ML-based proactive control of industrial processes

Edyta Kuk[1 [0000−0003−3298−7359]], Szymon Bobek[2 [0000−0002−6350−8405]], and Grzegorz J. Nalepa[2 [0000−0002−8182−4225]]

[1] AGH University of Science and Technology
[2] Jagiellonian University, Faculty of Physics, Astronomy and Applied Computer Science, Institute of Applied Computer Science, and Jagiellonian Human-Centered AI Lab (JAHCAI), and Mark Kac Center for Complex Systems Research, ul. prof. Stanisława Łojasiewicza 11, 30-348 Kraków, Poland

Abstract. This paper discusses the use of optimal control for improving the performance of industrial processes. Industry 4.0 technologies play a crucial role in this approach by providing real-time data from physical devices. Additionally, simulations and virtual sensors allow for proactive control of the process by predicting potential issues and taking measures to prevent them. The paper proposes a new methodology for proactive control based on machine learning techniques that combines physical and virtual sensor data obtained from a simulation model. A deep temporal clustering algorithm is used to identify the process stage, and a control scheme directly dependent on this stage is used to determine the appropriate control actions to be taken. The control scheme is created by an expert human, based on the best industrial practices, making the whole process fully interpretable. The performance of the developed solution is demonstrated using a case study of gas production from an underground reservoir. The results show that the proposed algorithm can provide proactive control, reducing downtime, increasing process reliability, and improving performance.

Keywords: machine learning · artificial intelligence · simulations · optimal control

1 Introduction

Proper process control is crucial to avoid mistakes and downtime that are costly in the case of real-world processes. Optimal control is a key approach that can help industrial processes achieve their maximum potential and improve their overall performance. It involves designing a control algorithm that adjusts the inputs to the process in real-time to achieve the desired output while satisfying certain constraints. Optimal control aims to identify control actions that maximize the desired objective function. The use of Industry 4.0 technologies [1], which offer real-time access to numerous parameters of the process and low-level operational status of the physical devices, can greatly enhance optimal control. Such information can be used to optimize the process in ways that were not possible before. Moreover, simulations play a critical role in the operation of complex industrial processes. They can be used to develop and test control algorithms before implementation helping engineers and operators to make informed decisions, optimize the performance of the process, improve efficiency, and ensure safety [20].

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

In addition to Industry 4.0, simulations can give the possibility to make use of virtual sensors. Virtual sensors provide estimates of process variables that are important for process control but cannot be directly measured by a physical sensor. Virtual sensors based on simulation models enable engineers to estimate process variables that would otherwise be difficult or impossible to measure, and allow for real-time monitoring and control of these variables [4]. The use of virtual sensors based on simulation models provides proactive control of the process, allowing engineers to predict potential issues and take measures to prevent them before they occur, reducing downtime and increasing the reliability of the process [6]. This is different from reactive process control which responds to problems after they have occurred.

Using simulation models vast amounts of synthetic data from industrial processes can be generated, including those that can be measured physically as well as virtual measurements. Machine learning (ML) techniques can be used to analyze this data and develop predictive models that enable proactive control and optimization of the process [17]. One approach is to use simulation models to generate training data for supervised learning algorithms that can predict process behavior and identify optimal control parameters. Unsupervised learning techniques can also be used to identify hidden patterns in the data and provide insights into process behavior [3]. Another approach is to use simulation models to create a digital twin of the process, which can be used to optimize control policies using reinforcement learning algorithms [12]. Overall, the combination of simulations with ML provides a powerful tool for process optimization and control. By leveraging the insights gained from simulation data analysis with ML engineers can make data-driven decisions that improve process performance.

1.1 Paper contribution

In this paper, we propose a new method for proactive control of industrial processes based on ML techniques. The idea is to combine both physical and virtual sensor data obtained from a simulation model through an ML algorithm to identify the current stage of the industrial process, which is then used to proactively determine appropriate control actions. The proposed approach can be applied to a single asset/process or to multiple assets if the simulation model includes a fleet of them. In this paper, we applied a deep temporal clustering algorithm that employs an autoencoder for temporal dimensionality reduction and a temporal clustering layer for cluster assignment, to combine simulation data and identify the process stage. The appropriate control actions to be taken are then determined according to the current process stage. Although stage identification is fully unsupervised, decision control is defined by a human expert, making the process more transparent and interpretable. This cooperative rather than substitutive role of AI in industrial applications follows the Industry 5.0 paradigm, which places more emphasis on human-AI collaboration rather than black-box automation [18].

1.2 Paper outline

The rest of the paper is organized as follows: In Section 2 we describe papers that cover ML approaches for determining industrial process control and we introduce our motivations. In Section 3 we introduce our methodology for establishing proactive process

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

control using ML methods. Next, we apply the developed method to define proactive control of wells on the gas reservoir which is shown in Section 4. In Section 5 we explore the possible applications of the proposed approach and discuss its potential and limitations. Finally, we summarise our work in Section 6.

2 Related works and motivation

In industrial practice, numerous sensors are installed at various locations to collect process data. If physical measurements are combined with synthetic data generated by virtual sensors from simulation models of the process the process's historical databases are massive. While there has been an increase in the availability of sensors, the raw data collected from these sensors will eventually become useless without proper data analysis, information extraction, and knowledge exploitation techniques. Data from industrial processes can be analyzed with machine learning techniques to develop predictive models for process control and optimization. In this section, we provide an overview of process control approaches that utilize machine learning methods.

In [7] the Authors discuss the importance of soft sensors, called virtual sensors, in the modern industry as an alternative to physical sensors. The development of soft sensors has been facilitated by achievements in data science, computing, communication technologies, statistical tools, and machine learning techniques. The article also discusses the significance of soft sensing in improving production safety and product quality management from the perspectives of system monitoring, control, and optimization.

In [5] the Authors discuss the challenges of interpreting large volumes of data collected from smart factories, which often contain information from numerous sensors and control equipment. Principal component analysis (PCA) is a common method for summarizing data, but it can be difficult to interpret in the context of fault detection and diagnosis studies. Sparse principal component analysis (SPCA) is a newer technique that can produce PCs with sparse loadings, and the article introduces a method for selecting the number of non-zero loadings in each PC using SPCA. This approach improves the interpretability of PCs while minimizing information loss.

The paper [13] begins by introducing the recent advancements in machine learning and artificial intelligence, which have been driven by advances in statistical learning theory and big data companies' commercial successes. They provide three attributes for process data analytics to make machine learning techniques applicable in process industries. The paper discusses the currently active topics in machine learning that could be opportunities for process data analytics research and development.

The article [15] discusses the most commonly used machine learning techniques in the context of production planning and control. The authors grouped the ML techniques into families to facilitate the analysis of results, and neural networks, Q-Learning, and decision trees were found to be the most frequently used techniques. The high usage of clustering techniques is likely due to the nature of data in manufacturing systems. The study also observed a strong growth in the use of neural networks since 2015, possibly due to the development of specialized frameworks and growing computing power. The

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

results also suggest a growing interest in ensemble learning techniques, possibly at the expense of decision trees since random forests can achieve better performance.

In [21] the Authors perform a diagnosis that uses supervised learning models to detect quality anomalies caused by abnormal process variations. This approach is different from diagnosis based on unsupervised learning models which can only analyze abnormal changes in process variables. The goal of supervised monitoring and diagnosis is to detect quality anomalies before they are measured and confirmed.

The authors of [16] emphasizes that deep learning is highly effective in modelling complex nonlinear processes, but it still faces several challenges in its application to process control. These challenges include difficulties with interpreting the features extracted from data and their relationship to outputs, sensitivity to network hyper-parameters, and the influence of the available training dataset's size.

The issue of optimal process control can be generalized to a wide range of statistical process control (SPC) approaches, such as the most commonly used in manufacturing, Six Sigma approach [19]. In the survey paper [14] the authors provide a comprehensive review of machine learning approaches in SPC tasks. While they notice that the applications of ML in that field is growing rapidly, they also point out several challenges that such approaches are facing nowadays. They define several requirements for Industry 4.0/5.0 technologies have to address in order to efficiently implement ML-based system in their process control setups. One of which is the requirement of interpretability of control charts for complex data and data fusion from multiple sources to improve the performance of the quality control mechanism. In following section we describe how our work fits in that requirements.

2.1 Motivation

The papers described above demonstrate the successful application of machine learning techniques in analyzing industrial process data. Prediction and interpretation are the two essential aspects of ML applications to process data analysis, with deep neural networks being better suited for accurate prediction and simpler models preferred for interoperability. However, interpretability is critical for industrial adoption as it helps to establish trust in the analytics algorithms. While deep learning is commonly applied to sensor data, achieving a balance between model complexity and maintainability is necessary to achieve a trade-off between accuracy and complexity. In our work we do not substitute human with black-box machine learning model, but rather include the human expert in the decision process. We approach this challenge by carefully separating technical aspect of the process control from human aspect. The technical aspect is related to data preprocessing, reduction, and unsupervised identification of higher-level concepts (i.e., clusters that represent different stages of the system). The human aspect is related to actual decision making, which is done by the expert in the field. In practice, expert knowledge was encoded in the form of simple rules and automatically executed given the stages of the process discovered in the previous step with a deep learning-based clustering algorithm.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

3 Methodology

This paper considers the practical limitations of typical industrial processes that enforce control changes within a discrete-time framework. As a result, we propose dividing the control determination process into smaller subproblems and creating a control scheme that can establish process control within a single time step and then use it in future steps. Furthermore, we assume that this control scheme should be generic enough to be applicable to all components.

We propose an algorithm that analyzes process parameters, such as temperature and pressure, not only at physical sensor locations but also at virtual sensors placed near them in order to proactively respond to process changes. Hence, the proposed procedure assumes the application of a process simulation model that is used to compute features to be analyzed. Then, data from both sources are combined by time series clustering algorithm to determine the current stage of the process, which can be unambiguously interpreted physically. In the proposed procedure these clusters representing the process stages are interpreted by experts based on domain knowledge and physical parameters of the process. The control actions to be taken at each time step are dependent on the determined process stage resulting in proactive process control. In the proposed approach the control scheme's structure is developed by engineers based on industrial practices. It makes it easily understandable for operators and adaptable to any specific case. An algorithm proposed in this work is presented as a block diagram in Fig. 1.

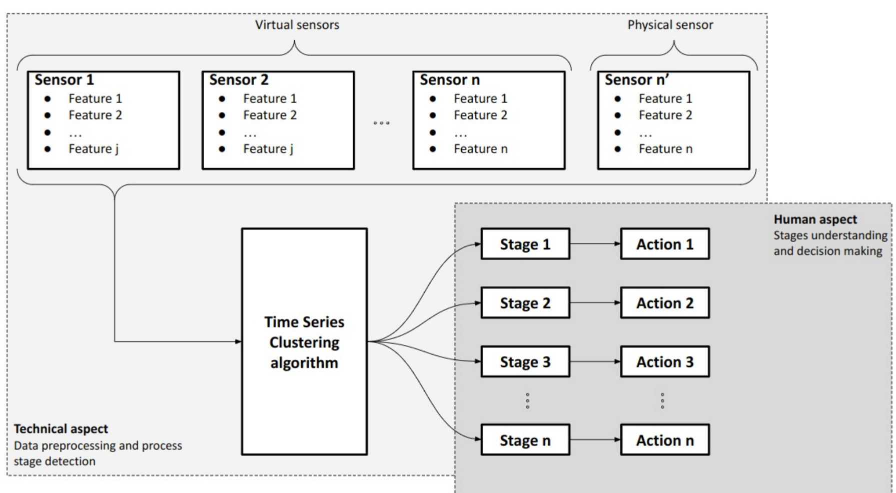

Fig. 1. Block diagram of the proposed ML-based proactive process control.

In this study, we apply deep temporal clustering (DTC) [11] method to define the process stage. The DTC algorithm is a new method for classifying temporal data. It uses advanced deep-learning techniques to reduce the complexity of the data and find patterns that allow the data to be split into different classes. As presented in Fig. 2, the

DOI: 10.1007/978-3-031-36021-3_56

method first uses an autoencoder to reduce the dimensionality of the data, a Convolutional Neural Network (CNN) to identify dominant short-term patterns, and finally a bi-directional Long Short Time Memory neural network (BI-LSTM) to identify longer-term temporal connections between the patterns. The clustering layer then groups the data into classes based on the identified patterns. The neural network architecture employs leaky rectifying linear units (L-ReLUs). The unique aspect of this method is its ability to handle complex temporal data without losing any time-related information, resulting in high classification accuracy.

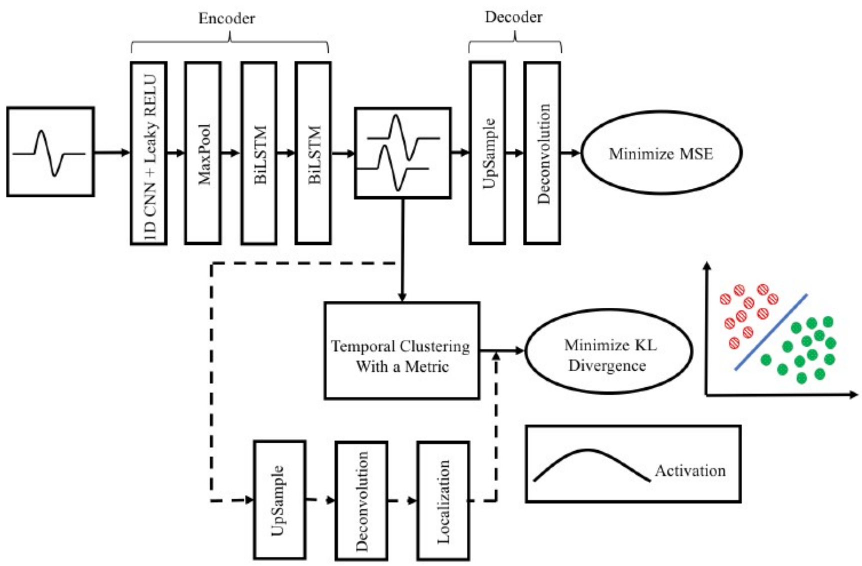

Fig. 2. The overview of the DTC algorithm [11].

The learning process in both the 1D CNN and BI-LSTM models relies on minimizing two cost functions in an interleaved manner [11]. The first cost function is based on the mean square error (MSE) of the reconstructed input sequence obtained from the BI-LSTM latent representation. This cost function ensures that the sequence remains well represented after undergoing dimensionality reduction. The second cost function is derived from the clustering metric, KL divergence. This cost function guarantees that the high-level features defining the subspace spanned by the cluster centroids successfully separate the sequences into distinct clusters exhibiting different spatio-temporal behavior. Optimization of the clustering metric modifies the weights of both the CNN and BI-LSTM models. Consequently, the high-level features encoded by the BI-LSTM model optimally segregate the sequences into clusters, effectively disentangling the spatio-temporal patterns.

To implement this deep learning method, it is necessary to divide the proposed algorithm into two phases: offline learning and online process stage assignment. During

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

the learning phase, the DTC algorithm learns to identify the process stage by analyzing historical values of virtual sensor measurements and physical characteristics. Consequently, the trained algorithm can classify the process stage based on the events that happen around the physical sensors, enabling proactive response.

In the online phase, the trained DTC algorithm analyzes parameter values provided by the simulation model and identifies the current process stage in each time step of the process simulation. Based on the decision scheme proposed, the appropriate control actions to be taken are determined in accordance with the current process stage. These actions are then transferred to the simulation model, resulting in a self-modifying control strategy.

4 Case study: gas production from underground reservoir

To illustrate the performance of the developed solution, it was tested on the problem of gas production from an underground reservoir.

4.1 Problem definition

The majority of gas reservoirs have a production profile that is closely linked to the presence of water. The production of water as a by-product is associated with the conning problem, which is the upward movement of water in the vicinity of the production well. If the critical rate is exceeded and water infiltrates the production well, gas and water will flow out concurrently. This problem is considered one of the most difficult challenges as it can significantly impact the productivity of the well, overall recovery efficiency, and increase production costs. Hence, the main focus of this paper is to identify the appropriate control mechanism that facilitates the reduction of water production.

Reducing well production to the critical rate is a clear solution, but it's not commercially viable. Therefore, proper production control is essential for gas wells that produce water. The production of hydrocarbons is regulated by maintaining a flow rate of wells situated in the reservoir. However, optimizing well control using classical methods is challenging as it is not feasible to establish a functional relationship between the decision variables and the quality indicator being optimized. This is because of the complicated geological structure of hydrocarbon reservoirs, which have complex geometry and spatial heterogeneity of petrophysical parameters. Moreover, the multidimensional physics of fluid flow makes it difficult to describe the system's dynamics. As a result, reservoir simulators are used to model the processes occurring in hydrocarbon reservoirs, which can predict the behavior of the reservoir.

Given the practical constraints of the gas production process that require discrete control changes, it is possible to break down the problem of well control determination into sub-problems. Therefore, modeling a control scheme that determines individual well control in a single time step is sufficient [8]. However, it remains challenging to accurately determine the actions that should be taken at a particular production stage to prevent water infiltration and optimize gas production. In industrial practice, gas wells producing water are typically managed using a reactive control approach that relies

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

on trial and error. However, a proactive control approach that takes preemptive control actions before undesirable changes in flow reach the well can improve the techno-economic efficiency of the production process. To go from reactive to proactive control this paper proposes a solution that utilizes reservoir simulation and virtual active sensors located near the production well. By monitoring reservoir parameters before undesired flow changes reach the well, these virtual sensors enable proactive inflow control and form the foundation for the proposed control system for gas wells conning water.

4.2 Methodology adjustment for the well control

In the case of well control, the idea of the proposed solution is to analyze reservoir parameters surrounding a production well measured by the virtual sensors combined with ground production data. Then, in each time step of the reservoir simulation, use the trained DTC algorithm to determine the current production stage of the well based on the reservoir parameters computed by the reservoir simulator. The control of gas wells that are affected by water influx is then automatically determined based on the production stage of the well.

4.3 Proposed solution application

The solution proposed in this paper was evaluated by its application on a real gas reservoir where a water coning problem occurs. The reservoir is modelled with the ECLIPSE reservoir simulator (Schlumberger Limited, Houston, TX, USA) that uses finite difference methods to model fluid flow in complex geological formations and well configurations. It allows for the simulation of a wide range of scenarios and is used for predicting reservoir performance, improving reservoir understanding and optimizing production strategies. The behavior of the reservoir is described by a system of non-linear partial differential equations that describe fluid flow in porous media [2]. The equations are solved numerically using Newton's iterative method.

In the case of the analyzed reservoir, there are 8 production wells, and 3 virtual active sensors are located in close proximity to each well as shown in Fig. 3. The ground production attributes, such as gas and water flow rates, are monitored in each well and the virtual sensors measure water saturation and reservoir pressure in the vicinity of the well to determine its production stage as shown in Fig. 4. These reservoir parameters are analyzed by the DTC model to detect clusters corresponding to the stages of the production process.

The proposed solution was applied to a historical production period to evaluate its effectiveness. In the offline stage, the DTC model was trained using multivariable historical data from the initial 20 years of production. The next 20 years of historical production data were used for validation. Both training and validation data are generated by the numerical reservoir model.

For the DTC model learning, data related to the well with the most complex reservoir situation with respect to water production (the highest water production) were used as it is the most representative. All the analyzed features measured by both physical and virtual sensors are represented as floating point values sampled at regular intervals (in the example provided, every hour). Hence, the input data include 8 attributes measured

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

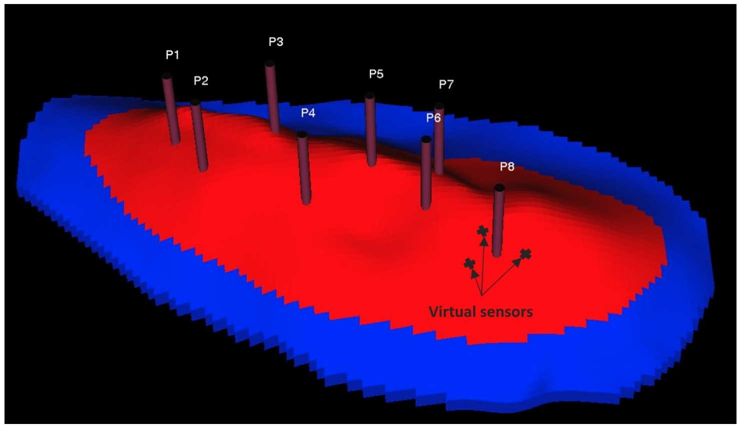

Fig. 3. Visualization of the analyzed reservoir (example of the virtual sensors location for P8 well).

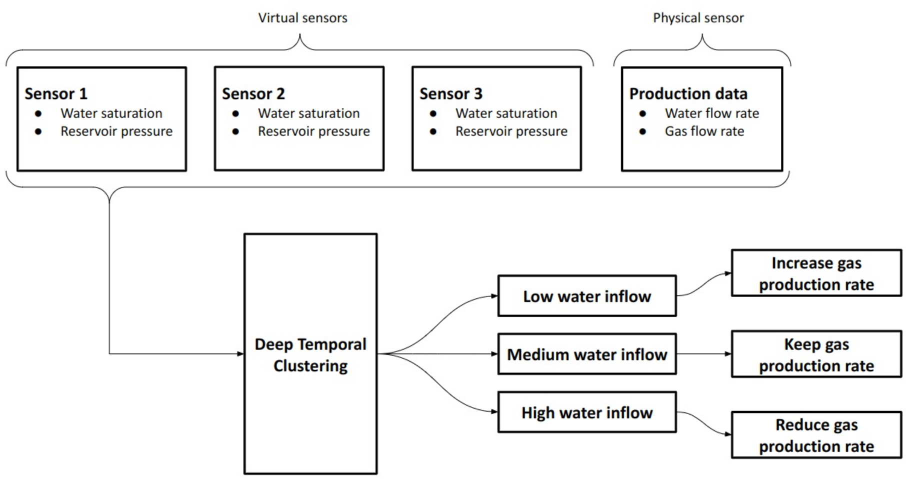

Fig. 4. Block diagram of the proposed control scheme for gas production for the underground reservoir.

in 175 200 time steps. The DTC model was trained using the Adam optimizer with a learning rate 0.01. Parameters of the neural network architecture were selected based on the values proposed in [11]. The output of the DTC model is an array with a class label assigned to each record of input data described above.

In the presented approach, the number of clusters was selected by the human expert based on the characteristics of the analyzed process. In the case of the water con-

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

ning problem, three distinct stages can be identified in the production process. The first stage involves low water inflow during the initial production period. The second stage is characterized by a gradual increase in water inflow. Finally, when water inflow reaches a high level, it indicates the final stage of production from the well, which is no longer economically viable. Taking into account this phenomenon, the proposed control scheme also includes 3 stages, to keep the solution easily understandable for the operator. When the trained DTC model assigned one of the three clusters to each time step, these clusters were interpreted by experts based on water saturation in the virtual sensor presented in Fig. 5. Cluster 0 represents low water inflow, cluster 1 denotes medium water inflow, and cluster 2 indicates high water inflow. This interpretation facilitates a clear physical understanding of the suggested gas well control.

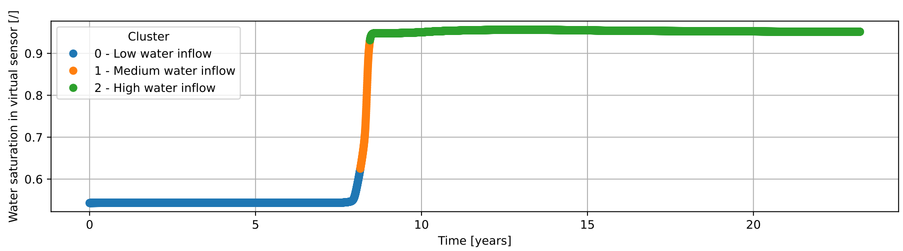

Fig. 5. Block diagram of the proposed ML-based proactive process control.

The online phase was tested using the last 20 years of production data. In this phase, the trained DTC model predicts the production stage of each well at every time step of the reservoir simulation. The well rates are then automatically adjusted according to the decision scheme presented in Fig. 4 that is developed based on engineering knowledge and simulation tests. The decision scheme involves categorizing production wells into three groups that correspond to the interpreted DTC clusters. If a well exhibits high water inflow, its flow rate is decreased. If water inflow is at a medium level, the well performance remains unchanged. However, for wells with low water inflow, the well rate is increased. This approach to determining the appropriate actions to be taken at each production stage results in self-modifying intelligent well control.

4.4 Results

The application of the developed solution resulted in proactive control of the production wells located on the analyzed gas reservoir. This intelligent control allows cumulative water production from the reservoir to be reduced by 15% compared to historical data based on operator experience (Fig. 6). Cumulative gas production remained at the same level, indicating increased production efficiency. The results demonstrate that the developed control strategy accurately manages the water coning phenomenon, reduces operational costs, and increases overall income without incurring extra expenses, as only the control system is changed. Furthermore, the difference between the results obtained using the developed control system and the historical data increases with time,

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

implying that the ML-based solution is more effective in complicated production situations as with time water saturation in the vicinity of the production wells is higher. Thus, the generated control outperforms human control, but the operator can still verify it.

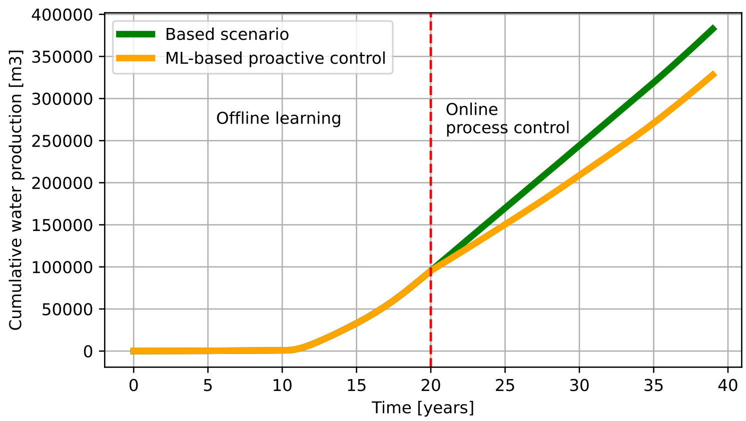

Fig. 6. Comparison of water production depending on the process control).

5 Discussion and potential applications

The approach proposed in this paper enables better management of the process under consideration. As data from virtual sensors are incorporated into this procedure, it allows for reacting proactively. However, the limitation of the proposed proactive control is that there is a need to use a process simulation model which may be difficult in the case of real-time applications. However, the trained ML model defining the process stage can be used directly (without simulation) as long as the process simulation model used to train it is valid.

Thanks to the use of advanced machine learning algorithms, the appropriate control can be determined also in situations that are not trivial for a human. As input features from both physical and virtual sensors are combined into a process stage that is physically interpretable it makes the whole decision procedure easily explainable what is the main differentiator of the proposed method.

However, the method presented here is not fully automatic as it requires interaction with a domain expert. In the proposed method, the time series clustering algorithm combines the simulation data into clusters corresponding to process stages, but the meaning of these clusters is given by an expert based on domain knowledge about the analyzed process and values of physical parameters.

In the developed solution, the final structure of the decision scheme is developed by domain experts based on the best industrial practices and domain knowledge. It

DOI: 10.1007/978-3-031-36021-3_56

makes the decision process understandable for the operator, which is especially important considering real applications. In addition, the decision scheme can be adapted to any particular process which makes the proposed method general.

As the proposed solution is generic, it can be applied to process control at different levels of granulation which confirms its scalability. It can be used to define the control of the process as a whole. The proposed approach can be applied also to each component of the process. For example, the temperature of each component of the wind turbine can be measured and the proposed actions can be enough general to be applicable to all components. Another possibility is to apply this approach to a fleet of assets, such as wind turbines for which the same measurements are available. Then, the scheme can be used to define the control of the individual wind turbine and utilized for all of the assets in the fleet as in the presented example of gas wells.

The proposed method can be also adapted as a predictive maintenance tool. In this case, the process stage defined by ML algorithm can be interpreted as the component wear. For an example of a bearing, its temperature can be measured by the physical sensor and virtual sensors can tell about the temperature distribution around this physical measurement to react proactively. Then, the process stage determined based on the ML algorithm, assuming three clusters, can present a component in good condition, a component in wear, or a faulty component. As virtual sensors are utilized the proposed approach makes it possible to proactively predict the moment when maintenance is needed. As the proposed procedure not only defines the process stage but also defines actions to be taken it can also be used as a prescriptive maintenance tool. These actions can regulate operating conditions which allows asset maintenance to be scheduled in order to achieve specific objectives.

Future work of the study reported here includes replacing the simple control actions directly dependent on the process stage with an auto-adaptive decision tree introduced in [9, 10] where control actions are parametrized and can be optimized to reach optimal control. Since the input for this parameterized decision tree comprises simulation data rather than the process stage, the decision-making process is entirely automated while remaining straightforward to interpret. Therefore, it would be reasonable to extend the proposed procedure with an auto-adaptive parameterized decision tree as an automated and directly interpretable method.

6 Conclusion

The paper proposes a new method for proactive control of industrial processes using machine learning techniques. The approach involves combining physical and virtual sensor data obtained from a simulation model to identify the current stage of the process and proactively determine appropriate control actions. The deep temporal clustering algorithm is used to combine simulation data and identify the process stage, and a decision scheme proposed by engineers is used to determine control actions dependent on the identified stage.

The performance of the developed solution was illustrated in the problem of gas production from an underground reservoir. In this case, the proposed solution allowed for managing the considered water coning phenomenon, reducing of unfavourable water

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56

inflow, and increasing overall income without incurring extra expenses, as only the control system is changed.

The proposed method is generic and scalable. It can be applied to a single asset or multiple assets if the simulation model includes a fleet of them. It can also be adapted as a predictive and prescriptive maintenance tool. Future work includes optimizing the control actions to reach optimal control.

Acknowledgements This paper is funded from the XPM (Explainable Predictive Maintenance) project funded by the National Science Center, Poland under CHIST-ERA programme Grant Agreement No. 857925 (NCN UMO-2020/02/Y/ST6/00070)

References

1. Angelopoulos, A., Michailidis, E.T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S., Zahariadis, T.: Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects. Sensors **20**(1), 109 (2019)
2. Busswell, G., Banergee, R., Thambynayagam, M., Spath, J.: Generalized analytical solution for solution for reservoir problems with multiple wells and boundary conditions (04 2006). https://doi.org/10.2118/99288-MS
3. Casolla, G., Cuomo, S., Schiano Di Cola, V., Piccialli, F.: Exploring unsupervised learning techniques for the internet of things. IEEE Transactions on Industrial Informatics **PP**, 1–1 (09 2019). https://doi.org/10.1109/TII.2019.2941142
4. Dimitrov, N., Göçmen, T.: Virtual sensors for wind turbines with machine learning-based time series models. Wind Energy **25**(9), 1626–1645 (2022)
5. Gajjar, S., Kulahci, M., Palazoglu, A.: Real-time fault detection and diagnosis using sparse principal component analysis. Journal of Process Control **67**, 112–128 (2018). https://doi.org/https://doi.org/10.1016/j.jprocont.2017.03.005, https://www.sciencedirect.com/science/article/pii/S0959152417300677, big Data: Data Science for Process Control and Operations
6. Garcia, E., Montés, N., Llopis, J., Lacasa, A.: Miniterm, a novel virtual sensor for predictive maintenance for the industry 4.0 era. Sensors **22**(16) (2022). https://doi.org/10.3390/s22166222, https://www.mdpi.com/1424-8220/22/16/6222
7. Jiang, Y., Yin, S., Dong, J., Kaynak, O.: A review on soft sensors for monitoring, control, and optimization of industrial processes. IEEE Sensors Journal **21**(11), 12868–12881 (2021). https://doi.org/10.1109/JSEN.2020.3033153
8. Kuk, E.: Application of Artificial Intelligence Methods to Underground Gas Storage Control. SPE Annual Technical Conference and Exhibition, vol. Day 2 Tue, October 01, 2019 (09 2019). https://doi.org/10.2118/200305-STU, https://doi.org/10.2118/200305-STU, d023S103R025
9. Kuk, E., Stopa, J., Kuk, M., Janiga, D., Wojnarowski, P.: Petroleum reservoir control optimization with the use of the auto-adaptive decision trees. Energies **14**(18) (2021). https://doi.org/10.3390/en14185702, https://www.mdpi.com/1996-1073/14/18/5702
10. Kuk, E., Stopa, J., Kuk, M., Janiga, D., Wojnarowski, P.: Optimal well control based on auto-adaptive decision treemdash;maximizing energy efficiency in high-nitrogen underground gas storage. Energies **15**(9) (2022). https://doi.org/10.3390/en15093413, https://www.mdpi.com/1996-1073/15/9/3413
11. Madiraju, N.S.: Deep temporal clustering: Fully unsupervised learning of time-domain features. Ph.D. thesis, Arizona State University (2018)

12. Malykhina, G.: Digital twin technology as a basis of the industry in future. pp. 416–428 (12 2018). https://doi.org/10.15405/epsbs.2018.12.02.45
13. Qin, S.J., Chiang, L.H.: Advances and opportunities in machine learning for process data analytics. Computers Chemical Engineering **126**, 465–473 (2019). https://doi.org/https://doi.org/10.1016/j.compchemeng.2019.04.003, https://www.sciencedirect.com/science/article/pii/S0098135419302248
14. Tran, P.H., Ahmadi Nadi, A., Nguyen, T.H., Tran, K.D., Tran, K.P.: Application of Machine Learning in Statistical Process Control Charts: A Survey and Perspective, pp. 7–42. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-030-83819-5_2, https://doi.org/10.1007/978-3-030-83819-5_2
15. Usuga Cadavid, J.P., Lamouri, S., Grabot, B., Pellerin, R., Fortin, A.: Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. Journal of Intelligent Manufacturing **31** (08 2020). https://doi.org/10.1007/s10845-019-01531-7
16. Vallejo, M., de la Espriella, C., Gómez-Santamaría, J., Ramírez-Barrera, A.F., Delgado-Trejos, E.: Soft metrology based on machine learning: a review. Measurement Science and Technology **31**(3), 032001 (dec 2019). https://doi.org/10.1088/1361-6501/ab4b39, https://dx.doi.org/10.1088/1361-6501/ab4b39
17. Vazan, P., Znamenak, J., Juhas, M.: Proactive simulation in production line control. In: 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT). vol. 1, pp. 52–55 (2018). https://doi.org/10.1109/STC-CSIT.2018.8526675
18. Xu, X., Lu, Y., Vogel-Heuser, B., Wang, L.: Industry 4.0 and industry 5.0—inception, conception and perception. Journal of Manufacturing Systems **61**, 530–535 (2021). https://doi.org/https://doi.org/10.1016/j.jmsy.2021.10.006, https://www.sciencedirect.com/science/article/pii/S0278612521002119
19. Yang, H., Rao, P., Simpson, T., Lu, Y., Witherell, P., Nassar, A., Reutzel, E., Kumara, S.: Six-sigma quality management of additive manufacturing. Proceedings of the IEEE **PP** (10 2020). https://doi.org/10.1109/JPROC.2020.3034519
20. Zerilli, J., Knott, A., Maclaurin, J., Gavaghan, C.: Algorithmic decision-making and the control problem. Minds and Machines **29** (12 2019). https://doi.org/10.1007/s11023-019-09513-7
21. Zhu, Q., Qin, S.: Supervised diagnosis of quality and process faults with statistical learning models. In: I&EC research revised (2019)

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36021-3_56