

# Quantum enhanced ensemble GANs for anomaly detection in continuous biomanufacturing<sup>†</sup>

Rajiv Kailasanathan<sup>a,d</sup>, William R. Clements<sup>b</sup>, Mohammad Reza Boskabadi<sup>a</sup>, Shawn M. Gibford<sup>a</sup>, Emmanouil Papadakis<sup>c</sup>, Christopher J. Savoie<sup>e</sup> and Seyed Soheil Mansouri<sup>\*a</sup>

The development of continuous biomanufacturing processes requires robust and early anomaly detection, since even minor deviations can compromise yield and stability, leading to disruptions in scheduling, reduced weekly production, and diminished economic performance. These processes are inherently complex and exhibit non-linear dynamics with intricate relationships between process variables, thus making advanced methods for anomaly detection essential for efficient operation. In this work, we present a novel framework for unsupervised anomaly detection in continuous biomanufacturing based on an ensemble of generative adversarial networks (GANs). We first establish a benchmark dataset simulating both normal and anomalous operation regimes in a continuous process for the production of a small molecule. We then demonstrate the effectiveness of our GAN-based framework in detecting anomalies caused by sudden feedstock variability. Finally, we evaluate the impact of using a hybrid quantum/classical GAN approach with both a simulated quantum circuit and a real photonic quantum processor on anomaly detection performance. We find that the hybrid approach yields improved anomaly detection rates. Our work shows the potential of hybrid quantum/classical approaches for solving real-world problems in complex continuous biomanufacturing processes.

## 1 Introduction

The global biomanufacturing industry is expected to be valued between 500B\$ - 1T\$ by 2035<sup>1</sup>, and one of the key technological advancements required to achieve sustainable bioprocesses is the shift from batch to continuous mode of operation. Several challenges lie in the way of developing continuous processes, some of which are specific to bioprocesses, such as variation in quality of feedstock, genetic instability after many generations, or contamination risk. Given the high cost of operation of these processes, it is necessary to detect any anomalies in the process at a very early stage in order to intervene and prevent loss. This is also further reinforced in the form of strict regulatory requirements by several agencies that place strict regulations not only on the final product quality, but also on the quality of the process<sup>2</sup>. These requirements motivate the need for robust anomaly detection tools that can inform the operator of any anomalies in the process at an early stage.

Through steady advancements of process monitoring tools and the growing availability of high-resolution data, industries across domains, including biomanufacturing, are making significant strides in the areas of model based process understanding and the development of digital twins<sup>3-5</sup>. The development of robust methods for handling multivariate process data with multiple features is an active area of scientific and engineering research. One

of the most important requirements of these methods is to detect anomalies in the process data, ideally as early as possible in the process, so that the process operator can make appropriate decisions about the changes in the process to prevent error and loss of yield or entire batches<sup>6</sup>.

Deep learning-based approaches have been explored for anomaly detection in other industrial processes. Particularly, unsupervised methods that attempt to learn the characteristics of normal operational data to generate realistic synthetic process data have been explored for anomaly detection in a wide variety of fields. These include medicine<sup>7,8</sup>, video surveillance<sup>9</sup>, and quality inspection<sup>10</sup>. In the field of industrial process monitoring, generative models, particularly Generative Adversarial Networks (GANs), for anomaly detection are being explored in various sectors including nuclear and conventional power plant operation<sup>11,12,13</sup>, steam turbine health<sup>14</sup>, Oil and gas production<sup>15</sup>, fused magnesium furnace<sup>16</sup>. Anomaly detection strategies based on generative modeling tackle a critical and practical challenge: imbalanced process data. In industrial processes, normal operation is usually ensured through strict operational practices and as a result anomalous operation is usually rare. Hence, there is a huge disparity between the data that is available between normal operation and anomalous operation<sup>17</sup>. The modeling strategy for anomaly detection thus must be agnostic of the size of the anomalous process dataset. Generative models adopt an unsupervised learning strategy, and thus can be developed using only normal operational data.

However, one of the biggest obstacles faced by academic researchers is the availability of industrial process data<sup>18</sup>. To overcome this challenge, simulation models of the real physical processes are used to create data whenever possible. Particularly in chemical processes, the Tennessee Eastman simulated process dataset<sup>19</sup> has been widely used to benchmark emerging methods in anomaly detection in industrial processes, particularly using

<sup>a</sup> Department of Chemical and Biochemical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark

<sup>b</sup> ORCA Computing, London, UK

<sup>c</sup> Novo Nordisk, Kalundborg, Denmark

<sup>d</sup> Novo Nordisk Foundation Centre for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark

<sup>e</sup> SiC Systems Inc., Franklin, TN, USA

These authors have contributed equally: Rajiv Kailasanathan, William Clements and Mohammad Reza Boskabadi

\*Corresponding author: Seyed Soheil Mansouri (seso@kt.dtu.dk)

<sup>†</sup> Data and code available: <https://github.com/RajivKailasanathan/ADCoB>

generative models<sup>20 21 22 23</sup>.

Meanwhile, quantum GANs (QGANs) have emerged at the intersection of quantum computing and generative modeling. In these architectures, a quantum circuit makes use of the fundamental principles of quantum mechanics, including superposition, interference, and entanglement, to generate samples from probability distributions that are intractable to simulate classically<sup>24,25</sup>. Recent studies combine classical generative models with parameterized quantum circuits to create hybrid quantum-classical algorithms and observe improved performance<sup>26–30</sup>. Though the mechanisms that lead to this improved performance are still being elucidated, it has been suggested that the highly correlated and non-separable nature of quantum distributions can help neural networks produce more diverse data<sup>31</sup>. In the context of industrial process monitoring, hybrid quantum classical algorithms are thus an attractive strategy for advancing anomaly detection methods. Early applications of hybrid quantum-classical methods for anomaly detection have shown promise<sup>32–34</sup>.

In this work, we develop a framework for anomaly detection in a continuous biomanufacturing process. Biomanufacturing is a complex process with multiple sequential unit operations, feedback loops, and control units that results in highly non-linear process dynamics and complex relationships between process variables. To generate process data, we employ KTB-1, an established dynamic simulation model of a continuous process for production of Lovastatin<sup>35</sup>, a commercially relevant Active Pharmaceutical Ingredient. Leveraging insight from domain knowledge, we also introduce faults in the simulation that result in process anomalies. Thus, we establish a benchmark dataset for normal and faulty operations of the continuous biomanufacturing process. Subsequently, we investigate the applicability of GANs for anomaly detection in a continuous biomanufacturing process. We also extend our study to explore the potential of complex latent distributions generated by a quantum simulator and a real photonic quantum circuit in enhancing the performance of the GANs. The key contributions of this paper are:

1. Development of a benchmark dataset for normal and faulty operations of a continuous biomanufacturing process. Sudden variation in the quality of feedstock is modelled as a fault, which aligns closely with real process deviations that are unique to bioprocesses.
2. Development of an early anomaly detection framework using an ensemble of GANs. While GANs have applied for anomaly detection in industrial processes, this is the first study employing GANs for anomaly detection in continuous biomanufacturing.
3. Novel application of quantum computers to enhance the performance of GANs. We explore the potential of hybrid quantum-classical algorithms in advancing the development of new processes in biomanufacturing.

To ensure transparency and reproducibility, we also make our dataset and code publicly available at <https://github.com/RajivKailasanathan/ADCoB>.

## 2 Methodology

### 2.1 Overview

The anomaly detection workflow proposed in this study consists of a decomposed solution strategy that is applicable for any industrial time-series process data set. Figure 1 shows an overview of the anomaly detection framework.

We use an unsupervised approach to train an ensemble of generative adversarial networks (GANs). This training section uses only normal operational data of the process, and does not require any anomalous data. As a case study, we demonstrate the use of this framework on a continuous biomanufacturing process. A dynamic simulation model of the process (KT Biologics-1) is used to generate data required to train and test the model. By providing standard operation conditions (including normal fluctuations in process parameters), 'normal' operational data are generated. Similarly, sudden faults in the operational conditions are introduced to create anomalous data. The normal operation data is exclusively used to train the GANs. We also investigate different variations of a GAN, where we change the type of input distribution used by the generator to seed the synthetic data generation process. We compare the performance of three different distributions: a Gaussian distribution, a distribution generated by a simulated quantum circuit, and a distribution generated by a real photonic processor (ORCA Computing PT-2).

After the training process, the discriminators of the GANs are stored and used in the anomaly detection module. The input process is classified as anomalous if any one of the discriminators detects that it lies out of the distribution of the normal (training) data.

### 2.2 Dataset generation

Our work uses the KTB1 simulation model, implemented in MATLAB Simulink, to generate synthetic anomalies across a plantwide biomanufacturing system. The simulation provides a dynamic and modular environment for modeling process behavior and fault conditions, which has for example been used in prior work evaluating machine learning approaches for parameter estimation<sup>36</sup>. The selected case study focuses on the production of lovastatin, an active pharmaceutical ingredient (API), using *Aspergillus terreus* as the fungal host organism. The simulated process captures both upstream and downstream operations, replicating the integrated nature of industrial-scale biomanufacturing.

In the upstream section, raw materials, adenine and lactose, serving as nitrogen and carbon sources, are introduced into a mixing unit (M-101), where they are homogenized into a uniform feed solution. This medium is continuously supplied to a Continuous Stirred-Tank Reactor (CSTR), R-101, where *Aspergillus terreus* ferments the feedstock to biosynthesize lovastatin. The resulting fermentation broth is directed to a hydro-cyclone separator (HC-101), which separates the suspension into two output streams: a biomass-rich stream (R), part of which is recycled to the reactor to sustain biomass levels, and a lean stream (stream 5), which advances to downstream purification. This recycling loop introduces nonlinear dynamic behavior, making the process sensitive to disturbances and faults, conditions ideal for testing

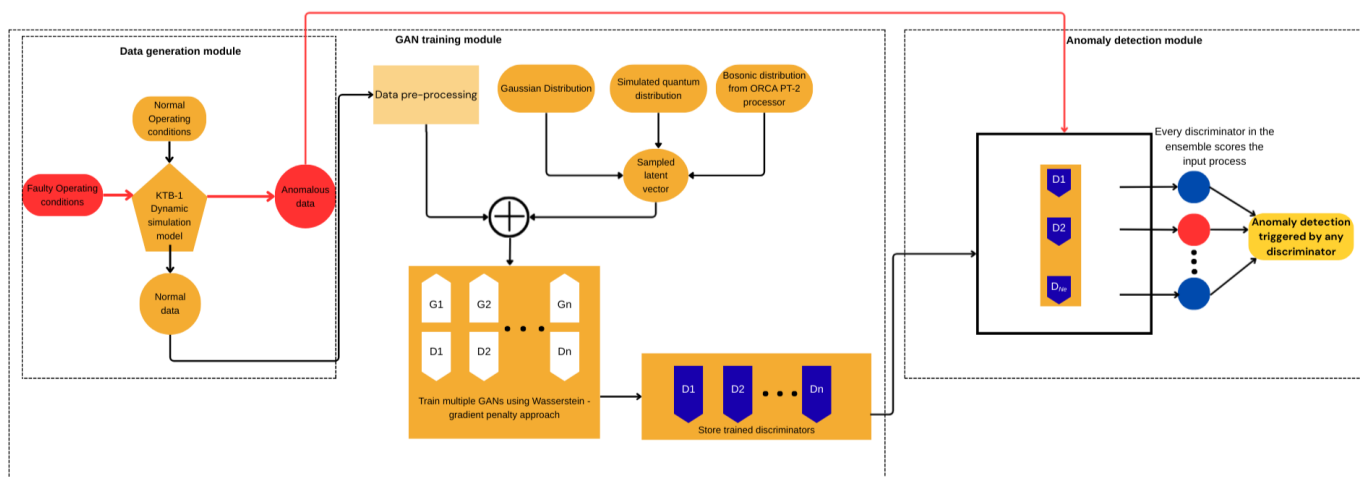


Fig. 1 The anomaly detection workflow using an ensemble of GANs proposed in this study. Left: first, using the KTB1 model for simulating the production process for lovastatin, a pharmaceutical ingredient, we produce a set of time series representing both normal and anomalous operating regimes. Center: we train an ensemble of generative adversarial networks (GANs), each consisting of a generator ( $G$ ) and discriminator ( $D$ ) neural network, on the data from the normal regime. The generator is trained to map samples from an input probability distribution to synthetic data. Our work compares GANs trained with three different types of distributions: a classical Gaussian distribution, a simulated quantum distribution, and a real quantum distribution produced by an ORCA Computing PT-2 system. Right: During the training process, each discriminator learns a decision boundary distinguishing between the normal operation data and other data. An ensemble of discriminators can thus be used as a detection anomaly module providing a set of learned criteria for flagging anomalies. If any discriminator in the ensemble flags a time series as anomalous, then the anomaly detection module labels the time series as an anomaly.

anomaly detection performance. The downstream section consists of a sequence of separation and purification stages. Centrifuges C-101 and C-102 remove remaining biomass and dense particulates. The clarified broth is then concentrated via nanofiltration (NF-101), which retains lovastatin molecules while allowing smaller solutes to pass through. Final purification is performed through a series of six reversed-phase chromatography columns (LC-101 to LC-106), which isolate and enrich lovastatin based on molecular interactions. This complete Simulink-based process model enables the controlled injection of faults, such as flow deviations, concentration changes, membrane fouling, or sensor drifts across multiple units, producing realistic anomaly scenarios that challenge conventional detection methods. Figure 2 presents a process flow diagram summarizing the full system configuration.

### Anomaly Dataset Generation

To develop and test fault detection and diagnosis algorithms, both normal and faulty datasets were generated using the KTB1 plant-wide dynamic simulator. Anomalies were introduced by implementing step changes in the lactose concentration ( $Clac,M$ ), which propagate through the process and affect the downstream behavior of the system. While the specific substrate differs across processes, concentration variations in feed streams are recognized disturbances in biomanufacturing, as also demonstrated in studies of continuous mAb production<sup>37</sup>. Each simulation run captures the dynamic response of the plant under different operating conditions using Monte Carlo simulations with randomized parameters.

For normal operating conditions, no fault was introduced. A total of 1000 Monte Carlo simulations were conducted by varying the mean of  $Clac,M$  using a normal distribution (mean = 20, variance = 0.01). This variation was implemented via a random number generator block, providing natural process variability while maintaining normal operation. Each simulation was run for 200 hours, and output data was sampled at each hour

For anomalous operating conditions, a fault was introduced by applying a step change in the lactose concentration ( $Clac,M$ ) at time step 10 of the simulation. The value was shifted from an initial baseline of 20 to a final value of 19.7. To make this fault more realistic and reflective of actual industrial conditions, the step signal was combined with a random noise generator, which simulates natural variability in the process. We used normally distributed noise with a mean of 0 and a standard deviation of 0.1. Our initial testing found that these parameters created a challenging anomaly detection dataset, where the anomalies could be visually identified by plotting the data over the entire time series but harder to identify at shorter time scales.

As a result, before time step 10, the signal fluctuates randomly around the nominal value of 20. After time step 10, the signal continued to fluctuate, but now around the new faulty value of 19.7. This composite signal allowed the simulation to reflect both a significant operational change and continuous environmental variability. A total of 100 time series were generated under this fault condition. These time series were used only for evaluating the performance of the anomaly detection algorithm trained on normal data.

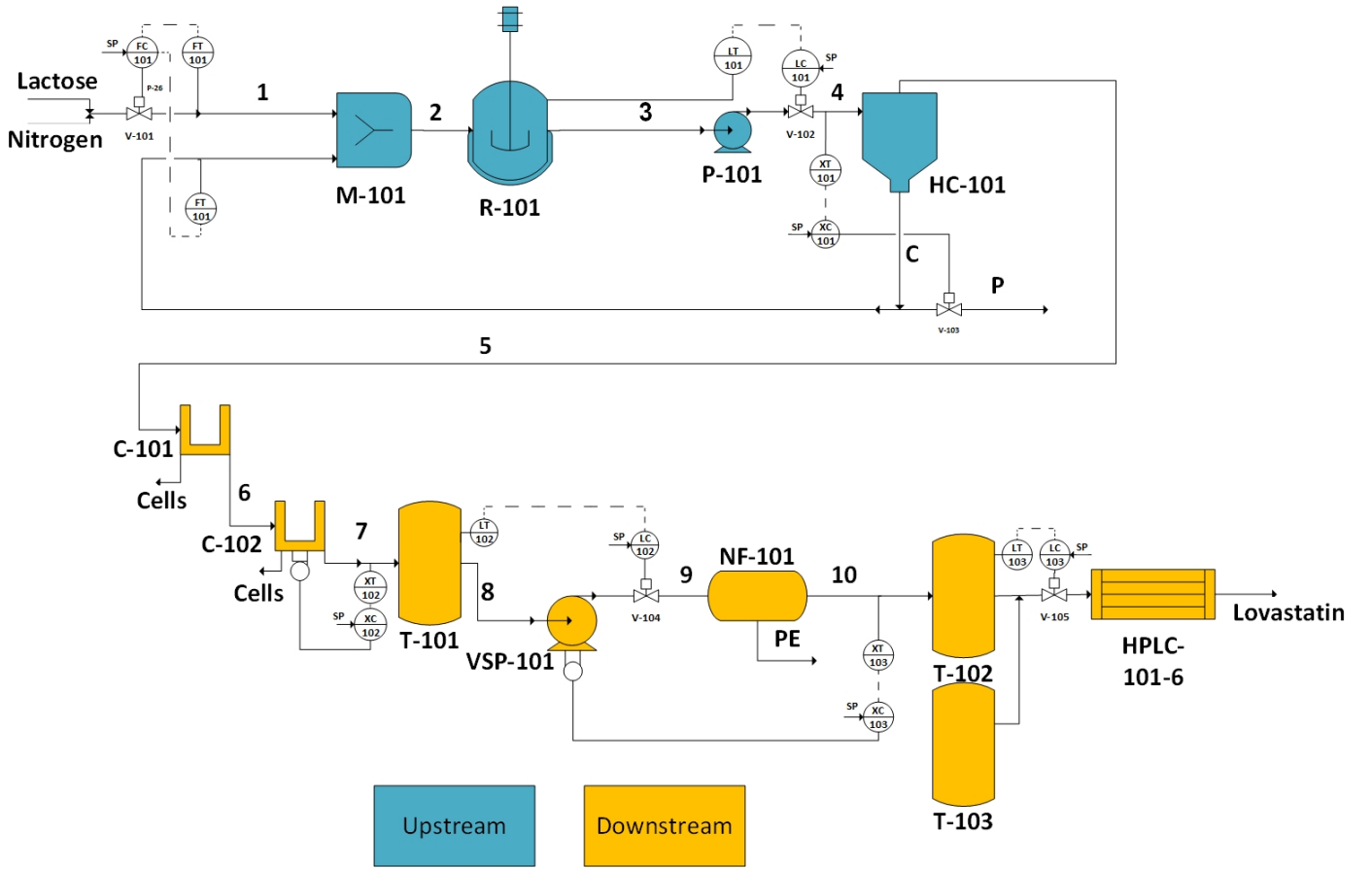


Fig. 2 Process flow diagram of the KTB-1 simulation model. The upstream section focuses on enhancing Lovastatin production (with V for valves, R for reactors, P for pumps, and HC for hydrocyclones), while the downstream section is dedicated to Lovastatin purification and includes the associated control system for dynamic process management (where C stands for centrifuge, T for tank, and NF for nanofiltration).

### 2.3 GANs for anomaly detection

We adopt an unsupervised anomaly detection method based on an ensemble of discriminator neural networks trained within a generative adversarial network (GAN) algorithm. Using GAN discriminators for unsupervised anomaly detection is a well-studied idea in machine learning<sup>8,38–40</sup>. For more robust anomaly detection, ensembles of discriminators have also been investigated<sup>41</sup>. GAN-based methods for unsupervised anomaly detection have been proposed within a smart factory context<sup>42</sup>, but ours is the first to investigate this approach for bio-manufacturing.

As shown in figure 3 in a GAN a generator and a discriminator are trained in an adversarial manner. The discriminator is trained to distinguish between real data and synthetic data produced by the generator, and the generator is trained to produce synthetic data that can fool the discriminator<sup>43</sup>. During training, the only real data that is used is the normal data, and no anomalous data is required. In this adversarial process, the discriminator effectively learns to distinguish between any out-of-distribution data produced by the generator and the normal data. After training, the generator is discarded and the discriminator is saved. When the discriminator is subsequently exposed to real anomalous data, the expectation is that it should flag it as such.

To make this method more robust, an ensemble of GANs can be used. A single discriminator learns the boundary between nor-

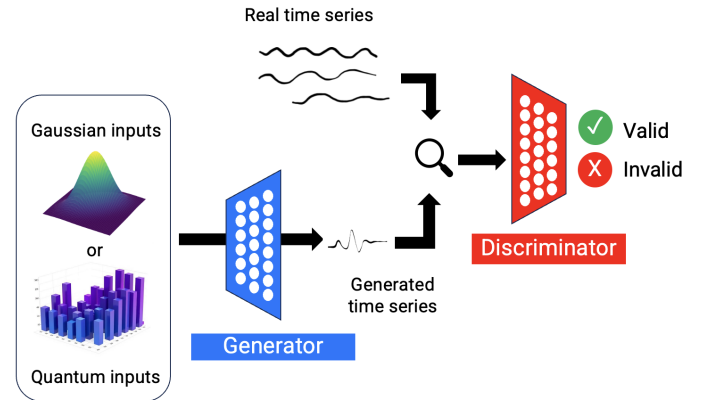


Fig. 3 Our anomaly detection scheme uses generative adversarial networks, in which two neural networks are trained in an adversarial way. The generator is trained to map samples from an initial distribution to realistic synthetic data. In this work, we investigate using either a classical Gaussian distribution or a quantum distribution. The discriminator learns to distinguish between the real data and the synthetic data. After training, we discard the generator and use the discriminator to determine whether an input time series is anomalous.

mal data and the specific out-of-distribution data produced by one generator. The discriminator thus learns one specific decision boundary in the problem space, but this is typically insufficient to detect a wide range of potential anomalies. Training multiple independent GANs can be an effective way of learning multiple such decision boundaries. Though training multiple GANs increases the computational cost of this method, in this work we consider lightweight models that can be trained in a reasonable amount of time.

In our specific implementation, we train an ensemble of Wasserstein GANs with gradient penalty<sup>44</sup>. These are known to be more stable to train and yield higher-quality results than the original GAN proposal. Both the generator and the discriminator consist of feed-forward neural networks with 2 hidden layers and LeakyReLU activation. The generator uses as input a latent vector drawn from a probability distribution known as the latent space, which in our work is either a Gaussian or a quantum probability distribution. We run experiments with 16-dimensional latent vectors.

After training, we use a "union rule" to flag anomalies, where if any one of the ensemble of discriminators flags the data as anomalous then the data is flagged as anomalous. This reflects the notion that each discriminator learns a different decision boundary between normal data and out-of-distribution data, such that anomalous data will generally lie outside the decision boundary for some but not all models. Specifically, since each discriminator output is a real number, we first determine the bottom- $x$  and top- $x$  quantiles (for a given value of  $x$ ) of this value for the normal data, then any data that is outside these thresholds is flagged as anomalous. Different values of  $x$  lead to different true positive and false positive rates, and as such we present our results as receiver operating characteristic (ROC) curves. We also present the associated areas under the curve (AUC).

## 2.4 Hybrid quantum/classical GANs

Anomaly detection methods based on ensembles of GANs require using generator architectures that can produce highly diverse data. Producing highly diverse data is expected to help the discriminator learn more useful decision boundaries between normal data and out-of-distribution data within the problem space. However, despite significant improvements in GAN models since they were first proposed, GANs often struggle with "mode collapse", where data diversity can be minimal<sup>45</sup>.

Recently, several studies using quantum processors as part of a hybrid quantum/classical GAN architecture have found that these methods can outperform purely classical GANs<sup>28–31,46</sup>. Quantum processors can sample from probability distributions that are intractable classically<sup>47</sup>. In these studies, samples drawn from these quantum distribution are used at the input of the generator instead of a classical distribution. This method has been observed to help the generator produce more diverse outputs<sup>31,46</sup>. We hypothesized that generators that produce more diverse synthetic data should in turn produce discriminators that learn a more informative decision boundary between normal and out-of-distribution data. As such, in this study we investigate the per-

formance of an ensemble of discriminators trained with hybrid quantum/classical generators, and compare it to the performance of an ensemble of discriminators trained using purely classical generators.

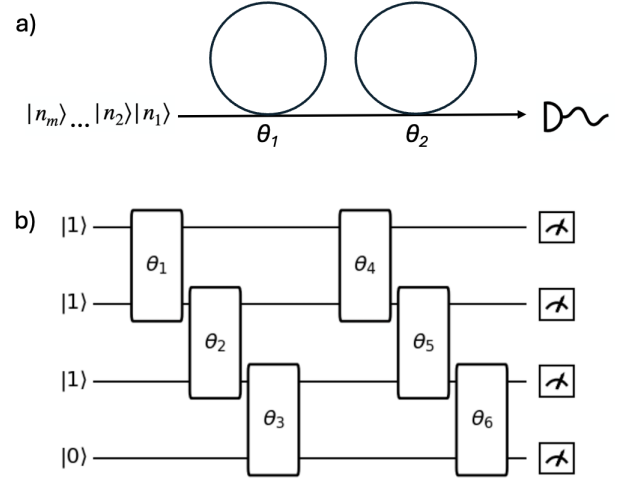


Fig. 4 Our experiments using a quantum processor use photonic quantum processors implemented with a time-bin architecture. As shown in a, in these architectures single photons are sequentially sent into a network of optical delay lines with programmable coupling coefficients. This creates an entangled state between different time bins which can be measured by a photon number resolving detector. The two-loop architecture shown in a) implements the quantum optical circuit shown in b), where each gate is a programmable beam splitter between two optical modes.

We use both simulated and real photonic quantum processors using the architecture shown in figure 4. In a photonic quantum processor, identical single photons are injected into an interference network. Interference between the photons creates a complex quantum state, which is collapsed by a measurement process that returns the output locations of the photons. Simulating this process classically requires computational resources that scale exponentially with the number of photons, such that simulations with more than a few tens of photons are intractable<sup>48</sup>. Our experiments consider small-scale quantum processors with 8 photons interfering in 16 channels.

Photonic quantum processors have several advantages compared to other quantum computing modalities that motivate using them for hybrid machine learning algorithms. Though near-term photonic systems are not universal for quantum computation, they naturally produce highly non-uniform distributions<sup>49</sup> that differ significantly from commonly used classical latent distributions. Moreover, photonic systems are not subject to decoherence, which pushes the outputs of quantum circuits implemented with other modalities closer to a uniform distribution<sup>50</sup>. Finally, their room-temperature operation and commercial availability make their integration into real-world workflows more straightforward<sup>51</sup>.

Our classical simulations of these processors use the algorithm from Clifford and Clifford<sup>52</sup>, whereas our experiments with a real quantum processor use a commercially available ORCA Computing PT-2. We note that classical simulations of these systems are

only tractable for the small-scale quantum systems considered in this work, and become intractable at larger sizes. We use randomly initialized and untrained photonic quantum circuits to provide an apples-to-apples comparison with the untrained Gaussian distribution used for the classical GANs. However, we note that these circuits could in principle be trained jointly with the classical neural networks in the generator and discriminator<sup>53</sup>.

### 3 Results

We perform experiments that aim to answer different research questions. First, we perform a range of experiments using ensembles of 30 models with different latent spaces. These experiments aim to validate our approach to anomaly detection, and elucidate how the choice and size of the latent space impact the overall model performance. Comparing the performance of the simulated quantum processor to that of the real quantum processor also allows us to examine the transferability of our method to real quantum hardware. Our second experiments use a larger ensemble of 60 models to investigate how model performance scales with the number of discriminators.

#### 3.1 Experiments with an ensemble of 30 discriminators

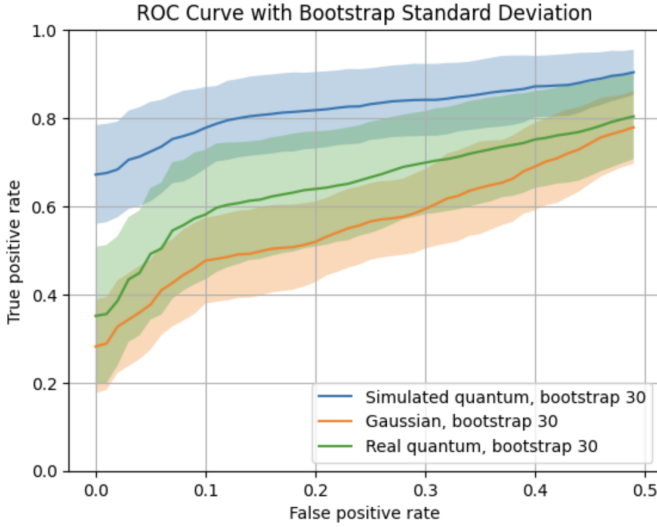


Fig. 5 Receiver operating characteristic (ROC) curves of our anomaly detection algorithm based on an ensemble of 30 GAN discriminators. We consider GANs that use a classical Gaussian latent space or a quantum distribution produced by the quantum interference between identical photons. The quantum distribution is generated either by a real or by a simulated quantum processor. The GANs differ only in the process that produced the latent vectors sent to the generator. The solid line and shaded areas correspond respectively to the mean and standard deviation calculated by sampling 30 models 20 times from an ensemble of 120 trained models.

Our first experiments focus on using ensembles of 30 discriminators for anomaly detection. To evaluate the mean and standard deviation of the performance of an ensemble for anomaly detection, for each experiment we train 120 GAN models and use 20 bootstrap ensembles of 30 models drawn from this larger set.

ROC curves from these different model ensembles are shown in

| Latent space                   | AUC               |
|--------------------------------|-------------------|
| Simulated quantum distribution | $0.409 \pm 0.012$ |
| Real quantum distribution      | $0.324 \pm 0.026$ |
| Gaussian distribution          | $0.280 \pm 0.023$ |

Table 1 Area under the curve (AUC) of the receiver operating characteristic curves of our anomaly detection algorithm for a size-30 bootstrap ensemble.

figure 5, with areas under the curve (AUC) shown in table 1. We observe that all approaches successfully detect anomalies, with some significant differences in performance between the models. These results validate the use of an ensemble of GAN discriminators as a promising method for unsupervised anomaly detection in a high-dimensional bio-manufacturing environment.

We also observe that the best results, with over 50% anomaly detection success at 0% false positives, are achieved with the quantum distributions. These results indicate that the greater diversity of the data produced by the hybrid quantum/classical discriminators helped the discriminators learn more useful decision boundaries for anomaly detection. We also find that the real quantum processor outperforms the classical Gaussian distribution but underperforms the simulated quantum distribution. This suggests that experimental imperfections such as photon loss have a material impact on performance, but using even an imperfect real quantum processor can still yield better results than the most commonly used Gaussian distribution. These results validate the use of a hybrid quantum/classical approach for detecting anomalies.

#### 3.2 Using a larger ensemble

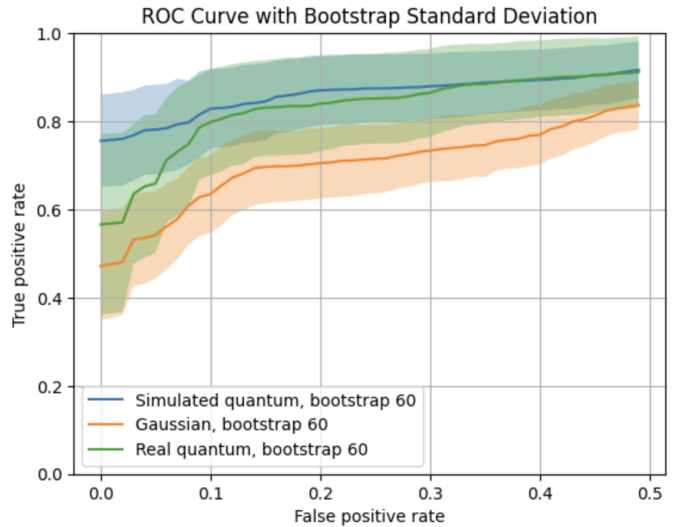


Fig. 6 ROC curves of our anomaly detection method based on an ensemble of 60 GAN discriminators using different latent spaces.

We also investigate the performance of our method using a larger ensemble of 60 models sampled from the size-120 ensemble, with results shown in figure 6 and table 2. We find that the larger ensemble yields improved results. Moreover, the hierar-

chy between models that was observed for the ensembles of 30 models is reflected here too. These results show that this method scales well with the amount of computation available. Whereas the use of the union rule method of flagging anomalies could have been expected to create a large number of false positives as the ensemble size increases, these results indicate that this is not a problem in practice.

| Latent space                   | AUC               |
|--------------------------------|-------------------|
| Simulated quantum distribution | $0.429 \pm 0.009$ |
| Real quantum distribution      | $0.412 \pm 0.019$ |
| Gaussian distribution          | $0.350 \pm 0.019$ |

Table 2 Area under the curve (AUC) of the receiver operating characteristic curves of our anomaly detection algorithm for a size-60 bootstrap ensemble.

## 4 Discussion

### 4.1 Advantage from quantum computing

Here, we further discuss the role of quantum computing in helping us achieve improved anomaly detection results. We consider two aspects of this question: whether the quantum computer is required to achieve higher-quality results, and the extent to which the quantum computer could be simulated classically.

First, though the quantum distribution provides a practical and well-studied method to improve the performance of a GAN, it is not the only method available. A wide range of other classical distributions can also be used in a GAN, from Bernoulli distributions<sup>54</sup> to trained distributions using an autoencoder<sup>55</sup>. Though (author?)<sup>31</sup> observe an improvement over a wide range of classical distributions, we cannot rule out that some of these methods may yield improved performance on this problem compared to the baseline Gaussian. However, it is impractical and computationally expensive to investigate a wide range of alternative distributions, while the increasing availability of quantum computers means that they are becoming an increasingly practical alternative. As such, in realistic settings where algorithm development time and computation is limited, adopting a quantum distribution as the go-to option for improving performance may be a more effective use of resources than investigating a wider range of classical distributions.

Second, we note that classical methods for simulating photonic quantum processors in the presence of noise have been developed<sup>56,57</sup>. The methods could in principle be used to derive the same improvement in performance that we observe with a real quantum processor using only classical resources. However, these methods are not practical for hybrid quantum/classical machine learning workflows. The mean-field approach of (author?)<sup>56</sup> reproduces the low-order statistics of a photonic quantum distribution with a quadratic overhead as opposed to a linear overhead at most with a real system. The approach of (author?)<sup>57</sup> was reported to require 9 minutes on 288 GPUs just to set up the simulation. As such, it is unlikely that either approach could run faster than a state-of-the-art photonic quantum processor. Though speed of generating samples is less of a concern when using a static, untrained latent space, this can be a major limi-

tation for more advanced variational settings. In those settings where both the quantum and the classical parts of the algorithm are regularly updated, neither approach thus provides a practical route to simulating these distributions.

### 4.2 Prospects for practical biomanufacturing applications

The convergence of quantum-enhanced machine learning and advanced control strategies presents a transformative opportunity for continuous biomanufacturing. As processes become increasingly complex and data-rich, traditional control systems struggle to maintain robustness in the face of nonlinear dynamics, rare anomalies, and evolving operational conditions. In this context, the integration of hybrid quantum Generative Adversarial Networks (GANs) with Model Predictive Control (MPC) offers a promising pathway toward intelligent, adaptive process management.

Hybrid quantum GANs leverage the expressive power of quantum-generated latent spaces to produce synthetic data that more accurately reflects the manifold of normal process behavior. This enhanced diversity enables discriminators to learn sharper decision boundaries, improving the sensitivity and specificity of anomaly detection. Importantly, these models operate in an unsupervised fashion, relying solely on normal operational data—an essential feature in industrial settings where labeled anomalies are scarce or unavailable.

Embedding this anomaly detection capability within an MPC framework creates a closed-loop system that not only monitors but also responds to deviations in real time. When the GAN ensemble flags an anomaly, the MPC module can dynamically adjust its prediction horizon, tighten constraints, or re-optimize control actions to mitigate potential risks. This dual-layered architecture—where anomaly detection informs control decisions—represents a shift from reactive to proactive process management.

Moreover, the inclusion of direct control signals from the anomaly detection module to the process simulation or plant adds a layer of safety and responsiveness. In scenarios where anomalies indicate critical faults, bypassing the MPC for immediate actuation can prevent cascading failures or product loss. This hybrid control logic, combining predictive optimization with anomaly-triggered overrides, aligns well with the stringent reliability and quality demands of pharmaceutical manufacturing.

From a broader perspective, this framework exemplifies how quantum computing can be pragmatically integrated into industrial workflows. Photonic quantum processors, with their room-temperature operation and commercial availability, offer a feasible route to enhancing machine learning models without requiring full-scale quantum infrastructure. As quantum hardware matures, its role in enabling more expressive generative models—and by extension, more intelligent control systems—is likely to expand.

In summary, the proposed Hybrid Quantum GAN-MPC framework is not merely a technical innovation but a conceptual advancement in how we think about process control. It bridges the gap between data-driven anomaly detection and model-based

control, paving the way for resilient, adaptive, and intelligent biomanufacturing systems.

## 5 Conclusions

In this study, we have introduced a novel framework for unsupervised anomaly detection in continuous biomanufacturing processes, leveraging an ensemble of generative adversarial networks (GANs) enhanced by hybrid quantum-classical approaches. By developing a benchmark dataset using the KTB-1 dynamic simulation model for lovastatin production, we simulated both normal operations and anomalies arising from sudden feedstock variability—a common challenge in bioprocesses. Our ensemble GAN method, trained exclusively on normal data, demonstrated robust early detection capabilities, with receiver operating characteristic (ROC) curves indicating effective discrimination between normal and anomalous regimes. A key innovation lies in the integration of quantum-generated latent distributions, sourced from both simulated and real photonic quantum processors (ORCA Computing PT-2). Comparative experiments revealed that these quantum distributions outperformed classical Gaussian baselines, yielding higher areas under the curve (AUC) values—up to 0.429 for simulated quantum ensembles of 60 discriminators. This improvement underscores the potential of quantum mechanics to enhance generative diversity, mitigating issues like mode collapse and enabling discriminators to learn more informative decision boundaries. Notably, even imperfect real quantum hardware provided advantages over classical methods, validating the practicality of hybrid architectures in near-term applications. This work represents the first application of hybrid quantum-classical GANs for anomaly detection in continuous biomanufacturing, addressing critical needs for robust process monitoring amid nonlinear dynamics and regulatory demands. The framework's unsupervised nature and scalability with ensemble size position it as a promising tool for integration with model predictive control (MPC) systems, facilitating proactive interventions to minimize yield losses and operational disruptions. While our results highlight quantum advantages, limitations include reliance on small-scale quantum processors and the computational cost of larger ensembles. Future research could explore trained quantum circuits, scalability to fault-tolerant quantum hardware, and real-world validation on industrial datasets. Overall, this study paves the way for quantum-enhanced machine learning to drive advancements in sustainable biomanufacturing, offering a pathway toward more resilient and efficient processes.

## Author contributions

**Rajiv Kailasanathan:** Methodology, Software, Investigation, Writing - Original Draft, **William Clements:** Methodology, Software, Investigation, Writing - Original Draft, **Mohammad Reza Boskabadi:** Methodology, Investigation, Writing - Original Draft, **Shawn M. Gibford:** Software, Investigation, Writing - Review & Editing, **Emmanouil Papadakis:** Writing - Review & Editing, **Christopher J. Savoie:** Conceptualization, Methodology, Writing - Review & Editing, **Sayed Soheil Mansouri** Conceptualization, Methodology, Supervision, Writing - Review & Editing, Project administration, Funding acquisition

## Conflicts of interest

There are no conflicts to declare.

## Data availability

All codes and datasets used in this study are available according to FAIR principles: <https://github.com/RajivKailasanathan/ADCoB>

## Acknowledgements

This work was funded by the Novo Nordisk Foundation through grant number NNF20CC0035580, and the Technical University of Denmark.

## References

- 1 D. Thomas, “Measuring the us biomanufacturing economy,” in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 88353, p. V02BT02A024, American Society of Mechanical Engineers, 2024.
- 2 Q. Su, S. Ganesh, M. Moreno, Y. Bommireddy, M. Gonzalez, G. V. Reklaitis, and Z. K. Nagy, “A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing,” *Computers & Chemical Engineering*, vol. 125, pp. 216–231, June 2019.
- 3 L. Lattanzi, R. Raffaeli, M. Peruzzini, and M. Pellicciari, “Digital twin for smart manufacturing: a review of concepts towards a practical industrial implementation,” *International Journal of Computer Integrated Manufacturing*, vol. 34, pp. 567–597, June 2021. Publisher: Taylor & Francis \_eprint: <https://doi.org/10.1080/0951192X.2021.1911003>.
- 4 M. R. Boskabadi, Y. Cao, B. Khadem, W. Clements, Z. N. Gerek, E. Reuthe, A. Sivaram, C. J. Savoie, and S. S. Mansouri, “Industrial agentic ai and generative modeling in complex systems,” *Current Opinion in Chemical Engineering*, vol. 48, p. 101150, 2025.
- 5 I. M. Hernández-Romero, S. M. Gibford, W. Clements, C. J. Savoie, A. Flores-Tlacuahuac, and S. S. Mansouri, “Exploring hybrid quantum-classical algorithms for multiscale bioprocess optimization,” *Industrial & Engineering Chemistry Research*, vol. 64, no. 19, pp. 9484–9499, 2025.
- 6 M. R. Boskabadi, M. Murugaiah, T. R. Nielsen, A. Sivaram, G. Sin, and S. S. Mansouri, “Virtual sensor for sustainable large-scale process monitoring,” *Industrial & Engineering Chemistry Research*, vol. 64, no. 7, pp. 3902–3917, 2025.
- 7 S. K. Lim, Y. Loo, N.-T. Tran, N.-M. Cheung, G. Roig, and Y. Elovici, “DOPING: Generative Data Augmentation for Unsupervised Anomaly Detection with GAN,” in *2018 IEEE International Conference on Data Mining (ICDM)*, pp. 1122–1127, Nov. 2018. ISSN: 2374-8486.
- 8 T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, “fanogan: Fast unsupervised anomaly detection with generative adversarial networks,” *Medical Image Analysis*, vol. 54, pp. 30–44, 2019.
- 9 K.-T. Nguyen, D.-T. Dinh, M. N. Do, and M.-T. Tran, “Anomaly Detection in Traffic Surveillance Videos with GAN-based Future Frame Prediction,” in *Proceedings of the 2020 Interna-*

- tional Conference on Multimedia Retrieval*, ICMR '20, (New York, NY, USA), pp. 457–463, Association for Computing Machinery, June 2020.
- 10 Y. T. Lai, J. S. Hu, Y. H. Tsai, and W. Y. Chiu, “Industrial Anomaly Detection and One-class Classification using Generative Adversarial Networks,” in *2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pp. 1444–1449, July 2018. ISSN: 2159-6255.
  - 11 Y. Choi, H. Lim, H. Choi, and I.-J. Kim, “GAN-Based Anomaly Detection and Localization of Multivariate Time Series Data for Power Plant,” in *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 71–74, Feb. 2020. ISSN: 2375-9356.
  - 12 S. Jintao, G. Chaoyue, S. Hui, S. Jiangang, and L. Zhe, “Data Expansion for Foreign Object Detection in Power Grid,” in *2019 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp. 1–4, Dec. 2019.
  - 13 X. Li, T. Huang, K. Cheng, Z. Qiu, and T. Sichao, “Research on anomaly detection method of nuclear power plant operation state based on unsupervised deep generative model,” *Annals of Nuclear Energy*, vol. 167, p. 108785, Mar. 2022.
  - 14 Z. J. Que, Y. Xiong, and Z. G. Xu, “A Semi-supervised Approach for Steam Turbine Health Prognostics Based on GAN and PF,” in *2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pp. 1476–1480, Dec. 2019. ISSN: 2157-362X.
  - 15 Q. Du, Y. Liu, H. Pan, Q. Ren, and H. Liu, “Multi-Dimensional Time Series Anomaly Detection Model Based on Transformer and GAN for Oil and Gas Production IoT Applications,” in *2024 IEEE 16th International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 688–694, Dec. 2024. ISSN: 2472-7555.
  - 16 S. Lu and H. Gao, “Deep Learning Based Fusion of RGB and Infrared Images for the Detection of Abnormal Condition of Fused Magnesium Furnace,” in *2019 IEEE 15th International Conference on Control and Automation (ICCA)*, pp. 987–993, July 2019. ISSN: 1948-3457.
  - 17 J. Kim, K. Jeong, H. Choi, and K. Seo, “GAN-Based Anomaly Detection In Imbalance Problems,” in *Computer Vision – ECCV 2020 Workshops* (A. Bartoli and A. Fusiello, eds.), (Cham), pp. 128–145, Springer International Publishing, 2020.
  - 18 S. S. Mansouri, A. Sivaram, C. J. Savoie, and R. Gani, “Models, modeling and model-based systems in the era of computers, machine learning and ai,” *Computers & Chemical Engineering*, vol. 194, p. 108957, 2025.
  - 19 J. J. Downs and E. F. Vogel, “A plant-wide industrial process control problem,” *Computers & Chemical Engineering*, vol. 17, pp. 245–255, Mar. 1993.
  - 20 X. Yang and D. Feng, “Generative Adversarial Network Based Anomaly Detection on the Benchmark Tennessee Eastman Process,” in *2019 5th International Conference on Control, Automation and Robotics (ICCAR)*, pp. 644–648, Apr. 2019. ISSN: 2251-2446.
  - 21 V. Pozdnyakov, A. Kovalenko, I. Makarov, M. Drobysheskiy, and K. Lukyanov, “Adversarial Attacks and Defenses in Fault Detection and Diagnosis: A Comprehensive Benchmark on the Tennessee Eastman Process,” *IEEE Open Journal of the Industrial Electronics Society*, vol. 5, pp. 428–440, 2024.
  - 22 F. Schoch, P. Graf, T. Schmiege, C. Wittenberg, C. Lanquillon, and N. C. Stache, “Deep Anomaly Detection with Extended Transformer-Based Model on Tennessee Eastman Process Dataset,” in *2024 IEEE 19th International Conference on Computer Science and Information Technologies (CSIT)*, pp. 1–4, Oct. 2024. ISSN: 2766-3639.
  - 23 F. Hartung, B. J. Franks, T. Michels, D. Wagner, P. Liznerski, S. Reithermann, S. Fellenz, F. Jirasek, M. Rudolph, D. Neider, H. Leitte, C. Song, B. Kloepper, S. Mandt, M. Bortz, J. Burger, H. Hasse, and M. Kloft, “Deep Anomaly Detection on Tennessee Eastman Process Data,” *Chemie Ingenieur Technik*, vol. 95, no. 7, pp. 1077–1082, 2023.
  - 24 Y. Du, M.-H. Hsieh, T. Liu, and D. Tao, “Expressive power of parametrized quantum circuits,” *Physical Review Research*, vol. 2, p. 033125, July 2020. Publisher: American Physical Society.
  - 25 M. P. Andersson, M. N. Jones, K. V. Mikkelsen, F. You, and S. S. Mansouri, “Quantum computing for chemical and biomolecular product design,” *Current Opinion in Chemical Engineering*, vol. 36, p. 100754, 2022.
  - 26 H. Jin and K. M. M. Jr, “Improving GANs by leveraging the quantum noise from real hardware,” July 2025. arXiv:2507.01886 [quant-ph].
  - 27 P.-Y. Kao, Y.-C. Yang, W.-Y. Chiang, J.-Y. Hsiao, Y. Cao, A. Aliper, F. Ren, A. Aspuru-Guzik, A. Zhavoronkov, M.-H. Hsieh, and Y.-C. Lin, “Exploring the Advantages of Quantum Generative Adversarial Networks in Generative Chemistry,” *Journal of Chemical Information and Modeling*, vol. 63, pp. 3307–3318, June 2023. Publisher: American Chemical Society.
  - 28 J. Li, R. O. Topaloglu, and S. Ghosh, “Quantum Generative Models for Small Molecule Drug Discovery,” *IEEE Transactions on Quantum Engineering*, vol. 2, pp. 1–8, 2021.
  - 29 M. S. Rudolph, N. B. Toussaint, A. Katarbarwa, S. Johri, B. Peropadre, and A. Perdomo-Ortiz, “Generation of High-Resolution Handwritten Digits with an Ion-Trap Quantum Computer,” *Physical Review X*, vol. 12, p. 031010, July 2022. Publisher: American Physical Society.
  - 30 M. Wilson, T. Vandal, T. Hogg, and E. G. Rieffel, “Quantum-assisted associative adversarial network: applying quantum annealing in deep learning,” *Quantum Machine Intelligence*, vol. 3, p. 19, June 2021.
  - 31 O. Bacarreza, T. Farnsworth, A. Makarovskiy, H. Wallner, T. Hicks, S. Sempere-Llagostera, J. Price, R. J. A. Francis-Jones, and W. R. Clements, “Quantum latent distributions in deep generative models,” *arXiv preprint arXiv:2508.19857*, 2025.
  - 32 D. Herr, B. Obert, and M. Rosenkranz, “Anomaly detection with variational quantum generative adversarial networks,” *Quantum Science and Technology*, vol. 6, p. 045004, Oct. 2021. arXiv:2010.10492 [quant-ph].
  - 33 B. Kalfon, S. Cherkaoui, J.-F. Laprade, O. Ahmad, and

- S. Wang, “Successive data injection in conditional quantum GAN applied to time series anomaly detection,” *IET Quantum Communication*, vol. 5, no. 3, pp. 269–281, 2024.
- 34 A. Ajagekar and F. You, “Quantum computing assisted deep learning for fault detection and diagnosis in industrial process systems,” *Computers & Chemical Engineering*, vol. 143, p. 107119, Dec. 2020.
  - 35 M. R. Boskabadi, P. Ramin, J. Kager, G. Sin, and S. S. Mansouri, “KT-Biologics I (KTB1): A dynamic simulation model for continuous biologics manufacturing,” *Computers & Chemical Engineering*, vol. 188, p. 108770, Sept. 2024.
  - 36 S. Shahhoseyni, A. Chakraborty, M. R. Boskabadi, V. Venkatasubramanian, and S. S. Mansouri, “Hybrid machine-learning for dynamic plant-wide biomanufacturing,” *Systems and Control Transactions*, pp. 1682–1687, July 2025.
  - 37 A. Schmidt, H. Helgers, L. J. Lohmann, F. Vetter, A. Juckers, M. Mouellef, S. Zobel-Roos, and J. Strube, “Process analytical technology as key-enabler for digital twins in continuous biomanufacturing,” *Journal of Chemical Technology & Biotechnology*, vol. 97, no. 9, pp. 2336–2346, 2022.
  - 38 T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, “Unsupervised anomaly detection with generative adversarial networks to guide marker discovery,” *arXiv preprint arXiv:1703.05921*, 2017.
  - 39 S. Akçay, A. Atapour-Abarghouei, and T. P. Breckon, “Ganomaly: Semi-supervised anomaly detection via adversarial training,” *arXiv preprint arXiv:1805.06725*, 2018.
  - 40 H. Zenati, C.-S. Foo, B. Lecouat, and V. Chandrasekhar, “Adversarially learned anomaly detection,” *arXiv preprint arXiv:1812.02288*, 2018.
  - 41 X. Han, X. Chen, and L.-P. Liu, “GAN ensemble for anomaly detection,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, pp. 4090–4097, 2021.
  - 42 T.-y. Kim, J. Lee, S. Gong, J. Lim, D. Kim, and J. Jeong, “A novel fs-gan-based anomaly detection approach for smart manufacturing,” *Machines*, vol. 13, no. 1, p. 21, 2024.
  - 43 I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” *Advances in neural information processing systems*, vol. 27, 2014.
  - 44 I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, “Improved training of wasserstein gans,” *Advances in neural information processing systems*, vol. 30, 2017.
  - 45 A. Srivastava, L. Valkov, C. Russell, M. U. Gutmann, and C. Sutton, “Veegan: Reducing mode collapse in gans using implicit variational learning,” *Advances in neural information processing systems*, vol. 30, 2017.
  - 46 T. Xiao, X. Zhai, J. Huang, J. Fan, and G. Zeng, “Quantum deep generative prior with programmable quantum circuits,” *Communications Physics*, vol. 7, no. 1, p. 276, 2024.
  - 47 D. Hangleiter and J. Eisert, “Computational advantage of quantum random sampling,” *arXiv preprint arXiv:2206.04079*, 2022.
  - 48 S. Aaronson and A. Arkhipov, “The computational complexity of linear optics,” in *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pp. 333–342, 2011.
  - 49 S. Aaronson and A. Arkhipov, “Bosonsampling is far from uniform,” *arXiv preprint arXiv:1309.7460*, 2013.
  - 50 S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, “Noise-induced barren plateaus in variational quantum algorithms,” *Nature communications*, vol. 12, no. 1, p. 6961, 2021.
  - 51 M. Słysz, P. Rydlichowski, K. Kurowski, O. Bacarreza, E. C. Gomez, Z. Chandani, B. Heim, P. Khalate, W. R. Clements, and J. Fletcher, “Hybrid classical-quantum supercomputing: A demonstration of a multi-user, multi-qpu and multi-gpu environment,” *arXiv preprint arXiv:2508.16297*, 2025.
  - 52 P. Clifford and R. Clifford, “The classical complexity of boson sampling,” in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 146–155, SIAM, 2018.
  - 53 G. Facelli, D. D. Roberts, H. Wallner, A. Makarovskiy, Z. Holmes, and W. R. Clements, “Exact gradients for linear optics with single photons,” *arXiv preprint arXiv:2409.16369*, 2024.
  - 54 A. Brock, J. Donahue, and K. Simonyan, “Large scale gan training for high fidelity natural image synthesis,” *arXiv preprint arXiv:1809.11096*, 2018.
  - 55 T. Hu, F. Chen, H. Wang, J. Li, W. Wang, J. Sun, and Z. Li, “Complexity matters: Rethinking the latent space for generative modeling,” *Advances in Neural Information Processing Systems*, vol. 36, pp. 29558–29579, 2023.
  - 56 B. Villalonga, M. Y. Niu, L. Li, H. Neven, J. C. Platt, V. N. Smelyanskiy, and S. Boixo, “Efficient approximation of experimental Gaussian boson sampling,” *arXiv preprint arXiv:2109.11525*, 2021.
  - 57 C. Oh, M. Liu, Y. Alexeev, B. Fefferman, and L. Jiang, “Classical algorithm for simulating experimental gaussian boson sampling,” *Nature Physics*, vol. 20, no. 9, pp. 1461–1468, 2024.

## A Data pre-processing

We pre-processed the data in the following way. First, though each time series in the original dataset was 200 hours long, our initial investigation found that using the entire dataset caused the anomalies to be too easy to detect, with near-perfect anomaly detection performance that provides little insight into the effectiveness of our method. We thus focused on the early anomaly detection setting and limited the dataset to the first 48 hours (with the anomaly starting at the 10th hour). Since changes to the measured quantities are slow, we further downselected to only use data collected every 8th hour. We tracked 5 features in the time series that were found to be affected by the anomaly, such that each time series was reduced down to  $(48/8) \times 5 = 30$  features after pre-processing. To help process this data with a neural network, we then standardized these features using the mean and standard deviation over time of each of the 5 quantities.