PROCESS OPERATIONS: FROM MODELS AND DATA TO DIGITAL APPLICATIONS

Constantinos C. Pantelides[*ab], Frances E. Pereira[a], Penelope J. Stanger[a], and Nina F. Thornhill[b],
[a]Siemens Process Systems Engineering Ltd., London W6 7HA, UK
[b]Imperial College London, London SW7 2AZ, UK

Abstract

Digital Applications are complex software systems for decision support in process operations and for process control. Each such application involves one or more computational modules being executed in an arbitrary real-time schedule, and communicating with each other and the external environment within which they are deployed. Each module may involve a mathematical computation based on a process model derived from first-principles or via machine learning applied to plant data; alternatively, it may have a purely statistical basis derived directly from plant data. There has been much progress in the use of such digital applications in industrial practice. However, achieving true scalability and sustainability in this direction will require general platforms that will allow the essentially code-free development of new applications and their large-scale deployments. We describe one such, recently developed, platform. We also consider the potential role of digital applications in the context of major trends in process operations, such as autonomous plant operation and process plant modularization.

Keywords

Process Operations, Digital Application, Digital Twin, Model-Predictive Control, Real-Time Optimization

Introduction

The digitalization of the process industries has gained substantial momentum in recent years. Some of the underlying technological basis has been available for a relatively long time in a form that could, and arguably should, have found industrial application earlier. However, there have been additional significant advances in key infrastructural areas of Operational Technology (OT) and Information Technology (IT). Of particular relevance are the hugely increased availability and accessibility of both plant data and computing power, and the algorithmic developments in machine learning and other areas of artificial intelligence.

Beyond technology, global economic, health and geopolitical developments have recently resulted in a visibly increased awareness by industry of the urgent need for step improvements in key indicators such as process profitability, sustainability, product quality, flexibility and time-to-market. This has led to significant financial investments.

The above considerations apply across the entire process lifecycle, from early-stage R&D to engineering design and process operations. This paper is concerned with the digitalization of the latter. One particular area of interest is advanced applications that make use of **first-principles mathematical models** for operations decision support and control. In a paper originally presented at the FOCAPO/CPC 2012 conference, Pantelides and Renfro (2013) discussed a range of such applications and analyzed

[*] To whom all correspondence should be addressed

their potential benefits but also the challenges associated with their development, deployment and maintenance – all of which represented non-negligible barriers to their wider adoption. A key objective of the present paper is to review the progress made since then towards addressing these problems and allowing complex applications to move from theoretical concepts to scalable industrial deployment.

The past decade has also witnessed significant developments in **data-driven techniques** for real-time process operations – either via the derivation of mathematical models using machine learning applied to plant data, or via statistical descriptions derived directly from those data. Despite their differences from methods based on first-principles models, data-driven techniques face many similar challenges, e.g. in terms of resilient and reliable operation and maintenance. Moreover, non-trivial applications may involve both model-based and data-driven computations. This paper will therefore aim to consider all of these within a common framework.

We start by introducing the general concept of a Digital Application (DA) as a complex software system and identifying its key aspects. We then review specific types of DAs with particular focus on their use in current industrial practice. Based on this, we argue that, because both of their intrinsic complexity and of their technological diversity, the wide adoption of DAs across the process industries will require the emergence of general platforms that allow the development of DAs of arbitrary complexity and their large-scale deployment in an essentially code-free manner. We outline the software architecture of one such general platform that has been developed in recent years. We also consider some aspects of the deployment of DAs within a plant's OT/IT system. We conclude with some perspectives on the potential role of DAs within major current developments in process operations.

Digital Applications

Digital Applications are integrated software systems that make use of mathematical models of process systems coupled with plant data to perform well-defined tasks[1]. The underlying models may be based on equations derived either from physical principles and other prior knowledge, or from plant data – or from combinations of the two approaches within a hybrid modelling framework. Alternatively, the models may be purely statistical in nature, based exclusively on current and past plant data.

In general, most non-trivial DAs involve multiple instances of one or more calculations being executed in parallel or sequentially to each other in real time and exchanging information with each other. They also involve bi-directional exchange of data with external data servers, such data being either inputs to the calculations or their results. As an illustration, consider real-time optimization (RTO), a well-established type of DA (Darby et al., 2011). In its conventional form, this involves an optimization calculation applied to a steady-state model of the process, with the aim of determining optimal operational settings for the plant's control system. As the optimization depends on the current inputs of the plant and the underlying model's parameters, it is typically preceded by a steady-state data reconciliation calculation which aims to obtain reliable up-to-date values of these quantities from available plant measurements. Moreover, since data reconciliation is performed on a steady state basis, it is essential to ensure that the plant is approximately at steady state before using the measurements. Overall, as illustrated in Figure 1, the RTO DA involves a sequence of instances of three different computations. Typically, one of these (steady-state detection) relies on statistical analysis of plant data (see, for example, Kelly and Hedengren, 2013) while the other two make use of numerical optimization techniques applied to a mathematical model of the process. The three computations are executed sequentially in a cycle, exchanging data with several external software systems such as the plant historian, commercial information systems and the plant control system.

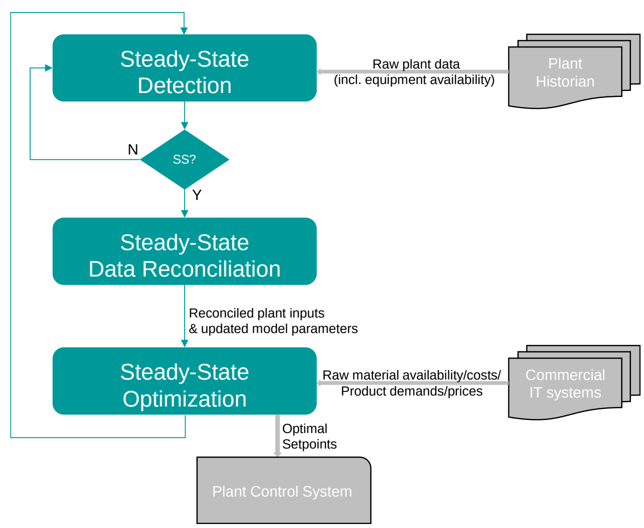

Figure 1: A typical Digital Application: Steady-state Real-Time Optimization

DAs in current industrial practice

The use of DAs for a wide variety of decision support and control tasks in process operations is becoming increasingly common in industrial practice. This section considers some of these recent developments. For ease of presentation, they are organized according to the size of the process envelope to which they are applied, starting from individual major equipment items, moving on to entire

[1] Although in recent times the term "Digital Twin" has been extensively used to describe such applications, it is somewhat misleading: the purpose of most DAs is not to mimic the behaviour of the process but to use model(s) that represent that behaviour to perform specific tasks.

plants, and finally to integrated networks of multiple plants within the same or different geographical sites.

DAs for major equipment items/plant units

Perhaps the most widespread types of DA are those used for advanced control of major equipment items or process units, such as reactors or distillation columns. The application of **model-predictive control** (MPC) based on linear models identified from plant data is now routine in this context. MPC based on *nonlinear* models is also finding industrial application, particularly for batch and/or highly nonlinear processes, or applications such as polymerization (see, for example, Bindlish, 2015) where the large number of distinct operating points and the frequent transitions between them pose challenges to the linear MPC approach. Although the earlier nonlinear MPC technology relied on surrogate models based on neural networks or bounded derivative networks (Turner and Guiver, 2005), more recently the increasing capability to solve nonlinear dynamic optimization problems reliably and efficiently, has allowed it to be applied directly to physics-based models incorporating a significant degree of detail in the description both of the distributed nature of the polymer in terms of molecular weight and degree of branching, and of the associated physical properties (Pfeiffer et al., 2020).

A closely related type of DA is that of **soft sensing** used to provide real-time estimates of the plant's key performance indicators (KPIs). In addition to allowing real-time monitoring of the plant's performance by the operators, these KPIs can act as controlled variables in the context of the plant regulatory and supervisory control system, often leading to much more direct control. Some soft sensors make use of physics-based models. An example is the real-time estimation of ethylene and propylene yields in olefin cracking furnaces based on detailed kinetic and heat transfer models; this allows yield to be controlled directly instead of relying on proxy quantities such as the coil outlet temperature. In the simplest case, the underlying computations are dynamic (or, in the case of processes with very fast dynamics, steady-state) simulations that attempt to emulate the plant performance by subjecting the model to measured inputs. More sophisticated methods make use of redundant measurements by applying state estimation techniques of varying degrees of sophistication (Tatjewski and Ławryńczuk, 2020). A significant recent development in this context is the incorporation of state estimation algorithms, such as Extended Kalman Filters (EKF) and Moving Horizon Estimators (MHE), within general process modelling environments such as gPROMS® in a manner that is applicable directly to wide classes of process models described by mixed sets of integral, partial and ordinary differential, and algebraic equations.

Alternatively, soft sensing may be performed using statistical models. An early example was presented by Casali et al. (1998) who reported results from soft sensing of particle size in an industrial grinding plant using an ARMAX model with measurements of water addition rate, pump speed, and solids concentration as inputs. Kadlec et al. (2009) gave a comprehensive review of data-driven soft sensing. They surveyed the main methods in use at the time, including principal component analysis, artificial neural networks, and support vector machines. Subsequent additions include Bayesian methods (Khatibisepehr et al., 2013), autoencoder models (Yuan et al., 2018), and deep learning methods (Sun et al., 2021).

As an example, Stanišić et al., (2015) created a nonlinear model using multi-layer perceptrons to estimate fineness of cement. They proposed a systematic way for selecting the inputs to the network from among all the process measurements by eliminating redundancies among correlated measurements. Tests of the implemented sensors showed good agreement between the soft-sensed estimates and laboratory measurement.

Data-driven soft sensors have their own challenges, including data preprocessing issues such as dealing with missing values, measurement delays and correlated input measurements. Industrial data sets that reflect these real-life issues are in short supply. As identified by Sun et al. (2021), the cost of generating and disseminating suitable data is not trivial, for instance in preparation of plant experiments and data collection to generate calibration data sets.

DAs can also be used to **monitor the condition of equipment** subject to slow degradation processes such as fouling in heat exchangers, coke deposition in thermal cracking furnaces, catalyst deactivation in catalytic reactors, and fouling and mechanical deterioration in rotating turbomachinery. In some cases, it is possible to incorporate descriptions of such degradation phenomena within a physics-based model, linking the instantaneous rate of degradation to the transient operating conditions within the equipment, and conversely determining the effect of cumulative degradation on the equipment's current performance. Such physical descriptions also allow **forecasting** of the future evolution of degradation under any number of potential operating scenarios, thus providing valuable input to future operating strategies and supporting predictive maintenance approaches whereby maintenance can be scheduled optimally rather than on a fixed schedule.

In practice, degradation phenomena are often much less well understood and more difficult to characterize than the main equipment physics, especially if they depend on factors (e.g. the condition of the metal surface) that are not normally taken into account by the model. Accordingly, a degree of online adaptation of the underlying model is often required in this context, typically using both current and historical data for the current run of the equipment since the time of its last maintenance. This is essentially a parameter estimation computation, often incorporating a Bayesian element to take account of prior information on the values of degradation model parameters obtained from previous runs.

An alternative class of techniques for equipment monitoring are those based entirely on process data. These are particularly useful, and often necessary, in view of the above-mentioned difficulty of establishing reliable physical descriptions. An application to the monitoring of the degradation in efficiency of off-shore turbomachines was presented by Zagorowska et al. (2020a). The method was based on regression fitting of a nonlinear function to measurements of efficiency, and its extrapolation into the future together with confidence bounds. The function was based on empirical knowledge of the form expected for degradation as a function of time. Knowledge of the level of degradation can also have an impact on the way the process is operated. For instance, Zagorowska et al. (2020b) presented a practical application from BASF in which degradation was inferred from measurements of pressure drop across a heat exchanger in a polymer production facility. The estimates were then used in supporting decisions about the scheduling of batches. Parameters in the calculation depended on the polymer recipe and were calibrated using measurements from previous batches (Wu et al., 2019).

DAs for supporting plant-wide operations

At the level of managing plant-wide operations, typical DAs include **data reconciliation** and **RTO** (cf. earlier discussion in this paper). Although the focus of RTO is often presented as optimizing continuous plant operational settings in the face of continuous disturbances (e.g. feed availability and/or composition; ambient temperature) or objectives (e.g. production rates), in industrial practice a major benefit is in handling discrete events (e.g. planned or unplanned changes in equipment availability), often by effecting discrete changes to the plant operation (e.g. switching equipment items or trains on or off; re-routing material flows). This poses computational challenges relating to the solution of the underlying large-scale mixed-integer optimization problems.

A more fundamental and well-known issue with RTO is its steady-state basis and its assumption that it is sufficient for the optimal set-points to be updated at periods that are much longer than the process time constants. This is unrealistic in many practical situations, especially as plants are increasingly being operated in a more dynamic fashion to respond more effectively to changes in the external environment. However, despite two decades of academic work on **dynamic RTO** (Kadam et al., 2003; Biegler, 2009), **hybrid RTO**[2] (Krishnamoorthy et al., 2018) and the related areas of **economic MPC** (Ellis et al., 2017) and **integrated production scheduling and control** (Tsay et al., 2019; Tsay and Baldea, 2020), currently full applications of this technology to industrial practice are mostly limited to those where the site-wide model is derived by integrating the linear models used for MPC of individual process units.

Another class of plant-wide DAs is that concerned with the **on-line analysis of alarm data**, particularly aiming at mitigating the impact of alarm floods during which the alarm rate is greater than what the operator can effectively manage. Some of those methods aim at rationalizing the alarm systems while others offer online support to the operator during the alarm floods. Lucke et al. (2019) provided a structured hierarchical review of this area.

Alarm rationalization involves removing redundant alarms to make sure that the operator receives only alarms that require an operator response. For instance, chattering alarms will be identified (Wang et al., 2016).

An important aspect of online support is the correct classification of an alarm flood. In online operation, probabilities are assessed that a developing alarm flood belongs to a class already seen by the off-line classifier, or to a left-over class that has been seen before but was not explained, or to a new type of alarm flood.

Other researchers have proposed methods for investigating the root causes of the alarm floods making use of alarms alongside or as an alternative to process measurements. Yu et al. (2015) and Hu et al. (2017) have used the methods of transfer entropy to infer causality.

A related area is that of **anomaly detection** which aims to identify that a fault is developing in the process. Such methods face challenges when applied to processes that have varying production regimes, for instance when making a variety of grades of polymer. In that case, an anomaly might be due either to a fault in the process or to a transition to a new operating mode.

Early methods for addressing this problem typically involved identification of each expected operating mode and construction of a linear model for monitoring it. More recently, researchers have been developing the theory for nonlinear methods such as kernel principal component analysis and its monitoring statistics. For industrial implementation, it is useful if the calibration method can detect the various modes in historical data even if they are not labelled according to the mode of operation. It is also necessary to distinguish between a fault and a new mode of operation. To this end, Tan et al. (2020) demonstrated a system for operator support that would recalibrate the model if the operator confirmed a new normal mode of operation, or report a fault if the anomalous results could not be so explained.

DAs beyond the single plant boundaries

DAs can also be used to coordinate the production in systems involving **integrated multiple plants** and to optimize the operations of **supply chains** potentially involving more than one organization. One example in this context is production optimization for groups of two or more olefin plants situated relatively near each other, sharing a common, and often limited, feedstock supply, and

---

[2] Hybrid RTO combines dynamic estimation techniques for determining the current plant state with a steady-state optimization step.

aiming to satisfy product demands that are aggregated across all plants. More complex situations may involve multiple plants distributed over wide geographical areas and connected via pipelines, something which introduces a significant temporal element in the optimization and an additional complexity in the design of the DA. Although in the past, such coordination problems were treated using simple, often linear, models of the individual plants, there is now substantial evidence that the use of more accurate plant models that would allow these large systems to be operated confidently nearer their constraints can result in large economic benefits. An early assessment of the potential of this approach was provided by Aluma et al. (2016). Since then, the emergence of general DA platforms (see below) has enabled the development of DAs that can handle the optimization of these ultra-large systems in a reliable and efficient manner. These have recently reached the stage of full-scale deployment, providing further confirmation that the expected benefits are realizable.

The need for general Digital Applications Platforms

The above discussion illustrates two key characteristics of DAs:

- **Complexity**: even ostensibly simple and well-understood DAs such as RTO (cf. Figure 1) are complex software systems combining model-based calculations with plant data in real time. The complexity is compounded by the need to handle abnormal, but all too common, situations such as temporary interruptions in the availability of data, erroneous data, and failures in model-based calculations.
- **Diversity**: there is a very wide range of existing and potential DAs which differ from each other in terms of intended function, underlying calculations and the external data servers with which they exchange information.

The complexity of DAs places stringent requirements on the testing and validation that is required to ensure that they can operate faultlessly, continuously and with little human intervention over extended periods of time. The complexity has also made many DAs difficult to maintain in the longer term despite their successful initial development and deployment.

In principle, DA complexity can be, and has been, addressed by systematic software engineering. However, this comes at a non-negligible cost in terms of both the money and the time required for the development of large amounts of customized computer code that is involved in even relatively simple applications such as RTO. Combined with the diversity of potentially useful DAs, this poses a serious impediment to their large-scale development. This difficulty is reflected by the current situation where commercially available DAs are limited to very few well-established classes (such as linear Model-Predictive Control and steady-state Real-Time Optimization), with each one of them being implemented on a customized code base.

The above considerations naturally lead to the idea of general Digital Applications Platforms (DAPs), i.e. integrated software frameworks for developing, validating, deploying and supporting DAs that conform to a given abstract model. The key objectives of a DAP would include minimizing the effort required for error-free development of new classes of DAs; ensuring efficient real-time DA execution within given software and hardware infrastructures; and ensuring resilience in abnormal situations such as those mentioned above. Overall, this would be analogous to the emergence of multipurpose process modelling environments (Pantelides and Britt, 1995): by supporting the efficient, code-free, development of process models within a particular class of mathematical problems (e.g. mixed sets of differential and algebraic equations), these tools have been instrumental in the rapid dissemination of advanced process modelling and process systems engineering methodologies within the process industries – including diverse sectors, such as pharmaceuticals and food & beverage, which did not have a strong tradition of making use of these technologies.

A software architecture for a Digital Applications Platform

One possible architecture for a DAP is outlined in Figure 2. A major element of any DA is a set of **computational modules** (CMs), each performing a different, usually but not always model-based, computation. For example, in the case of the RTO DA illustrated in Figure 1, there would be three such CMs, namely steady-state detection, data reconciliation, steady-state optimization. However, an important difference from the rigid architecture shown in Figure 1 is that neither the scheduling of the execution of these three CMs in time nor the exchange of data between them are hardcoded. Instead, this information is provided in a soft form in a **DA Metafile**. This allows the DA execution to be managed automatically by two generalized software modules, namely the **Scheduler** that orchestrates the execution of the modules, and the **Resource Manager** that allocates computational resources (e.g. computer cores or processors) to these executions. These modules allow significantly more sophistication in DA architecture than the conventional paradigm of a strict sequence of calculations being executed in a cycle, all on a single computer processor. For example, the occurrence or timing of each CM's execution may depend on the outcome of the execution of other CMs and/or events received from External Data Servers (see below); and multiple CMs may be executing asynchronously in parallel, making use of multiple computer cores or nodes.

As has already mentioned, another aspect of DAs is their communication with external software systems, such as the plant historian, commercial databases and plant control system (cf. the RTO example of Figure 1). To accommodate this requirement, the DAP architecture shown in Figure 2 incorporates the concept of generalized **External Data Servers** that can comprise an arbitrarily wide range of software systems and associated data exchange protocols, such as the Data Access (DA), Historical Data Access (HDA) and Unified Architecture (UA) standards of Open Platform Communications (OPC), SQL database access, and web services. All communication is handled via the platform's generalized **Data Manager** module.

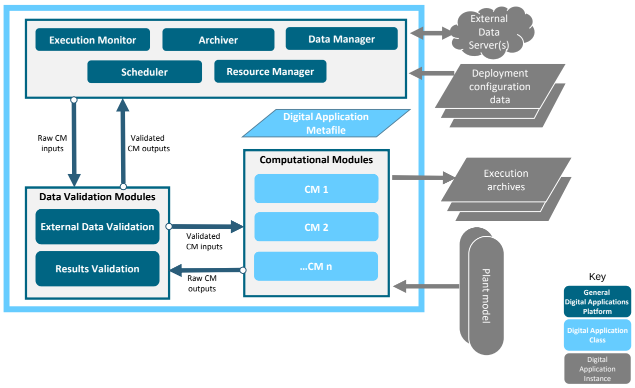

Figure 2: The gPROMS Digital Applications Platform

A major determining factor for DA resilience and robustness is ensuring the quality of any data used as inputs by the CMs. In general, neither the availability nor the correctness of data obtained from External Data Servers can be guaranteed. For example, data may become temporarily unavailable or corrupted due to communication problems; and even when available, the actual values may be wrong due to sensor faults. For the DA to continue operating reliably in all such circumstances, it needs to be able to identify faulty data (e.g. by considering the value or the rate of change over time of each individual item; and the relative values of multiple data items) and then take appropriate action (e.g. by replacing an invalid value with a previously obtained valid one, or by temporarily skipping the CM execution altogether). Within the DAP, this is the responsibility of the generalized **External Data Validation** module. The criteria used for validating data items of different types and the actions to be taken in handling invalid ones are not hardcoded but are also part of the DA Metafile. This is important as a large part of the software code of traditional DAs is dedicated to ensuring data integrity.

The Data Manager is also responsible for communicating the results of the DA to appropriate External Data Servers. For instance, in the RTO example of Figure 1, the optimal control set points need to be sent to the plant control system. Ensuring the integrity of such results is particularly important in the case of DAs operating in closed-loop mode. Accordingly, a separate **Results Validation** module has the responsibility of performing a set of final validity checks on these results, as a protection against errors arising from any unanticipated deficiencies in input data validation and model-based computations. Again, the rules and criteria used for this validation are part of the DA's Metafile. Finally, in view of the potential complexity of DAs, it is important to monitor their execution and take appropriate action in real time if this deviates from the expected performance (the **Execution Monitor** module) and also to maintain a record of this performance to enable later troubleshooting of any problems that may arise (the **Archiver** module).

Each class of DA (RTO, MPC, etc.) comprises the software components mentioned above, as delineated by the thick light blue line in Figure 2. Most of these components (shown in dark blue in Figure 2) are identical irrespective of the DA class and can, therefore, form part of a generalized software platform (DAP). The only DA class-specific elements (shown in light blue in Figure 2) are the DA Metafile and the CMs. In fact, most CMs required by typical DA classes correspond to common forms of model-based computations (e.g. dynamic simulation, state or parameter estimation, steady-state or dynamic optimization etc.) and can themselves be standardized. Overall, this significantly reduces or even eliminates the need for customized computer code, thereby facilitating both the development of new types of DAs and the maintenance of existing ones. Moreover, any improvement in functionality or performance in the underlying DAP is immediately reflected in all DA classes built on it.

Deployment of Digital Applications within the plant

The deployment of a DA at a given plant or other physical asset requires the provision of additional information relating to this specific instance, as illustrated by the elements shown in grey on the right part of Figure 2. These includes the mathematical or statistical plant models required by the CMs and an identification of the specific External Data Servers with which the DA will interact. Configuration information that maps a subset of the model's variables and other attributes onto quantities of relevance to the DA must also be provided. For example, in the case of a DA used for process control, it would typically be necessary to identify the manipulated, controlled, measured and disturbance variables among the (usually much more numerous) model variables. The mapping of some of these variables to data items on the External Data Servers must also be provided.

In general, configuring, validating and tuning a DA instance to be deployed at a specific plant or other system is a non-trivial engineering task. The generalized DA concepts and associated software framework presented here also allow the development of general-purpose software tools to support these off-line activities, as well as taking DA instances to online service and managing their execution over their lifetime. All this can be achieved without the need for any customized computer code. However, automatic techniques for tuning complex DAs involving multiple algorithmic parameters and adjustment factors (see, for example, Forgione et al., 2020 and Lu et al., 2021) can have a large impact in this context.

Finally, there are several options regarding the hardware and software environment used for DA execution. In conventional arrangements, DAs are deployed on dedicated industrial computers often situated within the plant control room, or in virtual machines within existing hardware. This provides maximum speed of response and arguably a higher guarantee of cyber security but may incur potentially higher costs of installation and maintenance. At the other extreme, DAs may be deployed on the Cloud, which provides maximum flexibility but may be unable to guarantee the required response times; it may also involve the transfer, processing and storage of very large amounts of data. A compromise between the two is provided by Edge Computing where the processing is performed locally, with only small amounts of data and/or results being communicated externally. By providing secure network connectivity, Edge devices allow DAs to be installed and maintained remotely by suitably authorized personnel (e.g. centrally-located corporate advanced control or digitalization teams). They also provide a degree of standardization for the external environments experienced by the DAs (e.g. in terms of access to plant data or mechanisms for interaction with plant personnel operator dashboards and other displays). These considerations are crucial for any large-scale development and deployment of DAs in the future.

Future perspectives for Digital Applications

We conclude this paper by considering DAs within the wider context of major ongoing trends and developments in process operations practice.

Autonomous plant operation

Autonomous plant operation is widely recognized as a major future milestone in the digital transformation of the process industries (see, for example, Gamer et al. (2019), Yokogawa Electric Co. (2020), and Georgiou and Sheth (2022)). According to the definition provided by Watson and Scheidt (2005), autonomy denotes the ability to develop a well-defined, yet modifiable, plan towards the achievement of high-level goals, making use of all available resources and taking account of all relevant constraints; execute the plan, modifying it if necessary; react appropriately, if not optimally, to unexpected events; and coordinate with human controllers. To achieve this objective, the physics- and data-based models described earlier in this paper need to be complemented by semantic models that encompass all available data and their relations in formal representations such as knowledge graphs. Another important pre-requisite is that any operational plans derived automatically (e.g. via artificial intelligence algorithms) can be described in a formal, high-level manner and then be executed reliably. We believe that the flexible Digital Applications Platforms described in this paper will be an essential ingredient of any such solution. Of particular importance in this context is the ability to define complex DAs "on-the-fly" (cf. the metafile in Figure 2) involving a potentially complex execution schedule of computational modules and other microservices.

Heterogeneous plant data

Data from operating plants take diverse forms, including process measurements sampled at various rates and intervals using a variety of online sensors, laboratory or at-line measurements of product quality, and logs of alarms, events and actions of operators. The enhanced availability of data in recent years has been an important enabling factor in the digital transformation of digital operations. Key improvements include more accurate determination of compositions via combinations of multiple spectroscopic techniques, complex measurements including video and acoustic imaging, electronic noses for detecting specific chemicals, and disposable sensors (Baldea and Zavala, 2022). Data have also become more accessible via the more extensive use of sophisticated plant historians and data aggregation software.

Bringing in data from a wide variety of sources can have a large impact on the scope and performance of the DAs discussed in this paper. For instance, accessing and analyzing a video image of a flare could be helpful in managing process alarms. However, for this to be realized, novel techniques need to be developed to generate actionable information from an analysis of heterogeneous data. In particular, an additional stage of feature extraction and feature classification is needed to render the information from very different data sources in a form that can be compared and fused. An example in the area of fault detection has been provided by Stief et al. (2019a) where measurements from disparate data sources were combined using Bayesian sensor fusion after preprocessing, feature extraction and classification.

Testing and comparison of advanced data analysis algorithms requires benchmark datasets. One such dataset is the PRONTO heterogeneous benchmark dataset (Stief et al., 2019b, 2019c). Data were collected from heterogeneous sources, including process measurements, alarm records, high-frequency ultrasonic flow and pressure measurements, and video.

Modular production

Modular production (ZVEI et al., 2019) represents a new paradigm for the process industries. Conventional plants are usually designed on a one-off basis to achieve specific production objectives in an optimal manner. In contrast, modular plants are assembled from standardized modules which are already designed, automated and fabricated. Although the final outcome may not necessarily deliver the (theoretically) best achievable performance, in practice significant benefits may be realized in terms of flexibility of production, reduced engineering effort and risk, and shorter time to market.

The modularization of production hardware is paralleled by the modularization of the associated automation and control systems. The information required to integrate an instance of a particular module within a modular plant is supplied in the form of a Module Type Package (MTP) conforming to a standardized vendor-independent format (NAMUR, 2013). Within the module itself, sophisticated DAs of the type considered in this paper can play a key role in ensuring safe operation and optimizing performance in the diverse contexts within which the module may be deployed. As in the case of the equipment hardware, standardization is a key consideration: the fact that the module is likely to be produced in large numbers, instead of being a one-of-a-kind design, provides the economic justification for the investment required for building reliable and efficient DAs.

References

Baldea, M., Zavala, V. (2022). Quo vadis process data? *Keynote presentation at FIPSE 5 conference*, Crete, Greece,

Biegler, L.T. (2009). Technology advances for dynamic real-time optimization. *In Proc. 10th International Symposium on Process Systems Engineering,* Salvador-Bahia, Brazil.

Bindlish, R. (2015). Nonlinear model predictive control of an industrial polymerization process. *Comput. Chem. Eng.*, *73*, 43-48

Casali, A., Gonzalez, G., Torres, F., Vallebuona, G., Castelli, L., Gimenez, P. (1998). Particle size distribution soft-sensor for a grinding circuit, *Powder Technology*, *99*, 15-21

Darby, M.L., Nikolaou, M., Jones, J., Nicholson, D. (2011). RTO: An overview and assessment of current practice. *J. Proc. Control*, *21*, 874-884.

Ellis, T., Liu, J., Christofides, P.D. (2017). Economic Model Predictive Control: Theory, Formulations and Chemical Process Applications, *Springer.*

Forgione, M., Piga, D., Bemporad, A. (2020). Efficient calibration of embedded MPC. *IFAC-Papers Online*, *53-2*, 5189-5194.

Gamer, T., Hoernicke, M., Kloepper, B., Bauer, R., Isaksson, A.J. (2019). The Autonomous Industrial Plant – Future of Process Engineering, Operations and Maintenance. *J. Proc. Control, 88*, 101-110.

Georgiou, A., Sheth, K. (2022). Autonomy in plant operations. *Keynote presentation at FIPSE 5,* Crete, Greece.

Hu, W., Wang, J., Chen, T., Shah, S.L. (2017). Cause-effect analysis of industrial alarm variables using transfer entropies, *Control Eng. Pract*. *64*, 205–214.

Kadlec, P., Gabrys, B., Sibylle Strandt, S., (2009). Data-driven Soft Sensors in the process industry, *Comput. Chem. Eng.*, *33*, 795-814.

Kadam, J. V., Marquardt, W., Schlegel, M., Backx, T., Bosgra, O. H., Brouwer, P-J., Dunnebier, G., van Hessem, D., Tiagounov, A., de Wolf, S. (2003). Towards integrated dynamic real-time optimization and control of industrial processes. *In Proc. FOCAPO 2003*, CACHE Co.

Kelly, J.D., Hedengren, J.D. (2013). A steady-state detection (SSD) algorithm to detect non-stationary drifts in processes. *J. Proc. Control,* 23, 326-331.

Khatibisepehr, S., Huang, B., Khare, S. (2013), Design of inferential sensors in the process industry: A review of Bayesian methods, *J. Proc. Control, 23*, 1575-1596.

Krishnamoorthy, D., Foss, B., Skogestad, S. (2018). Steady-state real-time optimization using transient measurements. *Comput. Chem. Eng*., 115, 34-45.

Lu, Q., González, L.D., Kumar, R., Zavala, V.M. (2021). Bayesian optimization with reference models: A case study in MPC for HVAC central plants. *Comput. Chem. Eng.*, *154*, 107491.

Lucke, M., Chioua, M., Grimholt, C., Hollender, M., and Thornhill. N.F., (2019). Advances in alarm data analysis with a practical application to online alarm flood classification. *J. Proc. Control, 79*, 56-71.

NAMUR (2013). *NE 148: Automation Requirements Relating to Modularisation of Process Plants*, Leverkusen, Germany.

Pantelides, C.C., Britt, H. (1995). Multipurpose process modeling environments. *In Proc. FOCAPD 1994*, CACHE Co., 128-141.

Pantelides, C.C., Renfro, J.G., (2013). The online use of first-principles models in process operations: review, current status and future needs. *Comput. Chem. Eng.*, *51*, 136-148.

Pfeiffer, B-M., Oppelt, M., Leingang, C., Pantelides, C.C., Pereira, F.E. (2020). Nonlinear model predictive control based on existing mechanistic models of polymerisation reactors. *IFAC-Papers Online*, *53-2*, 6076-6081.

Stanišić, D., Jorgovanović, N., Popov, N., Čongradac, V., (2015). Soft sensor for real-time cement fineness estimation, *ISA Transactions, 55*, 250-259.

Stief, A., Ottewill, J.R., Baranowski J., Orkisz, M. (2019a), A PCA and two-stage Bayesian sensor fusion approach for diagnosing electrical and mechanical faults in induction motors, *IEEE Trans. Industr. Electr., 66*, 9510-9520.

Stief, A., Tan, R., Cao, Y., Ottewill, J.R., Thornhill, N.F. Baranowski, J. (2019b). A heterogeneous benchmark dataset for data analytics: multiphase flow facility case study. *J. Proc. Control, 79*, 41-55.

Stief, A., Tan, R., Cao, Y., Ottewill, J.R. (2019c). PRONTO Heterogeneous Benchmark Dataset, *Zenodo 1341583*.

Sun, Q., Ge, Z. (2021). A survey on deep learning for data-driven soft sensors, *IEEE Transactions on Industrial Informatics*, *17*, 5853-5866.

Tan, R., Cong, C., Ottewill, J.R., Baranowski, J., and Thornhill, N.F. (2020). An on-line framework for monitoring nonlinear processes with multiple operating modes. *Journal of Process Control*, *89*, 119-130.

Tatjewski P., Ławryńczuk, M. (2020). Algorithms with state estimation in linear and nonlinear model predictive control. *Comput. Chem. Eng.*, 143, 107065.

Tsay, C., Kumar, A., Flores-Cerrillo, J., Baldea, M. (2019). Optimal demand response scheduling of an industrial air separation unit using data-driven dynamic models. *Comput. Chem. Eng. 126*, 22-34.

Tsay, C., Baldea, M. (2020). Integrating production scheduling and process control using latent variable dynamic models. *Control Engng. Pract*., *93*, 104201.

Turner, J., Guiver, P. (2005). Introducing the bounded derivative network—superseding the application of neural networks in control. *J. Proc. Control, 15*, 407-415.

Watson, D.P., Scheidt, D.H. (2005). Autonomous Systems. *Johns Hopkins APL Technical Digest, 26*, 368-376.

Wang, J., Yang, F., Chen , T., Shah, S.L., (2016). An overview of industrial alarm systems: main causes for alarm overloading, research status, and open problems. *IEEE Transactions on Automation Science and Engineering*, *13*, 1045-1061.

Wu, O., Bouaswaig, A., Imsland, L., Schneider, S.M., Roth, M., Leira, F.M., (2019) Campaign-based modeling for degradation evolution in batch processes using a multiway partial least squares approach. *Comput. Chem. Eng., 128*, 117-127.

Yu, W., Yang, F., (2015). Detection of causality between process variables based on industrial alarm data using transfer entropy. *Entropy 17*, 5868–5887.

Yuan, X., Huang, B., Wang, Y., Yang C., Gui, W., (2018). Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE. *IEEE Trans. Industr. Informatics*, *14*, 3235-3243.

Yokogawa Electric Co. (2020). *Digital Transformation in Process Industries: A Journey Towards Autonomous Operations*. EBOOK-DXPC-20200917

Zagorowska, M., Schulze-Spüntrup, F., Ditlefsen, A-M, Thornhill, N.F., Imsland, L., Lunde, E, (2020a). Adaptive detection and prediction of performance degradation in off-shore turbomachinery. *Applied Energy*, *268*, 114934.

Zagorowska, M., Wu, O., Ottewill, J.R., Reble, M., Thornhill, N.F., (2020b). A survey of models of degradation for control applications. *Ann. Rev. Control*, *50*, 150-173.

ZVEI, NAMUR, ProcessNet, VDMA (2019). *Process INDUSTRIE 4.0: The Age of Modular production*.