive network fuzzy inference system (ANFIS). As RBFNN uses hybrid learning (not hybrid modeling)—unsupervised learning for the middle layer and supervised learning (linear regression) for the last layer—it is usually faster than MLP, ANFIS, and WNN, which use iterative gradient descent algorithms. SVM is the slowest
355 of the kernel-based ML methods (SVM, GPR, and RVM). Bayesian networks rely on the expectation-maximization algorithm to optimize their parameters, which takes a little more training time than RTs. Dynamic ML methods, such as RNNs, involve more operations than their static ML counterparts,

⁴Soft sensors were developed based on industrial data using the ML algorithms surveyed in this paper. The soft sensors were tested on an independent validation dataset and the computation time for each soft sensor was recorded. It turned out that, among all of the soft sensors, the ELM-based soft sensor took the least computational time to produce estimates. The GRNN-based soft sensor was the second fastest one.

so they require more memory and computational power [62, 15]. Similarly,
360 deep neural networks (DNNs) often include several layers and hence, con-
tain many parameters. A large amount of training data is necessary to train
DNNs. Therefore, DNNs are recognized as the most computationally expen-
sive methods of all the data-driven techniques.

3.4. *Industry implementation of soft sensors*

365 In industries, soft sensors are developed by in-house control engineers
or third-party contractors (service engineers) from service providers such as
Honeywell or Yokogawa. These service providers use their own software to
build the soft sensors. When the existing technology used by service providers
is inadequate to handle a problem or in-house control engineers have no
370 knowledge of other soft sensing algorithms, the industries provide research
funding to universities, research organizations, or startups to develop sophis-
ticated soft sensors to model complex nonlinear processes. The following
steps outline how soft sensors are developed and implemented in industries.

1. After recognizing a need for a soft sensor application, a team consist-
375 ing of a panel operator, process engineer, control engineer, and project
manager is formed. The process engineer prepares a charter to define
the core objectives, scope, responsibilities, and timeline of the project.
This outlines the benefits that the soft sensor project can offer. All
the benefits are usually quantified in terms of how much money can be
380 saved. For example, this cost–benefit analysis typically involves weigh-
ing the upfront costs—hardware, software, consultants—and continued
costs—software licenses, in-house domain experts to handle support

and maintenance—against anticipated improved revenue and throughput, as well as reduced cost of the soft sensor. Once the team is satisfied with the benefits, the soft sensor project launches.

385

2. The next step in executing a soft sensor project involves obtaining process knowledge or expert experience knowledge to identify input variables that have a noteworthy influence on output variables [63]. The use of process knowledge or expert experience avoids the inclusion of redundant input variables in soft sensor modeling, leading to reduced model complexity and improved accuracy. In the absence of such knowledge, ML algorithms such as LASSO, hybrid LASSO, and ridge regression can be used to identify and remove input variables that have negligible impact on the output variable.

390

3. The third step entails process data collection and preprocessing. The process data are often abundant but poor in information. This is due to significant disturbances, outliers, and missing values. Soft sensors developed using these data may provide incorrect estimates for quality variables. The outliers and missing values from the raw industrial data should be removed to obtain clean data for developing the soft sensor. Although it may not be theoretically rigorous, the usual practice is to detect and delete samples with outliers [64]. Missing values are treated in the same fashion. This approach ensures that the clean data are free of outliers and missing values.

395

400

4. The data collection in industrial settings is often associated with multi-rate sampling. If the sampling frequency of the input variables is higher

405

than that of the output variable, then it is necessary to synchronize the variables. Down-sampling may be used to deal with the multi-rate sampling problem. In the down-sampling approach, samples of the input variables that do not have the respective measurements of the output variable are removed [54].

410

5. After the process data are preprocessed, they are split into training and validation subsets. The training subset is used to construct a soft sensor model whereas the validation subset is used to evaluate the prediction performance of the soft sensor model. This is called offline validation. The usual practice is to develop a linear model first. If the linear model cannot produce accurate estimates, then more complex statistical or ML algorithms are used.

415

6. If the soft sensor model delivers satisfactory performance in the offline validation, it is implemented in a distributed control system. Then the performance of the soft sensor is monitored for some time period. If the soft sensor exhibits poor performance, then modifications are made. This is online validation. For offline and online validation, metrics such as the correlation coefficient and root mean squared error are used to quantify the performance of soft sensors [54]. In addition, qualitative analysis is considered to see if soft sensor estimates follow the lab data trend. If the soft sensor estimates are poor, the input data are first examined for possible reasons, such as sensor failures, data transmission problems, outliers, plant shutdowns, and plant upsets. Poor estimates can be characterized by low correlation to lab data, estimates out of

420

425

430

the operational range, or significant deviation from lab data. If the input data are good, the following strategies are used to get accurate and reliable estimates:

- Retraining of the soft sensor using the latest data.
- 435 • Changing the soft sensor modeling algorithm.
- Using a different training algorithm.
- Changing the parameter initialization method.
- Using approaches that can avoid or reduce overfitting.

Regardless of the type of soft sensor, practicing engineers usually follow
440 the above approach to assess the performance of soft sensors.

7. If the online soft sensor consistently provides reasonable results, the soft sensor is used as a measuring device in a control loop. After successfully implementing the soft sensor-based control application, the soft sensor application is handed over to the panel operator. The human-in-the-
445 loop aspect described above is crucial in translating research results into practical applications.

3.5. Challenges in soft sensor development

Challenges that are often encountered in soft sensor developments are discussed below.

- 450 • Lack of labeled data is the main challenge that must be dealt with in order to build good soft sensor models. Quality variables are less frequently measured than easily measurable process variables, such as

temperature, pressure, flow rate, and level. A sample of a quality variable is collected once every shift (that is, 8 hours) or 24 hours. Because
455 of the long sampling interval, an insufficient amount of practical labeled data is available. A soft sensor trained with a limited amount of labeled data may not be able to capture the underlying relationship between the input variables and the output variable. To deal with
460 this problem, a virtual sample generation method may be used to obtain estimated output values for the corresponding input data [65]. As an alternative, semi-supervised learning may be used to construct the soft sensor. Unsupervised learning algorithms like PCA, autoencoders, stacked autoencoders, or deep belief networks can extract features from unlabeled input data. These features are related to the output variable
465 by any data-driven linear or nonlinear model [15].

- Operating conditions of the industrial process may change depending on the demand for products, prices of raw materials, and so on. A soft sensor developed using data from one operating condition may not perform well when the operating condition changes. In this situation,
470 multimode soft sensors can be used to get accurate estimates [66].
- Soft sensor maintenance is crucial to continuously attain reasonable estimates, as the performance of an online soft sensor may degrade over time. As a result, estimates obtained by a poorly performing soft sensor do not follow lab data trends. To circumvent this hurdle, the
475 soft sensor is retrained with recent data, and deployed online. A more popular approach to maintain the accuracy of the soft sensor is to adopt

a bias updating strategy. In the bias updating strategy, the soft sensor outputs are brought closer to the lab data [67].

4. Data-driven and hybrid modeling approaches for optimization and control

480

We revisit data-driven and hybrid modeling in the context of solving optimization and control problems. We further introduce reinforcement learning as an emerging paradigm for solving challenging control tasks. In the same way hybrid modeling represents a spectrum between knowledge-based
485 and data-based modeling, model-based optimization, model predictive control, and reinforcement learning all encompass model-based and model-free methodologies. Naturally, these techniques are also compatible with hybrid modeling approaches, offering new challenges and research opportunities.

4.1. Model-based optimization

490

A large number of hybrid modeling applications have been geared towards offline process optimization. Here, a hybrid model is appealing because key operational variables in terms of process performance may be included in the mechanistic part of the model. This is to retain sufficient extrapolation while capturing other parts of the process using data-driven techniques, for exam-
495 ple, to reduce the computational burden. Local (gradient-based) or stochastic search techniques have traditionally been applied to solve the resulting model-based optimization problems. But a recent trend has been using complete search techniques to overcome convergence to a local optimum and guarantee global optimality in problems with trained machine learning models embedded, such as multilayer perceptron (MLP) [68, 69, 70], Gaussian
500

process (GP) [71], or gradient-boosted trees [72]. Applications in chemical engineering include the optimization of simple reactor operations and process flowsheets [68] and optimal catalyst selection [72].

It should be noted that developing a data-driven or hybrid model to speed
505 up the optimization of a more fundamental model is akin to conducting a surrogate-based optimization. The latter constitutes an active research area in process flowsheeting, computational fluid dynamics, and molecular dynamics [73]. They can be broadly classified into local and global approaches. Global approaches proceed by constructing a surrogate model based on an
510 ensemble of mechanistic simulations before optimizing it, often within an iteration where the surrogate is progressively refined. Several successful implementations rely on MLPs [74], GPs [75, 76, 77], or a combination of various basis functions [40, 78] for the surrogate modeling. Practical applications have been for rigorous design of distillation columns [76, 77] and flowsheet or
515 superstructure optimization of chemical processes [75, 74]. By contrast, local approaches maintain an accurate surrogate of the mechanistic model within a trust region, whose position and size are adapted iteratively. This procedure entails reconstructing the surrogate model as the trust region moves around. Still, it can offer global convergence guarantees, for example, when the sur-
520 rogates meet the full linearity property [79]. Applications of this approach to chemical process optimization include solved-based CO₂ capture [80] and integrated carbon capture and conversion [81].

4.2. Model predictive control and real-time optimization

The real-time optimization (RTO) and nonlinear/economic model predic-
525 tive control (MPC) methodologies use a process model at their core. So far,

most successful implementations of RTO and MPC have relied on mechanistic models [82, 83, 84]. But there has been interest in data-driven approaches, which use surrogate models trained on historical data or mechanistic model simulations to drive the optimization. The type of surrogate models used in such data-driven MPC includes MLPs [85, 86] and GPs [87, 88]. However, 530 comparatively little work has been published on embedding hybrid models into MPC to reduce data dependency and infuse physical knowledge for better extrapolation capability [89, 90]. Teixeira et al. [91] applied batch-to-batch optimization to bioprocesses by relying on hybrid models where an adjustable mixture of nonparametric and parametric models represented the 535 cell population subsystem. In the RTO area, Cubillos et al. [92] investigated the use of parallel hybrid models with MLP embedded on the Williams benchmark plant, but then they had to use stochastic search methods to solve the resulting optimization problems. Recently, Zhang et al. [90] took the extra 540 step of using the same hybrid model simultaneously in the RTO and MPC layers and demonstrated the benefits for a simulated CSTR and distillation column. Notice that most of these applications consider serial hybrid models with embedded MLPs to approximate complex nonlinearities in the system. Nevertheless, there is a dearth of industrial or experimental implementations 545 of such technologies to date.

An RTO methodology that exploits the parallel approach of hybrid semi-parametric modeling at its core is *modifier adaptation* [93]. Unlike classical RTO, modifier adaptation does not adapt the mechanistic model but adds correction terms—the modifiers—to the cost and constraint functions 550 in the optimization model. The original work used process measurements

to estimate linear (gradient-based) corrections [94]. Gao et al. [95] proposed combining quadratic regression models trained on available plant data with a nominal mechanistic model to account for curvature information and filter out the process noise. Likewise, Singhal et al. [96] investigated data-driven approaches based on quadratic surrogates as modifiers for the predicted cost and constraint functions and devised an online adaptation strategy for the surrogates inspired by trust-region ideas. Implementations of this RTO methodology for industrial systems include load sharing for gas compressors [97] and solid-oxide fuel cells [98].

More recently, Ferreira et al. [99] were the first to consider GPs, trained from past measurement information, as the cost and constraint modifiers. Using nonparametric regression models to describe the plant-model mismatch in RTO applications makes sense insofar as the mismatch is generally structural. Del Rio Chanona et al. [100, 101] developed this strategy further by introducing modifier-adaptation schemes that rely on trust regions to capture the GPs' ability to capture the cost and constraint mismatch. Recently, Petsagkourakis et al. [102] proposed to use co-Kriging to drive the surrogate modeling, where a first (low-fidelity) GP emulating the mechanistic process model is integrated within a second (high-fidelity) GP that is trained using the process measurements. The benefits of using GPs in this context lie in their ability to perform real-time uncertainty quantification and allow chance constraints to be satisfied with high confidence. By and large, these developments share many common grounds with surrogate-based optimization techniques (see Section 4.1), with the added complexity that the process data are noisy and the process optimum might change over time. Finally, it

is worth noting that the potential benefits of this RTO technology have been mostly investigated through numerical simulation, which cannot substitute for both experimental and industrial validations and should be the subject of future research.

580 *4.3. Reinforcement learning*

Reinforcement learning (RL) is a class of numerical methods for the data-driven sequential decision-making problem [103]. The RL agent (algorithm) aims to find an optimal policy, or controller, based on industrial process data collected through interactions with its environment.

585 Note that RL represents a more general class of techniques from hybrid modeling-based optimization. Briefly, RL includes algorithms for synthesizing control policies without explicit reliance on a model of the process dynamics. The supplementary material contains a more precise background on RL; readers are also referred to Sutton and Barto [103].

590 Finding such a policy requires solving the Bellman equation based on the principle of optimality. However, the equation is often intractable as it ends up with a high-dimensional optimization problem [104]. Recent advances in machine learning (ML) enable feature analysis of raw sensory-level using deep neural networks (DNNs). The aid of DNNs facilitates efficient
595 numerical methods for approximately solving the Bellman equation. Therefore, the scalability of RL algorithms has been significantly improved. As a result, so-called deep RL is an emerging technology that has shown remarkable performance in real-world and simulated applications such as robotics, autonomous driving, and board games [105, 106, 107].

600 Deep RL has naturally gained attention from the process control community. In this section, we survey applications of RL in process control, and we discuss advances and challenges in RL as they potentially pertain to process control applications.

4.3.1. Reinforcement learning for process control

605 With high demands on the performance of process systems, efficient optimization is becoming increasingly essential. The ultimate dream goal of any process control system is to develop a controller capable of attaining optimality in large-scale, nonlinear, and hybrid models with constraints, fast online calculation, and adaptation. This ideal controller should be amenable
610 to a closed-loop solution and robust to online disturbances.

Mathematical programming-based control, such as MPC and direct optimization, are popular because they adequately address many of these requirements. Sections 4.1 and 4.2 discuss the mathematical programming paradigm in more detail. RL has been studied in parallel because it has contrasting features compared to mathematical programming methods [108]. According to
615 the review and perspective studies of Shin et al. [5], Nian et al. [8], Spielberg et al. [6], Yoo et al. [109], the advantages of RL are that: First, a closed-loop state feedback policy can be obtained for generic stochastic control problems, while an open-loop solution is obtained through mathematical programming
620 approaches. Most of the computation is done offline by learning the policy through offline data or simulation. Assuming that the environment used for offline training is identical to that of the online implementation, the policy is optimal. Second, the mathematical programming formulation for stochastic control problems often becomes prohibitively large to be solved within

625 a decision interval. On the other hand, uncertainties are implicitly or explicitly quantified by the value or policy functions in RL approaches. The trained RL policy can be implemented with minimal online computation required. Third, RL is flexible to varying levels of system knowledge, including model-free, partial model-free, and model-based RL. Table 6 summarizes the
630 comparison between RL and mathematical programming methods.

Several pioneering pieces of work due to Wilson and Martinez [110], Kaisare et al. [111], Peroni et al. [112] proposed applying model-free RL to process control problems over discretized state and action spaces. Q -learning was implemented for the tracking control of a fed-batch bioreactor
635 [110] and free-end maximization problem of a fed-batch bioreactor [111, 112]. Lee et al. [113], Lee and Lee [114] extended the concept of applying model-free RL to dual adaptive control and scheduling problems. It was shown that the approximation of the value function could provide robust control despite the presence of process noise and model changes. RL methods that
640 guarantee robustness in dynamic optimization were later studied in Nosair et al. [115], Yang and Lee [116].

Some recent applications of RL rely on a linear approximator to solve optimal control problems with a continuous state space model [117, 118, 119, 120]. Especially, Zhu et al. [117] applied a model-free RL variant called factorial
645 fast-food dynamic policy programming to a Vinyl Acetate monomer process. The algorithm improves scalability by breaking down the exponential size of the action space by action space factorization. In the meantime, model-free deep RL applications have become increasingly studied in the process control field. Table 7 summarizes some recent work in this area. In the remaining

Table 6: A comparison of RL and mathematical programming.

	Reinforcement learning	Mathematical programming approaches
Model knowledge	Flexible	Full model
Feedback	Trained policy function	Solution of optimization problems
Online computation	Negligible	High
Offline computation	High	Not required
Robustness	Backward propagation of uncertain scenario (value-based methods)	Forward propagation of uncertain scenario
Constraint handling	Immature (especially, state variable constraints)	Straightforward
Asymptotic stability	Ultimate upper-boundedness	Asymptotically stable
Scalability	High	Medium
Adaptation	Exploitation and exploration can be controlled. However, slow.	Fast. However, performance depends on estimators.

650 sections, we elaborate on the use of deep RL in process control.

4.3.2. Practical implementation of reinforcement learning

One promising application of RL is the synthesis of existing control structures [142, 143, 144, 145, 146]. For example, proportional-integral-derivative (PID) controllers constitute the lowest level of control structures, and augmenting these with RL methods immediately gives practical results. PID tuning is a suitable testbed for RL applications, as there exists a suite of tuning methods and industrial autotuners to benchmark against [137]. PID controllers are also standard in practice, meaning the base layer control is not substituted for a more complex strategy, for example, based on DNNs (see Figure 5).

Model-free RL was applied to schedule a set of PID gains obtained *a priori* [147] or from internal model control [148]. Berger and da Fonseca Neto [149] used a model-based RL method, called dual heuristic dynamic programming, to compute PID gains. Nian et al. [8] applied deep Q -network (DQN) to determine the gains of PID controllers and compared the performance with MPC. Lawrence et al. [137] conducted an experimental study on the auto-tuning of PID controllers using the twin-delayed DDPG (TD3) algorithm. Figure 5 depicts a feedback diagram in the RL setting: the actor is formulated as a PID controller for the flow rate to a two-tank system, while the agent processes data on a PC to update the actor-critic parameters.

Another application is to construct hierarchical control structures with RL methods. Shafi et al. [150] introduced a two-layer structure for optimizing the bitumen recovery rate of a primary separation vessel. A supervisory RL agent optimizes the recovery rate, while a low-level RL agent computes the

Table 7: Model-free deep RL applications in process control. Asterisk (*) indicates a model-based modification to the nominal algorithm. Highlighted rows indicate validation on a physical system.

	RL algorithm	Application/process
Pandian and Noel [121]	DQN*[122]	Quadruple tank system
Wang et al. [123]	PPO[124]	HVAC control
Ma et al. [125]	DDPG[126]	Polymerization system
Spielberg et al. [6]	DDPG[126]	HVAC control
Oh et al. [127]	DQN[122]	Moving bed process
Petsagkourakis et al. [128]	REINFORCE*[129]	Fed-batch bioreactor
Bao et al. [130]	TD3*[131]	Setpoint tracking
Dogru et al. [132]	A3C[133]	Hybrid three-tank system
Joshi et al. [134]	TD3[131]	Transesterification process
Mowbray et al. [135]	REINFORCE[129]	Setpoint tracking
Yoo et al. [136]	DDPG[126]	Polymerization process
Lawrence et al. [137]	TD3[131]	PID tuning
Zhu et al. [138]	SAC[139]	Polyol process
Janjua et al. [140]	GVF[141]	Water treatment

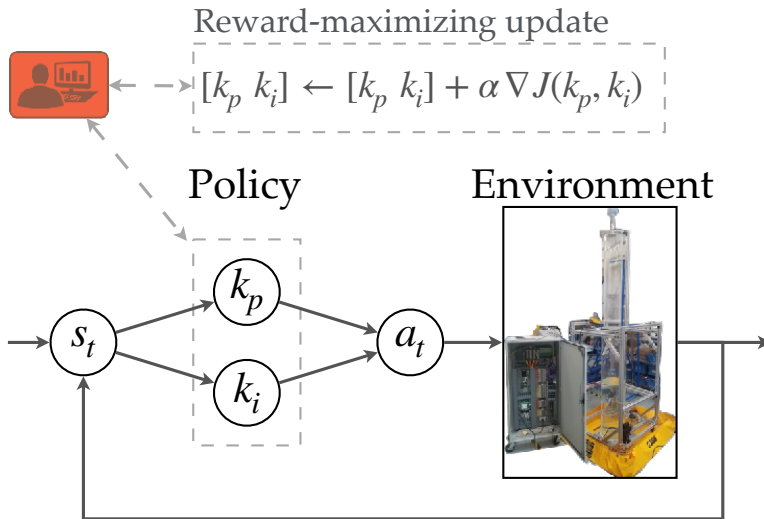


Figure 5: Application of RL for tuning PI controllers in a lab setting. The policy plays the role of a PI controller and receives updates towards improved performance. J is a general long-term cost function and k_p, k_i are controller gains. Adapted from [137].

675 interface level actuation. Kim et al. [151] proposed a different type of two-layer structure for a product maximization problem of a fed-batch bioreactor. A model-based RL agent solves the high-level optimization problem, and an MPC tracks the trajectory of the high-level optimizer, rejecting real-time disturbances.

680 Several studies make a comparison between RL methods based on practical performance criteria. Wang et al. [152] compared 14 model-free and model-based RL algorithms based on the following criteria: nominal performance, sample efficiency (total training time, training time per step), robustness against noise, and asymptotic performance. Lawrence et al. [137] 685 proposed nominal performance, stability, perturbation to the system, initialization, hyperparameters, training duration, practicality, and specialization as key criteria for evaluating RL methods for process control problems. In

addition, Dogru et al. [132] used the extent of exploration: the ratio of the visited over the total operational state and action spaces.

690 It is worth noting that RL implementations on physical systems are sparse. Some works in process control applications are validated on physical systems [148, 153, 121, 8, 137, 132]. These references tend to focus on PID tuning or low-dimensional state/action spaces. A cascaded tank system is also the most common environment. There are several plausible reasons
695 for the lack of real-world RL applications: The added engineering and software development is not always feasible to accommodate; the algorithmic complexity of RL algorithms exacerbates the issue; practical and theoretical problems, such as sample efficiency, convergence, and closed-loop stability, are pressing concerns. Indeed, most deep RL algorithms can achieve impressive final performance on complex tasks, but at the cost of extensive hyper-
700 parameter tuning and significant variation between implementations [154]. In the following section, we highlight a few methods that are geared towards making RL more reliable and scalable: Synthesis between model-based and model-free learning; transfer learning and meta-RL; offline RL.

705 4.3.3. *Challenges and advances in deep reinforcement learning*

Applying RL to industrial settings has many practical, technological, and theoretical challenges. We refer to Shin et al. [5], Nian et al. [8] for further reading. Here, we mainly focus on the sample efficiency of RL algorithms. Sample efficiency refers to the amount of data needed to train an RL agent.
710 The supplementary material contains a more general discussion about ML with limited data.

Classical algorithms for value-based methods, such as Q -learning, and

policy-based methods, such as REINFORCE⁵, enjoy theoretical convergence. However, convergence can be slow due to high variance in value estimates
715 or limited to the tabular setting or linear function approximation [103]. Nonetheless, these methods provide the foundation for deep RL algorithms. Deep RL attempts to scale up RL methods to high-dimensional problems as a synthesis with the deep learning framework. The first notable result is an extension of Q -learning, named DQNs, introduced by Mnih et al. [122].
720 DQNs are limited to discrete action spaces but showed impressive results in tasks with high-dimensional sensory input data, such as Atari games.

More recent algorithms, such as the deep deterministic policy gradient (DDPG) algorithm [126], allow for continuous action spaces. Despite the advances made by DDPG, it is notoriously difficult to use, for example, due to
725 sensitivity to hyperparameters and overestimation of Q -function values [154]. This limits the viability of DDPG for real-world applications such as process control, as a physical system cannot be extensively probed. However, the concurrent algorithms, TD3 [131] and soft actor-critic [139], built off DDPG to improve the overall training robustness and sample efficiency. Despite
730 these advances, model-free RL algorithms alone are not sufficiently data-efficient and, therefore, not yet useful in real industrial applications [155]. In the rest of this section, we identify several areas of RL research aimed at this issue.

Although formulating a dynamic model can be a bottleneck in the RL al-
735 gorithm, model-based methods require much fewer interactions with the plant

⁵*RE*ward Increment = *N*onnegative *F*actor \times *O*ffset *R*einforcement \times *C*haracteristic *E*ligibility

[155]. Several model-based RL algorithms have been developed, focusing on solving the continuous-time counterpart of the Bellman equation called the Hamilton-Jacobi-Bellman (HJB) equation. Since they aim to solve the HJB equation adaptively, the methods are called approximate dynamic programming (ADP) [156, 157, 158]. ADP algorithms vary with their levels of model
740 utilization, ranging from heuristic dynamic programming, dual heuristic programming, and globalized dual heuristic programming [159, 160]. Stochastic optimal control is an extension for handling stochastic differential equations, a continuous-time description for uncertainty. Policy improvement with path
745 integrals (PI²) is a sampling approach to solving the stochastic HJB equation [161]. PI² has shown remarkable data efficiency and performance for robot learning.

Another line of work has focused on unifying model-free and model-based approaches [162, 163]. The main motivation is that model-free algorithms
750 often achieve superior final (asymptotic) performance over model-based approaches but suffer from relatively weak sample complexity. Bao et al. [130] utilized ideas from D’Oro and Jaśkowski [162] wherein a dynamics model is used to improve the action gradient estimation of the critic network. While integrating dynamic models into traditionally model-free algorithms
755 has proved promising, these algorithms are designed to train an agent using online interactions on a system-by-system basis. More general strategies aim to reduce the cost of calibrating RL agents to novel environments by utilizing historical datasets, training over many related systems, or transferring previously trained agents to new ones.

760 *Offline RL* (sometimes called *batch RL*) aims to learn an optimal policy

from historical data alone [164]. Although off-policy algorithms like DDPG can theoretically learn from historical data, online exploration is critical unless constraints are imposed on the learned policy [165]. An offline strategy for pre-training RL agents with historical process data, followed by online
765 fine-tuning of the policy, is proposed by Mowbray et al. [135]. On the other hand, *transfer learning* is a framework for speeding up the training of RL agents. By pre-training a policy, such as in a simulation environment, one can use this as the initial policy on the true system of interest. This idea is demonstrated for batch bioprocess optimization [128]. One can efficiently
770 mitigate plant-model mismatch by fine-tuning the initial policy on the real system.

Meta-learning, or *learning to learn*, is a ML strategy for leveraging prior training experience to learn a new “task” quickly [166]. *Meta-RL* is a strategy for training a “meta agent” to synthesize experience from many related
775 systems to adapt its policy to novel systems rapidly. For example, Finn et al. [167] develop a simple and highly influential algorithm for any neural network architecture that directly optimizes for initial parameters such that they can quickly be adapted to new tasks with a small amount of data, showing superior performance over standard transfer learning in classifica-
780 tion and RL tasks. Duan et al. [168] propose strategies for learning a latent context variable as part of the meta-policy architecture, thereby capturing the “task” structure and enabling the meta-RL agent to adapt its policy with new process data. This framework is appealing in process control applica-
785 tions because many systems may have a known structure, making training over a distribution of related systems feasible. Consequently, this end-to-end

framework removes a model identification step during the online implementation of the RL agent by leveraging prior training experience. Meta-RL has also seen recent applications to process control [169].

While significant strides have been made to make these algorithms more sample-efficient, they are not yet practical. Motivated by this challenge, we have outlined different ways in which models can be integrated into otherwise model-free algorithms. Moreover, meta-RL, offline RL, and transfer learning, while still emerging, are promising avenues for MPC applications. These areas have tremendous potential for applications that can redefine automation in the process industries.

5. Discussion

Soft sensing and process control encompass statistical learning, machine learning, deep learning, and reinforcement learning to varying degrees. Table 8 shows the respective high-level prominence in these two application areas. Although Table 2 indicates significant interest in the soft sensing literature around deep learning, Table 5 shows methods like PLS and SVM have received the most industrial use. However, the prominent use of industrial

Table 8: Method-application pairs covered in this survey. ✓: significant emphasis, ✗: sparse emphasis.

	Soft sensing	Process control
Statistical learning	✓	✗
Machine learning	✓	✓
Deep learning	✓	✓
Reinforcement learning	✗	✓

data is still promising. Meanwhile, our survey of process control indicates a more significant emphasis on deep learning and reinforcement learning in the literature. Simulation-based studies are commonplace in this context, as discussed in Section 4.3.2.⁶

Table 8 and the above discussion show a duality between sensing and control in the context of machine learning methods. To fully capture the benefits of modern machine learning methods, a unified framework that encompasses modeling, sensing, and control is required. Reinforcement learning is well-suited to bridge the gap between sensing and control through a global reward-based objective (rather than treating prediction and control performance as independent goals). Applications in sensing do not necessarily contradict the model-free nature of reinforcement learning, which is most appealing. Rather, this characteristic makes it versatile for processing and optimizing real system data. To illustrate this point, Xie et al. [172] propose using reinforcement learning for sensing, even though it has typically been described in the context of control. Moreover, Esfahani et al. [173] utilize reinforcement learning for both state estimation and control under a single closed-loop performance objective.

On the other hand, Section 4.3 discussed the complexity of reinforcement learning algorithms. More broadly, deep learning and reinforcement learning algorithms are rife with complexity and hyperparameters, making it difficult to parse their fundamental inner workings [154, 174]. A promising

⁶Of course, existing implementations of MPC utilize soft sensors and, therefore, statistical/machine learning methods. It is also commonplace in academic studies to augment mathematical programming or reinforcement learning with autoencoders or RNNs to learn a state representation [170, 171].

825 avenue toward unifying sensing and control is distilling reinforcement learn-
ing pipelines and reimagining techniques from other branches of machine
learning. Truly robust and powerful methods will follow from such a critical
rapprochement of the longstanding statistical learning methods in Table 1
and newer concepts in deep learning and reinforcement learning. An instance
830 of this aspiration in action is by Eysenbach et al. [175], where they show a
novel use of binary classification and policy iteration is capable of achieving
state-of-the-art performance.

5.1. Conclusions

Recent advances in machine learning give us renewed optimism for achiev-
835 ing higher levels of automation in the process industries. To distill this broad
aspiration, we have surveyed soft sensing and process control through a prac-
tical lens. Soft sensing represents the most dominant area regarding indus-
trial applications of statistical and machine learning techniques. On the other
hand, considerable research attention has been given to deep learning appli-
840 cations, but with limited industrial successes. Through synthesizing research
trends and industrial requirements, we have strived to enable academics and
practitioners alike to develop sophisticated yet practical methods for building
better models and controllers.

Acknowledgements

845 We are grateful to the anonymous reviewers for their detailed and con-
structive feedback. NPL & RBG gratefully acknowledge the financial sup-
port of the Natural Sciences and Engineering Research Council of Canada
(NSERC) and Honeywell Process Solutions. JML gratefully acknowledges

the research facilities for this work provided by the Institute of Engineering
850 Research at Seoul National University. BC gratefully acknowledges funding
by the Engineering and Physical Sciences Research Council (EPSRC) un-
der grants EP/T000414/1 and EP/W003317/1. BH, FA and SKD gratefully
acknowledge financial supports from the Natural Sciences and Engineering
Research Council of Canada (NSERC) under grants IRCPJ 417793-15 and
855 ALLRP 561080-20.

References

- [1] Venkat Venkatasubramanian. The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE Journal*, 65(2):466–478, 2019.
- [2] Leo H Chiang, Evan L Russell, and Richard D Braatz. *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media, 2000.
860
- [3] S Joe Qin and Leo H Chiang. Advances and opportunities in machine learning for process data analytics. *Computers & Chemical Engineering*, 126:465–473, 2019.
- [4] Zhiqiang Ge, Zhihuan Song, Steven X Ding, and Biao Huang. Data mining and analytics in the process industry: The role of machine learning. *IEEE Access*, 5: 20590–20616, 2017.
865
- [5] Joohyun Shin, Thomas A Badgwell, Kuang-Hung Liu, and Jay H Lee. Reinforcement Learning—Overview of recent progress and implications for process control. *Computers & Chemical Engineering*, 127:282–294, 2019.
- [6] Steven Spielberg, Aditya Tulsyan, Nathan P. Lawrence, Philip D. Loewen, and R. Bhushan Gopaluni. Toward self-driving processes: A deep reinforcement learning approach to control. *AIChE Journal*, 65, 2019.
870
- [7] Aditya Tulsyan, Tony Wang, Gregg Schorner, Hamid Khodabandehlou, Myra Cofal, and Cenk Undey. Automatic real-time calibration, assessment, and maintenance of generic Raman models for online monitoring of cell culture processes. *Biotechnology and Bioengineering*, 117(2):404–416, 2020.
875
- [8] Rui Nian, Jinfeng Liu, and Biao Huang. A review on reinforcement learning: Introduction and applications in industrial process control. *Computers & Chemical Engineering*, page 106886, 2020.

- 880 [9] Xiaotian Bi, Ruoshi Qin, Deyang Wu, Shaodong Zheng, and Jinsong Zhao. One step forward for smart chemical process fault detection and diagnosis. *Computers & Chemical Engineering*, 164:107884, 2022.
- [10] Thomas Gamer, Mario Hoernicke, Benjamin Kloepper, Reinhard Bauer, and Alf J Isaksson. The autonomous industrial plant—future of process engineering, operations and maintenance. *Journal of Process Control*, 88:101–110, 2020.
- 885 [11] R. Bhushan Gopaluni, Aditya Tulsyan, Benoit Chachuat, Biao Huang, Jong Min Lee, Faraz Amjad, Seshu Kumar Damarla, Jong Woo Kim, and Nathan P. Lawrence. Modern machine learning tools for monitoring and control of industrial processes: A survey. *IFAC-PapersOnLine*, 53(2):218–229, 2020.
- 890 [12] Xiaotian Bi, Ruoshi Qin, Deyang Wu, Shaodong Zheng, and Jinsong Zhao. One step forward for smart chemical process fault detection and diagnosis. *Computers & Chemical Engineering*, 164:107884, 2022.
- [13] Joel Sansana, Mark N Joswiak, Ivan Castillo, Zhenyu Wang, Ricardo Rendall, Leo H Chiang, and Marco S Reis. Recent trends on hybrid modeling for industry 4.0. *Computers & Chemical Engineering*, 151:107365, 2021.
- 895 [14] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer Series in Statistics, New York, 2nd edition, 2009.
- [15] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.
- 900 [16] Moritz von Stosch, Rui Oliveira, Joana Peres, and Sebastião Feyeo de Azevedo. Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Computers & Chemical Engineering*, 60:86 – 101, 2014.
- [17] Benjamin Peherstorfer, Karen Willcox, and Max Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60(3):550–591, 2018.
- 905 [18] Dimitris C. Psychogios and Lyle H. Ungar. A hybrid neural network-first principles approach to process modeling. *AIChE Journal*, 38(10):1499–1511, 1992.
- [19] Hong-Te Su, N. Bhat, P. A. Minderman, and T. J. McAvoy. Integrating neural networks with first principles models for dynamic modeling. *IFAC Proceedings Volumes*, 25(5):327–332, 1992.
- 910 [20] Michael L. Thompson and Mark A. Kramer. Modeling chemical processes using prior knowledge and neural networks. *AIChE Journal*, 40(8):1328–1340, 1994.

- 915 [21] L. Chen, O. Bernard, G. Bastin, and P. Angelov. Hybrid modelling of biotechnological processes using neural networks. *Control Engineering Practice*, 8(7):821–827, 2000.
- [22] Dörte Solle, Bernd Hitzmann, Christoph Herwig, Manuel Pereira Remelhe, Sophia Ulonska, Lynn Wuerth, Adrian Prata, and Thomas Steckenreiter. Between the poles of data-driven and mechanistic modeling for process operation. *Chemie Ingenieur Technik*, 89(5):542–561, 2017.
- 920 [23] A. Schuppert and T. Mrziglod. Hybrid model identification and discrimination with practical examples from the chemical industry. In J. Glassey and M. von Stosch, editors, *Hybrid Modeling in Process Industries*, pages 63–88, Boca Raton, 2018. CRC Press.
- 925 [24] Sohrab Zendehboudi, Nima Rezaei, and Ali Lohi. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. *Applied Energy*, 228:2539–2566, 2018.
- [25] I. Ahmad, A. Ayub, M. Kano, and I.I. Cheema. Gray-box soft sensors in process industry: Current practice, and future prospects in era of big data. *Processes*, 8(2):243, 2020.
- 930 [26] William Bradley, Jinhyeun Kim, Zachary Kilwein, Logan Blakely, Michael Eydenberg, Jordan Jalvin, Carl Laird, and Fani Boukouvala. Perspectives on the integration between first-principles and data-driven modeling. *Computers & Chemical Engineering*, 166:107898, 2022.
- 935 [27] Mukul Agarwal. Combining neural and conventional paradigms for modelling, prediction and control. *International Journal of Systems Science*, 28(1):65–81, 1997.
- [28] Mohammed Saad Faizan Bangi and Joseph Sang-Il Kwon. Deep hybrid modeling of chemical process: Application to hydraulic fracturing. *Computers & Chemical Engineering*, 134:106696, 2020.
- 940 [29] Donovan Chaffart and Luis A. Ricardez-Sandoval. Optimization and control of a thin film growth process: A hybrid first principles/artificial neural network based multiscale modelling approach. *Computers & Chemical Engineering*, 119:465–479, 2018.
- 945 [30] D. Ghosh, E. Hermonat, P. Mhaskar, S. Snowling, and R. Goel. Hybrid modeling approach integrating first-principles models with subspace identification. *Industrial & Engineering Chemistry Research*, 58(30):13533–13543, 2019.
- 950 [31] Pau Cabaneros Lopez, Isuru A. Udugama, Sune T. Thomsen, Christian Roslander, Helena Junicke, Miguel Mauricio-Iglesias, and Krist V. Gernaey. Towards a digital twin: a hybrid data-driven and mechanistic digital shadow to forecast the evolution of lignocellulosic fermentation. *Biofuels, Bioproducts & Biorefining*, 14(5):1046–1060, 2020.

- [32] Dongda Zhang, Ehecatl Antonio Del Rio-Chanona, Panagiotis Petsagkourakis, and Jonathan Wagner. Hybrid physics-based and data-driven modeling for bioprocess online simulation and optimization. *Biotechnology & Bioengineering*, 116(11):2919–2930, 2019.
- 955 [33] Belmiro Duarte, P. M. Saraiva, and C. C. Pantelides. Combined mechanistic and empirical modelling. *International Journal of Chemical Reactor Engineering*, 2(1), 2004.
- [34] C. de Prada, D. Hose, G. Gutierrez, and J. L. Pitarch. Developing grey-box dynamic process models. *IFAC-PapersOnLine*, 51(2):523–528, 2018.
- 960 [35] Tim Hesterberg, Nam Hee Choi, Lukas Meier, and Chris Fraley. Least angle and ℓ_1 penalized regression: A review. *Statistics Surveys*, 2:61–93, 2008.
- [36] Mark J. Willis and Moritz von Stosch. Simultaneous parameter identification and discrimination of the nonparametric structure of hybrid semi-parametric models. *Computers & Chemical Engineering*, 104:366–376, 2017.
- 965 [37] Dongda Zhang, Thomas R. Savage, and Bovinille A. Cho. Combining model structure identification and hybrid modelling for photo-production process predictive simulation and optimisation. *Biotechnology & Bioengineering*, 117(11):3356–3367, 2020.
- [38] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016.
- 970 [39] Zachary T. Wilson and Nikolaos V. Sahinidis. The ALAMO approach to machine learning. *Computers & Chemical Engineering*, 106:785–795, 2017.
- [40] Alison Cozad, Nikolaos V. Sahinidis, and David C. Miller. A combined first-principles and data-driven approach to model building. *Computers & Chemical Engineering*, 73:116–127, 2015.
- 975 [41] K. M. Nauta, S. Weiland, A. C. Backx, and A. Jokic. Approximation of fast dynamics in kinetic networks using non-negative polynomials. In *2007 IEEE International Conference on Control Applications*, pages 1144–1149, 2007.
- 980 [42] J. L. Pitarch, A. Sala, and C. de Prada. A systematic grey-box modeling methodology via data reconciliation and SOS constrained regression. *Processes*, 7(3):170, 2016.
- [43] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686 – 707, 2019.
- 985

- [44] G. F. Carey and B. A. Finlayson. Orthogonal collocation on finite elements. *Chemical Engineering Science*, 30(5):587–596, 1975.
- 990 [45] Lorenz T Biegler. An overview of simultaneous strategies for dynamic optimization. *Chemical Engineering and Processing: Process Intensification*, 46(11):1043–1053, 2007.
- [46] H. Liu, J. Cai, and Y. S. Ong. Remarks on multi-output Gaussian process regression. *Knowledge-Based Systems*, 144:102–121, 2018.
- 995 [47] L. Le Gratiet and J. Garnier. Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4(5):365–386, 2014.
- [48] P. Perdikaris, M. Raissi, A. Damianou, N. D. Lawrence, and G. E. Karniadakis. Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2198), 2017.
- 1000 [49] K. Cutajar, M. Pullin, A. Damianou, N. Lawrence, and J. González. Deep Gaussian processes for multi-fidelity modeling. *Advances in Neural Information Processing Systems*, 32, 2018.
- [50] Amir Ali Ahmadi and Bachir El Khadir. Learning dynamical systems with side information, 2020.
- 1005 [51] Shima Khatibisepehr, Biao Huang, and Swanand Khare. Design of inferential sensors in the process industry: A review of Bayesian methods. *Journal of Process Control*, 23(10):1575–1596, 2013.
- [52] Luigi Fortuna, Salvatore Graziani, Alessandro Rizzo, and Maria Gabriella Xibilia. *Soft sensors for monitoring and control of industrial processes*. Springer Science & Business Media, 2007.
- 1010 [53] Fei Chu, Xu Zhao, Yuan Yao, Tao Chen, and Fuli Wang. Transfer learning for batch process optimal control using LV-PTM and adaptive control strategy. *Journal of Process Control*, 81:197–208, 2019.
- [54] Petr Kadlec, Bogdan Gabrys, and Sibylle Strandt. Data-driven soft sensors in the process industry. *Computers & Chemical Engineering*, 33(4):795–814, 2009.
- 1015 [55] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.
- 1020 [56] Laurenz Wiskott and Terrence J Sejnowski. Slow feature analysis: Unsupervised learning of invariances. *Neural computation*, 14(4):715–770, 2002.

- [57] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430, 2000.
- 1025 [58] Zhiqiang Ge. Supervised latent factor analysis for process data regression modeling and soft sensor application. *IEEE Transactions on Control Systems Technology*, 24(3):1004–1011, 2015.
- [59] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.
- 1030 [60] Le Yao and Zhiqiang Ge. Scalable semisupervised GMM for big data quality prediction in multimode processes. *IEEE Transactions on Industrial Electronics*, 66(5):3681–3692, 2018.
- [61] John D Kelleher, Brian Mac Namee, and Aoife D’arcy. *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies*. MIT press, 2020.
- 1035 [62] Kai Wang, Bhushan Gopaluni, Junghui Chen, and Zhihuan Song. Deep learning of complex batch process data and its application on quality prediction. *IEEE Transactions on Industrial Informatics*, 2018.
- [63] S Joe Qin, Siyi Guo, Zheyu Li, Leo H Chiang, Ivan Castillo, Birgit Braun, and Zhenyu Wang. Integration of process knowledge and statistical learning for the Dow data challenge problem. *Computers & Chemical Engineering*, 153:107451, 2021.
- 1040 [64] Christophe Leys, Christophe Ley, Olivier Klein, Philippe Bernard, and Laurent Licata. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of experimental social psychology*, 49(4):764–766, 2013.
- 1045 [65] Yalin Wang Ling Li, Seshu Kumar Damarla and Biao Huang. A Gaussian mixture model based virtual sample generation approach for small datasets in industrial processes. *Information Sciences*, 581:262–277, 2021.
- [66] Zhihuan Song Weiming Shao and Le Yao. Soft sensor development for multimode processes based on semisupervised Gaussian mixture models. *IFAC PapersOnline*, 51:614–619, 2018.
- 1050 [67] Kuilin Chen, Ivan Castillo, Leo H Chiang, and Jie Yu. Soft sensor model maintenance: A case study in industrial processes. *IFAC-PapersOnLine*, 48(8):427–432, 2015.
- [68] Artur M. Schweidtmann and Alexander Mitsos. Deterministic global optimization with artificial neural networks embedded. *Journal of Optimization Theory & Applications*, 180(3):925–948, 2019.
- 1055

- [69] R. Anderson, J. Huchette, W. Ma, C. Tjandraatmadja, and J. P. Vielma. Strong mixed-integer programming formulations for trained neural networks. *Mathematical Programming*, 183:3–39, 2020.
- 1060 [70] C. Tsay, J. Kronqvist, A. Thebelt, and R. Misener. Partition-based formulations for mixed-integer optimization of trained ReLU neural networks. *Advances in Neural Information Processing Systems*, 34:3068–3080, 2021.
- [71] Artur M. Schweidtmann, Dominik Bongartz, Daniel Grothe, Tim Kerkenhoff, Xiaopeng Lin, Jaromil Najman, and Alexander Mitsos. Deterministic global optimization with Gaussian processes embedded. *Mathematical Programming Computation*, 13:553–581, 2021.
- 1065 [72] M. Mistry, D. Letsios, G. Krennrich, R. M. Lee, and R. Misener. Mixed-integer convex nonlinear optimization with gradient-boosted trees embedded. *INFORMS Journal on Computing*, 33(3):1103–1119, 2021.
- 1070 [73] Lorenz T. Biegler, Yi-dong Lang, and Weijie Lin. Multi-scale optimization for process systems engineering. *Computers & Chemical Engineering*, 60:17–30, 2014.
- [74] Carlos A. Henao and Christos T. Maravelias. Surrogate-based superstructure optimization framework. *AIChE Journal*, 57(5):1216–1232, 2011.
- [75] José A. Caballero and Ignacio E. Grossmann. An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE Journal*, 54(10):2633–2650, 2008.
- 1075 [76] Natalia Quirante, Juan Javaloyes, and José. Caballero. Rigorous design of distillation columns using surrogate models based on kriging interpolation. *AIChE Journal*, 61(7):2169–2187, 2015.
- [77] Tobias Keßler, Christian Kunde, Kevin McBride, Nick Mertens, Dennis Michaels, Kai Sundmacher, and Achim Kienle. Global optimization of distillation columns using explicit and implicit surrogate models. *Chemical Engineering Science*, 197:235–245, 2019.
- 1080 [78] Fani Boukouvala and Christodoulos A. Floudas. ARGONAUT: Algorithms for Global Optimization of coNstrained grey-box compUTational problems. *Optimization Letters*, 11(5):895–913, 2017.
- 1085 [79] Andrew R. Conn, Katya Scheinberg, and Luis N. Vicente. *Introduction to Derivative-Free Optimization*. MOS-SIAM Series on Optimization, 2009.
- [80] John P. Eason and Lorenz T. Biegler. Advanced trust region optimization strategies for glass box/black box models. *AIChE Journal*, 64(11):3934–3943, 2018.
- 1090 [81] Ishan Bajaj, Shachit S. Iyer, and M. M. Faruque Hasan. A trust region-based two phase algorithm for constrained black-box and grey-box optimization with infeasible initial point. *Computers & Chemical Engineering*, 116:306–321, 2018.

- 1095 [82] Max Schwenzer, Muzaffer Ay, Thomas Bergs, and Dirk Abel. Review on model predictive control: An engineering perspective. *The International Journal of Advanced Manufacturing Technology*, 117(5-6):1327–1349, 2021.
- [83] S Joe Qin and Thomas A Badgwell. An overview of nonlinear model predictive control applications. *Nonlinear model predictive control*, pages 369–392, 2000.
- 1100 [84] Michael G Forbes, Rohit S Patwardhan, Hamza Hamadah, and R Bhushan Gopaluni. Model predictive control in industry: Challenges and opportunities. *IFAC-PapersOnLine*, 48(8):531–538, 2015.
- [85] Stephen Piche, Bijan Sayyar-Rodsari, Doug Johnson, and Mark Gerules. Nonlinear model predictive control using neural networks. *IEEE Control Systems Magazine*, 20(3):53–62, 2000.
- 1105 [86] Zhe Wu, Anh Tran, David Rincon, and Panagiotis D. Christofides. Machine learning-based predictive control of nonlinear processes. Part I: Theory. *AIChE Journal*, 65(11):e16729, 2019.
- [87] Juš Kocijan, Roderick Murray-Smith, Carl Edward Rasmussen, and Agathe Girard. Gaussian process model based predictive control. In *Proceeding of American Control Conference*, volume 3, pages 2214–2219, 2004.
- 1110 [88] L. Hewing, J. Kabzan, and M. N. Zeilinger. Cautious model predictive control using Gaussian process regression. *IEEE Transactions on Control Systems Technology*, 28(6):2736–2743, 2020.
- [89] Casimir C. Klimasauskas. Hybrid modeling for robust nonlinear multivariable control. *ISA Transactions*, 37(4):291–297, 1998.
- 1115 [90] Zhihao Zhang, Zhe Wu, David Rincon, and Panagiotis D. Christofides. Real-time optimization and control of nonlinear processes using machine learning. *Mathematics*, 7(10):890, 2019. doi: 10.3390/math7100890.
- 1120 [91] Ana P. Teixeira, João J. Clemente, António E. Cunha, Manuel J. T. Carrondo, and Rui Oliveira. Bioprocess iterative batch-to-batch optimization based on hybrid parametric/nonparametric models. *Biotechnology Progress*, 22(1):247–258, 2006.
- [92] F. A. Cubillos, G. Acuña, and E.L. Lima. Real-time process optimization based on grey-box neural models. *Brazilian Journal of Chemical Engineering*, 24:433–443, 2007.
- 1125 [93] B. Chachuat, B. Srinivasan, and D. Bonvin. Adaptation strategies for real-time optimization. *Computers & Chemical Engineering*, 33(10):1557–1567, 2009.
- [94] A. Marchetti, B. Chachuat, and D. Bonvin. Modifier-adaptation methodology for real-time optimization. *Industrial & Engineering Chemistry Research*, 48(13):6022–6033, 2009.

- 1130 [95] Weihua Gao, Simon Wenzel, and Sebastian Engell. A reliable modifier-adaptation strategy for real-time optimization. *Computers & Chemical Engineering*, 91:318–328, 2016.
- [96] Martand Singhal, Alejandro G. Marchetti, Timm Faulwasser, and Dominique Bonvin. Real-time optimization based on adaptation of surrogate models. *IFAC-PapersOnLine*, 49(7):412–417, 2016.
- 1135 [97] P. Milosavljevic, A. G. Marchetti, A. Cortinovis, T. Faulwasser, M. Mercangöz, and D. Bonvin. Real-time optimization of load sharing for gas compressors in the presence of uncertainty. *Applied Energy*, 272:114883, 2020.
- 1140 [98] T. de Avila Ferreira, Z. Wuillemin, A.G. Marchetti, C. Salzmann, J. Van Herle, and D. Bonvin. Real-time optimization of an experimental solid-oxide fuel-cell system. *Journal of Power Sources*, 429:168–179, 2019.
- [99] T. d. A. Ferreira, H. A. Shukla, T. Faulwasser, C. N. Jones, and D. Bonvin. Real-time optimization of uncertain process systems via modifier adaptation and Gaussian processes. In *2018 European Control Conference (ECC)*, pages 465–470, 2018.
- 1145 [100] Ehecatl Antonio del Rio Chanona, JE Alves Graciano, Eric Bradford, and Benoit Chachuat. Modifier-adaptation schemes employing Gaussian processes and trust regions for real-time optimization. *IFAC-PapersOnLine*, 52(1):52–57, 2019.
- [101] Ehecatl Antonio del Rio Chanona, Panagiotis Petsagkourakis, Eric Bradford, JE Alves Graciano, and Benoît Chachuat. Real-time optimization meets Bayesian optimization and derivative-free optimization: A tale of modifier adaptation. *Computers & Chemical Engineering*, 147:107249, 2021.
- 1150 [102] P. Petsagkourakis, B. Chachuat, and E. A. del Rio-Chanona. Safe real-time optimization using multi-fidelity Gaussian processes. In *60th IEEE Conference on Decision and Control (CDC)*, pages 6734–6741, 2021.
- 1155 [103] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
- [104] Dimitri Bertsekas. *Dynamic programming and optimal control: Volume I*, volume 1. Athena scientific, 2012.
- 1160 [105] Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. End-to-end training of deep visuomotor policies. *The Journal of Machine Learning Research*, 17(1):1334–1373, 2016.
- [106] Grady Williams, Paul Drews, Brian Goldfain, James M Rehg, and Evangelos A Theodorou. Aggressive driving with model predictive path integral control. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1433–1440, 2016.

- 1165 [107] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of Go without human knowledge. *Nature*, 550(7676):354, 2017.
- [108] Lucian Buşoniu, Tim de Bruin, Domagoj Tolić, Jens Kober, and Ivana Palunko. Reinforcement learning for control: Performance, stability, and deep approximators. 1170 *Annual Reviews in Control*, pages 8–28, 2018.
- [109] Haeun Yoo, Ha Eun Byun, Dongho Han, and Jay H Lee. Reinforcement learning for batch process control: Review and perspectives. *Annual Reviews in Control*, 52: 108–119, 2021.
- 1175 [110] JA Wilson and EC Martinez. Neuro-fuzzy modeling and control of a batch process involving simultaneous reaction and distillation. *Computers & Chemical Engineering*, 21:S1233–S1238, 1997.
- [111] Niket S Kaisare, Jong Min Lee, and Jay H Lee. Simulation based strategy for nonlinear optimal control: Application to a microbial cell reactor. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 13(3-4):347–363, 1180 2003.
- [112] Catalina Valencia Peroni, Niket S Kaisare, and Jay H Lee. Optimal control of a fed-batch bioreactor using simulation-based approximate dynamic programming. *IEEE Transactions on Control Systems Technology*, 13(5):786–790, 2005.
- 1185 [113] Jong Min Lee, Niket S Kaisare, and Jay H Lee. Choice of approximator and design of penalty function for an approximate dynamic programming based control approach. *Journal of Process Control*, 16(2):135–156, 2006.
- [114] Jong Min Lee and Jay H Lee. An approximate dynamic programming based approach to dual adaptive control. *Journal of process control*, 19(5):859–864, 2009.
- 1190 [115] Hussam Nosair, Yu Yang, and Jong Min Lee. Min–max control using parametric approximate dynamic programming. *Control Engineering Practice*, 18(2):190–197, 2010.
- [116] Yu Yang and Jong Min Lee. A switching robust model predictive control approach for nonlinear systems. *Journal of Process Control*, 23(6):852–860, 2013.
- 1195 [117] Lingwei Zhu, Yunduan Cui, Go Takami, Hiroaki Kanokogi, and Takamitsu Matsubara. Scalable reinforcement learning for plant-wide control of vinyl acetate monomer process. *Control Engineering Practice*, 97:104331, 2020.
- [118] Bei Sun, Mingfang He, Yalin Wang, Weihua Gui, Chunhua Yang, and Quanmin Zhu. A data-driven optimal control approach for solution purification process. *Journal of Process Control*, 68:171–185, 2018.

- 1200 [119] Yulei Ge, Shurong Li, and Peng Chang. An approximate dynamic programming method for the optimal control of Alkali-Surfactant-Polymer flooding. *Journal of Process Control*, 64:15–26, 2018.
- [120] Jong Woo Kim, Go Bong Choi, and Jong Min Lee. A POMDP framework for integrated scheduling of infrastructure maintenance and inspection. *Computers & Chemical Engineering*, 112:239–252, 2018.
- 1205 [121] B Jaganatha Pandian and Mathew Mithra Noel. Control of a bioreactor using a new partially supervised reinforcement learning algorithm. *Journal of Process Control*, 69:16–29, 2018.
- [122] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, and Georg Ostrovski. Human-level control through deep reinforcement learning. *Nature*, 518 (7540):529, 2015.
- [123] Yuan Wang, Kirubakaran Velswamy, and Biao Huang. A novel approach to feedback control with deep reinforcement learning. *IFAC-PapersOnLine*, 51(18):31–36, 2018.
- 1215 [124] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.
- [125] Yan Ma, Wenbo Zhu, Michael G Benton, and José Romagnoli. Continuous control of a polymerization system with deep reinforcement learning. *Journal of Process Control*, 75:40–47, 2019.
- 1220 [126] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.
- [127] Tae Hoon Oh, Jong Woo Kim, Sang Hwan Son, Hosoo Kim, Kyungmoo Lee, and Jong Min Lee. Automatic control of simulated moving bed process with deep Q-network. *Journal of Chromatography A*, 1647:462073, 2021.
- 1225 [128] Panagiotis Petsagkourakis, Ilya Orson Sandoval, Eric Bradford, Dongda Zhang, and Ehecatl Antonio del Rio-Chanona. Reinforcement learning for batch bioprocess optimization. *Computers & Chemical Engineering*, 133:106649, 2020.
- [129] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8:229–256, 1992.
- 1230 [130] Yaoyao Bao, Yuanming Zhu, and Feng Qian. A Deep Reinforcement Learning Approach to Improve the Learning Performance in Process Control. *Industrial & Engineering Chemistry Research*, page acs.iecr.0c05678, 2021.

- 1235 [131] Scott Fujimoto, Herke van Hoof, and David Meger. Addressing function approximation error in actor-critic methods. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1587–1596. PMLR, 10–15 Jul 2018.
- 1240 [132] Oguzhan Dogru, Nathan Wiecezorek, Kirubakaran Velswamy, Fadi Ibrahim, and Biao Huang. Online reinforcement learning for a continuous space system with experimental validation. *Journal of Process Control*, 104:86–100, 2021.
- [133] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *International conference on machine learning*, pages 1928–1937. PMLR, 2016.
- 1245 [134] Tanuja Joshi, Shikhar Makker, Hariprasad Kodamana, and Harikumar Kandath. Twin actor twin delayed deep deterministic policy gradient (TATD3) learning for batch process control. *Computers & Chemical Engineering*, 155:107527, 2021.
- 1250 [135] Max Mowbray, Robin Smith, Ehecatl A Del Rio-Chanona, and Dongda Zhang. Using process data to generate an optimal control policy via apprenticeship and reinforcement learning. *AIChE Journal*, page e17306, 2021.
- [136] Haeun Yoo, Boeun Kim, Jong Woo Kim, and Jay H Lee. Reinforcement learning based optimal control of batch processes using Monte-Carlo deep deterministic policy gradient with phase segmentation. *Computers & Chemical Engineering*, 144:107133, 2021.
- 1255 [137] Nathan P. Lawrence, Michael G. Forbes, Philip D. Loewen, Daniel G. McClement, Johan U. Backström, and R. Bhushan Gopaluni. Deep reinforcement learning with shallow controllers: An experimental application to PID tuning. *Control Engineering Practice*, 121:105046, 2022.
- 1260 [138] Wenbo Zhu, Ivan Castillo, Zhenyu Wang, Ricardo Rendall, Leo H Chiang, Philippe Hayot, and Jose A Romagnoli. Benchmark study of reinforcement learning in controlling and optimizing batch processes. *Journal of Advanced Manufacturing and Processing*, 4(2):e10113, 2022.
- 1265 [139] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1861–1870. PMLR, 10–15 Jul 2018.
- 1270 [140] Muhammad Kamran Janjua, Haseeb Shah, Martha White, Erfan Miahi, Marlos C Machado, and Adam White. Gvfs in the real world: making predictions online for water treatment. *Machine Learning*, pages 1–31, 2023.

- 1275 [141] Richard S Sutton, Joseph Modayil, Michael Delp, Thomas Degris, Patrick M Pilarski, Adam White, and Doina Precup. Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 761–768, 2011.
- [142] M Sedighzadeh and A Rezazadeh. Adaptive PID controller based on reinforcement learning for wind turbine control. In *Proceedings of World Academy of Science, Engineering and Technology*, volume 27, pages 257–262. Citeseer, 2008.
- 1280 [143] Ignacio Carlucho, Mariano De Paula, Sebastian A. Villar, and Gerardo G. Acosta. Incremental Q-learning strategy for adaptive PID control of mobile robots. *Expert Systems with Applications*, 80:183–199, 2017.
- [144] William J. Shipman and Loutjie C. Coetzee. Reinforcement learning and deep neural networks for PI controller tuning. *IFAC-PapersOnLine*, 52(14):111–116, 2019.
- 1285 [145] Athindran Ramesh Kumar and Peter J. Ramadge. DiffLoop: Tuning PID controllers by differentiating through the feedback loop. In *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6, 2021.
- [146] Ayub I. Lakhani, Myisha A. Chowdhury, and Qiugang Lu. Stability-preserving automatic tuning of PID control with reinforcement learning. *Complex Engineering Systems*, 2(1):3, 2022.
- 1290 [147] Jay H Lee and Jong Min Lee. Approximate dynamic programming based approach to process control and scheduling. *Computers & Chemical Engineering*, 30(10-12):1603–1618, 2006.
- [148] Lena Abbasi Brujeni, Jong Min Lee, and Sirish L Shah. *Dynamic tuning of PI-controllers based on model-free reinforcement learning methods*. IEEE, 2010.
- 1295 [149] Marcus AR Berger and JoÃo Viana da Fonseca Neto. Neurodynamic programming approach for the PID controller adaptation. *IFAC Proceedings Volumes*, 46(11):534–539, 2013.
- 1300 [150] Hareem Shafi, Kirubakaran Velswamy, Fadi Ibrahim, and Biao Huang. A Hierarchical Constrained Reinforcement Learning for Optimization of Bitumen Recovery Rate in a Primary Separation Vessel. *Computers & Chemical Engineering*, page 106939, 2020.
- [151] Jong Woo Kim, Byung Jun Park, Tae Hoon Oh, and Jong Min Lee. Model-based reinforcement learning and predictive control for two-stage optimal control of fed-batch bioreactor. *Computers & Chemical Engineering*, 154:107465, 2021.
- 1305 [152] Tingwu Wang, Xuchan Bao, Ignasi Clavera, Jerrick Hoang, Yeming Wen, Eric Langlois, Shunshi Zhang, Guodong Zhang, Pieter Abbeel, and Jimmy Ba. Benchmarking model-based reinforcement learning. *arXiv preprint arXiv:1907.02057*, 2019.

- 1310 [153] S. Syafie, F. Tadeo, E. Martinez, and T. Alvarez. Model-free control based on reinforcement learning for a wastewater treatment problem. *Applied Soft Computing*, 11(1):73–82, 2011.
- [154] Peter Henderson, Riashat Islam, Philip Bachman, Joelle Pineau, and David Meger. Deep reinforcement learning that matters. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
- 1315 [155] Benjamin Recht. A tour of reinforcement learning: The view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, 2:253–279, 2019.
- [156] Danil V Prokhorov and Donald C Wunsch. Adaptive critic designs. *IEEE transactions on Neural Networks*, 8(5):997–1007, 1997.
- 1320 [157] Frank L Lewis and Draguna Vrabie. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits and Systems Magazine*, 9(3), 2009.
- [158] Yu Jiang and Zhong-Ping Jiang. Robust adaptive dynamic programming and feedback stabilization of nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 25(5):882–893, 2014.
- 1325 [159] Jong Woo Kim, Byung Jun Park, Haeun Yoo, Tae Hoon Oh, Jay H Lee, and Jong Min Lee. A model-based deep reinforcement learning method applied to finite-horizon optimal control of nonlinear control-affine system. *Journal of Process Control*, 87:166–178, 2020.
- 1330 [160] Jong Woo Kim, Tae Hoon Oh, Sang Hwan Son, Dong Hwi Jeong, and Jong Min Lee. Convergence analysis of the deep neural networks based globalized dual heuristic programming. *Automatica*, 122:109222, 2020.
- [161] Evangelos Theodorou, Jonas Buchli, and Stefan Schaal. A generalized path integral control approach to reinforcement learning. *Journal of machine learning research*, 11(Nov):3137–3181, 2010.
- 1335 [162] Pierluca D’Oro and Wojciech Jaśkowski. How to learn a useful critic? Model-based action-gradient-estimator policy optimization. *Advances in Neural Information Processing Systems*, 33, 2020.
- [163] Michael Janner, Justin Fu, Marvin Zhang, and Sergey Levine. When to trust your model: Model-based policy optimization. In *Advances in Neural Information Processing Systems*, pages 12498–12509, 2019.
- 1340 [164] Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *arXiv preprint arXiv:2005.01643*, 2020.

- 1345 [165] Scott Fujimoto, David Meger, and Doina Precup. Off-policy deep reinforcement learning without exploration. In *International Conference on Machine Learning*, pages 2052–2062, 2019.
- [166] Mike Huisman, Jan N. van Rijn, and Aske Plaat. A survey of deep meta-learning. *Artificial Intelligence Review*, 2021.
- 1350 [167] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1126–1135, 2017.
- [168] Yan Duan, John Schulman, Xi Chen, Peter L Bartlett, Ilya Sutskever, and Pieter Abbeel. RL²: Fast reinforcement learning via slow reinforcement learning. *arXiv preprint arXiv:1611.02779*, 2016.
- 1355 [169] Daniel G. McClement, Nathan P. Lawrence, Johan U. Backström, Philip D. Loewen, Michael G. Forbes, and R. Bhushan Gopaluni. Meta-reinforcement learning for the tuning of PI controllers: An offline approach. *Journal of Process Control*, 118: 139–152, 2022.
- [170] Daniele Masti and Alberto Bemporad. Learning nonlinear state–space models using autoencoders. *Automatica*, 129:109666, 2021.
- 1360 [171] Matthew Kyle Schlegel, Volodymyr Tkachuk, Adam M White, and Martha White. Investigating action encodings in recurrent neural networks in reinforcement learning. *Transactions on Machine Learning Research*, 2022.
- [172] Junyao Xie, Oguzhan Dogru, Biao Huang, Chris Godwaldt, and Brett Willms. Reinforcement learning for soft sensor design through autonomous cross-domain data selection. *Computers & Chemical Engineering*, 173:108209, 2023.
- 1365 [173] Hossein Nejatbakhsh Esfahani, Arash Bahari Kordabad, Wenqi Cai, and Sebastien Gros. Learning-based state estimation and control using mhe and mpc schemes with imperfect models. *European Journal of Control*, page 100880, 2023.
- 1370 [174] Edward H. Bras, Tobias M. Louw, and Steven M. Bradshaw. Classical actor-critic applied to the control of a self-regulatory process. *IFAC-PapersOnLine*, 56(2):7172–7177, 2023. 22nd IFAC World Congress.
- [175] Benjamin Eysenbach, Tianjun Zhang, Sergey Levine, and Russ R Salakhutdinov. Contrastive learning as goal-conditioned reinforcement learning. *Advances in Neural Information Processing Systems*, 35:35603–35620, 2022.