

Kins, Ruben; Möbitz, Christian; Gries, Thomas

**Article — Published Version**

## Towards autonomous learning and optimisation in textile production: data-driven simulation approach for optimiser validation

Journal of Intelligent Manufacturing

*Suggested Citation:* Kins, Ruben; Möbitz, Christian; Gries, Thomas (2024) : Towards autonomous learning and optimisation in textile production: data-driven simulation approach for optimiser validation, Journal of Intelligent Manufacturing, ISSN 1572-8145, Springer US, New York, Vol. 36, Iss. 5, pp. 3483-3508,  
<https://doi.org/10.1007/s10845-024-02405-3>

This Version is available at:

<https://hdl.handle.net/10419/323389>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by/4.0/>



# Towards autonomous learning and optimisation in textile production: data-driven simulation approach for optimiser validation

Ruben Kins<sup>1</sup> · Christian Möbitz<sup>1</sup> · Thomas Gries<sup>1</sup>

Received: 18 January 2024 / Accepted: 12 April 2024 / Published online: 24 May 2024  
© The Author(s) 2024

## Abstract

The textile industry is a traditional industry branch that remains highly relevant in Europe. The industry is under pressure to remain profitable in this high-wage region. As one promising approach, data-driven methods can be used for process optimisation in order to reduce waste, increase profitability and relieve mental burden on staff members. However, approaches from research rarely get adopted into practice. We identify the high dimensionality of textile production processes leading to high model uncertainty as well as an incomplete problem formulation as the two main problems. We argue that some form of an autonomous learning agent can address this challenge, when it safely explores advantageous, unknown new settings by interacting with the process. Our main goal is to facilitate the adoption of promising research into practical applications. The main contributions of this paper include the derivation and formulation of a probabilistic optimisation problem for high-dimensional, stationary production processes. We also create a highly adaptable simulation of the textile carded nonwovens production process in Python that implements the optimisation problem. Economic and technical behavior of the process is approximated using both Gaussian Process Regression (GPR) models trained with industrial data as well as physics-motivated explicit models. This 'simulation first'-approach makes the development of autonomous learning agents for practical applications feasible because it allows for cheap testing and validation before physical trials. Future work will include the comparison of the performance of different agent approaches.

**Keywords** Textile Production · Safe Optimisation · Process Optimisation · Simulation Environment

## Introduction

The technical textiles industry remains highly relevant in Europe, despite the general outsourcing of production to low wage countries. As an example, the turnover of nonwovens, which are primarily used as technical and hygiene textiles, amounted to 9.2 bn € in 2019, with an increase of 16.3 % in produced weight from 2014 to 2019 (EDANA, 2022). The industrial production of nonwovens is highly complex due to a high number of intermediate process steps

with many degrees of freedom (Brydon & Brydon, 2007). Due to the complexity of the production process the process setup is largely done by trial and error (He et al., 2021). This leads to wasted material, energy and time if the product does not conform to the specifications or if the process setup is not optimal. As a consequence, there is a huge potential in making the production of nonwovens more economical and sustainable by reducing waste. Producing companies in high-wage countries are under pressure to operate their processes close to their optimal operating points to compensate for higher costs. This results in overhead costs for complex production management approaches (Büscher et al., 2014). Intelligent, cognitive production planning systems are a promising approach to reduce the tension in this 'Polylemma of Production' (Büscher et al., 2014). The process setup and optimisation is traditionally done by machine operators that have to manage several conflicting goals simultaneously. These goals include minimizing the cost per unit, maximising the overall productivity and adhering to the prod-

✉ Ruben Kins  
ruben.kins@ita.rwth-aachen.de

Christian Möbitz  
christian.moebitz@ita.rwth-aachen.de

Thomas Gries  
thomas.gries@ita.rwth-aachen.de

<sup>1</sup> Institut für Textiltechnik, RWTH Aachen University, Aachen, Germany

uct quality specifications (He et al., 2022; Vedpal, 2013). For this task the operators have to process a lot of data and information, which leads to an increased mental workload that results in a feeling of being overwhelmed and ineffective (Ledzińska & Postek, 2017). Recent advancements in research focused on the development of self-learning systems, predominantly utilizing techniques from Reinforcement Learning (RL) and Bayesian Optimization (BO). These approaches seek to increase system performance by finding an optimal Policy or an optimal set of parameters respectively (Brunke et al., 2022). Intelligent agents in the context of industrial textile production have the potential to not only optimise economical (He et al., 2022) and sustainability aspects of textile production processes but also to relieve operators of the mental burden to balance several optimisation goals. RL and BO are usually tested, validated and, in the case of RL, pre-trained using virtual simulation environments before being deployed in the real world. The reason for this is that real world experiments are prohibitively expensive while a simulation allows for cheap experiments with millions of time steps (Scheiderer et al., 2020; Ibarz et al., 2021).

Our contributions in this work include the following achievements:

1. Identification of the shortcomings of current optimisation approaches from research that hinder adoption into practice. We address both the issues of optimisation problem formulation as well as handling high dimensionality.
2. Proposal of an approach about how to overcome the shortcomings using Autonomous Learning Agents (ALA).
3. Explicit formulation of the probabilistic constrained optimisation problem for a technical textile production process.
4. Groundwork for future contributions by providing a simulation architecture including its implementation. This simulation of a technical textile production process allows cheap testing of promising approaches from research to facilitate adoption into practice.
5. High adaptability and modularity facilitate transfer of our simulation into other domains.

In this work, we identify the shortcomings of traditional optimisation approaches from industry and research and formulate steps to overcome them. We also want to encourage adoption of suitable techniques from research into practice. To reach this goal, we formulate the full optimisation problem with regards to industrial requirements. Additionally, we develop and provide a simulation of a textile nonwoven production process for cheap and quick testing and validation of intelligent agents.

## Theoretical background

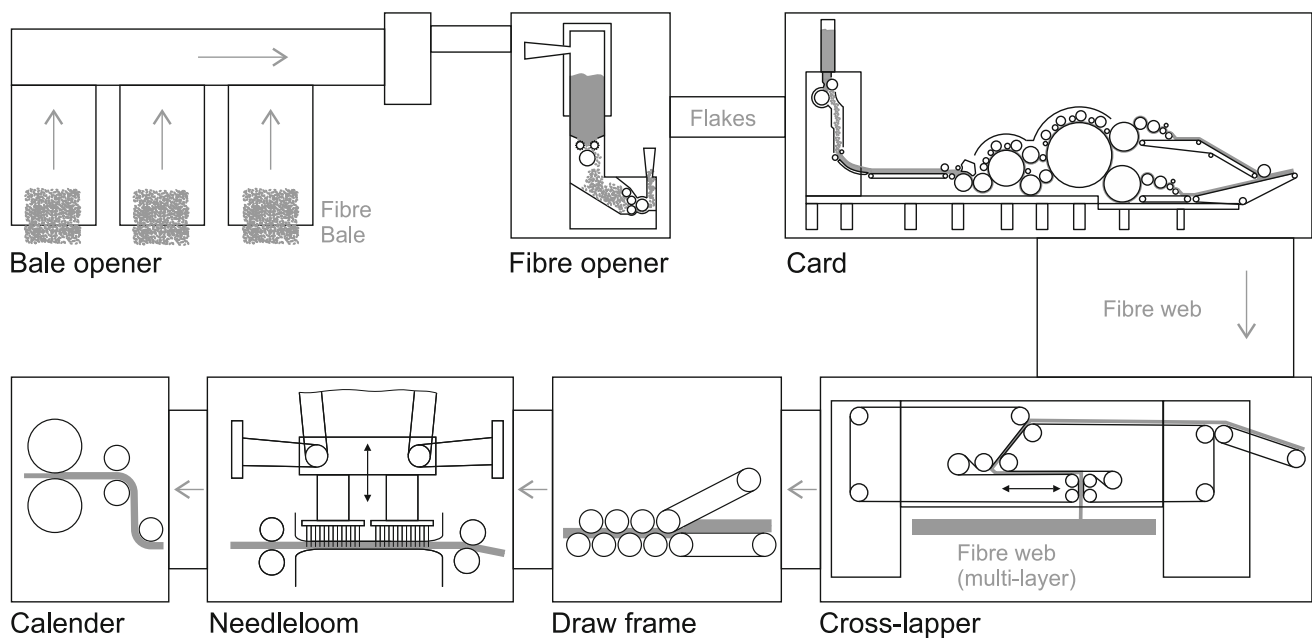
In order to establish a common understanding between both Production Engineers as well as ML Researchers, theoretical background for the process as well as the optimisation problem at hand is given. This chapter introduces the nonwovens production process, traditional process control methods as well as the usage of RL and BO as intelligent agents for process optimisation.

### Nonwoven production process

Nonwovens are widely used in many hygiene, construction, automotive, filtration and medical applications due to their wide range of possible properties (Wilson, 2007). Generally, nonwovens are produced by forming fibres or filaments directly into a surface structure without using methods like weaving or knitting (Wilson, 2007). The process is highly productive because there is no need for producing a yarn as an intermediate step. There are many ways to produce nonwovens, so we focus on the carded nonwovens production process as one of the most widespread methods (Brydon & Brydon, 2007). An overview of the process is presented in 1.

The process begins with opening staple fibre bales in parallel to achieve the desired fibre mixture. The properties of each of the processed fibre components include material, length, diameter, cross section geometry, crimp, tensile strength as well as lubrication on the surface, offering many degrees of freedom in material design. Because the fibre properties cannot be controlled by the nonwovens producer, deviations from the specification are an important disturbance on the process. The fibre flakes from the bale opener are further opened and mixed in subsequent steps to achieve a homogeneous material mixture. (Brydon & Brydon, 2007)

In the next step the fibres pass the carding machine, which consists of a sequence of at least 11 rollers equipped with toothed wires (called card clothing). There exists a multitude of possible cylinder numbers and arrangements (Schlichter et al., 2012). The card disentangles fibres and forms them into a homogeneous web that is uniform in weight per area and thickness (Brydon & Brydon, 2007) in both Machine/Production Direction (MD) and the perpendicular Cross Direction (CD). The speed of each cylinder can be set, influencing the fibre processing as well as their movement from cylinder to cylinder. The toothed wires, which are in direct contact with the fibres, each have a different geometry based on their purpose. Many times, the card represents the process bottleneck when it comes to fibre mass throughput (Schlichter et al., 2012). Cards have many different designs: Cloppenburg (2019) counts 12 card clothing degrees of freedom for each cylinder. Additionally, the position and velocity of each cylinder can be set. This results in 154 degrees of freedom for an 11 cylinder card for wire choice alone.



**Fig. 1** The carded nonwovens production process (Pohlmeyer et al., 2022; Brydon & Brydon, 2007)

The next process step is the Cross-lapper, which decouples the card web properties (e.g. width and weight per area) from the final product by stacking a desired amount of card web layers on top of each other with a given product width (Brydon & Brydon, 2007) (see figure 1). The Draw Frame reorients the fibres inside the web from CD to MD by drafting or stretching it. This is achieved by a series of cylinder pairs. The cylinder velocities increase with each pair, so that the fibre web is pulled apart with a constant ratio called draft ratio. Subsequent machines also have a draft ratio. The drafting process is an example for the complexity of process control: Changing the draft ratio influences the tensile strength in MD/CD, the weight per area, the fibre web evenness as well as the production speed (Brydon & Brydon, 2007).

During the next step, called needlepunching, the web is punctuated by hundreds to thousands of serrated needles that oscillate vertically inside a needleloom to entangle the fibres. The fibre web is consolidated and receives its mechanical strength due to the entanglement. The feed per stroke setting changes the length the fibre web gets pushed forward per oscillation cycle. This method presents a bottleneck on the possible production speeds because a higher speed increases mechanical stress on the needles, causing them to break. The geometry of the needles as well as the arrangement of the needles on the needle boards also have a huge effect on the final mechanical strength of the nonwoven. The entanglement of fibres contracts the fibre web, which has a stronger effect on the edges of the textile. This leads to a higher weight per area at the edges compared to the middle, which is the

so-called 'smile effect'. The smile effect can be counteracted by drafting the nonwoven periodically before cross-lapping, cancelling out the effect.

Thermoplastic nonwovens can be further strengthened in the calender under the influence of heat and pressure. Calendaring means to guide the nonwoven between a pair of heated cylinders while exerting a certain pressure. This causes the fibres to partially melt and consolidate, increasing the nonwoven strength, but also possibly damaging it. Finally, the nonwoven is finished and wound up on rolls at the winder. (Brydon & Brydon, 2007; Schlichter et al., 2012)

The nonwoven as a product needs to conform to requirements as stated by the contract between customer and producer. The requirements often include a tolerance for the weight per unit area, a minimum tensile strength and optical requirements (Mao & Russel, 2007). Testing the product requires costly lab measurements (like a tensile tester (ISO, 2007)) that cause manual labour and destroy the sample (Mao & Russel, 2007). If a nonwoven is found non-conforming, it cannot be sold, which leads to waste of material, time and energy (Mao & Russel, 2007; Vedpal, 2013).

The production of nonwovens is considered a continuous, stationary production process because the process properties stay mostly constant while the nonwoven is continuously produced. The process is controlled by applying setpoint values, e.g. cylinder speeds, for the lower control loops (Vedpal, 2013). In the production process many machines are chained together. Each machine has many degrees of freedom both in mechanical design (like card cylinder layout) and setup (like needle and toothed wire choice) (Brydon & Brydon, 2007).

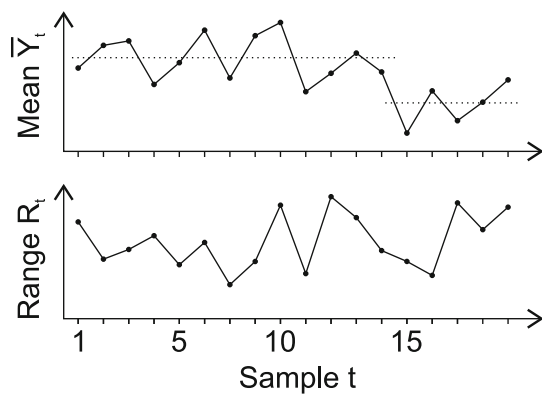


Fig. 2 Example of a continuous Shewhart chart

Also, there is a huge choice of fibre and nonwoven properties (Wilson, 2007). This leads to the conclusion that, due to high dimensionality, there are likely no two completely similar nonwoven production lines regarding layout and setup unless specifically designed like that.

### Process control in textile production engineering

A very popular method of improving production processes is called Lean Six Sigma. It aims to reduce costs, reduce defects and improve cycle times among other goals (Pongboonchai-Empl et al., 2023).

Process control in textile production has a more narrow focus: It aims to ensure that the process is under control, meaning that the process output reliably meets the desired specifications and does not produce defects. This is especially important because production environments are dynamic, meaning that production context like customer expectations, process and raw material properties change constantly (Vedpal, 2013).

The first activity in process control is sampling the process quality in order to evaluate it. If a quality defect is found, so-called internal failure costs are incurred: The defective product is either scrapped, meaning the time, material and energy for its production is lost or it is reworked so it meets specifications (Qiu, 2013). If the defect is not found, external failure costs result if the customer finds the defect. External failure costs outweigh internal costs long-term (Qiu, 2013). Sampling the quality also incurs so-called appraisal costs for labour and equipment (Aslam et al., 2021). Our goal is not to reduce these costs, but rather to use sampling as effectively as possible.

The most widely used method of process control in industrial textile production is called Statistical Process Control (SPC) which is a collection of methods from both descriptive as well as inferential statistics (Vedpal, 2013). One widely applied method of SPC is the Shewhart Chart (Aslam et al., 2021). An example can be seen in 2.

The top of the chart in Fig. 2 shows the mean of some quality characteristic  $\bar{Y}_{i,t}$  of the  $t$ -th sample of  $m$  products (Qiu, 2013) according to the following formula:

$$\bar{Y}_{i,t} = \frac{1}{m} \sum_{j=1}^m Y_{i,t,j} \quad \text{for } t = 1, 2, \dots, n \quad (1)$$

*Note that we changed the notation from the original authors for consistency within this paper.* The top  $\bar{Y}$  chart is used to detect systematic changes in the statistical properties of a process. When the control limits  $U$  or  $L$  (which are the confidence bounds of some form of test statistic) are crossed, measures need to be undertaken. The bottom  $R$  chart visualises the process variability using sample ranges (Qiu, 2013). An example for the nonwovens process is the tensile strength, which needs to be sampled at least five times (ISO, 2007) for one measurement. Since they are the actual measure according to the norm, we will refer to the sample means as  $Y_t$  instead of  $\bar{Y}_t$  from now on. The Shewhart chart is only concerned with changes of statistical properties of a process, but not its ability to produce conforming products. For the simple case of a univariate, normally distributed characteristic and only a Lower Specification Limit ( $LSL$ , the compliment being the Upper Specification Limit  $USL$ ) the lower process capability ratio  $C_{pl}$  is defined as (Qiu, 2013):

$$C_{pl,i} = \frac{\mu_i - LSL}{3\sigma_i}, \quad Y_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad (2)$$

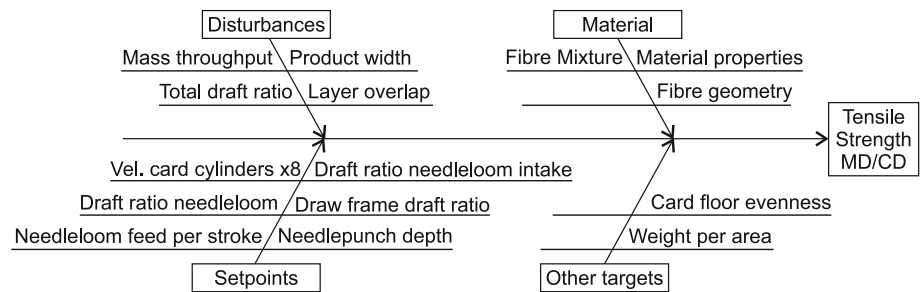
Equation 2 describes the capability of the process to produce outputs  $i$  within specified limits. The capability describes the distance of the process mean to the specification limits measured in three standard deviations. That means  $C_{pl,i} = 1$  is equal to a defect rate of 66.807 PPM. In order to improve process capability for an output (like a product quality criterion), we can increase its mean  $\mu_i$  or reduce its variability  $\sigma_i$  (Qiu, 2013). Increasing  $\mu$  can be costly, for example to increase the average weight per area in a nonwovens production process, one has to spend more raw material, creating a trade-off between additional material and failure costs.

SPC allows us to detect anomalies or systematic changes in a process. However, no action recommendations can be derived with this method as it has no predictive capabilities. Instead, brainstorming techniques and so-called cause-and-effect diagrams are used to derive corrective actions manually (Vedpal, 2013). Such a diagram is shown in 3 using expert interviews as data source.

To conclude, traditional methods for process control apply proven statistical methods for stabilizing a process. They are limited to operator knowledge for improvement and do not address economical aspects.



**Fig. 3** Cause-and-effect diagram for the tensile strength of a nonwoven in MD and CD



## ML in textile production engineering

Many works try to improve upon traditional statistical methods by training predictive ML models with lab sample data of process quality characteristics (He et al., 2021; Weichert et al., 2019). Additionally, better process setpoints are suggested by applying an optimisation algorithm on top of the models (Weichert et al., 2019).

Predicting the process outcome is an advantage over SPC methods, because corrective measures can be automatically simulated and/or optimised before being applied to the process (He et al., 2021). ML in textile production engineering has been addressed in research literature. He, Shu et al. conducted a literature review and found a total of 5 works about modelling the nonwovens production process as well as 32 works about other textile production processes (He et al., 2021). Our works include the model-driven optimisation of the pneumatic fibre transport (Möbitz, 2021). They found that, for the spinning process, a total of 48 Fibre properties as well as 71 process parameters were used as model inputs. This highlights what some call the 'curse of dimensionality' in multistage manufacturing processes (Ismail et al., 2022) where hundreds (Weichert et al., 2019) of potentially influential process parameters exponentially increase the model complexity and need for training data. For textile processes, 21 different process quality characteristics were modelled. The characteristics were predicted using several ML methods, Artificial Neural Networks (ANN) being by far the most popular one (He et al., 2021).

The general assumption of ML modeling is that the expected production outcome  $Y$  is systematically influenced by the process variables  $\mathbf{X}$ . The systematic behaviour can be predicted by a parameterised complex function  $f$  like an ANN (Goodfellow et al., 2016):

$$\mathbf{E}[Y|\mathbf{X} = \mathbf{x}] = f(\mathbf{x}; \theta) \quad (3)$$

where  $\mathbf{x} = [\mathbf{s}^T, \mathbf{d}^T]^T$

In the above equation, the process parameters  $\mathbf{X}$  consist of both setpoint variables  $\mathbf{s}$  and disturbances  $\mathbf{d}$  (like environmental conditions and fibre deviations from specification). The expectation operator indicates that the model provides

only a point estimate instead of a prediction that includes process output uncertainties. The parameters are found by minimising a loss function  $L$  that quantifies the error between prediction  $f(\mathbf{x}_i; \theta)$  and sample  $y_i$  (Russell & Norvig, 2016):

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{n} \sum_{i=1}^n L(y_i, f(\mathbf{x}_i; \theta)) \quad (4)$$

He et al. (2021) conclude their review by stating that current research of ML in Textile Production Engineering has not lead to industrial applications. This is despite being around since at least 1993 and a steadily growing number of publications (He et al., 2021). The authors criticise that all studies did not address the true complexity of textile production processes (He et al., 2021). Other meta-studies of ML-enabled Lean Six Sigma research found that many authors mention a lack of empirical data from practical implementations of their research (Pongboonchai-Empl et al., 2023).

We conclude by stating two central disadvantages that hinder practical application of ML research in textile production engineering into practice.

**Disadvantage 1: Models generalise poorly** Textile production processes are subject to what is called 'curse of dimensionality'. With a higher number of process states  $\mathbf{X}$  the problem complexity and amount of data needed rises exponentially (Sutton et al., 2018). He et al. (2021) are optimistic that with the rise of the Industrial Internet of Things (IIoT) more high quality data from the textile industry will be available. However, the curse of dimensionality is still a fundamental issue. Data from a lab setting will also not be able to address this issue. Lab machines allow for systematic variation of some setpoints and disturbances. However, many process layout, machine design and machine setup parameters are different between lab and industry settings, especially the scale (see Section "Nonwoven Production Process"). The difference in process setup means that models learned from lab experiments do not generalise well to industrial settings. It is highly unlikely that there will be a universal model that accurately predicts the process behaviour across all dimensions. That means that there needs to be a specialised model for each combination of machine setup, process layout and fibre mixture. What follows is that an accurate model of any

production process needs to be learned with data that has been sampled on that specific process. To account for differences between different processes, we add process layout, machine design and machine setup variables  $l$  to the process states:  $\mathbf{x} = [\mathbf{s}^T, \mathbf{d}^T, \mathbf{l}^T]^T$ . Should a model disregard important disturbances, we have high external model uncertainties. Should a model disregard important, mostly unchangeable process layout or machine design influences, we deal with high constant uncertainties in our system model. Both lead to unreliable predictions and thus possible damage to equipment or economic losses.

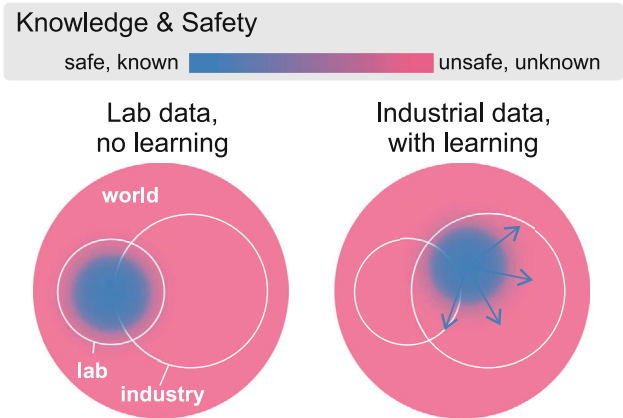
**Disadvantage 2: Unsuitable problem formulation for real-world textile production** He et al. (2021) criticise that available literature is mainly concerned with process modeling of singular quality aspects, but not the full optimisation problem at hand. It remains unclear how the model can be used to derive corrective actions without the full production context, not offering advantages over SPC.

As Thomassey and Zeng (2018) note, there is a high uncertainty in decision making in the garment industry. However, traditional methods offer only point estimates (as seen in 3), no quantification of uncertainty. As seen in section 2.2 the quantification of process performance in SPC requires an uncertainty estimate of the random uncertainty due to the inherent nature of the process. Setting  $\mu = LSB$  would lead to many quality defects, highlighting the need for random uncertainty estimates. In ML literature, inherent randomness or unpredictability in a process is called *Aleatoric Uncertainty*. It cannot be reduced by additional sampling. SPC literature calls this phenomenon *Process Noise* (Qiu, 2013). A suitable problem formulation for industrial process optimisation needs to address the full production context (including economic and quality aspects) instead of singular aspects as well as random uncertainty from process noise.

## Hypothesis formulation

While *disadvantage 2* can be solved by a proper problem formulation and choice of ML method, *disadvantage 1* poses a more fundamental problem. If one model made in a lab or industrial setting loses its predictive power once e.g. fibre properties change, the usefulness of such models is questionable. When there is no data for a certain set of process variables available, the model uncertainty will be high because it needs to extrapolate from other examples (Hüllermeier & Waegeman, 2021). A high uncertainty leads to wrong predictions and conclusions. In ML literature, uncertainty due to unavailable data is called *Epistemic Uncertainty*. The act of reducing Epistemic Uncertainty by additional, deliberate sampling is called *exploration* (Hüllermeier & Waegeman, 2021).

A possible answer to *disadvantage 1* is to not try to make a model that generalises across all process configurations. A



**Fig. 4** Comparison of lab and industry based data driven learning approaches (Based on Brunke et al. (2022))

model could be trained only for a specific process at hand, ignoring all variables that change rarely. Generating data for this task from a full Design of Experiments is economically infeasible due to the high dimensionality and a high likelihood of incurring failure costs. There needs to be careful exploration by interaction with the process. During exploration, only economically favourable process settings are sampled. Additionally, safety aspects need to be considered to avoid failure costs by crashing the process or producing products with insufficient quality. Any learning system that employs exploration directly on a physical process needs to automatically address safety (Brunke et al., 2022).

The agent needs to trade off between exploiting known information about the process and exploring new solutions that have the potential of being better than previously known solutions. This trade-off is known as the *Exploration vs. Exploitation Dilemma* (Sutton et al., 2018). A research area that additionally considers safety is called *Safe Optimisation* (Brunke et al., 2022). We will call an algorithm that exhibits the previously stated properties an *Autonomous Learning Agent (ALA)*. We form the hypothesis that is the motivation for this work:

*Autonomous Learning Agents that balance Exploration and Exploitation, while consistently prioritizing safety, are viable concepts for optimizing industrial textile production processes.*

Our hypothesis is illustrated in figure 4, where the knowledge of the data-driven approach without learning does not expand whereas the learning-based approach is able to expand knowledge in sensible directions.

## Autonomous agents in production engineering

In this section we give a brief overview over possible ALA implementations. (Guan et al., 2023; Stojanovic, 2023; Tao et al., 2023; Zhuang et al., 2023)

**Reinforcement Learning** Autonomous agents in production environments have been getting much attention in the research community recently (Panzer & Bender, 2022). We will give a short overview, leaving an in-depth comparison to future works. Panzer and Bender (2022) performed a systematic literature review of the usage of Reinforcement Learning (RL) as autonomous agents in production systems. They found that RL is being applied in many different domains, often outperforming traditional methods. Most research focuses on dynamic process control, scheduling and dispatching, not on economical optimisation and quality control. They conclude that, while very promising, RL needs more validation in practical settings as well as more focus on finding suitable algorithms for a given problem (Panzer & Bender, 2022).

Nian et al. (2020) also reviewed the application of RL for industrial process control. They highlight the flexibility and adaptability of RL for solving many different problem settings. The difficulty of designing an RL agent is the need for an accurate simulation environment. Both stability as well as state constraints are the most difficult problems to handle in industrial settings, according to the authors (Nian et al., 2020). Scheiderer et al. (2020) address the need for a simulation environment for RL agents during the design phase by developing a 'Simulation as a Service' architecture. This architecture allows simulation providers to offer their software for the training and testing of RL agents through an interface at the example of heavy plate rolling (Scheiderer et al., 2020). This design allows cheap testing and validation of the agents. He et al. (2022) employed RL to the textile ozonation process. The authors chose the approach due to the high dimensionality of textile production processes. The optimisation goal was formulated as a stochastic Markov Game in which multiple RL agents operate in a multi-objective optimisation problem (He et al., 2022). The solution was found to outperform Genetic Algorithms like NSGA-II in a simulated production environment (He et al., 2022). The paper does not provide the code for the simulation environment, limiting the reproducibility for other research. Also, only Deep-Q-Networks were used without investigation into alternatives (He et al., 2022).

**Data-driven optimal control techniques** Recently, much attention has been paid to the field of data-driven optimal control, e.g. in Stojanovic (2023). Görges (2017) performed a review of the similarities between RL and optimal control techniques such as Model Predictive Control (MPC). Discrete Optimal Control techniques are mostly concerned with a sequence of control inputs that minimise a quadratic value function. The author argues that while the learning of RL is slow, imposing structure on system dynamics models, value function and control policy through prior knowledge helps speed up the learning process. (Stojanovic, 2023) is a good example of the convergence between RL and opti-

mal control as they learn unknown system dynamics and disturbances using system feedback data. Control policy and policy evaluation are split up into an Actor-critic architecture. Structure is imposed through a discrete-time algebraic Riccati equation (DARE) that is solved using adaptive dynamic programming. The structure of the problem also allows for theoretical stability guarantees. Another example of adaptive control is Iterative Learning Control (ILC), which is concerned with optimizing control sequences for repetitive tasks. The approach includes learning from past errors and updating the control sequence Zhuang et al. (2023); Guan et al. (2023). An example for the incorporation of system dynamics uncertainty and disturbances can be found in Tao et al. (2023).

It is important to note that dynamics play a minor role in mostly stationary textile production processes. We are mostly concerned with finding optimal values for a set of setpoints considering the current production context. Still, optimal control theory offers a principled way of modeling and analysis of industrial production systems.

**(Safe) Bayesian optimisation** According to Panzer and Bender (2022) and Brunke et al. (2022) RL is often used for dynamic and combinatoric problems. As explained in "Nonwoven production process", the nonwovens production process is mostly stationary. Also, RL is widely criticised for being sample inefficient, which is a problem in production environments where samples are expensive. Another approach for ALA can be Bayesian Optimisation (BO). BO aims to 'allocate resources to identify optimal parameters as efficiently as possible' by employing Bayesian inference on an unknown, uncertain objective function (Garnett, 2023). It has been applied to Engineering applications in the form of an adaptive Design of Experiments in order to only sample regions with a high probability of improvement (Greenhill et al., 2020). Safe BO extends the BO concept with the possibility of enforcing probabilistic constraints on the decisions. There are many different approaches that implement this general approach like (Berkenkamp et al., 2023; Sui et al., 2018) and (Kirschner et al., 2019).

## Optimisation problem formulation

Our second disadvantage formulated the need for a complete problem formulation that addresses economic and quality aspects of the textile production as well as probabilistic effects on process performance. Thus, in this section we formulate the optimisation problem to be solved by the ALA independently of the solution approach.



## Objective function

Manufacturing companies in low-wage countries primarily achieve their advantage through economy of scale (Brecher et al., 2012). Companies in high-wage countries try to compensate their disadvantage of higher unit costs by focusing on synergy, fast adaptation to market needs and sophisticated production planning approaches (Brecher et al., 2012). The dilemma arises from additional overhead costs due to focusing on planning instead of producing. ? argue that autonomous systems can compensate the dilemma of high wage countries by quickly achieving economy of scale for smaller lot sizes by quickly increasing production volume. This effect is shown in Fig. 5 on the left. We define the production volume  $P_V$  in [ $\text{m}^2\text{min}^{-1}$ ] of a nonwovens production process as the amount of unit areas produced over time:

$$P_V = b_P v_P \quad (5)$$

where  $b_P$  is the production width that is multiplied with  $v_P$  as the production speed in [ $\text{mmin}^{-1}$ ].

Hax and Majluf (1982) note that economies of scale also result in a lower price per product, meaning that the price per product is an indicator of how effectively that company can manage its resources. The sinking price per unit with increasing total produced product volume is called the *Experience Curve* (See figure 5 on the right, Hax and Majluf (1982)). Among the five reasons the authors identify for the sinking costs are

1. **Learning:** With each produced unit, the operators learn to use resources more efficiently
2. **Process improvements:** Large production volumes give more opportunity to modify and optimise the production process.

The Experience Curve leads us to another important conclusion: *When Supervised Learning is used in a production environment, it can only learn from experiences already made, not generating any improvements.* According to the experience curve, we describe the variable costs per unit area  $C_U$  in [ $\text{€m}^{-2}$ ] in a nonwovens production process as the sum of energy and material costs per unit area:

$$\begin{aligned} C_U &= C_{E,U} + C_{F,U} \\ &= \frac{Y_P C_{E,kWh}}{P_V} + \frac{\dot{m} C_{F,kg}}{P_V} \end{aligned} \quad (6)$$

In the formula above the energy and material costs per unit area are denoted  $C_{E,U}$  and  $C_{F,U}$ . The electrical power consumption of the process  $Y_P$  is multiplied with the energy price per kW to model process electricity costs. The costs per unit area  $C_{E,U}$  are then calculated by dividing by the production

rate  $P_V$ . The material costs are defined by the product of the total fibre mass flow into the process  $\dot{m}$  and the fibre price per kg  $C_{F,kg}$ . Again, we divide by the production rate for conversion from process costs to unit area costs. The contribution margin per unit area  $M_U$  in € denotes the difference between earnings  $C_{S,U}$  and variable costs. It is defined as:

$$M_U = C_{S,U} - C_U \quad (7)$$

Per the formula above we seek to reduce our variable costs to increase profitability as we assume the earnings per unit area cannot be influenced. To optimise total profit over time instead of over unit areas produced we reformulate equation 7 so that it reflects the process contribution margin over time as our objective function  $f_{obj}$  in [ $\text{€h}^{-1}$ ]:

$$\begin{aligned} f_{obj} &= M_U P_V \\ &= C_{S,U} b_P v_P - Y_P C_{E,Wh} - \dot{m} C_{F,kg} \end{aligned} \quad (8)$$

The above equation not only considers lowering production costs as a factor to improve economic efficiency but also considers increasing productivity as a factor. Including productivity is justified because producing with low costs per unit area and low productivity is not a feasible solution. High fix costs due to maintenance and personnel might still cause the process to be unprofitable if productivity is disregarded (Cloppenburg, 2019).

We obtained an objective function to be maximised that represents the economic efficiency of our production process. The objective function depends on setpoint parameters  $\{b_P, v_P\} \in \mathbf{s}$ , as well as the black box process output  $Y_P$ . The important mass throughput variable  $\dot{m}$  is not independent of other factors, since the process mass balance needs to be satisfied. The needed mass throughput is calculated internally by the process control system. The goal of the upcoming process modelling chapter is to express the objective function only as a function of  $\mathbf{x}_t$  to capture systematic influences on process performance. For now, we express non-probabilistic process outputs like the mass throughput  $\dot{m}$  as a regular function:

$$\dot{m} = f_m(\mathbf{x}_t) \quad (9)$$

All probabilistic process outputs are modeled as a non-stationary stochastic process whose statistical properties can be expressed as a function of process state  $\mathbf{x}_t$  (see also 2):

$$\begin{aligned} Y_{i,t} &= f_{\mu,i}(\mathbf{x}_t) + \epsilon_{i,t} \\ \forall i &\in \{p, w, MD, CD\} \\ \epsilon_{i,t} &\sim \mathcal{N}(0, f_{\sigma,i}^2(\mathbf{x}_t)) \end{aligned} \quad (10)$$

In the equation above,  $f_{\mu,i}(\mathbf{x}_t)$  denotes the systematic process behavior, whereas  $\epsilon_{i,t}$  denotes the process noise or

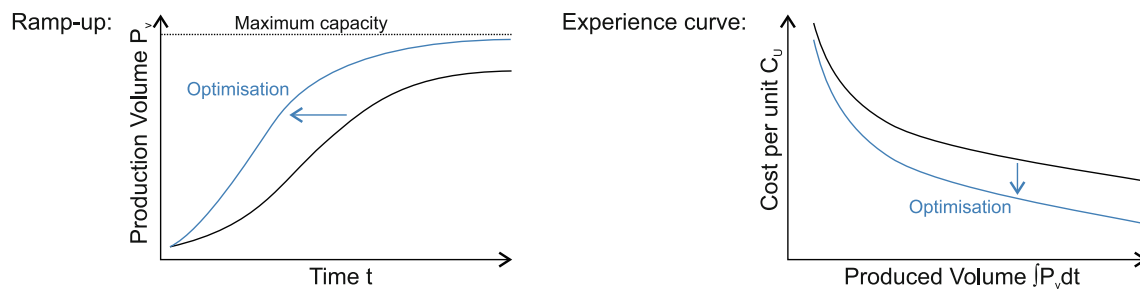


Fig. 5 Ramp-up and experience curve of a production system ? (Hax & Majluf, 1982)

aleatoric uncertainty. The process noise is assumed to be a heteroskedastic normal distribution whose variance  $f_{\sigma,i}^2(\mathbf{x}_t)$  depends on the process state. The indices  $w$ ,  $CD$  and  $MD$  represent an incomplete list of process outputs that can be subject to quality constraints.  $CD$  and  $MD$  represent the tensile strength of the textile in  $MD$  and  $CD$  respectively. The term  $w$  represents the product weight per area.

In the next section we formulate the full optimisation problem by introducing a way to handle the constraints.

## Optimisation problem

According to Brunke et al. (2022), the term safe optimisation refers to 'sampling inputs that do not violate a given safety threshold'. There are three basic strategies to enforce constraints that are subject to uncertainty according to the authors Brunke et al. (2022):

1. **Soft constraints:** If a constraint is violated, a penalty is given instead of the objective function value. This encourages the agent to avoid violating constraints.
2. **Probabilistic constraints:** The agent needs to satisfy all constraints with a certain, fixed probability.
3. **Hard constraints:** The agent satisfies constraints at all times. This robust approach leads to conservative, sub-optimal solutions and is therefore disregarded (Bertsimas & Sim, 2004).

Therefore we focus on the problem formulation for the first two approaches.

**Probabilistic constraints** We first formulate the optimisation problem for the probabilistic constraints. We seek to find a vector of setpoint variables  $\mathbf{s}_t^*$  that maximises our objective. We used the general approach as described by Garnett (2023). The solution also needs to satisfy the probabilistic output constraints as described by Li (2007). The probability of the process output being outside of the quality bounds  $q$  needs

to be above a threshold  $\alpha$ .

$$\begin{aligned} \mathbf{s}_t^* &\in \operatorname{argmax}_{\mathbf{s}_t \in \mathbf{S}} \mathbf{E}[f_{obj}|\mathbf{x}_t] \\ \text{s.t. } &P(q_{i,low,t} \leq Y_i \leq q_{i,high,t}|\mathbf{x}_t) \geq \alpha_i \\ \text{s.t. } &\dot{m}_{low} \leq \dot{m}_t \leq \dot{m}_{high} \\ &\forall i \in \{w, MD, CD\} \end{aligned} \quad (11)$$

Note that  $\mathbf{x} = [\mathbf{s}^T, \mathbf{d}^T, \mathbf{l}^T]^T$ . The fibre mass throughput is constrained in order to avoid clogging of card cylinders and other damage to equipment. Note that the manipulable vector needs to be an element of the safe set  $\mathbf{S}$  that does not violate any static actor constraints. Brunke et al. (2022) call this approach 'safe exploration with single time step state constraints'. The problem does not only depend on setpoint, but also disturbance and layout variables. These variables as well as quality constraints might change over time, therefore the optimal solution depends on the time step  $t$ . Most quality characteristics are not measured in real-time but rather by using lab equipment (ISO, 2007), therefore quality samples are only available sparsely and with a significant delay. Therefore, methods that deal with objective maximisation and constraint satisfaction separately will have a significant advantage because they do not need quality samples to complete a full optimisation step.

**Soft constraints** The second approach is to enforce satisfaction of constraints by penalising the agent for violations. For the penalty, we subtract the sales price from the objective function because we assume the product cannot be sold. We formulate an augmented objective function:

$$f_{aug} = \begin{cases} f_{obj} & \text{if constr. satisfied} \\ f_{obj} - C_{S,UBVP} & \text{otherwise} \end{cases} \quad (12)$$

The optimisation problem now becomes:

$$\mathbf{s}_t^* \in \operatorname{argmax}_{\mathbf{s}_t \in \mathbf{S}} \mathbf{E}[f_{aug}|\mathbf{x}_t, q_{i,high/low,t}] \quad (13)$$

Note that we essentially relaxed the constraints of equation 11. The optimal solution depends on the situation, making our problem a *contextual* problem. The agent needs to learn to associate the optimal solution with certain situations (Garnett, 2023). The agent's actions only affect the immediate reward, making the problem a *bandit* problem (Sutton et al., 2018).

**Conclusion of problem formulation** In this subsection we formulated the optimisation problem of maximising the economic and resource-wise efficiency of a nonwovens production process. We formulated the problem as both a stochastic optimisation problem well as a contextual bandit. To conclude this section we summarise the most important optimisation problem characteristics:

- **Decision space:** The problem has a continuous and high dimensional decision space.
- **Context:** The optimal solution  $s_t^*$  depends on current production context, which makes the problem instationary. Due to high dimensionality, not all context can be taken into account by the agent, resulting in high external uncertainties. If some context is not learned by the agent, the agent needs to consider additional safety margins. Due to wear of components, the agent performance decreases over time due to high constant uncertainties. The agent needs to compensate this by continuously adapting.
- **Noisy sampling:** The process outputs are noisy due to its stochastic nature, resulting in random uncertainties.
- **Black box problem:** We lack an explicit expression, making our process a 'black box'. Solving the problem requires a careful balance between exploitation and exploration. (Sutton et al., 2018)
- **Sparse samples:** Some process outputs are continuously sampled with inline measurement equipment (like weight per area and power consumption) while others are sampled rarely if they require manual work (like tensile strength).
- **Safety-critical:** Not only are there hard constraints on setpoints to avoid catastrophic crashes, there also needs to be some guarantee to produce with sufficient quality.
- **No dynamics:** The optimal solution does not depend on previous system states.

## Simulator architecture

In the last section we formulated the optimisation problem and outlined its key characteristics. We outlined that testing and validating any ALA in a simulation before implementing it on the physical process is advantageous. Implementing the ALA in a simulation first allows cheap testing and validation. Only in a simulation is the comparison of different ALA algorithms, architectures and hyperparameters economically

feasible. The best performing agents can finally be validated on the physical process. In this section, we derive a simulation architecture that is able to implement the optimisation problem we formulated. This allows researchers to test and validate different approaches in a reproducible manner and encourages practical applications.

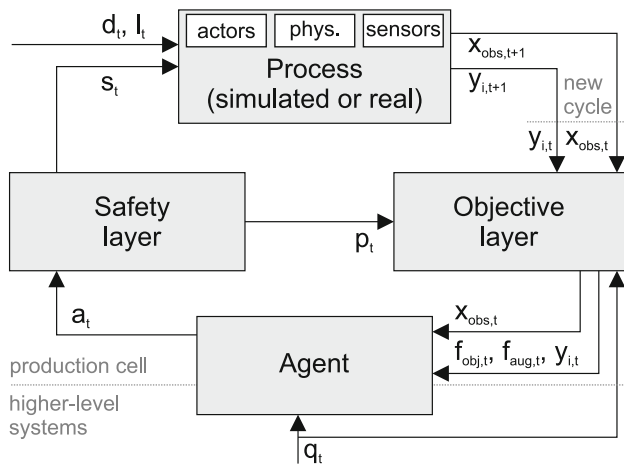
At first we define the scope of the simulator: We aim to optimise a single nonwovens production line according to the current context. Hoffmann (2019) postulates that for the control of a single manufacturing unit the agent needs to be located on the edge between the manufacturing unit and higher-level systems. This approach ensures low-level interaction with both sensors and actor control loops in the process as well as communication with higher-level systems (e.g. for setting production goals or quality requirements) (Hoffmann, 2019). The calculation of the objective function is usually seen as part of the environment and not the agent (Sutton et al., 2018), but it is also not part of the process. We model it as an intermediate entity.

In the previous chapter we derived that the optimisation is safety-critical. There are hard constraints on manipulable variables as well as derived quantities like mass throughput to avoid machine crashes. These constraints can be checked a priori, in which case Brunke et al. (2022) mention several works that implement a *safety layer* which transforms an action  $a_t$  into a safe setpoint vector  $s_t$ . In our case we simply do not execute an unsafe action and set a penalty flag for the agent to receive:

$$s_t = \begin{cases} a_t & \text{if } (a_t \in S) \wedge (\dot{m}_{low} \leq \dot{m}_t \leq \dot{m}_{high}) \\ s_{t-1} & \text{otherwise} \end{cases} \quad (14)$$

We leave more sophisticated methods of guaranteeing safety (like solving a quadratic program (Pham et al., 2018)) to the agent. The other aspect of safety is the satisfaction of the product quality characteristics. This is part of the objective layer. The resulting architecture is shown in figure 6.

At the beginning of a time step, the process output as well as the partial process state are recorded. The performance layer then calculates the objective function (which depends on the quality requirements) which provides all information needed to the agent. The agent receives the information as well as additional context from higher-level systems and calculates an action. The safety layer transforms the action into a safe action and overrides the performance function with a penalty if it is found to be unsafe. The action is transmitted to the process actors for execution. The physical process behavior then depends on the system state as well as non-deterministic influences. Finally, the sensors record the new process outputs as well as the partial process state and a new cycle begins.



**Fig. 6** Control loop architecture with agent and process

In the next step we derive the simulator architecture for implementation using Unified Modeling Language (UML). The simulator needs to fulfill the following requirements as per the previous sections:

- Enable cyclic agent-environment interaction
- Simulate different contexts or scenarios
- Implement a safety layer
- Simulate the continuous black box process outputs as a noisy stochastic process.
- Simulate dependent variables
- Implement the objective function calculation

Beside the technical aspects, usability as well as adaptability are also important. The simulator needs to be adaptable to multiple agent architectures and frameworks. For example, the *stable baselines* reference implementations for common RL algorithms (Hill et al., 2018) rely on the widely used *gymnasium* interface for interaction between agent and environment (Towers et al., 2023). As previously stated, textile production processes are highly individualised, which shows the need for a configurable design. New process outputs, constraints or quality requirements should be easy to integrate. Also, to enable reproducible experiments, an experiment tracker needs to log all relevant data points. The environment needs to be able to reset itself to allow for multiple consecutive experiment runs. Finally, we should aim for computational efficiency to decrease experiment run time.

We developed an architecture, which is shown in figure 7, that implements the requirements stated above. We chose a highly modular structure that can easily be changed and adapted later. All class properties are initialised with plain-text \*.yaml configuration files so that the simulator behavior can be changed without the need for changing the source code. Dictionaries with key-value pairs are the standard data

exchange format, minimising the risk of reading the wrong position in an array when the configuration has been changed.

The class *SpecificEnvironment* is the main entry point for interaction with the simulator. It is abstract, meaning the specific implementation depends on the agent and its requirements. The specific implementation only has to have methods for executing a new *step* for cyclic interaction as well as a *reset* method for doing multiple runs without re-initialisation. *SpecificEnvironment* needs to adapt our *Environment* class, which is the central hub for all other components. *Environment* coordinates the calculation of *setpoints*, *dependent variables* and *disturbances*, also referred to as the environment *state*. Additionally, the process *outputs* get stored as well as the time step. This class also manages the execution of the *step* and *reset* methods.

We require the possibility of changing the production scenario, which is implemented by the *ScenarioManager* class. It can deterministically or randomly (according to a specified uniform distribution) change disturbances or quality requirements at predetermined time steps. This behavior allows domain randomization so the agent learns to adapt to different contexts and doesn't overfit to one singular scenario.

The *ActionManager* class transforms the agent *actions* into *setpoints*, mainly by checking them against the static actor and dependent variable constraints and setting penalty flags. It also coordinates the calculation of dependent variables and thus acts as a basic process control system.

The *OutputManager* allocates one process for each process output. This achieves parallel processing for the execution of the ML models as they are the most computationally expensive part of a step execution. The class receives the process *state* on calling the *step* method. It calls its *call\_models* method for getting a prediction on the process output distribution modeled as a normal distribution. It then samples the predicted distribution and returns the samples as process *outputs*. The *ModelAdapter* is meant to provide abstraction over different ML frameworks and adapts any framework to be compatible with the simulator. The *predict\_y* method uses the associated ML *Model* class to get a prediction of the process output distribution.

After all quantities of *Environment* have been calculated, *ObjectiveManager* first checks for quality constraint satisfaction and then uses *ObjectiveFunction* for the calculation of the objective value. If the penalty flag has been set, the penalty gets returned instead. All quantities get logged by the *ExperimentTracker* class so the run can be analysed after completion. Its concrete implementation depends on the logging framework being used. Lastly, the *Environment* returns values as stated in *observables* to *SpecificEnvironment* to finalise the *step* execution. A simplified overview of the step execution sequence can be found in figure 8.

For the sake of clarity, logging and scenario management are not shown in the figure.

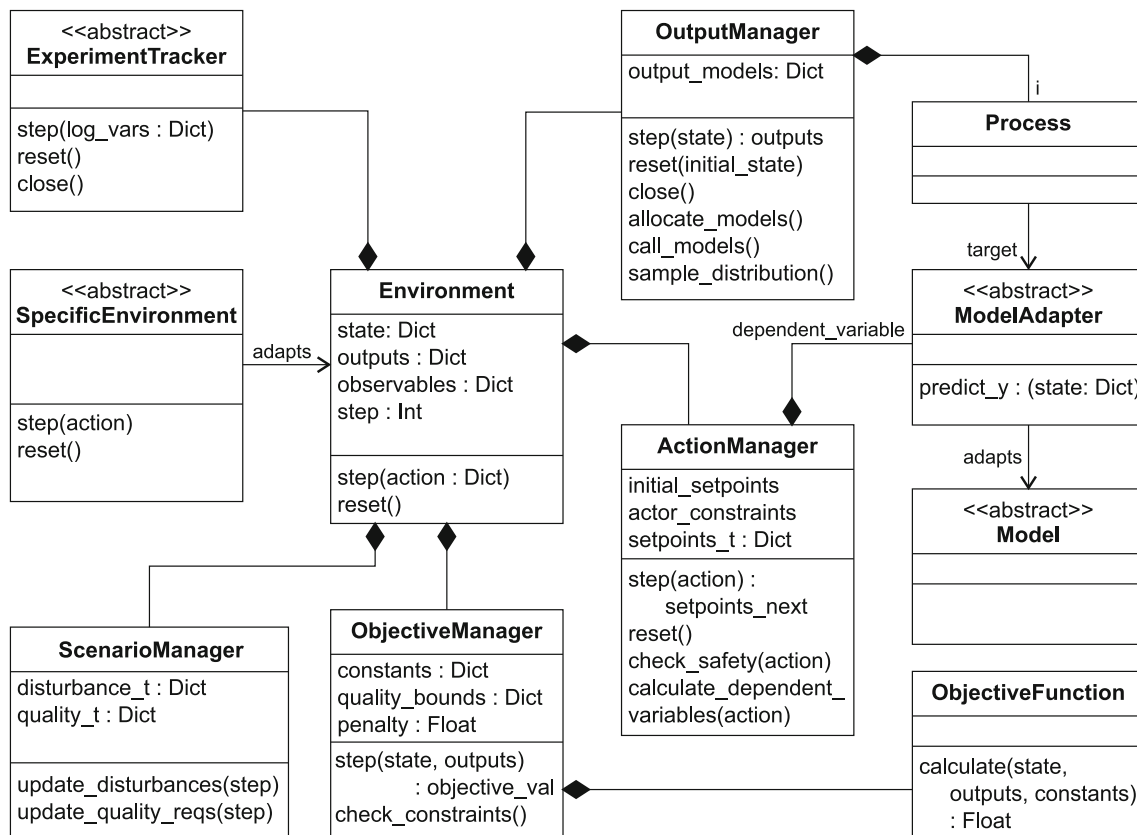


Fig. 7 Simulator architecture expressed in a UML class diagram

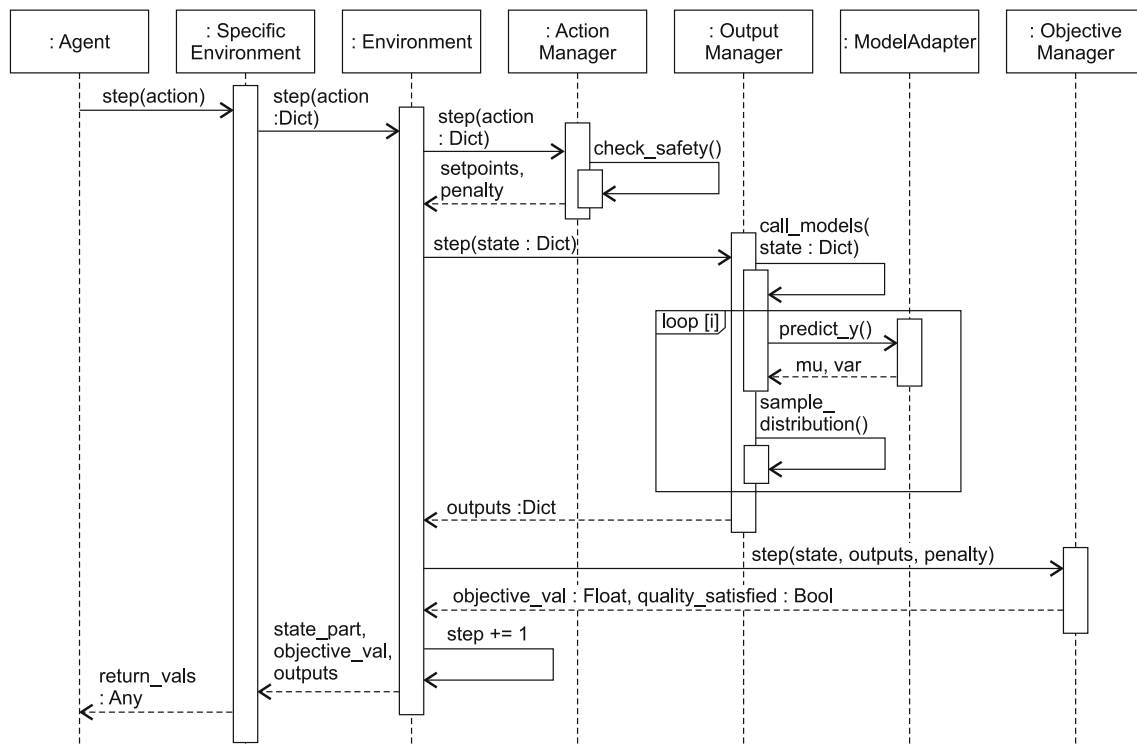


Fig. 8 UML sequence diagram of a (partial) time step execution



## Conclusion of architecture design

In this chapter we derived a highly modular and adaptable simulator architecture. It allows the simulation of different production scenarios with a cyclic interaction between agent and environment.

## Process modeling

In the last chapters we provided an explicit formulation of the optimisation problem as well as the simulator architecture. In this chapter, we provide models for the missing quantities (process outputs  $Y_i$  and mass throughput  $\dot{m}$ ). These models are also subject to the shortcomings formulated in *disadvantage 1*, meaning they are only valid for a very specific production process. Since we want to test the learning behavior of an ALA, generalisability is not an issue as long as we capture the influence of the most important setpoints and disturbances.

Relevant process outputs are all quality characteristics that are demanded for a specific product by the customer. Among these quantities can be weight per area, tensile strength in MD and CD, flame-retardant properties, air permeability, thickness or optical characteristics (Mao & Russel, 2007). We focus on the case where the nonwoven is required to have a minimum weight per area as well as tensile strength in MD and CD. In equation 7 we also introduced the power consumption, resulting in a total of five variables to be modeled:

- Tensile strengths  $Y_{MD}$  and  $Y_{CD}$
- Power consumption  $Y_p$
- Weight per area  $Y_w$
- Mass throughput  $\dot{m}$

## Data acquisition

The challenge of modelling the process outputs is that for most outputs there are no first principle models, therefore we need to rely on an empirical approach using data sampling and ML modeling. In our approach we sampled data from an industrial nonwovens production process since we aim to simulate an industrial production process. The data was recorded using per-minute averages

For a description of the output values, refer to table 1. For a description of the recorded process states or columns, please refer to table 2

As shown in table 1, there is much less data about the tensile strength than the power consumption. That is because sampling the tensile strength is a manual process whereas the other was done automatically using sensors. This observation helps estimate the time steps the agent should need to learn: If one optimisation step is only completed after a lab sample,

**Table 1** Description of the process output data

Targets	Acquisition time span	Valid rows
Power consumption	454d 3 h	293,172
Tensile Strength	374d 5 h	1,091

the agent should make significant progress within hundreds of time steps. As can be seen in table 2, some process setpoints are classified as disturbances, because many are set due to reasons other than process optimisation, so to the agent they become a disturbance.

## Biases and causality in industrial data sets

There are important variables missing from the data set, most importantly the actual fibre properties as opposed to the specification. Also, machine setup and condition are not recorded. This leads to omitted-variable bias, where a statistical model falsely attributes changes in the omitted variables to variables that were recorded. (Kocak, 2022)

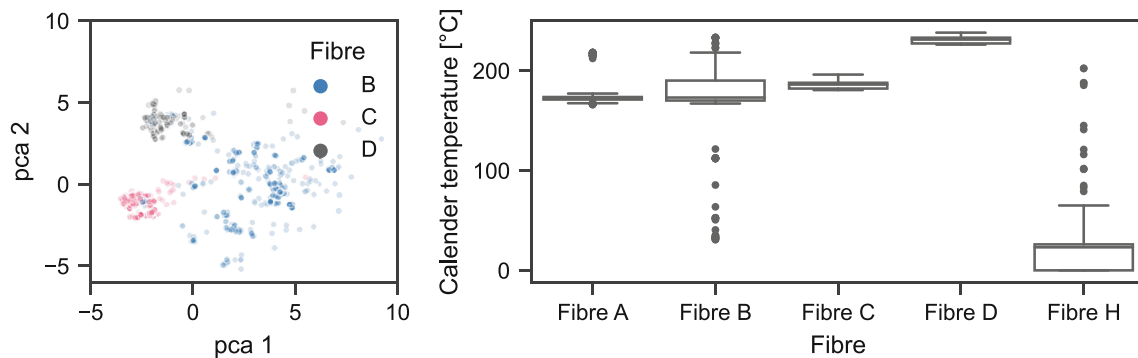
For the tensile strength, the available observations are in the same order of magnitude as the number of observed parameters. Because of high dimensionality with little available data, this might lead to overconfident models that generalize poorly if they attribute small changes in the target value to random variations in the process states (Kocak, 2022). Because of this, a highly unstable feature selection process is to be expected.

The recorded process states also include the fibre mixture. Different fibres might require different process settings, especially calender temperatures. Also, customer requirements might differ between fibre groups because they are chosen for a specific purpose. Figure 9 shows the data distribution using the first two axes of a Principle Component Analysis (PCA) of the data as well as as a boxplot of the calender temperature broken down by fibre mixture. The figure clearly shows that there are complex interrelations between fibre groups and process settings. If these interrelations are not accounted for, they might again lead to wrong conclusions and biases.

If we aim towards data-driven optimisation of textile production processes, we derive actions based on process models. These actions are also called *interventions* in the field of causal inference (Peters et al., 2017). It is very important to note the difference between prediction and causal inference. Prediction aims to find correlations between observations and a target to predict the target value based on hypothetical observations (Peters et al., 2017). However, if we impose an intervention on a system we might get significantly different probability distributions than what can be expected from field observations (Becker et al., 2023). This is because the

**Table 2** Description of the recorded process states

Process states (columns)	Number of columns	Type
Environment conditions	16	Disturbance
Process setpoints	60	Setpoint, disturbance
Actual values	94	Dependent, disturbance
Fibre mixture	1 (54 categories)	Disturbance

**Fig. 9** PCA diagram and Boxplot showing data clusters depending on fibre mixture

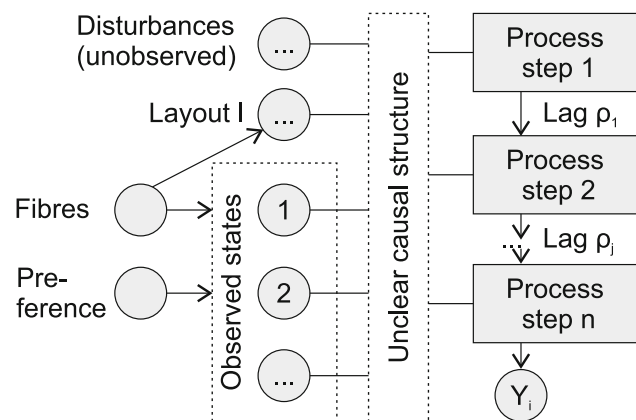
assumption that the data is independent and identically distributed (IID) is violated if the data does not come from a carefully designed experiment (Schölkopf, 2022). As shown in this chapter, there are a multitude of biases in the data set that lead to a process state not being independent of other mechanisms.

**Example:** We observe the process power consumption  $Y_p$  as well as the fibre mass throughput  $\dot{m}$  and environment temperature. We obtain two models with reasonably strong predictive capabilities, one uses environment temperature ( $P(Y_p|T_{env})$ ), the other fibre mass throughput ( $P(Y_p|\dot{m})$ ) as feature. To reduce power consumption we could intervene and install a climate control system to reduce the temperature or reduce mass throughput. We would find that reducing mass throughput will reduce power consumption because processing less mass leads to less strain on the motors. We would find that climate control has no effect, because the higher power consumption caused an increase in temperature due to dissipated energy. Our wrong assumption was that temperature is an independent state.

We conclude that while industrial data has the advantages of more data volume and more practical relevance, the significant disadvantages include biases that might lead to causally incorrect models.

## Data preprocessing and feature selection

Using only observational data without additional assumptions, we can not distinguish between models that are robust towards interventions and those that are not (Schölkopf, 2022). Humans on the other hand are very skilled at intuition

**Fig. 10** SCM of the process outputs and observed states

about causal structure (Peters et al., 2017). If we want to optimise a production process using an intelligent agent, such an agent should have control over or be informed about the Independent Causal Mechanism (ICM) of the process. ICM are defined as the "The causal generative process of a system's variables [...] composed of autonomous modules that do not inform or influence each other" (Peters et al., 2017). During the modeling stage we need to be careful to only include ICM as states in order to achieve realistic process behavior. Our current understanding about the process causal structure is shown in Fig. 10 in the form of a Structural Causal model (SCM) graph. An SCM is a directed graph where the nodes represent one system variable (Peters et al., 2017). The edge direction represents the direction of causality.

Figure 10 shows the omitted-variable bias because of unobserved process states. Also, many states are not inde-

pendent of fibre mixtures, operator preference or operating standard procedures. It also shows the lags  $\rho_j$  between different process steps that results from the nonwoven carrying the signal has to physically travel in between machines. We account for the lag by offsetting the timestamp of observations at one process step by  $\tau = s_j/v_j$ , which is the time the nonwoven takes to travel in between process steps.

We account for complex fibre group inter-dependencies by only including comparable rows with a calender temperature of at least 150° C. We also only include fibres with at least 30 available samples, leaving us with seven fibre groups (Fibre A-G) as well as 332 samples for tensile strength. We encode the fibre group in the data using one hot encoding, where each fibre group has its own column with values being either 1 or 0). The states are then scaled using the RobustScaler method of the library Scikit-Learn (Pedregosa et al., 2011).

We then select features for each output model using Lasso Regression as the basis, which is robust against multicollinearity because it penalises high weights (Dangeti, 2017). We use Lasso Regression in conjunction with a stepwise approach with forward feature selection. Each iteration the state with the highest additive explanatory power is added until the adjusted  $R^2$  value does not increase. We customise the stepwise algorithm by having an expert agree with each addition. The feature in question is only added to the selected features if the expert agrees to its causal correctness. We capture expert opinion with Ishikawa diagrams as seen in Fig. 3. For additional robustness we use Lasso Regression with 10-fold cross-validation.

## Regression model architecture

In this section we explore the choice of suitable methods for modeling the process outputs. We previously stated the need for outputting both a mean as well as some form of confidence bound to capture the probabilistic process behavior. Another advantageous model property is smooth interpolation between data points to achieve realistic process outputs.

We chose the non-parametric, non-linear Gaussian Process Regression (GPR) as an ML model due to several advantageous properties. Firstly, the exact GPR method models the expected value of a process as well as its confidence intervals and an estimate of process noise. Also, GPR allows smooth interpolation in between sparse data points. Lastly, GPR gives much control over the shape of predicted values and allows the integration of prior knowledge (Duvenaud, 2014). The disadvantage is that the computational complexity is  $O(n^3)$  with data points due to expensive matrix inversions (Duvenaud, 2014). More scalable versions of GPR are available (Titsias, 2009). The popular feedforward ANN model can also be extended to output process noise (Hüllermeier & Waegeman, 2021). Our framework can easily be

adapted to accommodate this approach due to its modular structure.

Given a systematic process behavior where the actual output  $Y_i$  is corrupted by additive homoskedastic noise  $\epsilon_i \sim \mathcal{N}(0, \eta^2)$  we get the marginal posterior distribution  $Y_i(x') \sim \mathcal{N}(\mu(x'), \sigma^2(x'))$  (meaning the predicted distribution of the process output) for a query location  $\mathbf{x}'$  (Duvenaud, 2014):

$$\begin{aligned}\mu(x') &= k(x', \mathbf{X})^\top \mathbf{C}^{-1} \mathbf{y} \\ \sigma^2(x') &= k(\mathbf{x}', \mathbf{x}') + \eta^2 \\ &\quad - k(\mathbf{x}', \mathbf{X})^\top \mathbf{C}^{-1} k(\mathbf{x}', \mathbf{X})\end{aligned}\quad (15)$$

Where the letter  $\mathbf{C} = k(\mathbf{X}, \mathbf{X}) + \eta^2 \mathbf{I}$  denotes the covariance matrix with the noise term. Since we are only interested in the process noise, we modify the model prediction to only return the noise  $\eta^2$ . The kernel function design  $k$  has a huge influence on the model behavior (Duvenaud, 2014). The kernel function can be interpreted as a measure of similarity between two data points. A popular Kernel function is the Polynomial (Poly) Kernel (Duvenaud, 2014)

$$k(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^\top \mathbf{x}' + c)^d \quad (16)$$

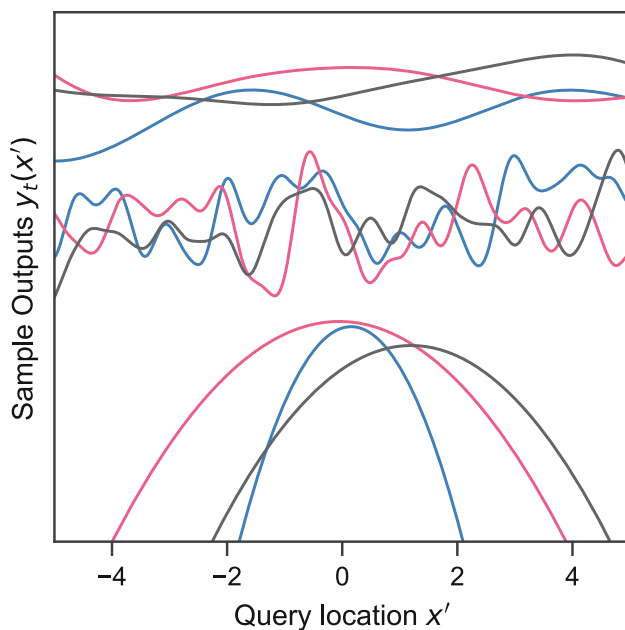
between two locations  $\mathbf{x}$  and  $\mathbf{x}'$ . Another example is the Radial Basis Function (RBF) Kernel (Duvenaud, 2014)

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right) \quad (17)$$

which is a non-stationary Kernel, which means that similarity of two data points is measured only by their distance in space. This Kernel leads to high uncertainties when extrapolating from known data points (Duvenaud, 2014). Low lengthscale parameters  $\sigma$  lead to complex models due to low similarities, while high values lead to smooth model behavior (Duvenaud, 2014). Lengthscales, noise and other parameters are fitted using gradient-based optimisers like ADAM on the Negative Log Likelihood (NLL) as a loss function (Gardner et al., 2023). The influence of kernel design on the model shape is highlighted in figure 11.

For a full introduction into GPR, please refer to Garnett (2023) or Duvenaud (2014).

The problem at hand is characterised by high dimensionality with a low amount of available data. Duvenaud (2014) provides general guidelines about how data-efficient kernels can be designed. If permissible, non-stationary kernels like the Poly Kernel can be used because of their non-local nature. Another technique is called Automatic Relevance Detection (ARD). With this technique, each model dimension is assigned its own kernel. This way, the model can exhibit different amounts of complexity across each dimension because



**Fig. 11** Influence of kernel design on prior (meaning before considering observations) sample outputs. Top: RBF Kernel, high lengthscale; Middle: low lengthscale; Bottom: 2nd-degree Poly kernel

each dimension has its own lengthscale. We can apply lengthscale constraints to limit complexity and enforce smooth predictions (Duvenaud, 2014). The application of lengthscale constraints also encourages the model to attribute influences from missing variables to noise rather than falsely attribute them to one of the features.

### Tensile strength MD model

We performed Lasso regression ( $\alpha = 1$ ) for feature selection and compared the standard approach using all features with our proposed stepwise approach. The results are shown in table 3

The results show that our custom model using expert feedback outperforms the baseline model as can be seen in table 3. Based on e.g. the alpha value, Lasso regression chose different feature sets, which highlights the unstable feature selection we expected earlier. The weight per area has a huge influence on tensile strength because it determines the amount of fibres in the cross-section of the nonwovens. The baseline model preferred quantities that are a direct consequence of weight per area like the cross-lapper height (which is causally incorrect since the height is adjusted to accommodate different weights per area). The stepwise model with expert opinion chose causally correct features that determine weight per area (Cross-lapper layers count, weight per area card floor as well as draft ratios). The feed per stroke as well as the calender temperature setting are two additional states that are directly linked to fibre web consolidation.

Figure 12 shows the performance of our model. There are some outliers in the data set where the prediction is much worse than the average. This is likely due to an error in the manual data recording process, so we removed all data points where the absolute residual is higher than 180 N. This drastically improved our model performance as shown in 3.

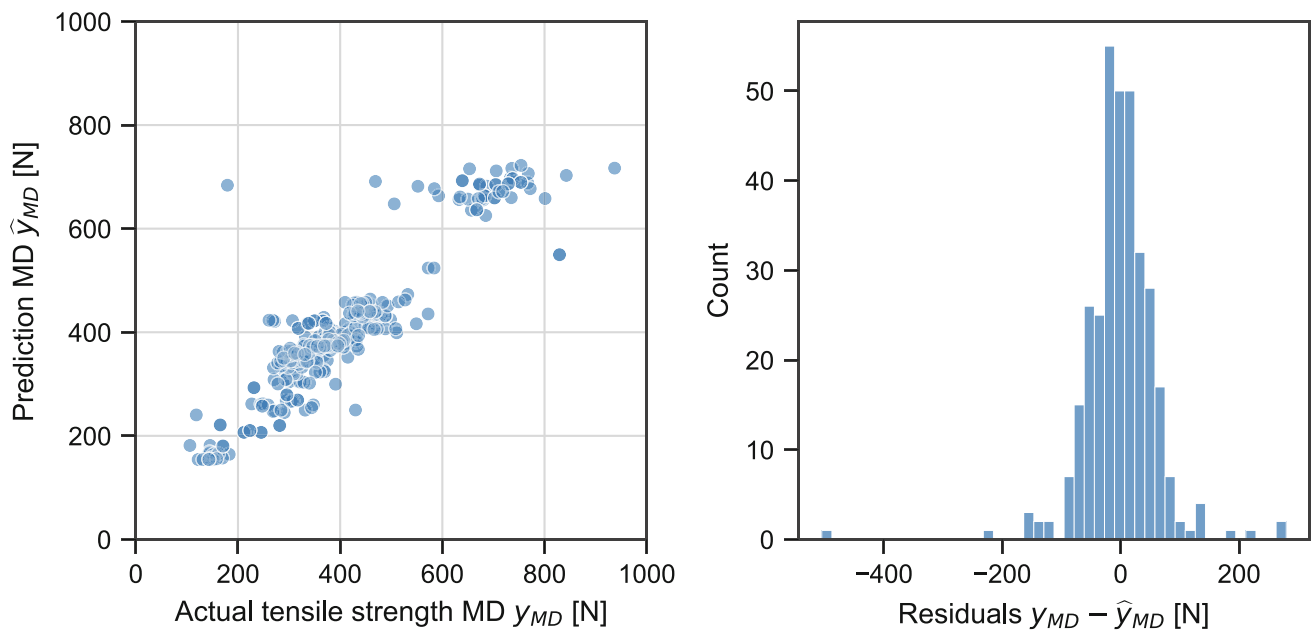
For the GPR models, we normalised the data using Scikit-Learn's RobustScaler so that the lengthscale values are comparable. We used the GPyTorch framework for the computations (Gardner et al., 2023). All kernels are multiplied with each other according to the formula  $k_1 k_2 \dots k_n$ . All lengthscale values except for the fibre mixtures were constrained to be above 0.5 to limit model complexity. We chose the polynomial kernel for the cross-lapper layers count as well as the weight per area card floor. The reason for our choice is that we assume the connection to be mostly linear since both variables determine the amount of fibres in the nonwovens cross-section. Note that the selected fibres are different from the Lasso model. That is because we included all fibres at first and then removed all fibres with a lengthscale above 20.

The lengthscales were not effected by the constraints with the exception of *needleloom draft ratio*, as shown in 4. The highest lengthscale value is *calender temperature*, indicating little complexity along this dimension. A comparison of model performance with and without constraints is shown in 5. The table shows that the NLL loss is unaffected by the constraints. The process noise  $\eta^2$  is 1108, which translates to a process with a standard deviation regarding the Tensile Strength in MD of 33 N.

Figure 13 shows the model behavior along one dimension while all other dimension values are kept constant at their respective median. For clarity, the prediction variance is shown without the noise variance. The grey dashed lines are the 0.05- and 0.95-quantiles of the data that we use as upper and lower actor setpoint constraints. The constraints for the calender temperature are missing since we classify this state as a disturbance. The calender temperature is chosen for criteria other than process optimisation. Most dimensions show polynomial-like behavior independently of kernel choice. Increasing the layers by a factor of two also roughly increases the tensile strength by a factor of two. This makes sense since double the fibres are in the cross-section of the nonwoven. The state 'cross-lapper layers count' determines how many card web layers are stacked on top of each other and therefore can only assume positive integer values. Therefore, the model takes continuous inputs, but rounds this dimension to Integers. Increasing the card floor by a factor of two does not lead to the same increase, likely because of floor defects acting as a weak point in the textile. The feed per stroke graph does not show a distinct peak, indicating that the technically possible maximum strength of consolidation has not been reached in the process at hand. All draft ratios show that high ratios decrease the tensile strength because it reduces

**Table 3** Comparison of our tensile strength MD Lasso models

Baseline Lasso Model	Our stepwise Lasso Model	Outliers removed
Adjusted $R^2$ : 0.80	Adjusted $R^2$ : 0.84	Adjusted $R^2$ : 0.89
Selected Features	Selected Features	Selected Features
1. Needlepunch depth	1. Cross-lapper layers count	Same as stepwise
2. Cross-lapper carriage speed	2. Draw frame draft ratio	Lasso model
3. Cross-lapper height front	3. Fibre C	
4. Bale opener weight setpoint	4. Fibre B	
5. Needleloom feed per stroke	5. Calender temperature	
6. Card belt scale speed	6. Weight per area card floor	
7. Feeder velocity	7. Needleloom feed per stroke	
8. Cross-lapper height back	8. Fibre A	
9. Needleloom scraper roll position	9. Needleloom draft ratio	

**Fig. 12** Prediction errors and residuals of the tensile strength MD Lasso model**Table 4** Kernel Specification for the tensile strength MD GPR model

Number	Dimension Name	Kernel	Constraints	Hyperparameters
–	Constant scale	Const	–	$c = 0.000835$
1	Cross-lapper layers count	Poly	$d = 2$	$c = 4.79$
2	Weight per area card floor	Poly	$d = 2$	$c = 8.61$
3	Needleloom feed per stroke	RBF	$\sigma > 0.5$	$\sigma = 8.63$
4	Needleloom draft ratio	RBF	$\sigma > 0.5$	$\sigma = 0.73$
5	Draw frame draft ratio	RBF	$\sigma > 0.5$	$\sigma = 2.02$
6	Calender temperature	RBF	$\sigma > 0.5$	$\sigma = 14.65$
7	Fibre A	RBF	$\sigma > 0.0$	$\sigma = 1.76$
8	Fibre D	RBF	$\sigma > 0.0$	$\sigma = 15.97$



**Table 5** Comparison our GPR models

	GPR model with constraints	Without constraints
Noise Variance $\eta^2$ in [N <sup>2</sup> ]	1108	1111
NLL Loss	0.106	0.106

the weight per area ( $\downarrow$ ). Fibre mixture A differs from other mixtures mostly in a generally higher tensile strength, but it is also more sensitive to changes in needleloom draft ratios.

### Tensile strength CD model

The process of building the Tensile Strength CD model is analogous to the last section. At first we perform feature selection with our stepwise Lasso algorithm. The comparison of the standard approach as well as our approach is given in table 6. The standard approach chooses features that correlate directly with fibre mass throughput (features 2, 3 and 5), but do not cause the weight per area of the nonwoven. The stepwise feature set remains the same as with the MD model, but with different fibre group selections.

The GPR kernel design is the same as before, but does not include calender temperature. We increased the lengthscale constraints to 1.0 to enforce more smooth predictions. As shown in table 8, this negatively affects model performance and forces the model to explain the variance with the noise term. This is permissible, as our goal is to achieve realistic models over good predictive capabilities.

Figure 14 shows the model behavior for the tensile strength in cross direction. The model behavior along the dimensions is similar to the MD model. The main difference lies in the behavior of the tensile strength when the nonwoven is subjected to drafting. The slope of the tensile strength is steeper for the draw frame draft ratio. The principle that applies is again that high ratios decrease the tensile strength because it reduces the weight per area ( $\downarrow$ ). The re-orientation of fibres from CD to MD direction works in the same direction ( $\downarrow$ ). Because both effects work in the same direction, the result is a steeper slope.

### Power consumption model

For the prediction of the power consumption, we followed the same process as before. The standard Lasso model chose features related to production speeds and the calender nip force, which correlates to weight per area. Due to convergence issues, the model performance is poor. Our model with expert feedback chose the final production speed, fibre mass throughput as well as the needleloom feed per stroke. A higher mass throughput leads to higher strain on the motors while the other two features combined determine the

motor rpm of the needleloom. The experts rejected many features containing temperatures as well as 'switching states'. Many process steps before the card process fibres discontinuously, resulting in on-off-switches that cannot be influenced directly.

For the GPR model, we randomly sampled 1,500 data rows from our data set with seven fibre mixtures in order to limit computational complexity. We also removed the production speed because it was almost constant with a lengthscale of over 20 (see table 10). As can be seen in table 11 the higher constraints for enforcing smoothness have a significant influence on model performance.

As can be seen in Fig. 15, the mass throughput leads to a steady increase in power consumption whereas a higher feed per stroke leads to fewer motor rpm of the needleloom and therefore a decrease in power consumption. The changes due to the two states are in the range of 40 kW, indicating a high base load of approx. 290 kW that cannot be influenced directly.

### Weight per area model

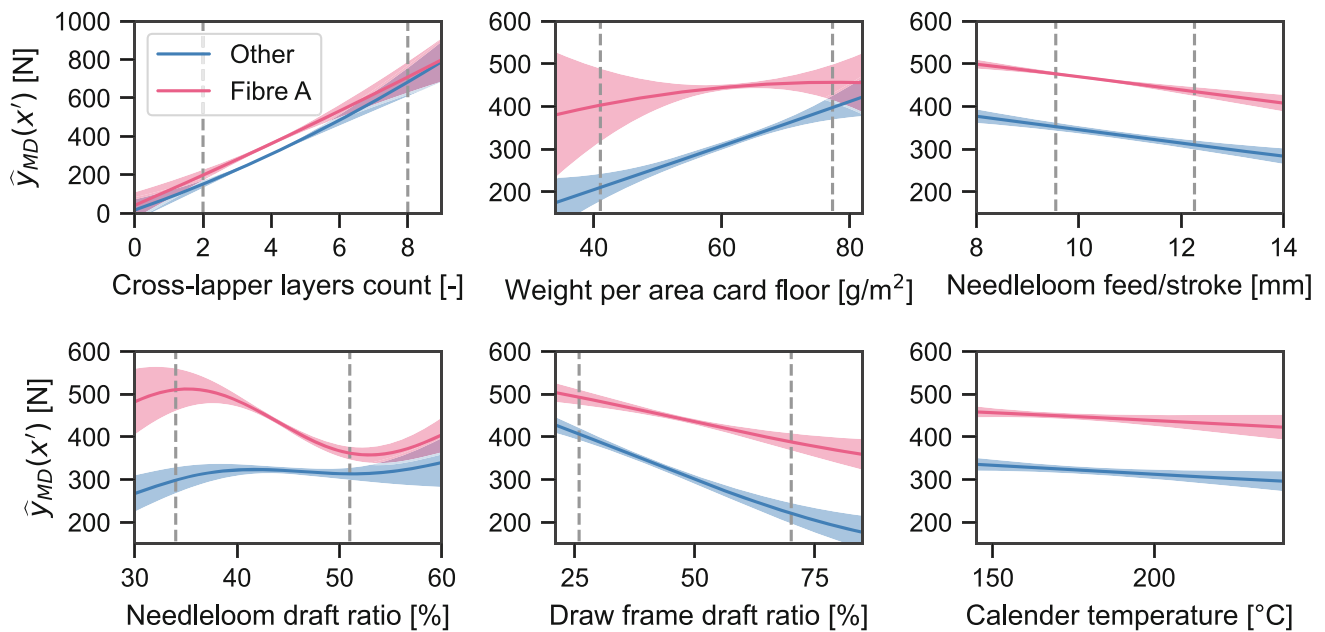
In contrast to the previous ML models, the weight per area is calculated using mass balance equations:

$$E[Y_{w,l}] = 2w_{card\ web}n_{layers} \sum_m \frac{100}{d_{m,l}+100} \quad (18)$$

The final weight per area  $Y_w$  of the product is a result of stacking multiple layers of the card web inside the cross-lapper. The factor of two before  $n_{layers}$  comes from the fact that the cross-lapper always needs to return. Afterwards, the fibre web is subject to multiple draft ratios that reduce the weight per area. This is illustrated in Fig. 16. We used the expectation operator because the mass balance does not account for local deviations in mass distribution.

We now consider the draft ratios of the process that appear in equation 18. There are four draft ratios  $d_m$  in total:

1. Profiling  $d_{profiling,l}$  that periodically drafts the nonwoven to counteract the so-called smile-effect;
2. Drafting inside the draw frame  $d_{drawframe}$ ;
3. Drafting at the needleloom intake  $d_{intake}$ ;
4. Drafting inside the needleloom  $d_{needleloom}$ .



**Fig. 13** Predictions of the tensile strength MD GPR model

**Table 6** Comparison of our tensile strength CD Lasso models

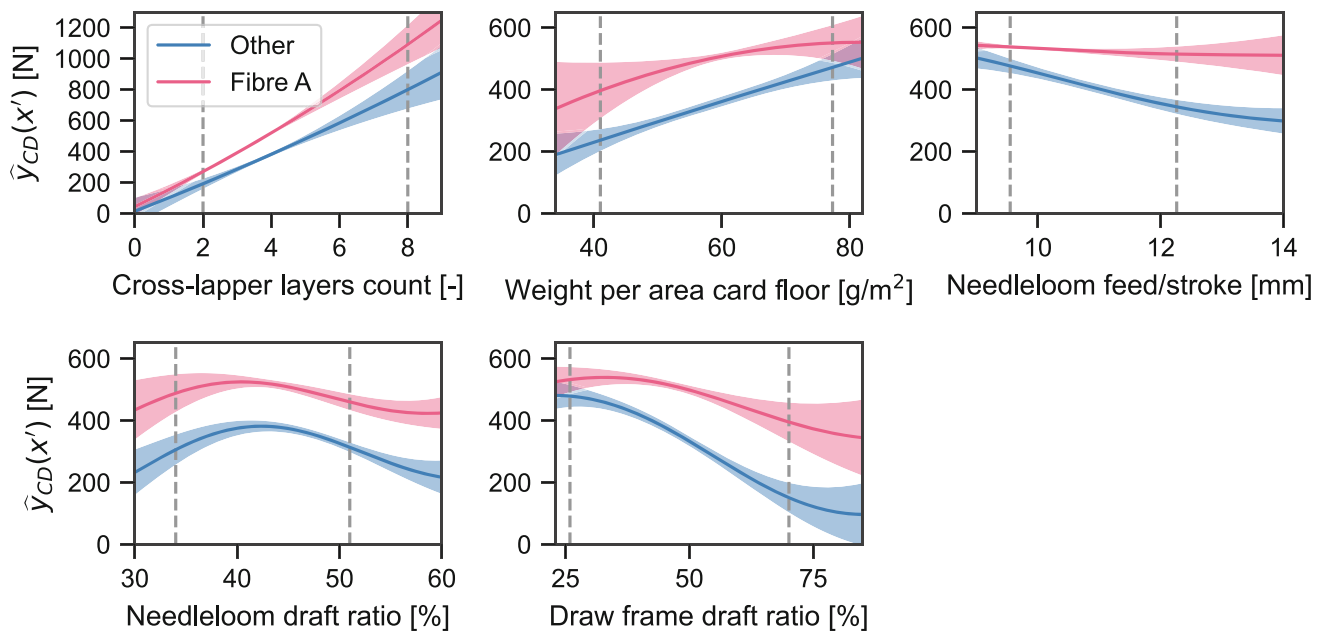
Baseline Lasso model	Our stepwise Lasso model	Outliers removed
Adjusted $R^2$ : 0.71	Adjusted $R^2$ : 0.73	Adjusted $R^2$ : 0.83
Selected features	Selected features	Selected features
1. Fibre F	1. Cross-lapper layers count	Same as stepwise
2. Feeder velocity	2. Draw frame draft ratio	Lasso model
3. Mixing chamber belt velocity	3. Needleloom draft ratio	
4. Fibre C	4. Fibre A	
5. Bale opener weight setpoint	5. Needleloom feed per stroke	
6. Needleloom scraper roll position	6. Weight per area card floor	
7. Cross-lapper layer overlap	7. Calender temperature	

**Table 7** Kernel Specification for tensile strength CD GPR model

Number	Dimension Name	Kernel	Constraints	Hyperparameters
–	Constant scale	Const	–	$c = 0.001322$
1	Cross-lapper layers count	Poly	$d = 2$	$c = 3.45$
2	Weight per area card floor	Poly	$d = 2$	$c = 14.50$
3	Needleloom feed per stroke	RBF	$\sigma > 1.0$	$\sigma = 2.35$
4	Needleloom draft ratio	RBF	$\sigma > 1.0$	$\sigma = 1.003$
5	Draw frame draft ratio	RBF	$\sigma > 1.0$	$\sigma = 1.000$
6	Fibre A	RBF	$\sigma > 0.0$	$\sigma = 1.64$

**Table 8** Comparison of the two GPR models

	GPR model with constraints	without constraints
Noise Variance	1295	1182
$\eta^2$ in $[N^2]$		
NLL Loss	0.088	0.056



**Fig. 14** Predictions of the tensile strength CD GPR model

**Table 9** Comparison of our power consumption Lasso models

Baseline Lasso model	Our stepwise Lasso model
Adjusted $R^2$ : 0.57	Adjusted $R^2$ : 0.80
Selected features	Selected features
1. Cross-lapper delivery speed	1. Mass throughput
2. Calender nip force	2. Production speed
3. Needleloom draft ratio	3. Needleloom feed per stroke

**Table 10** Kernel Specification for the power consumption GPR model

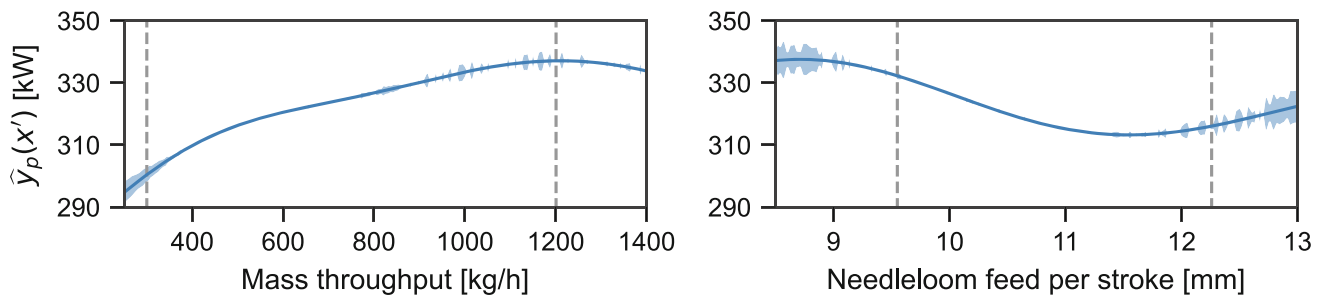
Number	Dimension Name	Kernel	Constraints	Hyperparameters
–	Constant scale	Const	–	$c = 0.6817$
1	Cross-lapper layers count	RBF	$\sigma > 2.0$	$c = 2.01$
2	Needleloom feed per stroke	RBF	$\sigma > 2.0$	$\sigma = 2.01$

To explain the profiling we first have to address the smile-effect. To model local deviations from the expected weight per area we have to account for both systematic as well as probabilistic deviations. During needlepunching and calendering the fibre web is subject to the systematic smile effect  $w_{smile}$  (Russell & Norvig, 2016). This is the phenomenon of a higher contraction on the edges of the nonwoven, leading to a weight distribution in the form of a smiley. We account for

this by dividing the floor into five lanes with weight per area  $Y_{wl}$ , lane three being the middle lane. To get a clear picture of systematic changes in weight per area please refer to 16. The weight per area is increased by half of the disturbance  $w_{smile}$  whereas the middle lane is decreased by the full  $w_{smile}$  so the overall mass does not change. To account for probabilistic deviations we added noise to each lane  $\epsilon$  that scales with

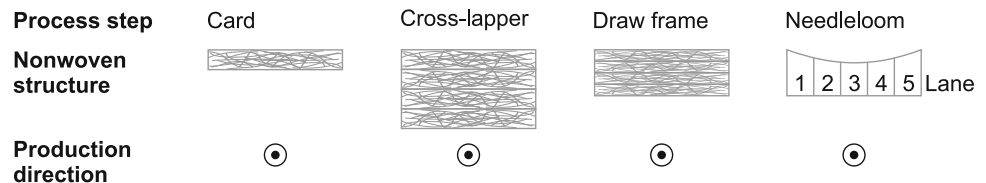
**Table 11** Comparison of the two GPR models

	GPR model with constraints	without constraints
Noise Variance	13.90	6.36
$\eta^2$ in $[kW^2]$		
NLL Loss	–0.097	–0.330



**Fig. 15** Predictions of the power consumption GPR model

**Fig. 16** Schematic of the nonwoven transformation along the process



the expected weight per area.

$$\begin{aligned}
 Y_{w1} &= \mathbf{E}[Y_{w,1}](1 + c_w \epsilon_1) + 0.5 w_{\text{smile}} \\
 Y_{w2} &= \mathbf{E}[Y_{w,2}](1 + c_w \epsilon_2) \\
 Y_{w3} &= \mathbf{E}[Y_{w,3}](1 + c_w \epsilon_3) - w_{\text{smile}} \\
 Y_{w4} &= \mathbf{E}[Y_{w,4}](1 + c_w \epsilon_4) \\
 Y_{w5} &= \mathbf{E}[Y_{w,5}](1 + c_w \epsilon_5) + 0.5 w_{\text{smile}}
 \end{aligned} \quad (19)$$

We chose the constant scale  $c_w$  so that the variance is equal to 12 at  $100 \text{ gm}^{-2}$ .

Profiling periodically drafts the nonwoven to counteract the smile-effect. We model its influence by drafting the edge lanes 1 and 5 by half the profiling setting while applying the negative value to the middle lane to satisfy the mass balance:

$$d_{\text{profiling},l} = \begin{cases} 0.5 d_{\text{profiling}} & \text{if } l = 1 \vee l = 5 \\ -d_{\text{profiling}} & \text{if } l = 3 \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

The discussion of the profiling concludes the probabilistic weight per area models for each lane.

In order to calculate the dependent variable *mass throughput* in  $[\text{kg h}^{-1}]$ , we need to multiply how many units are produced over time ( $v_P b_P$ ) with how much this area weights per area. The expected weight per area is given by  $\mathbf{E}[Y_{w,1}]$ , disregarding profiling. The formula we get, including conversion factors, is:

$$\dot{m} = \frac{6}{100} w_0 v_P b_P \left( \frac{\text{kg}}{\text{g}} \right) \left( \frac{\text{min}}{\text{h}} \right) \quad (21)$$

## Conclusion of model implementation

In this section we built models from first principles as well as industrial data. We showed that for industrial data, the IID assumption is usually violated and derived a practical strategy to deal with this drawback. Even though we applied many strategies to derive causally correct models, we cannot guarantee correctness without a controlled experiment.

## Test of simulation behavior

In the last chapters we discussed the motivation for the architecture as well as the components of the simulator. In this chapter we show the overall simulator behavior using a simple pre-programmed agent. We also summarize the most important characteristics of the simulation.

### Simulator characteristics

In this subsection we summarize the degrees of freedom that the simulator has. All degrees of freedom are listed in table 12.

According to table 12, there are seven setpoints that serve as the degrees of freedom for the agent actions. Context is expressed through dependent variables as well as disturbances. This increases the dimensionality of the system states to 14. Additional information can optionally be provided to the agent by returning the process output values as well as the scenario-dependent output bounds/constraints. The full system state dimensionality then amounts to 32 dimensions. Both disturbances and process output constraints can be changed using the *ScenarioManager* class. This domain randomisation enables the agent to learn many arbitrary pro-

**Table 12** Overview of the simulator degrees of freedom

Number	Variable	Type	Initial Value	Constraints	Unit
1	Card delivery weight per area	Setpoint	63.0	(41.0, 77.28)	$\text{gm}^{-2}$
2	Cross-lapper layers count	Setpoint	4	(2, 8)	–
3	Needleloom draft ratio	Setpoint	43.0	(34.0, 51.0)	%
4	Draw frame draft ratio	Setpoint	44.7	(25.9, 70.1)	%
5	Cross-lapper profiling	Setpoint	1.0	(0.0, 15.0)	%
6	Needleloom feed per stroke	Setpoint	10.5	(9.55, 12.26)	mm
7	Production speed	Setpoint	14.26	(7.0, 14.72)	$\text{m m}^{-1}$
1	Mass throughput	Dependent	–	(300.0, 1201.0)	$\text{kg m}^{-1}$
1	Needleloom draft ratio intake	Disturbance	10.0	–	%
2	Smile effect strength	Disturbance	10.0	–	$\text{gm}^{-2}$

**Table 12** continued

Number	Variable	Type	Initial Value	Constraints	Unit
3	Calender temperature	Disturbance	174.0	–	$^{\circ}\text{C}$
4	Fibre A	Disturbance	0	–	–
5	Fibre D	Disturbance	0	–	–
6	Product width	Disturbance	5.62	–	m
1	Area weight lane 1	Process output	–	Scenario	$\text{gm}^{-2}$
2	Area weight lane 2	Process output	–	Scenario	$\text{gm}^{-2}$
3	Area weight lane 3	Process output	–	Scenario	$\text{gm}^{-2}$
4	Tensile strength MD	Process output	–	Scenario	N
5	Tensile strength CD	Process output	–	Scenario	N
6	Line power consumption	Process output	–	–	kW

duction scenarios and contexts. The optimal configuration of returned states that are available to the agent is highly dependent on the Autonomous Learning Agent (ALA) approach and its implementation and thus needs to be investigated experimentally.

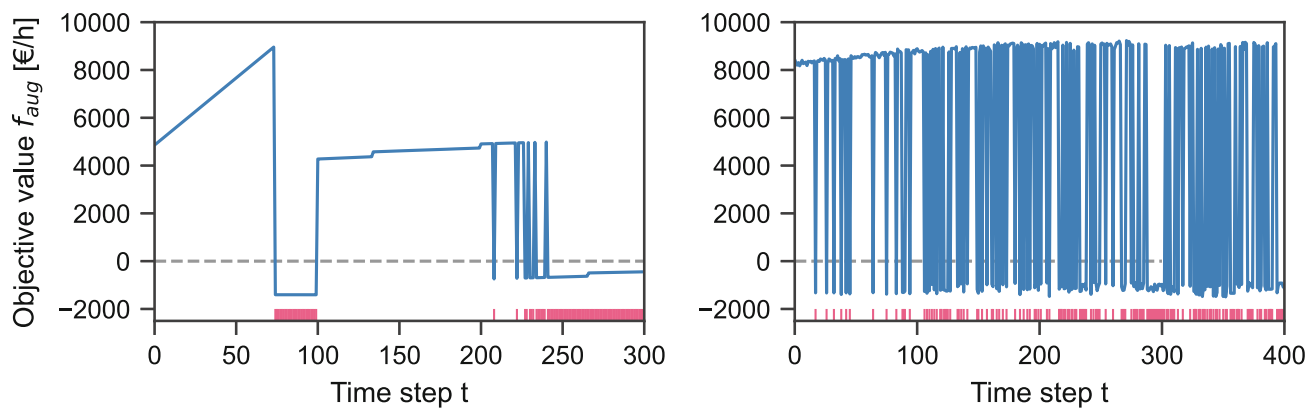
### Interaction between environment and agent

We now provide insights into the simulator behavior using a simple pre-programmed agent that runs through a test scenario. We also tested the popular, unsafe, gradient-free optimiser CMA-ES (Hansen, 2006) in a second run. We do not propose any solution method for our optimisation problem as this will be part of later works. Using CMA-ES is supposed to show the limits of unsafe optimisation methods as well as the general simulation behavior. For simplicity, the only output constraints during our test runs are lower bounds of  $180 \text{ gm}^{-2}$  on all weight per area values. Our pre-programmed agent will start with the initial values and steadily increase the production speed for the first 100 time steps. Then it will reset and steadily decrease both the cross-lapper layers count as well as the card delivery weight per

area for another 200 time steps. The resulting objective curve is shown in figure 17 on the left. The right part shows the performance of the CMA-ES optimiser.

As shown in Fig. 17, the objective value increases at first because of the higher productivity. At time step 75, the objective value becomes negative because a penalty has been triggered (pink marker). That is because the maximum mass throughput has been reached. The unsafe actions are not executed and the line continues to produce, even though the penalty has been returned. After time step 100, the agent decreases the weight per area of the product. At first, this leads to an increase in the objective value because less fibre material is wasted. The steps in the objective value occur when the models round the layers count to the next integer value. After step 209, the process goes out of control, which is indicated by singular weight per area violations. The expected weight per area decreases until all values are out of bounds. Note that the penalty decreases because the actions are still executed, but less fibres are wasted on faulty products. To reference Fig. 5: In the first 100 time steps the production volume increases whereas afterwards the cost per unit decreases. The test of CMA-ES shows that the optimiser





**Fig. 17** Simulator behavior during our test runs with pre-programmed agent (left) and CMA-ES agent (right)

is able to steadily increase the objective value. However, as an unsafe algorithm it continuously violates safety constraints and is therefore not a suitable approach. The simulator successfully interacted with two simple agents and shows logical behavior, so we conclude by stating that the simulator implementation has been successful.

## Conclusion and outlook

The findings and results of this paper are as follows:

1. We found that most research focuses on Supervised Learning methods that do not properly address the challenge of high dimensionality in textile production processes.
2. We also found that most research only focuses on single process aspects without addressing process uncertainties. This approach does not capture real-life process behavior.
3. We postulated that ALA are a promising approach to perform industrial-scale process optimisation in the textile industry, more specifically the nonwovens production.
4. We formulated a probabilistic constrained optimisation problem for stationary textile production processes that maximises profitability while ensuring sufficient product quality. We also formulated a relaxed version using soft constraints.
5. We developed a highly modular and adaptable simulation architecture that facilitates cheap testing of ALA for encouraging transition of research into practical applications. The adaptability and modularity also facilitate transfer to other domains.
6. We populated the architecture with realistic GPR models to mimic real-world process behavior of a carded nonwovens production. Our simulator approximates the realistic behavior of a generic nonwovens process along its most important states.
7. We outlined the challenges that arise from working with industrial data sets that do not originate from controlled experiments.

We hope our contributions from this work will facilitate transfer of approaches from research into practical applications. Our simulation allows researchers to quickly assess the performance of their algorithms in an industrial setting. It also aids in reproducible comparisons of different ALA approaches or sets of hyperparameters.

In our work we incorporated both classic methods like SPC and cost curves as well as more modern ML methods based on industrial data and GPR. We bridged the gap between classical methods and modern ML methods, hoping to give new perspectives on industrial use cases at the intersection of Production Engineering, Operations Research as well as ML Research.

Since we seek to facilitate the transfer of data-driven ALA approaches into practice, it is important to discuss weaknesses and advantages of such approaches. The central advantage of using ALA, according to our hypothesis, is that they are a feasible alternative to using models derived from offline Supervised Learning. While the latter fails to generalise for the high-dimensional textile production, the former ALA are able to learn a specific process through interaction, ideally only sampling economically favourable states. In Chapter 5 we showed how the derivation of causally correct models using industrial data is challenging. We also showed how disregarding important setpoint and context variables will lead to high system model uncertainties. We conclude that any ALA approach should be carefully designed to include the most important system states and should only have control over the Independent Causal Mechanisms (ICM) of the process. The risk is that the agent will perform ineffective or unsafe actions if it has not been designed carefully. Another risk is that the Agent learns too

slow to make significant contributions to the producer's economic success.

In the future, we will compare different ALA approaches against each other. The best performing approach will be tested on a physical nonwovens production line at our institute. However, a real-world agent will likely not act in a closed loop manner, but still have a human operator act as a safeguard. That is because not all safety criteria can be objectively measured with sensors in a practical manner (e.g. card cylinders are prone to clog with fibres, so every cylinder would need to be optically monitored).

Our approach can be enhanced by coupling simulators (and agents) across multiple companies within the textile value chain because the value chain is split into many steps with intermediate products (Kins & Gries, 2023). This approach would lead to more collaboration along the value chain and offer more flexibility in the optimisation.

**Acknowledgements** We want to thank our partners BNP Brinkmann GmbH & Co. KG., Hörstel and MAT Solutions GmbH, Haltern am See for their contributions to making the sampling of industrial data possible.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This work was financially supported by the German Ministry of Education and Research (grant number 01IS22002A).

**Code Availability** The code that has been developed in this work is available in this public repository. <https://git.rwth-aachen.de/ruben.kins/adanowo-simulator>.

## Declarations

**Conflict of interest** The authors declare that they have no competing financial interests or personal relationships that could have influenced the work reported in this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aslam, M., Saghir, A., & Ahmad, L. (2021). *Introduction to statistical process control*. Hoboken, NJ: Wiley. <https://doi.org/10.1002/9781119528425>
- Becker, M., Brockmann, M., & Niemietz, P., et al. (2023). The internet of production as the foundation of data utilization in production. In: Trauth D, Bergs T, Prinz W (eds) *The monetization of technical data - Innovations from Industry and Research*. Springer, Berlin, p 247–264. [https://doi.org/10.1007/978-3-662-66509-1\\_15](https://doi.org/10.1007/978-3-662-66509-1_15)
- Berkenkamp, F., Krause, A., & Schoellig, A. P. (2023). Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics. *Machine Learning*, 112(10), 3713–3747. <https://doi.org/10.1007/s10994-021-06019-1>
- Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, 52, 35–53. <https://doi.org/10.1287/opre.1030.0065>
- Brecher, C., Jeschke, S., & Schuh, G., et al. (2012). Integrative production technology for high-wage countries. In: Brecher C (ed) *Integrative Production Technology for High-Wage Countries*. Springer, Berlin, p. 17–76. [https://doi.org/10.1007/978-3-642-21067-9\\_2](https://doi.org/10.1007/978-3-642-21067-9_2)
- Brunke, L., Greeff, M., Hall, A. W., et al. (2022). Safe learning in robotics: From learning-based control to safe reinforcement learning. *Annual Review of Control, Robotics, and Autonomous Systems*, 5, 411–444. <https://doi.org/10.1146/annurev-control-042920-020211>
- Brydon, A.G., & Pourmohammadi, A. (2007) Dry-laid web formation. In: Russell, S.J. (ed). *Handbook of nonwovens*. CRC Press, Boca Raton, p 16–111. <https://doi.org/10.1533/9781845691998.16>
- Büscher, C., Hauck, E., Schilberg, D., & Jeschke, S. (2014). Key performance indicators for the impact of cognitive assembly planning on ramp-up process. In S. Jeschke, I. Isenhardt, F. Hees, & K. Henning (Eds.), *Automation, communication and cybernetics in science and engineering 2013/2014* (pp. 557–577). Springer International Publishing.
- Cloppenburg, F. (2019) Wirtschaftliche und technische Modellierung und Selbstoptimierung von Vliesstoffkreppelein. Shaker
- Dangeti, P. (2017). *Statistics for Machine Learning*. Birmingham, Mumbai: Packt Publishing Limited.
- Duvenaud, D. (2014) Automatic model construction with gaussian processes. PhD thesis, University of Cambridge, <https://doi.org/10.17863/CAM.14087>
- EDANA (2022) Nonwovens markets, facts and figures. <https://www.edana.org/nw-related-industry/nonwovens-markets>, accessed: 26/09/2023
- Gardner, JR., Pleiss, G., & Bindel, D., et al. (2023) Gpytorch: Black-box matrix-gaussian process inference with gpu acceleration. <https://github.com/cornellius-gp/gpytorch>, <https://doi.org/10.48550/arXiv.1809.11165>, accessed: 26/09/2023
- Garnett, R. (2023). Bayesian optimization. *Cambridge University Press, Cambridge, United Kingdom*. <https://doi.org/10.1017/9781108348973>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. Cambridge, Massachusetts: MIT Press.
- Görges, D. (2017). Relations between model predictive control and reinforcement learning. *IFAC-PapersOnLine*, 50(1), 4920–4928. <https://doi.org/10.1016/j.ifacol.2017.08.747>
- Greenhill, S., Rana, S., Gupta, S., et al. (2020). Bayesian optimization for adaptive experimental design: A review. *IEEE Access*, 8, 13937–13948. <https://doi.org/10.1109/ACCESS.2020.2966228>
- Guan, S., Zhuang, Z., Tao, H., et al. (2023). Feedback-aided pd-type iterative learning control for time-varying systems with non-uniform trial lengths. *Transactions of the Institute of Measurement and Control*, 45(11), 2015–2026. <https://doi.org/10.1177/01423312221142564>
- Hansen, N. (2006) *The CMA Evolution Strategy: A Comparing Review*, Springer, Berlin, Heidelberg, pp 75–102. [https://doi.org/10.1007/3-540-32494-1\\_4](https://doi.org/10.1007/3-540-32494-1_4)
- Hax, A. C., & Majluf, N. S. (1982). Competitive cost dynamics: The experience curve. *Interfaces*, 12(5), 50–61. <https://doi.org/10.1287/inte.12.5.50>
- He, Z., Tran, K. P., Thomassey, S., et al. (2022). Multi-objective optimization of the textile manufacturing process using deep-q-network based multi-agent reinforcement learning. *Journal of*

- Manufacturing Systems*, 62, 939–949. <https://doi.org/10.1016/j.jmsy.2021.03.017>
- He, Z., Xu, J., Tran, K. P., et al. (2021). Modeling of textile manufacturing processes using intelligent techniques: a review. *The International Journal of Advanced Manufacturing Technology*, 116(1–2), 39–67. <https://doi.org/10.1007/s00170-021-07444-1>
- Hill, A., Raffin, A., & Ernestus, M., et al (2018) Stable baselines. <https://github.com/hill-a/stable-baselines>, accessed: 26/09/2023
- Hoffmann, M. (2019). Smart Agents for the Industry 4.0. Springer Fachmedien, Wiesbaden, <https://doi.org/10.1007/978-3-658-27742-0>
- Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning*, 110(3), 457–506. <https://doi.org/10.1007/s10994-021-05946-3>
- Ibarz, J., Tan, J., Finn, C., et al. (2021). How to train your robot with deep reinforcement learning: lessons we have learned. *The International Journal of Robotics Research*, 40(4–5), 698–721. <https://doi.org/10.1177/0278364920987859>
- Ismail, M., Mostafa, N. A., & El-assal, A. (2022). Quality monitoring in multistage manufacturing systems by using machine learning techniques. *Journal of Intelligent Manufacturing*, 33(8), 2471–2486. <https://doi.org/10.1007/s10845-021-01792-1>
- ISO (2007) Textiles - test methods for nonwovens: Part 18: Determination of breaking strength and elongation of nonwoven materials using the grab tensile test. ISO Standard ISO 9073-18:2007, International Organization for Standardization, Geneva, Switzerland
- Kins, R., & Gries, T. (2023) Added value through linking of product and process data on the example of a textile process chain. In: Trauth D (ed) *The Monetization of Technical Data: Innovations from Industry and Research*. Springer, Berlin, Heidelberg, p 315–328, [https://doi.org/10.1007/978-3-662-66509-1\\_18](https://doi.org/10.1007/978-3-662-66509-1_18)
- Kirschner, J., Mutný, M., & Hiller, N., et al. (2019) Adaptive and safe bayesian optimization in high dimensions via one-dimensional subspaces. In: *Proceedings of the 36th International Conference on Machine Learning*, vol 97. PMLR, Long Beach, CA, pp 3429–3438 <https://doi.org/10.3929/ethz-b-000385951>
- Kocak, B. (2022). Key concepts, common pitfalls, and best practices in artificial intelligence and machine learning: focus on radiomics. *Diagnostic and Interventional Radiology*, 28(5), 450–462. <https://doi.org/10.5152/dir.2022.211297>
- Ledzińska, M., & Postek, S. (2017). From metaphorical information overflow and overload to real stress: Theoretical background, empirical findings, and applications. *European Management Journal*, 35(6), 785–793. <https://doi.org/10.1016/j.emj.2017.07.002>
- Li, P. (2007). Prozessoptimierung unter Unsicherheiten. *Oldenbourg, München und Wien.*, <https://doi.org/10.1524/9783486842500>
- Mao, N., & Russel, J. (2007). Characterisation, testing and modelling of nonwoven fabrics. In S. J. Russell (Ed.), *Handbook of nonwovens* (pp. 401–514). Boca Raton: CRC Press.
- Möbitz, C. (2021). *Optimisation of energy efficiency of pneumatic fibre conveying*. Düren: Shaker.
- Nian, R., Liu, J., & Huang, B. (2020). A review on reinforcement learning: Introduction and applications in industrial process control. *Computers & Chemical Engineering*, 139, 106886. <https://doi.org/10.1016/j.compchemeng.2020.106886>
- Panzer, M., & Bender, B. (2022). Deep reinforcement learning in production systems: a systematic literature review. *International Journal of Production Research*, 60(13), 4316–4341. <https://doi.org/10.1080/00207543.2021.1973138>
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12, 2825–2830. <https://doi.org/10.5555/1953048.207819>
- Peters, J., Janzing, D., & Schölkopf, B. (2017). *Elements of causal inference: Foundations and learning algorithms*. Cambridge, Massachusetts: The MIT Press.
- Pham, T.H., de Magistris, G., & Tachibana, R. (2018). Optlayer - practical constrained optimization for deep reinforcement learning in the real world. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, Brisbane, pp 6236–6243, <https://doi.org/10.1109/ICRA.2018.8460547>
- Pohlmeyer, F., Kins, R., Cloppenburg, F., et al. (2022). Interpretable failure risk assessment for continuous production processes based on association rule mining. *Advances in Industrial and Manufacturing Engineering*, 5, 100095. <https://doi.org/10.1016/j.aime.2022.100095>
- Pongboonchai-Empl, T., Antony, J., & Garza-Reyes, J.A. et al (2023) Integration of industry 4.0 technologies into lean six sigma dmaic: a systematic review. *Production Planning & Control* pp 1–26. <https://doi.org/10.1080/09537287.2023.2188496>
- Qiu, P. (2013). *Introduction to Statistical Process Control*. CRC Press, Hoboken, NJ. <https://doi.org/10.1201/b15016>
- Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3rd ed.). Boston: Pearson.
- Scheiderer, C., Thun, T., Idzik, C., et al. (2020). Simulation-as-a-service for reinforcement learning applications by example of heavy plate rolling processes. *Procedia Manufacturing*, 51, 897–903. <https://doi.org/10.1016/j.promfg.2020.10.126>
- Schlichter, S., Rübenach, B., Morgner, B., et al. (2012). Trockenverfahren. In H. Fuchs & W. Albrecht (Eds.), *Vliesstoffe* (pp. 123–228). Weinheim: Wiley-VCH.
- Schölkopf, B (2022) Causality for machine learning. In: Geffner, H., Dechter, R., Halpern, J.Y. (eds). *Probabilistic and Causal Inference - The Works of Judea Pearl*, vol 27. ACM, New York, p. 765–804, <https://doi.org/10.1145/3501714.3501755>
- Stojanovic, V. (2023). Fault-tolerant control of a hydraulic servo actuator via adaptive dynamic programming. *Mathematical Modelling and Control*, 3(3), 181–191. <https://doi.org/10.3934/mmc.2023016>
- Sui, Y., Zhuang, V., & Burdick, J.W., et al. (2018) Stagewise safe bayesian optimization with gaussian processes. In: Dy, J., Krause, A. (eds). *International Conference on Machine Learning*, pp 7602–7613, <https://api.semanticscholar.org/CorpusID:49318805>
- Sutton, R. S., Barto, A. G., & Bach, F. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). Massachusetts: MIT Press Ltd.
- Tao, H., Zheng, J., Wei, J., et al. (2023). Repetitive process based indirect-type iterative learning control for batch processes with model uncertainty and input delay. *Journal of Process Control*, 132, 103112. <https://doi.org/10.1016/j.jprocont.2023.103112>
- Thomassey, S., & Zeng, X. (2018). Introduction: Artificial intelligence for fashion industry in the big data era. In: Thomassey S, Zeng X (eds) *Artificial Intelligence for Fashion Industry in the Big Data Era*. Springer, Singapore, <https://doi.org/10.1007/978-981-13-0080-6>
- Titsias, M.K. (2009). Variational learning of inducing variables in sparse gaussian processes. In van Dyk, D., Welling, M. (eds). *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, Clearwater Beach, FL, pp 567–574, <https://api.semanticscholar.org/CorpusID:7811257>
- Towers, M., Terry, J.K., & Kwiatkowski, A., et al. (2023) Gymnasium. <https://zenodo.org/record/8127025>
- Vedpal, J. (2013). Basics of process control in textile manufacturing. In A. Majumdar (Ed.), *Process control in textile manufacturing* (Vol. 127, pp. 3–14). Oxford: Woodhead publishing.
- Weichert, D., Link, P., Stoll, A., et al. (2019). A review of machine learning for the optimization of production processes. *The International Journal of Advanced Manufacturing Technology*, 104(5–8), 1889–1902. <https://doi.org/10.1007/s00170-019-03988-5>

- Wilson, A. (2007). Development of the nonwovens industry. In S. J. Russell (Ed.), *Handbook of nonwovens* (pp. 1–16). CRC Press, Boca Raton: Woodhead publishing series. in textiles.
- Zhuang, Z., Tao, H., Chen, Y., et al. (2023). An optimal iterative learning control approach for linear systems with nonuniform trial lengths under input constraints. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(6), 3461–3473. <https://doi.org/10.1109/TSMC.2022.3225381>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.