TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Dong, L., Becha, M., Köhler, J., Zagorowska, M., & Mercangöz, M. (2027). Safety-aware load sharing in compressor stations using constrained Bayesian optimization. *Computers and Chemical Engineering*, *216*, Article 109865. https://doi.org/10.1016/j.compchemeng.2026.109865

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Computers and Chemical Engineering 216 (2027) 109865

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/cace

Safety-aware load sharing in compressor stations using constrained Bayesian optimization

Liqiu Dong [a,*,1], Mohamed Becha [b,1], Johannes Köhler [c], Marta Zagorowska [d], Mehmet Mercangöz [a]

[a] Department of Chemical Engineering, Imperial College London, SW7 2AZ, UK
[b] Department of Mechanical and Process Engineering, ETH Zürich, 8092 Zürich, Switzerland
[c] Department of Mechanical Engineering, Imperial College London, SW7 2AZ, UK
[d] Delft Center for Systems and Control, TU Delft, Delft, The Netherlands

ARTICLE INFO

Dataset link: https://github.com/AISL-at-Imperial-College-London/CBO_comp

Keywords:
Compressor stations
Load-sharing optimization
Real-time optimization
Gaussian process
Safety-aware optimization
Energy efficiency

ABSTRACT

We present a unified, self-learning framework for real-time load-sharing optimization of compressor stations in parallel and series configurations. The proposed optimizer operating at the supervisory layer models the station as a data-driven steady-state mapping from shaft torque inputs to measured performance variables, thereby avoiding explicit use of compressor maps within the optimization layer. Constrained Bayesian optimization is used to learn Gaussian Process surrogate models for the economic objective and surge-related operability constraints directly from plant-response data. Safety is incorporated through measured distance-to-surge information, allowing the optimizer to favor energy-efficient operating points while accounting for uncertainty in the learned surge-margin predictions. The framework is initialized using only three operating points and is evaluated in a high-fidelity dynamic simulation environment for both parallel and series compressor stations. Across the reported simulation scenarios and initialization cases, the method consistently converged to improved load-sharing solutions while maintaining positive steady-state surge margins and without observed surge control line crossings or anti-surge controller activation. Relative to an equal-load baseline, the proposed framework achieved energy savings of up to 2.6% in the parallel configuration and 3.1% in the series configuration. These results demonstrate the potential of the proposed CBO-based load-sharing framework to improve compressor-station efficiency while respecting operability constraints in a data-driven optimization setting.

1. Introduction

Gas compressors are widely used in industry, including gas storage (Silva and Camponogara, 2014; Kurz and Brun, 2010; Liu et al., 2019), air separation units (Widell and Eikevik, 2010), long-distance pipeline transport (Borraz-Sánchez and Haugland, 2013), liquefied natural gas (LNG) processing (Mokhatab et al., 2013; Shaw, 1994) and industrial pressurized air networks (Lassahn et al., 2026). They are also expected to remain central to emerging decarbonization applications, including CO_2 and H_2 transport for carbon capture and storage and hydrogen infrastructure (Kurz et al., 2023). Because compressor stations operate continuously and at large scale, they represent a major source of industrial energy consumption. For instance, compressor stations within the UK natural gas network alone consume approximately 1.65 TWh annually and are the single largest energy user within the National Transmission System (National Gas Transmission, 2025). In large-scale pipeline transport, it has been estimated that nearly 5% of the total transported gas is consumed by the turbine drivers of the compression process (de Marco and Elias, 2011). Consequently, even modest improvements in operating efficiency can have significant economic and environmental benefits (Kapp et al., 2024; Variny et al., 2021; Viholainen et al., 2015). To put this scale into perspective, an efficiency improvement of only 1%–3% applied to an annual compressor-station energy demand of 1.65 TWh would correspond to approximately 16.5–49.5 GWh of savings per year. Depending on the driver technology, such reductions translate directly into lower electricity procurement costs or lower fuel consumption, together with a corresponding reduction in operating emissions.

* Corresponding author.
E-mail address: liqiu.dong22@imperial.ac.uk (L. Dong).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.compchemeng.2026.109865
Received 18 April 2026; Received in revised form 20 August 2026; Accepted 21 August 2026
Available online 29 August 2026
0098-1354/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

In practice, compressors in compressor stations can be arranged in parallel or series configurations to meet the demand for delivered mass flows and pressures which change throughout the day and seasonally over a year (Kurz et al., 2012). A combination of parallel and series configurations is also possible (Zhao et al., 2021). Parallel configurations are used to meet high mass-flow requirements at moderate pressure ratios, whereas series configurations are used to achieve high discharge pressures beyond the capability of a single unit (Boyce, 2003). In many industrial settings, a dispatch or supervisory layer provides an overall mass-flow or pressure target to the station, and the individual compressors are then operated using equal or near-equal load distribution (Li et al., 2021). Although such strategies are simple and operationally robust, they do not account for differences in individual compressor efficiencies and therefore can lead to excess energy use (Zagorowska et al., 2020). In order to improve energy efficiency, Load-Sharing Optimization (LSO) methods have been developed to distribute the required load across multiple compressors while satisfying process and equipment constraints (Milosavljevic et al., 2020; Xenos et al., 2015; Kumar and Cortinovis, 2017). Classical formulations typically manipulate compressor rotational speeds and rely on compressor maps and thermodynamic relations to determine suitable operating points. A major practical limitation of this approach is its dependence on accurate plant models. Compressor characteristics change over time because of fouling, erosion, maintenance history, and manufacturing variability (Kurz and Brun, 2012). As a result, an optimizer based on, initially accurate, compressor maps may gradually become suboptimal and, in the worst case, may recommend operating points that approach unsafe regions of the compressor operating envelope.

The most important safety concern is compressor surge, an aerodynamic instability that can occur when a compressor operates at low flow and high head or pressure ratio, leading to flow reversal, strong vibration, and possible equipment damage (Du et al., 2025; Dixon and Hall, 2013). In compressor stations, this makes the LSO problem inherently safety-critical. If the compressor map or surge boundary used by the optimization layer does not reflect the current physical state of the machines, plant-model mismatch can degrade both economic performance and operational safety (Xenos et al., 2015; Tirnovan et al., 2008; Cortinovis et al., 2016). Previous research has attempted to address this mismatch in several ways. Early work relied on gray-box Real-Time Optimization (RTO), a supervisory optimization paradigm in which a process model is periodically updated using plant measurements and then re-optimized to compute improved steady-state operating targets. While RTO can correct some mismatch, it remains vulnerable to structural model errors, since updating a limited set of parameters cannot compensate for inadequacies in the model equations themselves. To mitigate this problem, two main research directions emerged: Modifier Adaptation (MA) and Gaussian Process (GP) augmentation of physical models (Milosavljevic et al., 2020; Zagorowska et al., 2023). MA compensates for mismatch by correcting the optimization problem directly from process data, but it is typically data-intensive, requires repeated steady-state evaluations, and can be sensitive to noise and disturbances (del Rio Chanona et al., 2021). GP-augmented approaches improve local model accuracy by learning residual mismatch, including within classical RTO structures (Ahmed et al., 2022), but this type of learning is only driven by the mismatch in calculated optimum operating points with updated model information and lacks guided exploration.

Research gaps. These developments point to two main limitations of existing approaches. First, the existing approaches do not explicitly promote active exploration of the operating region. As a result, model accuracy remains local, which can limit performance when the station must operate efficiently over a broad range of loads. Second, surge constraints are often treated in a certainty-equivalent fashion, as if the aerodynamic stability boundary were perfectly known. In practice, however, the surge control line is typically available only through operational safety margins and measured distance-to-surge (DTS) information. A more suitable optimization framework should therefore both learn from data across the operating region and account explicitly for uncertainty in safety-related constraints as a function of the optimization degrees of freedom. Bayesian optimization (BO) is well suited to this setting because it is designed for optimization of expensive-to-evaluate black-box functions (Wang et al., 2023; Kudva et al., 2022; Krishnamoorthy and Doyle III, 2023), which extends to evaluations or experiments using physical assets such as compressors. BO uses a probabilistic surrogate model, commonly a GP, to represent the objective function together with predictive uncertainty (Rasmussen and Williams, 2005). This allows a systematic trade-off between exploration of insufficiently sampled regions and exploitation of promising operating points. Acquisition functions such as upper confidence bound (UCB) and expected improvement (EI) are commonly used to guide this search (Shahriari et al., 2016; Ungredda and Branke, 2024). However, direct application of standard BO to compressor stations may be unsafe. An exploratory move into a low-flow and high-pressure ratio region may bring the compressor close to surge before the boundary is adequately learned. This motivates the use of constrained Bayesian optimization (CBO), in which constraint functions are also learned and incorporated into the acquisition step (Amini et al., 2025). CBO has been applied successfully in a range of safety-critical problems, including autonomous motion control tuning in robotics (Berkenkamp et al., 2023), yield optimization in chemical reactors (Desimpel et al., 2026), reactor design (Kudva et al., 2022), and structural optimization in aerospace engineering (Maathuis et al., 2025). Nevertheless, its application to compressor stations with coupled units and surge-related constraints remains an open problem.

1.1. Contributions of this work

In this work, we propose a unified self-learning framework for real-time LSO of compressor stations based on CBO. Unified here refers to the treatment of parallel and series compressor arrangements under the same problem formulation. The framework operates at the supervisory layer and models the station as a data-driven mapping from shaft torque inputs to measured performance variables. By formulating the optimization problem directly in terms of torque rather than rotational speed, the method avoids explicit use of compressor map models in the optimization layer, requiring only measured plant responses and DTS information to guide safe and energy-efficient operation. The framework is implemented for both parallel and series compressor configurations and is evaluated in dynamic simulation studies using previously validated compressor models (Cortinovis et al., 2015). The main contributions of this paper are as follows:

(i) We develop a self-learning framework for LSO of compressor stations, interchangeably applicable to both parallel and series compressor configurations using CBO.
(ii) We formulate the compressor station load-sharing problem directly in terms of shaft torque and solve it using a load-sharing optimizer operating at the supervisory layer, thereby avoiding explicit use of compressor maps within the optimization layer while retaining measured DTS information as a safety signal.
(iii) We demonstrate high sample efficiency in the reported case studies, where the learning process is initialized using only three initial operating points.
(iv) We validate the framework through dynamic simulation studies using two acquisition strategies, namely constrained expected improvement (CEI) and upper confidence bound with a log-barrier term (UCB-LB). In the reported simulations, both strategies reduce energy consumption relative to equal-load operation, with savings of up to 2.60% in the parallel configuration and 3.06% in the series configuration. Across all simulations considered, safe operation is maintained with no instance of crossing of the surge control line or anti-surge controller activation.

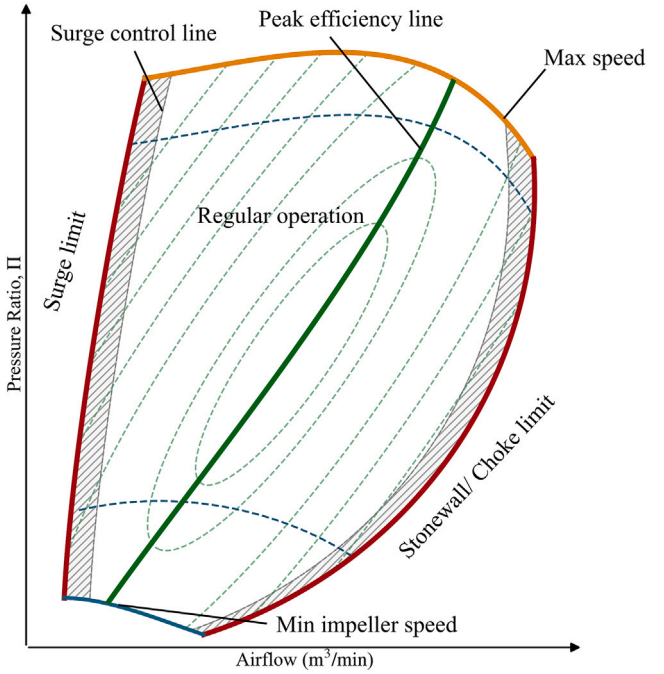


Fig. 1. A compressor map example illustrating operation between choke/stonewall and the surge control line.
Source: Reconstructed from Chaudhari et al. (2023).

The remainder of the paper is organized as follows. Section 2 provides background and preliminaries, including the load-sharing formulations for parallel and series compressor configurations and the CBO tools used in this work. Section 3 then presents the proposed CBO-based load-sharing optimization framework for compressor stations, including the unified torque-based optimization formulation, surrogate modeling strategy, and implementation details. Section 4 reports the simulation results and discusses the main findings. Finally, Section 5 concludes the paper.

2. Background and preliminaries

This section introduces the technical background required for the proposed framework. First, the load-sharing problems for parallel and series compressor configurations are formulated and the corresponding control architectures are described. Next, the main elements of constrained Bayesian optimization are summarized, including Gaussian Process surrogate modeling, acquisition functions, and constrained search mechanisms. Together, these preliminaries establish the problem setting and methodological foundation used in the subsequent development of the proposed load-sharing optimization framework.

2.1. System description and problem formulation

In this section, we define the compressor station load-sharing optimization problems considered in this work and describe the control architectures used in the case studies. The case studies use the dynamic compressor model summarized in Appendix A, including compressor map and surge characterizations. This model is used only to emulate real compressors and is not supplied to the load-sharing optimizer, which instead learns data-driven surrogates of the station's steady-state behavior from simulated input–output data. Fig. 1 illustrates a representative compressor map, including the surge control line and the stable operating region. The map is shown only to describe the simulated plant; it is not supplied to the load-sharing optimizer.

The implemented control architecture comprises a single load-sharing optimizer operating at the supervisory layer and a local anti-surge protection layer. The load-sharing optimizer distributes the load among the compressors to minimize total power consumption while satisfying the station demand and maintaining operation away from surge. The local anti-surge protection layer comprises one anti-surge controller (ASC) for each compressor. Each ASC monitors the corresponding compressor's surge margin and actuates its recycle valve if the operating point enters the anti-surge protection region. The individual ASCs are therefore distinct from the load-sharing optimizer. In the formulations below, the compressor steady-state relations and the surge constraint function are inherited from the simulation model summarized in Appendix A.

Although the parallel and series configurations have different physical interconnections and controlled station variables, their steady-state load-sharing problems share the same mathematical structure. Let $\mathbf{z} \in \mathcal{Z}$ denote the vector of compressor operating targets. Both configurations can be represented by:

$$\begin{aligned} \min_{\mathbf{z}\in\mathcal{Z}} \quad & \sum_{i=1}^{n_c} P_i(\mathbf{z}) \\ \text{s.t.} \quad & h(\mathbf{z}) = h_{\text{ref}}, \\ & S_i(\mathbf{z}) \geq 0, \qquad i = 1, \dots, n_c. \end{aligned} \tag{1}$$

Here, n_c denotes the number of compressors in the station. The term heterogeneous refers to differences in compressor performance characteristics, represented by compressor map and efficiency parameters. These differences remain fixed throughout the simulations and produce different power consumption profiles, which form the basis for load-sharing optimization. $P_i(\mathbf{z})$ denotes the shaft power of compressor i, $h(\mathbf{z})$ is the configuration-dependent station demand variable, h_{ref} is its prescribed reference, and $S_i(\mathbf{z})$ is the surge-margin indicator. The separate descriptions in Section 2.1.1 define $h(\mathbf{z})$ and describe the physical interconnections and control-loop implementation for each configuration.

2.1.1. Parallel configuration

Consider a station with n_c heterogeneous compressors operating in parallel, indexed by $i = 1, \dots, n_c$. In this configuration, all compressors share common suction and discharge headers, so that they operate under the same boundary pressures while collectively supplying the required station throughput. The load-sharing optimizer distributes the load among the individual compressors so as to minimize total power consumption while satisfying the required total outlet mass flowrate and maintaining safe operation of every unit.

For the parallel configuration, the demand-related station variable is the total outlet mass flowrate:

$$h(\mathbf{z}) = M_{\text{out,tot}}(\mathbf{z}), \qquad h_{\text{ref}} = M_{\text{ref}}.$$

Here, $M_{\text{out,tot}}(\mathbf{z})$ is the total outlet mass flowrate delivered by the station and M_{ref} is the required total flowrate.

Underlying control loops for the parallel configuration. The parallel operation case study considers a station comprising three centrifugal compressors (C_1, C_2, and C_3), as illustrated in Fig. 2.

For the parallel station, the primary station demand variable is the total outlet mass flowrate, $M_{\text{out,tot}}$. The load-sharing optimizer receives this measurement from the common outlet header together with the compressor power measurements (P_1, P_2, P_3) and the DTS values computed for each unit. Based on this information, it computes the desired operating targets and transmits corresponding torque references (τ_1, τ_2, τ_3) to the variable-speed drives.

The anti-surge protection layer operates at a faster timescale and uses high-frequency pressure and flow measurements to estimate the instantaneous DTS for each compressor. Specifically, pressure transmitters at the suction and discharge sides together with flow measurements

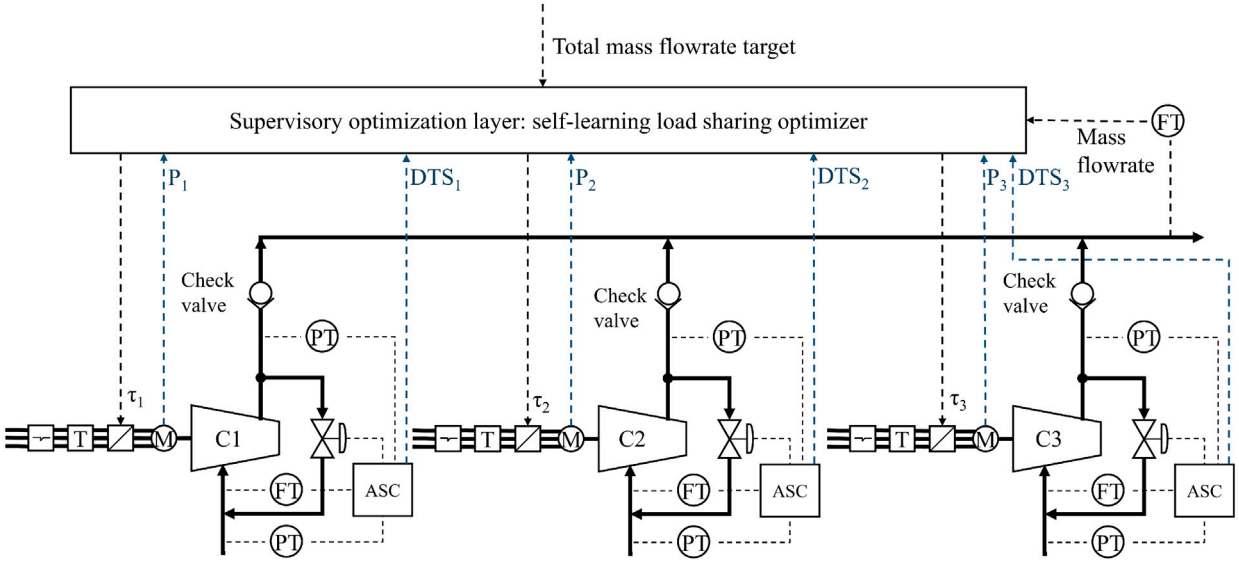


Fig. 2. Compressor station in parallel configuration with the self-learning load-sharing optimizer. Solid lines represent gas-flow connections, while blue dashed paths labeled P_i and DTS_i represent the measured power and distance-to-surge signals, respectively, transmitted from compressor i to the load-sharing optimizer. The upper block represents the single load-sharing optimizer operating at the supervisory layer of the control hierarchy. The individual ASC blocks constitute the separate local anti-surge protection layer. Here, M denotes the electric motor, T the transformer, ASC the anti-surge controller, PT a pressure transmitter, FT a flow transmitter, and τ_i the torque reference for compressor i.

are used to calculate the surge margin, which is then supplied both to the load-sharing optimizer and to the corresponding local anti-surge controller. This layered architecture enables the load-sharing optimizer to optimize load allocation without replacing the local protection mechanisms required for safe compressor operation.

2.1.2. Series configuration

Consider next a station with n_c heterogeneous compressors arranged in series, indexed by $i = 1, \dots, n_c$. In this configuration, the discharge of compressor i feeds the suction of compressor $i+1$, so that the compressors collectively deliver the required overall pressure rise. For this configuration, the load-sharing optimizer distributes the compression load across the stages so as to minimize total power consumption while achieving the required final discharge pressure and maintaining safe operation of every unit.

For the series configuration, the demand-related station variable is the final discharge pressure:

$$h(\mathbf{z}) = P_{\mathrm{dis}}(\mathbf{z}), \qquad h_{\mathrm{ref}} = P_{\mathrm{ref}}.$$

Thus, the load-sharing optimizer distributes the compression duty across the stages while tracking the prescribed final discharge pressure.

Underlying control loops for the series configuration. The series operation case study considers two centrifugal compressors (C_1 and C_2) connected in series, as shown in Fig. 3. The same architecture, comprising a load-sharing optimizer operating at the supervisory layer and a separate local anti-surge protection layer, is used for the series configuration.

For the series configuration, the station demand variable is the final discharge pressure, P_{dis}, rather than the total outlet mass flowrate. The load-sharing optimizer receives the discharge pressure measurement together with compressor power measurements and DTS indicators, and computes the operating targets for the individual compressor drives. These targets are applied as torque references to the compressor drives.

Each compressor is protected by its own local anti-surge controller, which evaluates the instantaneous operating margin based on local pressure and flow measurements. In addition, the series configuration includes an intercooler between the two compressors. This unit reduces the temperature of the gas leaving the first stage before it enters the second stage, thereby affecting the operating condition of the downstream compressor and the overall distribution of compression duty across the station. It is modeled by setting a fixed outlet temperature of 293 K, corresponding to the inlet temperature of the downstream compressor reported in Table 2. The coolant side and its dynamics are not modeled explicitly. Consequently, the intercooler cooling duty and auxiliary power consumption are not included in the optimization objective.

2.2. Constrained Bayesian optimization (CBO)

2.2.1. Gaussian process (GP)

A GP is a non-parametric, probabilistic model that defines a distribution over functions (Rasmussen and Williams, 2005). A GP is fully specified by a mean function $\mu(\mathbf{u})$ and a covariance (or kernel) function $k(\mathbf{u}, \mathbf{u}')$:

$$f(\mathbf{u}) \sim \mathcal{GP}(\mu(\mathbf{u}), k(\mathbf{u}, \mathbf{u}')) \tag{2}$$

The mean function $\mu(\mathbf{u}) = \mathbb{E}[f(\mathbf{u})]$ represents the expectation of the function's value at $\mathbf{u}$. The kernel function $k(\mathbf{u}, \mathbf{u}') = \mathbb{E}[(f(\mathbf{u}) - \mu(\mathbf{u}))(f(\mathbf{u}') - \mu(\mathbf{u}'))]$ models the covariance between the function values at any two points $\mathbf{u}$ and $\mathbf{u}'$.

In general, the latent function $f(\mathbf{u})$ is not directly accessible; instead, one obtains noisy measurements $y = f(\mathbf{u}) + \epsilon$. In GP regression, noise is commonly modeled as Gaussian, i.e. $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$, although GPs can also provide suitable confidence bounds under more general noise conditions.

Given a set of observations $\mathcal{D}_n = \{(\mathbf{u}_i, y_i)\}_{i=1}^n$, where $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_n]^\top$ and $\mathbf{y} = [y_1, \dots, y_n]^\top$, we can compute the posterior predictive distribution f_* at an arbitrary test point $\mathbf{u}_*$:

$$p(f_* | \mathbf{u}_*, \mathcal{D}_n) \sim \mathcal{N}(\mu_n(\mathbf{u}_*), \sigma_n^2(\mathbf{u}_*)) \tag{3}$$

Assuming a zero-mean prior, the posterior mean $\mu_n(\mathbf{u}_*)$ and posterior variance $\sigma_n^2(\mathbf{u}_*)$ are computed:

$$\mu_n(\mathbf{u}_*) = \mathbf{k}_*^\top (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{y} \tag{4}$$

$$\sigma_n^2(\mathbf{u}_*) = k(\mathbf{u}_*, \mathbf{u}_*) - \mathbf{k}_*^\top (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{k}_* \tag{5}$$

where $\mathbf{K} = k(\mathbf{U}, \mathbf{U})$ is the $n \times n$ covariance matrix of all training inputs, and $\mathbf{k}_* = k(\mathbf{U}, \mathbf{u}_*)$.

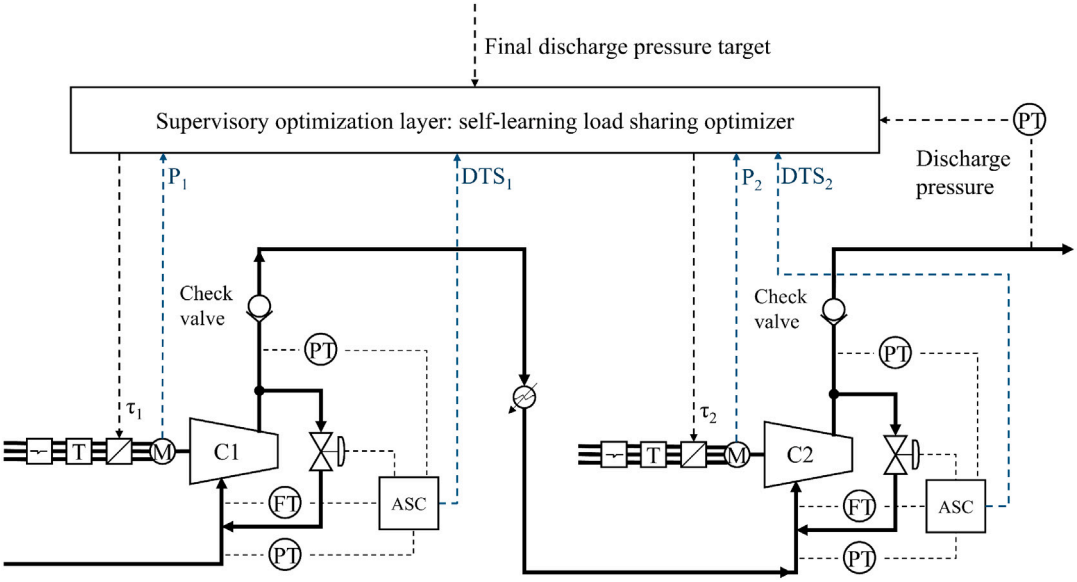


Fig. 3. Compressor station in series configuration with the self-learning load-sharing optimizer. Solid lines represent gas-flow connections, while blue dashed paths labeled P_i and DTS_i represent the measured power and distance-to-surge signals, respectively, transmitted from compressor i to the load-sharing optimizer. The upper block represents the single load-sharing optimizer operating at the supervisory layer of the control hierarchy. The individual ASC blocks constitute the separate local anti-surge protection layer. Here, M denotes the electric motor, T the transformer, ASC the anti-surge controller, PT a pressure transmitter, FT a flow transmitter, and τ_i the torque reference for compressor i.

2.2.2. Bayesian optimization (BO)

The focus of this work is to solve complex optimization problems where the objective function, $f(\mathbf{u})$, is a "black box". This scenario is common in engineering, where an evaluation of $f(\mathbf{u})$ may correspond to a costly physical experiment or an expensive computer simulation. In this context, the analytical form of $f(\mathbf{u})$ is unknown. The optimization problem is:

$$\mathbf{u}^* = \arg\min_{\mathbf{u}\in\mathcal{U}} f(\mathbf{u}) \tag{6}$$

where $\mathcal{U} \subset \mathbb{R}^d$ is the compact search space of the d-dimensional decision variables, $\mathbf{u}$.

The high evaluation cost restricts our ability to explore the search space $\mathcal{U}$. To address this, our goal is to find the best solution $\mathbf{u}^*$ with a small number of function evaluations using BO.

BO consists of two coupled components: fitting a probabilistic surrogate model to the observed data and optimizing an acquisition function $\alpha : \mathcal{U} \to \mathbb{R}$ using that surrogate model. The acquisition function uses the surrogate model's predictive mean and uncertainty to balance exploitation and exploration when selecting the next evaluation point.

The BO algorithm iteratively builds a dataset of observations $\mathcal{D}_n = \{(\mathbf{u}_i, y_i)\}_{i=1}^n$. At each new iteration $n+1$, it updates the surrogate model based on $\mathcal{D}_n$. In this framework, GPs are employed as the surrogate model. It then optimizes the (computationally cheap) acquisition function to select the next query point, $\mathbf{u}_{n+1} = \arg\max_{\mathbf{u}\in\mathcal{U}} \alpha(\mathbf{u}|\mathcal{D}_n)$. This new point is then evaluated by setting the physical system to $\mathbf{u}_{n+1}$ and measuring the corresponding output $y_{n+1} = f(\mathbf{u}_{n+1}) + \epsilon$, and the resulting pair $(\mathbf{u}_{n+1}, y_{n+1})$ is added to the dataset, $\mathcal{D}_{n+1} = \mathcal{D}_n \cup \{(\mathbf{u}_{n+1}, y_{n+1})\}$. This process repeats for a fixed budget of evaluations. In a real-time optimization setting, the process repeats continuously without a budget limit and the true expensive function evaluation corresponds to moving the physical system to the query point and measuring the resulting system outputs.

2.2.3. Acquisition function

The acquisition function serves as the decision mechanism of the BO framework, determining the next sampling point by quantifying the utility of exploring the search space $\mathcal{U}$ (Krause and Hübotter, 2025). It leverages the GP posterior to balance the exploitation of known low-power regions with the exploration of high-uncertainty zones. We focus on two of the most widely used and established acquisition strategies in Bayesian optimization, namely upper confidence bound (UCB) and expected improvement (EI), because they are particularly well suited for use in an RTO setting. In this setting, each new operating point must balance the search for more efficient load distributions against the need to limit unnecessary exploration of poorly characterized regions of the operating envelope. UCB addresses this trade-off explicitly through uncertainty-aware exploration, whereas EI prioritizes operating points that are expected to improve upon the best observed performance. Together, they provide two complementary and practically relevant decision mechanism candidates to test for data-efficient optimization of compressor operation.

Upper confidence bound (UCB). The UCB family of acquisition functions is based on the principle of optimism in the face of uncertainty (Frazier, 2018). It balances exploration and exploitation by combining the GP posterior mean with a multiple of the posterior standard deviation. For minimization problems, this principle is implemented through the lower confidence bound (LCB), defined as:

$$\alpha_{\mathrm{LCB}}(\mathbf{u}) = \mu_n(\mathbf{u}) - \kappa_n \sigma_n(\mathbf{u}), \tag{7}$$

and the next evaluation point is selected as:

$$\mathbf{u}_{n+1} = \arg\min_{\mathbf{u}\in\mathcal{U}} \alpha_{\mathrm{LCB}}(\mathbf{u}). \tag{8}$$

Here, the parameter $\kappa_n > 0$ governs the exploration–exploitation trade-off: larger values place greater emphasis on uncertain regions, whereas smaller values favor exploitation of regions predicted to have low objective values.

Expected improvement (EI). Expected improvement (EI) provides a more direct measure of the potential benefit of sampling a candidate point by quantifying the expected improvement over the best objective value observed so far (Jones et al., 1998), denoted by:

$$y_n^* = \min\{y_1, \ldots, y_n\}. \tag{9}$$

For a candidate input $\mathbf{u}$, the improvement random variable is defined as:

$$I(\mathbf{u}) = \max\big(0,\, y_n^* - f(\mathbf{u})\big). \tag{10}$$

Since $f(\mathbf{u}) \mid \mathcal{D}_n$ is random under the GP posterior, the acquisition function is defined as the posterior expectation of this improvement:

$$\alpha_{\mathrm{EI}}(\mathbf{u}) = \mathbb{E}\left[I(\mathbf{u}) \mid \mathcal{D}_n\right]. \tag{11}$$

Considering the Gaussian GP posterior at $\mathbf{u}$, i.e.:

$$f(\mathbf{u}) \mid \mathcal{D}_n \sim \mathcal{N}\left(\mu_n(\mathbf{u}), \sigma_n^2(\mathbf{u})\right), \tag{12}$$

then EI admits the closed-form expression:

$$\alpha_{\mathrm{EI}}(\mathbf{u}) = \left(y_n^* - \mu_n(\mathbf{u})\right)\Phi\left(\frac{y_n^* - \mu_n(\mathbf{u})}{\sigma_n(\mathbf{u})}\right) + \sigma_n(\mathbf{u})\phi\left(\frac{y_n^* - \mu_n(\mathbf{u})}{\sigma_n(\mathbf{u})}\right), \tag{13}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the cumulative distribution function (CDF) and probability density function (PDF) of the standard normal distribution, respectively. The next evaluation point is selected as:

$$\mathbf{u}_{n+1} = \arg\max_{\mathbf{u}\in\mathcal{U}} \alpha_{\mathrm{EI}}(\mathbf{u}). \tag{14}$$

2.2.4. Bayesian optimization with constraints

Industrial applications rarely permit unconstrained exploration. Consequently, the corresponding constrained optimization problem is formulated as:

$$\begin{aligned} &\min_{\mathbf{u}\in\mathcal{U}} \quad f(\mathbf{u}) \\ &\text{s.t.} \quad c_q(\mathbf{u}) \geq m_{\mathrm{safe}}, \qquad q = 1, \ldots, Q, \end{aligned} \tag{15}$$

where $f(\mathbf{u})$ is the objective function and $c_q(\mathbf{u})$, $q = 1, \ldots, Q$ are the constraint functions, and m_{safe} denotes the prescribed minimum admissible constraint value. They are treated as unknown black-box mappings that can only be evaluated through plant operation or simulation.

Since the true objective $f(\mathbf{u})$ and constraint functions $c_q(\mathbf{u})$ are unknown, the optimizer uses their GP surrogate models. In the surrogate-based architecture adopted here, exact pointwise satisfaction of the true unknown constraints during acquisition optimization is not guaranteed because feasibility is available only through the uncertain GP posterior. Alternative feedback-based RTO architectures for constraint handling are discussed in Section 4.5. The constrained problem is handled through soft surrogate-based reformulations of the acquisition function, in which predicted constraint violation is discouraged using probability-of-feasibility factors, barrier terms, or penalty terms derived from the posterior mean and variance of the learned constraint models. In this way, the optimizer can balance improvement of the objective against uncertainty-aware satisfaction of the learned constraints while still permitting informative exploration near the estimated feasibility boundary.

2.2.5. Constrained expected improvement (CEI)

Constrained expected improvement modifies the standard expected improvement criterion by weighting it with the posterior probability that a candidate point is feasible (Gardner et al., 2014). By modeling the objective and each constraint with separate, independent GP models, the CEI acquisition function is defined as:

$$\alpha_{\mathrm{CEI}}(\mathbf{u}) = \alpha_{\mathrm{EI}}(\mathbf{u}) \prod_{q=1}^{Q} \mathbb{P}\left(c_q(\mathbf{u}) \geq m_{\mathrm{safe}} \mid \mathcal{D}_n\right). \tag{16}$$

Under the GP posterior, the probability of feasibility (PoF) for constraint q is:

$$\mathbb{P}\left(c_q(\mathbf{u}) \geq m_{\mathrm{safe}} \mid \mathcal{D}_n\right) = \Phi\left(\frac{\mu_{q,n}(\mathbf{u}) - m_{\mathrm{safe}}}{\sigma_{q,n}(\mathbf{u})}\right), \tag{17}$$

where $\Phi(\cdot)$ denotes the CDF of the standard normal distribution. Hence,

$$\alpha_{\mathrm{CEI}}(\mathbf{u}) = \alpha_{\mathrm{EI}}(\mathbf{u}) \prod_{q=1}^{Q} \Phi\left(\frac{\mu_{q,n}(\mathbf{u}) - m_{\mathrm{safe}}}{\sigma_{q,n}(\mathbf{u})}\right). \tag{18}$$

This construction favors candidate points that are both likely to improve the objective and likely to satisfy the safety constraints.

2.2.6. Upper confidence bound with log-barrier function (UCB-LB)

As an alternative, motivated by confidence-bound-based approaches to safe GP optimization (Sui et al., 2015), we construct a conservative acquisition function by combining a lower confidence bound for the objective with a log-barrier term based on conservative lower bounds of the constraint surrogates. For each constraint q, we define:

$$\underline{c}_{q,n}(\mathbf{u}) = \mu_{q,n}(\mathbf{u}) - \beta_q \sigma_{q,n}(\mathbf{u}), \tag{19}$$

where $\beta_q > 0$ controls the degree of conservativeness. The resulting acquisition function is:

$$\alpha_{\mathrm{UCB-LB}}(\mathbf{u}) = -\left[\mu_{f,n}(\mathbf{u}) - \kappa_n \sigma_{f,n}(\mathbf{u}) - b\sum_{q=1}^{Q} \log\left(\underline{c}_{q,n}(\mathbf{u}) - m_{\mathrm{safe}}\right)\right], \tag{20}$$

defined over the region where $\underline{c}_{q,n}(\mathbf{u}) > m_{\mathrm{safe}}$ for all q. Here, $b > 0$ is the barrier parameter. As $\underline{c}_{q,n}(\mathbf{u}) \to m_{\mathrm{safe}}^{+}$, the barrier term tends to $+\infty$ in the minimization form, thereby discouraging candidate points that approach the estimated constraint boundary. In this way, the optimizer is guided by conservative surrogate-based constraint information without requiring explicit knowledge of the true constraint.

2.2.7. Optimizing the acquisition function

The acquisition function optimization problem arising in the proposed framework is generally non-convex and may contain multiple local optima, particularly because the GP-based acquisition functions combine nonlinear mean, uncertainty, and constraint-related terms. For this reason, we employ a genetic algorithm (GA) to perform the inner acquisition optimization (Mitchell, 1999).

This choice is motivated by the need for a robust search procedure for acquisition landscapes that may be multimodal and nonsmooth in practice. In the relatively low-dimensional compressor load-sharing problems considered here, the GA provided a reliable and straightforward way to optimize the acquisition function without requiring gradient information.

3. Constrained Bayesian optimization of compressor stations

This section presents the proposed load-sharing optimization framework. As defined in Section 2.1, the single load-sharing optimizer operates at the supervisory layer of the control hierarchy. Rather than relying on explicit compressor maps or detailed first-principles relations by the load-sharing optimizer, the framework treats the compressor station as a data-driven input–output system and uses constrained Bayesian optimization to identify improved operating conditions from measured plant responses.

3.1. Unified LSO formulation for parallel and series configurations

The formulations in Section 2.1 define the load-sharing objectives for parallel and series compressor arrangements depending on either the mass flows or the final discharge pressure, respectively. In both cases, the supervisory task is to minimize the total compressor power required to satisfy the prescribed station demand while maintaining a positive surge margin for every unit.

In the implementation proposed here, this load-sharing problem is realized through a black-box mapping from shaft torque inputs to measured plant variables. Specifically, instead of relying on explicit speed-based thermodynamic models inside the optimizer, we represent the station through measured steady-state responses as functions of the applied torque vector $\mathbf{T}$. This choice is motivated by the fact that shaft torque is a directly actionable variable in the considered electrically driven compressor systems and provides a convenient interface between the load-sharing optimizer and the local drive and anti-surge protection components, and will be discussed in detail in Section 4.5.

The resulting formulation is independent of whether the compressors are arranged in parallel or in series. In both cases, the optimizer

operates only on measurable plant quantities, without requiring the load-sharing optimizer to use explicit compressor maps or internal model equations. The framework is evaluated using two high-fidelity simulation environments as described in Appendix A (Cortinovis et al., 2015).

3.1.1. Optimization objective

We consider a compressor station composed of n_c compressors arranged either in parallel or in series. The load-sharing optimizer operating at the supervisory layer treats the station as a steady-state nonlinear mapping from the applied shaft-torque vector to measurable plant variables,

$$\mathbf{y} = \mathbf{f}(\mathbf{T}), \qquad \mathbf{T} \in \mathcal{T} \subset \mathbb{R}^{n_c}, \tag{21}$$

where $\mathbf{T} = [T_1, T_2, \dots, T_{n_c}]^\top$ denotes the vector of compressor shaft torques and $\mathcal{T}$ denotes the admissible torque set imposed by actuator limits.

The corresponding steady-state output vector is written as:

$$\mathbf{f}(\mathbf{T}) = \begin{bmatrix} P(\mathbf{T}) \\ h(\mathbf{T}) \\ S_1(\mathbf{T}) \\ \vdots \\ S_{n_c}(\mathbf{T}) \end{bmatrix}, \tag{22}$$

where $P(\mathbf{T}) = \sum_{i=1}^{n_c} P_i(\mathbf{T})$ denotes the total power consumption of the compressor station, $h(\mathbf{T})$ denotes the demand-related station variable, namely the final discharge pressure in the series configuration and the total outlet mass flowrate in the parallel configuration, and $S_i(\mathbf{T})$ denotes the DTS of compressor i. Consistent with Appendix A.3, the surge margin is defined as:

$$S_i(\mathbf{T}) = m_{\text{comp},i}(\mathbf{T}) - m_{\text{surge},i}\big(\Pi_i(\mathbf{T})\big), \tag{23}$$

with the surge control line represented by the following form as:

$$m_{\text{surge},i}(\Pi_i) = s_{0,i} + s_{1,i}\Pi_i, \tag{24}$$

where Π_i is the pressure ratio of compressor i, and $s_{0,i}$ and $s_{1,i}$ are the fixed intercept and slope, respectively, of the reduced surge control line representation. In the simulation model, the coefficients are $s_{0,i} = -0.347$ and $s_{1,i} = 0.463$. The same coefficients are used for all compressors and remain fixed throughout the simulations.

The load-sharing problem is therefore posed as:

$$\min_{\mathbf{T}\in\mathcal{T}} \quad \sum_{i=1}^{n_c} P_i(\mathbf{T}) \tag{25a}$$

$$\text{s.t.} \quad h(\mathbf{T}) = h_{\text{ref}}, \tag{25b}$$

$$S_i(\mathbf{T}) \geq m_{\text{safe}}, \qquad i = 1, \dots, n_c. \tag{25c}$$

where h_{ref} denotes the required station demand, the mass flow in the parallel setup and the final discharge pressure in the series configuration, and $m_{\text{safe}} > 0$ is a prescribed minimum surge-margin threshold.

Problem (25) is the torque based realization of the generic load-sharing problem in Eq. (1): the generic operating target vector $\mathbf{z}$ is replaced by the shaft torque vector $\mathbf{T}$, and the surge constraint is tightened from $S_i(\mathbf{z}) \geq 0$ to $S_i(\mathbf{T}) \geq m_{\text{safe}} > 0$.

Problem (25) is unified in the specific sense that both compressor arrangements use the same torque-based optimization structure, GP-surrogate workflow, and CEI and UCB-LB acquisition designs. Only the dimension of $\mathbf{T}$ and the definition of the station demand h differ between the two arrangements. In classical load-sharing optimization, the decision variables are typically compressor speeds and the optimization relies directly on compressor maps and first-principles thermodynamic relations. In the present work, by contrast, the detailed dynamic and thermodynamic relations summarized in Appendix A are used only to construct the high-fidelity simulation environment and are not supplied to the load-sharing optimizer. Instead, the optimizer treats the station as a black-box steady-state mapping from torque inputs to measured outputs.

This formulation is motivated by practical implementation considerations. Although the full compressor maps are not assumed to be available to the optimizer, the surge control line and the corresponding DTS signals are typically available from the plant protection system. The proposed framework therefore uses measured DTS values as the primary safety feedback and learns, from data, how torque allocation influences both station performance and distance to surge.

By operating directly on shaft-torque setpoints and using measured operability indicators rather than explicit compressor maps by the load-sharing optimizer, the proposed framework reduces dependence on potentially inaccurate plant models while remaining compatible with the instrumentation and protection structures commonly available in industrial compressor stations.

3.1.2. Learning data-driven surrogate models using GPs

Given the "black-box" nature of the plant, we replace the unknown functions with independent probabilistic GP surrogates. Using the GP definitions presented in Section 2.2.1, let:

$$\mathcal{F} = \left\{ P, h, S_1, \dots, S_{n_c} \right\}$$

denote the set containing the station power, demand variable, and DTS constraint functions.

The GP surrogates are trained separately. Mathematically, this defines a factorized surrogate model conditional on the torque vector $\mathbf{T}$. Each output is learned as a function of the common torque vector, and, under the fixed external conditions considered here, the effect of compressor speed on surge behavior is reflected indirectly in the learned torque-to-DTS mapping.

To ensure numerical robustness, each torque input is normalized using its admissible bounds, and the resulting normalized torque vector is denoted by $\tilde{\mathbf{T}} \in [0,1]^{n_c}$, where:

$$\tilde{T}_i = \frac{T_i - T_i^{\min}}{T_i^{\max} - T_i^{\min}}, \qquad i = 1, \dots, n_c.$$

In the prediction equations below, $\mathbf{T}_*$ denotes a test point in physical torque coordinates, and $\tilde{\mathbf{T}}_*$ denotes its normalized representation.

We utilize a Squared Exponential kernel with Automatic Relevance Determination (ARD). For both parallel and series configurations, identical formulations are adopted for the GP surrogate modeling framework.

Model-specific output preprocessing. For a function $f \in \mathcal{F}$, consider a dataset:

$$\mathcal{D}_f = \left\{(\mathbf{T}_k, y_{f,k})\right\}_{k=1}^{N}, \qquad y_{f,k} = f(\mathbf{T}_k).$$

The physical output is expressed in terms of a deterministic centering or trend function $c_f(\mathbf{T})$, a positive output scale a_f, and a standardized latent function:

$$f(\mathbf{T}) = c_f(\mathbf{T}) + a_f \tilde{f}(\tilde{\mathbf{T}}).$$

After output preprocessing, the standardized latent function is modeled as:

$$\tilde{f}(\tilde{\mathbf{T}}) \sim \mathcal{GP}\left(0, k_f(\tilde{\mathbf{T}}, \tilde{\mathbf{T}}')\right), \qquad f \in \mathcal{F}.$$

Equivalently, the corresponding physical-output surrogate is:

$$f(\mathbf{T}) \sim \mathcal{GP}\left(c_f(\mathbf{T}), a_f^2 k_f(\tilde{\mathbf{T}}, \tilde{\mathbf{T}}')\right).$$

The scaled observations are:

$$y_{f,k}^{\text{scaled}} = \frac{y_{f,k} - c_f(\mathbf{T}_k)}{a_f},$$

and are collected in:

$$\mathbf{y}_{f,\text{scaled}} = \begin{bmatrix} y_{f,1}^{\text{scaled}} & \cdots & y_{f,N}^{\text{scaled}} \end{bmatrix}^\top.$$

The model-specific choices of $c_f(\mathbf{T})$ and a_f are:

$$\left(c_f(\mathbf{T}), a_f\right) = \begin{cases} \left(\bar{y}, \sigma_y\right), & f = P, \\ \left(\widehat{h}_{\text{lin}}(\mathbf{T}), \sigma_r\right), & f = h, \\ \left(0, \sigma_{S_i,\text{scale}}\right), & f = S_i. \end{cases}$$

For the power model,

$$\bar{y} = \frac{1}{N}\sum_{k=1}^{N} y_{P,k}, \qquad \sigma_y = \text{std}(y_{P,1}, \ldots, y_{P,N}),$$

so the power observations are centered at their empirical mean and scaled by their standard deviation.

For the demand variable, $h(\mathbf{T})$ corresponds to the final discharge pressure in the series configuration or the total outlet mass flowrate in the parallel configuration. Since this variable often exhibits an approximately affine dependence on torque, an affine trend is fitted using least-squares regression. We denote the fitted trend by $\widehat{h}_{\text{lin}}(\mathbf{T})$. This function provides the deterministic trend component of the demand-variable surrogate, while the zero-mean GP models deviations from this trend:

$$\widehat{h}_{\text{lin}}(\mathbf{T}) = \hat{\beta}_0 + \sum_{i=1}^{n_c} \hat{\beta}_i T_i, \qquad \hat{\boldsymbol{\beta}} = (\boldsymbol{\Phi}^\top \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^\top \mathbf{y}_h,$$

where:

$$\mathbf{y}_h = \begin{bmatrix} y_{h,1} & \cdots & y_{h,N} \end{bmatrix}^\top,$$

and

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & T_{1,1} & \cdots & T_{1,n_c} \\ \vdots & \vdots & & \vdots \\ 1 & T_{N,1} & \cdots & T_{N,n_c} \end{bmatrix} \in \mathbb{R}^{N \times (n_c+1)}.$$

This corresponds to the ordinary least-squares solution under the assumption that $\boldsymbol{\Phi}$ has full column rank. The regression coefficients are re-estimated and the residual GP is re-trained whenever a new observation is appended to the dataset.

The residuals are:

$$r_k = y_{h,k} - \widehat{h}_{\text{lin}}(\mathbf{T}_k), \qquad \sigma_r = \text{std}(r_1, \ldots, r_N).$$

Rather than modeling $h(\mathbf{T})$ directly with a GP, the demand-variable surrogate incorporates the approximately affine torque–demand relationship as a deterministic trend and allows the GP to model the remaining nonlinear deviations. This procedure extends conventional centering by the sample mean. Rather than subtracting a constant mean, we subtract the fitted affine function $\widehat{h}_{\text{lin}}(\mathbf{T})$.

Because the affine regression includes an intercept, the least-squares residuals have zero sample mean, up to numerical precision:

$$\bar{r} = \frac{1}{N}\sum_{k=1}^{N} r_k = 0.$$

Further subtraction of the residual mean is therefore unnecessary. For the demand-variable model, the common scaling expression reduces to:

$$r_k^{\text{scaled}} = \frac{r_k - \bar{r}}{\sigma_r} = \frac{r_k}{\sigma_r}.$$

For each compressor $i \in \{1, \ldots, n_c\}$, an independent GP model is constructed for the DTS margin $S_i(\mathbf{T})$. The DTS observations are scaled about the fixed physical reference $S_i = 0$, rather than centered at their empirical mean, because this preserves the location and sign of the surge-control boundary in the scaled output space. Given observations $y_{S_i,k} = S_i(\mathbf{T}_k)$, the outputs are scaled as:

$$\sigma_{S_i,\text{scale}} = \text{std}(y_{S_i,1}, \ldots, y_{S_i,N}), \qquad y_{S_i,k}^{\text{scaled}} = \frac{y_{S_i,k}}{\sigma_{S_i,\text{scale}}}.$$

Common GP training and prediction. For every $f \in \mathcal{F}$, the regularized covariance matrix and its Cholesky factorization are:

$$\mathbf{K}_f = k_f\left(\tilde{\mathbf{T}}_{1:N}, \tilde{\mathbf{T}}_{1:N}\right) + \sigma_\epsilon^2 \mathbf{I} = \mathbf{L}_f \mathbf{L}_f^\top,$$

where $\sigma_\epsilon^2 = 10^{-4}$ is the fixed observation-noise variance used to regularize the covariance matrix in the standardized output space, and $\mathbf{L}_f$ is its lower-triangular Cholesky factor.

For a test point $\mathbf{T}_*$ with normalized representation $\tilde{\mathbf{T}}_*$, define:

$$\mathbf{k}_{*,f}(\tilde{\mathbf{T}}_*) = \begin{bmatrix} k_f(\tilde{\mathbf{T}}_1, \tilde{\mathbf{T}}_*) & \cdots & k_f(\tilde{\mathbf{T}}_N, \tilde{\mathbf{T}}_*) \end{bmatrix}^\top,$$

and

$$k_{**,f}(\tilde{\mathbf{T}}_*) = k_f(\tilde{\mathbf{T}}_*, \tilde{\mathbf{T}}_*).$$

The predictive mean and variance in the original output space are:

$$\mu_f(\mathbf{T}_*) = c_f(\mathbf{T}_*) + a_f \mathbf{k}_{*,f}(\tilde{\mathbf{T}}_*)^\top \boldsymbol{\alpha}_f,$$

and

$$\sigma_f^2(\mathbf{T}_*) = a_f^2\left(k_{**,f}(\tilde{\mathbf{T}}_*) - \mathbf{v}_f(\tilde{\mathbf{T}}_*)^\top \mathbf{v}_f(\tilde{\mathbf{T}}_*)\right),$$

where:

$$\boldsymbol{\alpha}_f = \mathbf{L}_f^{-\top} \mathbf{L}_f^{-1} \mathbf{y}_{f,\text{scaled}}, \qquad \mathbf{v}_f(\tilde{\mathbf{T}}_*) = \mathbf{L}_f^{-1} \mathbf{k}_{*,f}(\tilde{\mathbf{T}}_*).$$

Equivalently, the posterior prediction at the test point can be written as:

$$f(\mathbf{T}_*) \mid \mathcal{D}_f \sim \mathcal{N}\left(\mu_f(\mathbf{T}_*), \sigma_f^2(\mathbf{T}_*)\right).$$

In practice, $\boldsymbol{\alpha}_f$ and $\mathbf{v}_f(\tilde{\mathbf{T}}_*)$ are computed by forward and backward substitution rather than by forming matrix inverses explicitly.

3.1.3. Initialization of the surrogate models

Before surrogate-based optimization is activated, an initial dataset is collected to provide the minimum amount of spatial variation required to fit the GP models. In the present study, this initialization dataset contains exactly three steady-state operating points, i.e. $N_{\min} = 3$.

For each run, the procedure begins from a safe nominal torque vector, denoted by $\mathbf{T}_{\text{nom}} \in \mathcal{T}$. This point serves as the center of the initial design. Once $\mathbf{T}_{\text{nom}}$ has been selected, two additional initialization points are generated by applying prescribed local perturbations around it. The resulting initialization set is therefore:

$$\mathcal{T}_{\text{INIT}} = \left\{\mathbf{T}_{\text{nom}},\ \mathbf{T}_{\text{nom}} + \Delta\mathbf{T}^{(1)},\ \mathbf{T}_{\text{nom}} + \Delta\mathbf{T}^{(2)}\right\}, \tag{26}$$

where $\Delta\mathbf{T}^{(1)}$ and $\Delta\mathbf{T}^{(2)}$ are fixed perturbation vectors chosen a priori. Specifically, the parallel configuration uses $\Delta\mathbf{T}^{(1)} = [0, 0.25, 0.25]^\top$ and $\Delta\mathbf{T}^{(2)} = [0.25, 0.25, 0]^\top$ N m, whereas the series configuration uses $\Delta\mathbf{T}^{(1)} = [0.25, 0]^\top$ and $\Delta\mathbf{T}^{(2)} = [0, 0.25]^\top$ N m. Thus, every nonzero perturbation component has a fixed magnitude of 0.25 N m. These perturbations are selected to remain within the admissible torque domain and to keep the compressors in a safe operating region during the initialization phase.

Operationally, the station is first brought to $\mathbf{T}_{\text{nom}}$ and allowed to settle to steady state. The corresponding measurements of total power, demand variable, and compressor DTS values are stored as the first entry of the dataset. The remaining two initialization points are then applied sequentially in the same manner, so that the initial dataset is populated as:

$$\mathcal{D}_0 = \left\{\left(\mathbf{T}^{(j)}, P^{(j)}, h^{(j)}, \mathbf{S}^{(j)}\right)\right\}_{j=1}^{3}, \tag{27}$$

with $\mathbf{T}^{(j)} \in \mathcal{T}_{\text{INIT}}$.

This initialization strategy deliberately uses a very small number of measurements. Rather than relying on a large exploratory DoE, it provides only the minimum information needed to start training the GP surrogates, after which further data are collected online through the CBO procedure. In this way, the optimizer is initialized from a safe local neighborhood of a nominal operating condition while retaining the ability to learn improved load-sharing policies directly from subsequent plant-response data.

3.1.4. Real-time optimization

Once the minimum dataset requirement is satisfied ($|D| \geq N_{\min}$), the algorithm transitions to the active learning phase. The GP surrogate models for power, demand tracking, and surge-margin constraints are updated using the current dataset, and the corresponding acquisition function $J(\mathbf{T})$ is constructed. The acquisition function is maximized using a genetic algorithm (GA), and the resulting candidate $\mathbf{T}_{k+1}$ is selected as the next torque vector to be evaluated.

After initialization, the real-time optimizer outputs are updated only when a new steady-state operating point has been detected; otherwise, the previously applied torque is maintained.

Steady-state detection is based on the primary station variable, namely the total outlet mass flowrate in the parallel configuration and the final discharge pressure in the series configuration. A steady-state operating point is declared only when three conditions are satisfied simultaneously over a sliding temporal window: limited peak-to-peak variation, low statistical dispersion measured by the median absolute deviation, and a negligible linear trend. Requiring these conditions to persist over the detection window reduces the influence of transient fluctuations and ensures that the measurements used for model updates are representative of the underlying steady-state input–output mapping.

Once steady state has been confirmed, the corresponding measurements of power, demand variable, and DTS are evaluated for possible inclusion in the dataset. Novelty is calculated using Euclidean distance in the normalized torque space:

$$d_{\min}(\tilde{\mathbf{T}}_k) = \min_{j=1,\dots,N} \left\| \tilde{\mathbf{T}}_k - \tilde{\mathbf{T}}_j \right\|_2 .$$

A subsequent steady-state observation is appended only when $d_{\min}(\tilde{\mathbf{T}}_k) > r_{\text{nov}}$, where $r_{\text{nov}} > 0$ denotes the corresponding dimensionless novelty radius. Its numerical values and selection are described in Section 3.3.

The novelty criterion controls only whether an observed steady-state point is added to the GP training dataset. It does not constrain the acquisition-function search space or exclude a candidate torque command from being selected and applied.

The overall procedure therefore alternates between steady-state data collection and surrogate-based optimization. During transients, the torque command is held constant; at each accepted steady-state point, the dataset may be expanded, the GP models may be updated, and a new torque command may be computed by the load-sharing optimizer. The proposed real-time optimization framework is summarized in Algorithm 1.

For each optimization run, one acquisition strategy is selected a priori and retained throughout the run: J_1 for CEI or J_2 for UCB-LB. The following two subsections define these acquisition functions, each combining power reduction, reference tracking, and operational safety.

3.1.5. Implementation of CEI

We construct an acquisition function that maximizes the expected reduction in power while simultaneously accounting for the likelihood of satisfying safety and tracking constraints. The acquisition function is formulated as a multiplicative product (Marler and Arora, 2004) and is a more conservative formulation of the CEI criterion introduced in Section 2.2.5:

$$J_1(\mathbf{T}) = \text{EI}(\mathbf{T}) \cdot \prod_{i=1}^{n_c} \text{PoF}_i(\mathbf{T}) \cdot \delta_{\text{track}}(\mathbf{T}) \tag{28}$$

- **Expected improvement** (EI($\mathbf{T}$)): This term quantifies the anticipated power reduction relative to the best feasible observation found so far, $P_{\min}$. Let $\Delta P(\mathbf{T}) = P_{\min} - \mu_P(\mathbf{T})$ denote the difference between the best observed power and the GP power prediction. The analytic formulation for EI is given by:

$$\text{EI}(\mathbf{T}) = \Delta P(\mathbf{T})\Phi(Z) + \sigma_P(\mathbf{T})\phi(Z)$$

Algorithm 1: CBO for real-time load-sharing optimization

Input: Nominal torque $\mathbf{T}_{\text{nom}}$, torque bounds $\mathcal{T}$, initial torque set $\mathcal{T}_{\text{INIT}}$, minimum dataset size $N_{\min}$, novelty radius r_{nov}, reference h_{ref}, and acquisition strategy $\mathcal{A} \in \{\text{CEI}, \text{UCB-LB}\}$

Initialize: $D \leftarrow \emptyset$, where each element of D is a tuple $(\mathbf{T}_k, P_k, h_k, \mathbf{S}_k)$ with $\mathbf{S}_k = [S_{1,k}, \dots, S_{n_c,k}]^\top$; initialize GP priors for P, h, and $\{S_i\}_{i=1}^{n_c}$; $\mathbf{T}_0 \leftarrow \mathbf{T}_{\text{nom}}$

for $k = 0, 1, 2, \dots$ **do**
  Measure $\mathbf{T}_k$, P_k, h_k, and $\mathbf{S}_k = [S_{1,k}, \dots, S_{n_c,k}]^\top$;
  if *steady-state conditions are satisfied* **then**
    if $D = \emptyset$ **then**
      Append $(\mathbf{T}_k, P_k, h_k, \mathbf{S}_k)$ to D;
    else
      Compute $d_{\min} \leftarrow \min_{(\mathbf{T}_j, \cdot) \in D} \|\tilde{\mathbf{T}}_k - \tilde{\mathbf{T}}_j\|_2$;
      if $d_{\min} > r_{\text{nov}}$ **then**
        Append $(\mathbf{T}_k, P_k, h_k, \mathbf{S}_k)$ to D;
    if $|D| < N_{\min}$ **then**
      Select the next unused initialization point: $\mathbf{T}_{k+1} \leftarrow$ next point in $\mathcal{T}_{\text{INIT}}$;
    else
      Update GP posteriors using D: $\mu_P, \sigma_P^2, \mu_h, \sigma_h^2, \mu_{S_i}, \sigma_{S_i}^2$;
      Construct the selected acquisition function: $J(\mathbf{T}) \leftarrow J_1(\mathbf{T})$ if $\mathcal{A} = \text{CEI}$, or $J(\mathbf{T}) \leftarrow J_2(\mathbf{T})$ if $\mathcal{A} = \text{UCB-LB}$;
      Solve $\mathbf{T}_{k+1} \leftarrow \arg\max_{\mathbf{T} \in \mathcal{T}} J(\mathbf{T})$;
  else
    $\mathbf{T}_{k+1} \leftarrow \mathbf{T}_k$;
  Apply torque command $\mathbf{T}_{k+1}$;

where $Z = \frac{\Delta P(\mathbf{T})}{\sigma_P(\mathbf{T})}$, while $\Phi(\cdot)$ and $\phi(\cdot)$ represent the standard normal CDF and PDF, respectively. By focusing on improvement rather than absolute value, it naturally balances the exploration of highly uncertain efficiency regions with the exploitation of known optima.

- **Surge feasibility** ($\text{PoF}_i(\mathbf{T})$): To protect the machinery, we calculate the PoF for each compressor i. Rather than using the predictive mean directly, we utilize a conservative LCB. The LCB and the corresponding probability of feasibility are defined as:

$$\text{LCB}_i(\mathbf{T}) = \mu_{S_i}(\mathbf{T}) - \beta\sigma_{S_i}(\mathbf{T}), \qquad \text{PoF}_i(\mathbf{T}) = \Phi\left(\frac{\text{LCB}_i(\mathbf{T}) - m_{\text{safe}}}{\sigma_{S_i}(\mathbf{T})}\right),$$

where $\beta > 0$ determines the degree of conservativeness and $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. Smaller PoF values therefore reduce the acquisition value of candidate points whose conservatively estimated surge margins are close to or below m_{safe}.

- **Weighted reference tracking** ($\delta_{\text{track}}(\mathbf{T})$): The demand variable tracking is handled via an inverse quadratic weighting function centered at the reference setpoint, h_{ref}:

$$\delta_{\text{track}}(\mathbf{T}) = \frac{1}{\left(\frac{\mu_h(\mathbf{T}) - h_{\text{ref}}}{\Delta_h}\right)^2 + 1}$$

where Δ_h is a scaling factor defined by the physical units of the demand variable. This formulation guarantees a maximum utility of 1 when the predicted output perfectly matches the demand reference. Setpoints that fail to meet the required mass flow or pressure yield a decaying penalty governed by Δ_h, facilitating smooth convergence toward the feasible demand manifold by avoiding numerical instabilities.

Under the standard regularity assumptions used in GP confidence-bound analyses (Chowdhury and Gopalan, 2017), and provided that β is chosen sufficiently large, $\mathrm{LCB}_i(\mathbf{T})$ may be interpreted as a conservative approximation of the unknown true surge margin with high probability (Fiedler et al., 2021; Tokmak et al., 2025). Consequently, operating points whose conservatively estimated surge margins lie close to the threshold m_{safe} receive small probability-of-feasibility values, and their overall acquisition value is strongly reduced through the multiplicative CEI structure. In this way, the proposed CEI formulation promotes risk-aware exploration and conservative operation with respect to the learned steady-state surge constraints. Unlike the UCB-LB formulation introduced in the next subsection, this mechanism acts as a soft probabilistic filter rather than a stronger exclusion rule, and therefore should not be interpreted as a formal end-to-end safety guarantee for the full closed-loop system. Transient dynamics, steady-state detection delays, and model-calibration errors may still affect practical operation.

3.1.6. Implementation of UCB-LB

As described in Section 2.2.6, as an alternative to the multiplicative CEI formulation, we directly minimize a composite objective that aggregates a conservative estimate of power (via a UCB criterion), a quadratic tracking penalty, and a log-barrier safety term enforcing conservative feasibility of all surge-margin constraints:

$$J_2(\mathbf{T}) = -\left(w_P \mathcal{J}_{\mathrm{power}} + w_{\mathrm{track}} \mathcal{J}_{\mathrm{track}} + w_{\mathrm{surge}} \mathcal{J}_{\mathrm{safe}}\right) \tag{29}$$

To maintain compatibility with the maximization step in the GA, the composite cost is negated and written as an acquisition function. Its main components are as follows:

- **Overall power consumption term** ($\mathcal{J}_{\mathrm{power}}$): The power contribution is based on a LCB estimate from the GP surrogate. This ensures that candidate operating points are favored when they are predicted to yield low power consumption with high confidence:

  $$\mathcal{J}_{\mathrm{power}}(\mathbf{T}) = \mu_P(\mathbf{T}) - \kappa \sigma_P(\mathbf{T})$$

  where $\kappa > 0$ is the exploration parameter. By subtracting the scaled standard deviation, the optimizer is mathematically encouraged to explore regions of the torque space where the surrogate model is highly uncertain, preventing convergence to local optima.
- **Quadratic tracking term** ($\mathcal{J}_{\mathrm{track}}$): Demand satisfaction is enforced through a quadratic penalty on the deviation between the predicted demand variable, $\mu_h(\mathbf{T})$, and its target reference, h_{ref}:

  $$\mathcal{J}_{\mathrm{track}}(\mathbf{T}) = (\mu_h(\mathbf{T}) - h_{\mathrm{ref}})^2$$

  Depending on the station configuration, this variable corresponds to either the total outlet mass flowrate or the final discharge pressure. The quadratic nature of this penalty heavily discourages large deviations from the setpoint while allowing minor, acceptable fluctuations near the optimum.
- **Log-barrier safety term** ($\mathcal{J}_{\mathrm{safe}}$): Safety is handled through an interior-point barrier formulation applied to the surge-margin constraints for all n_c compressors. The logarithmic barrier explicitly penalizes operating points that approach the learned constraint boundary too closely:

  $$\mathcal{J}_{\mathrm{safe}}(\mathbf{T}) = -\sum_{i=1}^{n_c} \log\left(\max(\mathrm{LCB}_i(\mathbf{T}) - m_{\mathrm{safe}}, \epsilon)\right)$$

  where m_{safe} represents the safety margin, and $\epsilon > 0$ is a numerical floor that prevents singularities when the barrier is violated. This formulation discourages candidate points that fall outside the conservatively estimated feasible region by driving the penalty to a high ceiling value.

As in the CEI case, $\mathrm{LCB}_i(\mathbf{T})$ may be interpreted as a conservative approximation of the unknown surge margin under the GP confidence-bound assumptions discussed above. In the UCB-LB formulation, however, this conservative estimate is incorporated through an explicit log-barrier term, so that candidate torque vectors with predicted surge margins close to the constraint boundary are penalized more aggressively. As with CEI, this mechanism promotes conservative operation but does not constitute a formal end-to-end safety guarantee for the full closed-loop system.

3.2. Configuration-specific load-sharing problems

The proposed CBO framework is applied to the parallel and series compressor configurations shown in Figs. 2 and 3. The parallel station comprises three heterogeneous compressors ($n_c = 3$) discharging into a common header, whereas the series station is modeled as a two-stage compression system ($n_c = 2$). In both configurations, the objective is to minimize the total station power while tracking the prescribed station demand and maintaining safe operation with respect to surge. The configuration-specific decision variables, tracking requirements, and safety constraints are summarized in Table 1.

It is important to recall that the optimizer is not provided with the underlying compressor models and that the torque values are the only inputs manipulated by the optimization algorithm. In the parallel configuration, the remaining inputs are fixed and identical for all three compressors: the inlet valve opening is $u_{\mathrm{in}} = 0.4541$, the outlet valve opening is $u_{\mathrm{out}} = 0.56$, the recycle valve opening is $u_{\mathrm{rec}} = 0$, and the inlet temperature is $T_{\mathrm{in}} = 293$ K.

In the series configuration, the remaining inputs are also fixed. These parameters are summarized in Table 2, where u_{in} is the inlet valve opening, u_{out} is the outlet valve opening, u_{rec} is the recycle valve opening, and T_{in} is the inlet temperature.

To evaluate the response of the algorithm to changes in the station demand, both configurations are simulated for $30,000$ s. The reference is changed at intervals of $5,000$ s according to the step-change profiles summarized in Table 3.

3.3. Implementation details and hyperparameters

To ensure full reproducibility of the proposed CBO framework, the main algorithmic hyperparameters and steady-state detection settings are explicitly defined here. The framework was implemented in MATLAB 2025b, with the physical plant simulated via Simulink. Unless otherwise stated, all pressure, flow, and torque variables used by the load-sharing optimizer are normalized to the interval $[0, 1]$ using the corresponding minimum and maximum admissible operating bounds. Consequently, all hyperparameters and steady-state detection thresholds reported in this section are given in normalized units, except for the window length W, which is reported in seconds.

GP initialization and fitting. To initialize the learning process, the initial Design of Experiments (DoE) dataset for each run was constructed from three operating points: one safe nominal torque vector and two prescribed local perturbations around it. The initialization perturbation vectors are specified in Section 3.1.3. To assess sensitivity to initialization, the algorithm was evaluated over five independent runs. In each run, a different safe nominal torque vector $\mathbf{T}_{\mathrm{nom}}^{(s)}$, $s = 1, \ldots, 5$, was selected, and the corresponding three-point initialization dataset was generated using the same local perturbation pattern.

To ensure computational efficiency during real-time online execution, all GP surrogate models used a squared exponential kernel with automatic relevance determination (ARD) and fixed hyperparameters. All GP hyperparameters were selected empirically and subsequently held fixed throughout each run. Because the model outputs are scaled before GP fitting, the signal variance was set to $\sigma_f^2 = 1.0$, and the observation noise variance was fixed at $\sigma_\epsilon^2 = 10^{-4}$. The characteristic length scale was set to $\ell = 0.50$ for the power and surge models, and

Table 1
Configuration-specific load-sharing objectives and constraints.

Element	Parallel configuration	Series configuration
Decision variables and power objective	The optimizer manipulates $\mathbf{T} = [T_1, T_2, T_3]^\top$ and minimizes the total measured power, $P(\mathbf{T}) = \sum_{i=1}^{3} P_i(\mathbf{T})$.	The optimizer manipulates $\mathbf{T} = [T_1, T_2]^\top$ and minimizes $P(\mathbf{T}) = \sum_{i=1}^{2} P_i(\mathbf{T})$. The torque split must achieve the required final discharge pressure without over-compressing the first stage or choking the second.
Tracking constraint	The demand variable is the total discharge mass flowrate, $h(\mathbf{T}) = M_{\text{out,tot}}$, which must track $h_{\text{ref}} = M_{\text{ref}}$.	The demand variable is the final discharge pressure, $h(\mathbf{T}) = P_{\text{dis}}$, which must track $h_{\text{ref}} = P_{\text{ref}}$.
Safety constraints	Safety is monitored through three DTS GPs (S_1, S_2, S_3). Each compressor must remain above the surge control line.	Safety is monitored through two DTS GPs (S_1, S_2). Both stages must remain above the surge control line because a surge event in the first stage can induce a flow collapse in the second.

Table 2
Static operational parameters for the series compressor configuration.

	Compressor 1 (C_1)	Compressor 2 (C_2)
u_{in}	0.5 [–]	0.4541 [–]
u_{out}	0.46 [–]	0.56 [–]
u_{rec}	0.0 [–]	0.0 [–]
T_{in}	293 [K]	293 [K]

Table 3
Step-change station-demand profiles for the parallel and series configurations.

Timespan [s]	Mass-flow demand [kg/s]	Pressure demand [kPa]
0–5000	2.1	248
5000–10,000	1.8	244
10,000–15,000	1.9	246
15,000–20,000	1.7	243
20,000–25,000	1.6	242
25,000–30,000	1.5	241

to $\ell = 0.25$ for the demand-variable residual model in order to reduce excessive extrapolation. The hyperparameters were held fixed and were not re-optimized during the workflow. When a new observation was accepted, the GP posterior was recomputed using the expanded dataset while retaining the fixed hyperparameters.

Steady-state detection. To promote the collection of data that closely approximates thermodynamic equilibrium, a multi-criteria steady-state detector was implemented upstream of the load-sharing optimizer. A sliding temporal window of $W = 20$ s was evaluated continuously. Steady-state was declared only when the peak-to-peak amplitude variation remained below $\Delta_{\max} = 0.015$, the robust statistical dispersion measured by the median absolute deviation (MAD) remained below 0.004, and the linear trend (slope) of the windowed data remained below 5×10^{-4}.

Novelty criterion. The physical novelty radius of 0.25 N m was selected empirically through preliminary simulation trials in which several candidate radii were examined. Among the tested values, it provided a practical balance between avoiding near-duplicate steady-state observations, limiting dataset growth, and retaining sufficiently distinct observations for GP posterior updates. For implementation, the physical radius was converted using the torque span, resulting in normalized radii of 0.0208 and 0.025 for the parallel and series configurations, respectively. These values were held fixed across both acquisition functions and all initialization runs.

GA configuration. The acquisition functions were maximized using a custom GA implementation. To balance computational efficiency with global search capability, the population size was set to 25 individuals and the algorithm was run for 20 generations. The main evolutionary settings, together with the GP and steady-state detection parameters, are summarized in Table 4.

Acquisition-function weight selection. The power, tracking, and surge weights were tuned empirically through preliminary closed-loop simulation trials. Several candidate weight combinations were evaluated separately for the parallel and series configurations. The values reported in Table 4 provided the best overall trade-off among the tested combinations between reducing power consumption, maintaining reference tracking, and preserving a positive surge margin. These are therefore the final selected values rather than the initial values examined during tuning. Once selected, the weights were held fixed across all five initialization runs.

4. Results and discussion

In this section, we evaluate the performance of the proposed CBO framework across parallel and series compressor configurations. We provide a comparative analysis of the two primary strategies: constrained expected improvement (CEI) and upper confidence bound with log-barrier (UCB-LB), against the baseline (Equal-Load distribution) and the theoretical optimum determined via exhaustive search.

The evaluation focuses on four critical dimensions of performance: the algorithmic robustness to various initial conditions; the ability to maintain mass flow or discharge pressure at the setpoint; total energy savings and the operational safety (surge avoidance).

4.1. Performance across initialization scenarios

To assess robustness with respect to initialization, each configuration was simulated over five distinct initialization scenarios. The nominal torque vectors were constructed as:

$$\mathbf{T}_{\text{nom}}^{(s)} = t_s \mathbf{1}_{n_c}, \qquad s = 1, \ldots, 5, \qquad (t_1, \ldots, t_5) = (13, 14, 15, 16, 17) \text{ N m}.$$

where $n_c = 3$ for the parallel configuration and $n_c = 2$ for the series configuration. Thus, the series cases used $[t_s, t_s]^\top$, whereas the parallel cases used $[t_s, t_s, t_s]^\top$. The same five nominal vectors were used for CEI and UCB-LB. Each nominal vector was used to construct the corresponding three-point initialization dataset using the same prescribed local perturbation pattern described in Section 3.1.3. This allows the sensitivity of the framework to cold-start conditions to be assessed while keeping the initialization procedure itself consistent across runs.

Across all scenarios, both the CEI and the UCB-LB strategies demonstrated a consistent trajectory toward high energy efficiency outcomes. Although the nominal torque vector differed across the five scenarios, CEI and UCB-LB used the same vector within each corresponding scenario. Both strategies successfully identified improved load distributions within a limited number of steady-state iterations, while tracking the station targets for flow or pressure. Each closed-loop simulation was run for a fixed duration of 30,000 s, comprising six 5,000 s demand intervals. The overall optimization terminated at the end of the fixed simulation horizon. Across the tested scenarios, all DTS

Table 4
Algorithmic hyperparameters. Unless otherwise noted, values are given in normalized units.

Category	Parameter	Value
Gaussian Process	Kernel Function	Squared Exponential with ARD
	Observation Noise (σ_e^2)	10^{-4}
	Length Scale (ℓ)	0.50 (Power/Surge) 0.25 (Demand Residuals)
	Novelty Radius (r_{nov})	0.0208 (Parallel), 0.025 (Series)
Genetic Algorithm	Population Size	25
	Maximum Generations	20
	Crossover Probability	0.60
	Mutation Rate	0.40
	Selection Mechanism	Tournament (Size = 2)
Steady-state Detector	Sliding Window Size (W)	20 s
	Max Peak-to-Peak Amplitude ($\Delta_{\max}$)	0.015
	Max Median Absolute Deviation (MAD)	0.004
	Max Linear Trend (Slope)	5×10^{-4}
Acquisition Function	Surge Safety Margin (m_{safe})	0.01
	Conservative Safety Bound (β)	3.0
	Tracking Scaling Factor (Δ_h)	0.001 (Parallel) 1 (Series)
	UCB Exploration Factor (κ)	0.01
	Power Weight (w_P)	0.01 (Parallel) 1 (Series)
	Tracking Weight (w_{track})	2×10^7 (Parallel) 10 (Series)
	Surge Weight (w_{surge})	0.01 (Parallel) 1×10^4 (Series)

values remained positive during the learning phase, including the initial exploration period, and no activation of the local anti-surge controllers was observed. In CEI, feasibility enters through a soft multiplicative probability term; therefore, the method remains more exploratory and can select candidates closer to the estimated constraint boundary, with a correspondingly greater risk of constraint violation under surrogate uncertainty. In contrast, UCB-LB applies an explicit log-barrier penalty to conservative lower confidence bounds, thereby prioritizing feasibility and yielding a more conservative operating profile. The simulations showed no surge control line crossings for either method. However, this observation is only an empirical result rather than a formal safety guarantee for either acquisition function.

4.2. Setpoint tracking performance

An essential requirement of the proposed load-sharing optimizer is that it improves load allocation without degrading tracking of the prescribed station demand variable h. This analysis therefore evaluates whether the generated torque commands maintain accurate demand tracking while the optimizer searches for more energy-efficient operating points. Figs. 4 and 5 show the corresponding transient tracking responses under the UCB-LB and CEI strategies.

Beyond the initialization and early exploration phase, both acquisition functions exhibit good tracking performance, with only small steady-state deviations from the reference. The framework also responds effectively to step changes in station demand: as the reference changes, the load-sharing optimizer updates the torque allocation and the closed-loop system settles to the new operating condition without loss of tracking quality. This behavior over the full simulation horizon indicates that the proposed data-driven framework can coordinate load sharing and demand satisfaction in a consistent manner.

To ensure a fair comparison, the first 2500 s of each simulation were excluded from the tracking-performance assessment, since this interval includes the initialization procedure and the early exploratory transients associated with surrogate-model construction. As shown in Figs. 4 and 5, the tracking responses exhibit larger transient deviations during the first 2500 s interval than those observed later in the simulations. Tracking performance was then quantified using the mean absolute error (MAE) and root mean square error (RMSE), computed across the five independent initialization runs. Table 5 reports the ensemble mean (μ) and standard deviation (σ) of these metrics for the series configuration, while Table 6 reports the corresponding results for the parallel configuration. Overall, both acquisition strategies achieved comparable tracking performance, with only minor differences in MAE and RMSE across the two compressor-station configurations.

Table 5
Tracking performance metrics (Mean ± SD) for series configuration.

Metric	CEI	UCB-LB	Units
MAE	$4.86 \times 10^1 \pm 5.65$	$4.62 \times 10^1 \pm 8.83$	[Pa]
RMSE	$2.90 \times 10^2 \pm 2.12 \times 10^1$	$2.81 \times 10^2 \pm 1.54 \times 10^1$	[Pa]

Table 6
Tracking performance metrics (Mean ± SD) for parallel configuration.

Metric	CEI	UCB-LB	Units
MAE	$5.77 \times 10^{-3} \pm 8.29 \times 10^{-4}$	$5.83 \times 10^{-3} \pm 8.66 \times 10^{-4}$	[kg/s]
RMSE	$2.22 \times 10^{-2} \pm 8.23 \times 10^{-4}$	$2.23 \times 10^{-2} \pm 4.88 \times 10^{-4}$	[kg/s]

4.3. Analysis of energy efficiency gains

To establish a realistic benchmark, the performance of the proposed framework is first compared against a conventional equal-load distribution baseline. In this baseline configuration, the load-sharing optimizer is replaced by a decentralized control structure in which a primary PID controller regulates the overall process demand (total mass flow-rate setpoint for the parallel configuration and discharge pressure setpoint for the series configuration). The resulting control effort is then distributed equally among the compressors, such that each unit receives the same torque setpoint:

$$T_i(t) = \frac{T_{\text{total}}(t)}{n_c}, \quad \forall i \in \{1, \dots, n_c\}. \tag{30}$$

Although this strategy ensures stable reference tracking through standard feedback mechanisms, it is inherently insensitive to differences in compressor efficiency, health, and local operating conditions, and therefore cannot exploit more economical load-sharing policies.

In addition to this baseline, a second benchmark was introduced to represent a more conventional real-time optimization (RTO) workflow. Classical two-stage RTO schemes typically rely on a parametric mechanistic model, which is first updated through a parameter-estimation

Table 7
Parallel configuration steady-state power comparison: Oracle is the theoretical optimum.

Target $M_{out,tot}$	Steady-state Power [kW]					Savings vs. Baseline [%]			
[kg/s]	Baseline	Oracle	2-stage	CEI	UCB-LB	Oracle	2-stage	CEI	UCB-LB
2.1	324.700	315.055	317.032	318.296	316.273	2.970	2.362	1.972	2.595
1.8	197.977	197.283	197.586	197.314	197.613	0.351	0.197	0.335	0.184
1.9	225.703	223.294	223.729	223.342	223.780	1.067	0.876	1.046	0.852
1.7	178.226	178.151	178.212	178.173	178.161	0.042	0.008	0.030	0.037
1.6	164.493	164.027	164.346	164.218	164.402	0.283	0.089	0.167	0.055
1.5	155.995	153.722	154.806	153.867	153.942	1.457	0.762	1.364	1.316

Table 8
Series configuration steady-state power comparison: Oracle is the theoretical optimum.

Target P_{dis}	Steady-state Power [kW]					Savings vs. Baseline [%]			
[kPa]	Baseline	Oracle	2-stage	CEI	UCB-LB	Oracle	2-stage	CEI	UCB-LB
248.0	185.293	184.386	185.289	184.387	184.573	0.490	0.002	0.489	0.389
244.0	119.241	117.365	119.063	118.710	117.621	1.573	0.149	0.445	1.359
246.0	143.728	142.187	142.190	142.307	142.715	1.072	1.070	0.989	0.705
243.0	111.690	109.287	110.008	109.330	109.397	2.152	1.510	2.113	2.053
242.0	107.541	104.128	104.047[a]	104.246	104.824	3.174	3.249[a]	3.064	2.527
241.0	106.529	103.190	104.164	103.741	104.117	3.134	2.220	2.617	2.264

[a] This specific power reduction was observed despite a failure to reach the target pressure setpoint; consequently, the point has been excluded from efficiency comparisons to account for the tracking offset.

step using current plant measurements and then used in a deterministic optimization problem to compute the economically optimal operating point. To ensure a fair comparison with the proposed framework, the present study adopted a fully data-driven analog of this two-stage structure. Specifically, the same GP-based surrogate modeling framework used in the CBO formulations was retained, but the optimization step was carried out without an explicit exploration incentive. In practice, this corresponds to a deterministic exploitation policy analogous to the UCB-LB formulation with the exploration term removed, while still enforcing the surge constraint through a conservative prediction confidence bound. This benchmark therefore isolates the effect of active exploration and allows direct assessment of whether the proposed CBO policies outperform a non-exploring adaptive optimization alternative.

In addition, we established a theoretically optimal benchmark, denoted as *oracle*, which assumes full knowledge of the compressor maps and thermodynamic relations as discussed in Appendix B. This benchmark provides the minimum achievable steady-state power consumption for each reference target and therefore serves as an upper bound on attainable performance.

Tables 7 and 8 report the steady-state power consumption for the parallel and series configurations, respectively. To ensure a transparent comparison, the reported values were computed as the average power over the final 500 s of each 5000 s setpoint window, after the transient effects associated with the initial design-of-experiments phase and subsequent closed-loop settling had decayed. The resulting efficiency metrics therefore reflect the converged load-sharing equilibrium rather than temporary excursions or non-equilibrium operating periods.

The corresponding steady-state torque allocations obtained by CEI and UCB-LB are reported in Table 9. The torque values were calculated using the same procedure as the reported power values.

Table 9 shows that both CEI and UCB-LB produce asymmetric, demand-dependent torque allocations. In the series configuration, the second compressor consistently receives a slightly higher torque ($T_2 > T_1$), consistent with the learned power surface discussed in Section 4.5. In the parallel configuration, the load distribution changes with the demand target, reflecting the compressor performance characteristics across the operating range. The two acquisition strategies generally produce similar, but not identical, torque allocations.

The results show that the proposed CBO strategies consistently outperform the equal-load baseline and, in most cases, also improve upon the non-exploring two-stage RTO benchmark. Across all feasible operating points, the highest energy savings relative to equal-load operation reached 3.064%, while the best additional improvement over the two-stage RTO benchmark was 1.21 percentage points. This latter result is particularly important, as it demonstrates that the exploration mechanism embedded in the CBO policies is not merely a means of model refinement, but can translate directly into improved steady-state economic performance compared with a purely exploitative adaptive optimizer. At the same time, both CEI and UCB-LB remain close to the theoretical optimum over a broad range of demand setpoints, indicating that high-quality load-sharing solutions can be recovered without prior access to compressor performance maps.

Table 9
Steady-state torque allocations selected by CEI and UCB-LB. For the parallel configuration, $\mathbf{T} = [T_1, T_2, T_3]^\top$, whereas for the series configuration, $\mathbf{T} = [T_1, T_2]^\top$. All torque values are in N m.

Configuration	Demand target	CEI torque vector	UCB-LB torque vector
Parallel	2.1 [kg/s]	[19.85, 21.00, 20.78]	[19.68, 21.00, 21.00]
	1.8 [kg/s]	[15.51, 16.61, 15.86]	[15.54, 16.69, 15.73]
	1.9 [kg/s]	[16.33, 18.17, 17.19]	[16.14, 18.74, 16.92]
	1.7 [kg/s]	[14.45, 15.17, 14.84]	[14.84, 15.17, 14.48]
	1.6 [kg/s]	[14.42, 13.43, 13.33]	[14.31, 13.24, 13.60]
	1.5 [kg/s]	[13.89, 11.89, 12.27]	[13.98, 11.56, 12.55]
Series	248 [kPa]	[18.64, 20.10]	[18.70, 19.54]
	244 [kPa]	[14.04, 14.63]	[14.62, 14.79]
	246 [kPa]	[16.58, 17.10]	[16.56, 17.13]
	243 [kPa]	[13.45, 13.91]	[13.51, 13.82]
	242 [kPa]	[12.45, 12.69]	[12.47, 12.73]
	241 [kPa]	[11.23, 11.74]	[11.34, 11.57]

Overall, these results support the view that the proposed framework offers two distinct advantages over conventional supervisory alternatives: first, it can systematically improve upon static equal-load policies by redistributing torque according to learned compressor-specific behavior; second, it can outperform non-exploring two-stage adaptive optimization by actively probing the operating space in a constraint-aware manner, thereby discovering more efficient feasible equilibria.

4.4. Operational safety and surge integrity

The final evaluation criterion concerns the framework's ability to maintain operational safety by avoiding aerodynamic surge within the simulated envelope. As demonstrated in Figs. 6 and 7, the DTS for all compressors remained positive across the entire 30,000 s simulation using both the UCB-LB and CEI strategies for series and parallel

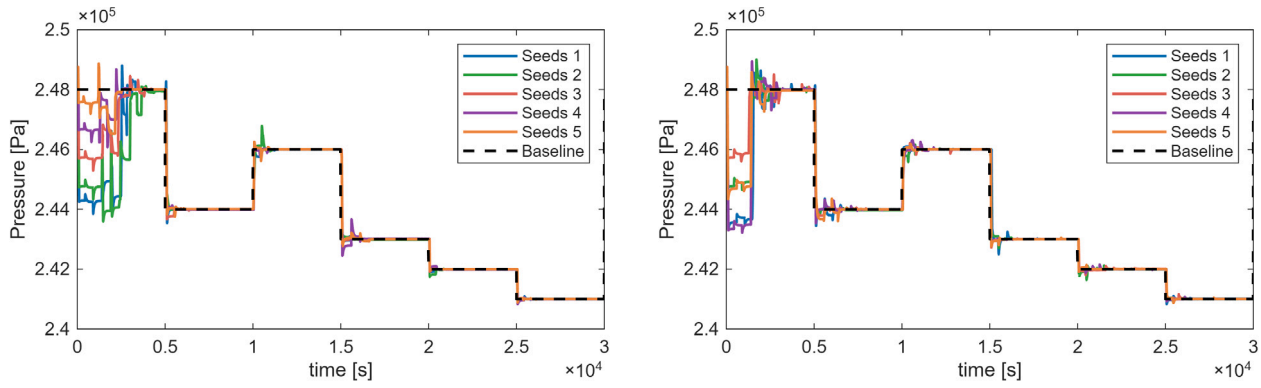


Fig. 4. Pressure-tracking performance for the series configuration: (a) UCB-LB and (b) CEI.

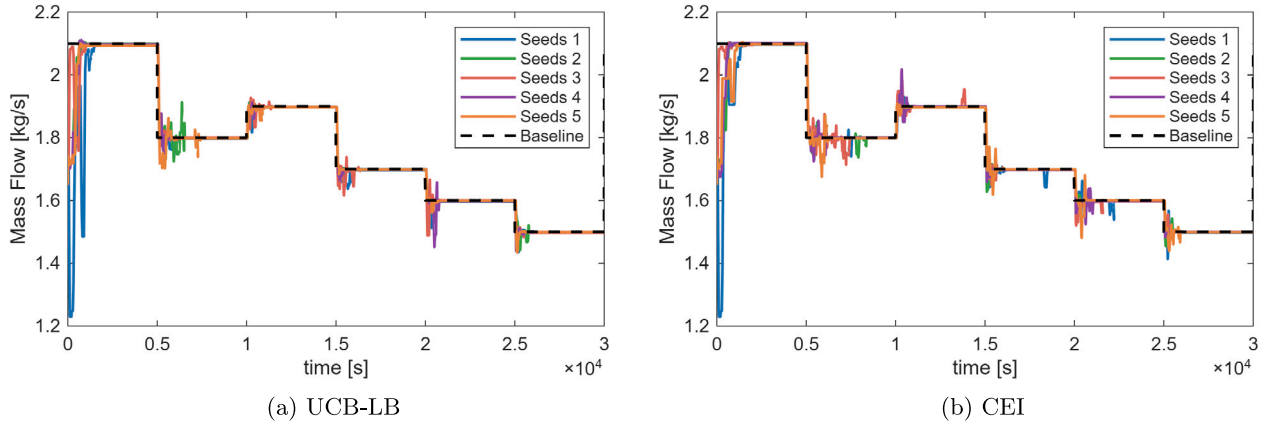


Fig. 5. Mass-flow-tracking performance for the parallel configuration: (a) UCB-LB and (b) CEI.

configurations, respectively. Notably, while our CBO formulation mathematically evaluates safety only at steady-state intervals, this positive safety margin was empirically preserved in our simulations even during rapid setpoint transitions.

The results validate the effectiveness of our surrogate-based safety masks. In the CEI formulation, the PoF successfully avoided torque setpoints that posed a risk of surge. Similarly, the Log-Barrier term in the UCB-LB formulation created a steep numerical penalty that steered the GA away from the stability boundary. This "safety-first" architecture is designed to prioritize hardware reliability alongside the pursuit of energy efficiency, supporting the potential of the proposed framework for industrial application.

Figs. 8 and 9 illustrate the projection of the simulated operating trajectories onto the compressor maps for both the CEI and UCB-LB strategies for the series and parallel scenarios, respectively. The solid black lines, representing the real-time evolution of the torque-driven states, demonstrate that all compressors operate within the stable regime throughout the simulation. By consistently maintaining the operating points to the right of the surge limit line throughout the 30,000-second simulation, the framework demonstrates its capacity to navigate complex efficiency relationships without ever encroaching upon the aerodynamic instability zone.

To assess the contribution of the safety components, we repeated the simulations after removing the PoF terms from CEI and the log-barrier term from UCB-LB, while retaining the same demand sequences and initialization conditions. In the series configuration, both ablated formulations crossed the surge control line in all five initialization runs during the final demand interval, for which the discharge pressure reference was 241 kPa.

4.5. Discussion and practical implications

Implications of the acquisition-function choice. The comparison between the two acquisition strategies highlights a practical trade-off between ease of deployment, conservativeness, exploration behavior, and, to a lesser extent, energy efficiency. From an implementation standpoint, the CEI-based formulation is attractive because it does not require additional weight tuning, whereas the UCB-LB formulation with the log-barrier requires configuration-specific adjustment of the relative weights assigned to power, tracking, and safety. At the same time, the operating behavior induced by the two strategies is noticeably different. In the reported simulations, CEI tended to operate closer to the surge control line at low-demand conditions, reflecting a less conservative use of the feasible region, while the log-barrier formulation maintained a larger margin and therefore behaved more conservatively. This difference is also consistent with the observed sampling patterns: CEI was more exploratory, collecting on average 65 samples in the parallel configuration and 18 samples in the series configuration, compared with 56 and 15 samples, respectively, for the log-barrier formulation. From an energy-efficiency perspective, the difference between the two strategies is present but small, and depends on the compressor configuration. In the parallel case, CEI and UCB-LB achieve very similar steady-state savings overall, with UCB-LB showing a slightly higher average saving across the reported operating points, largely because of its stronger performance at the highest-demand condition, while CEI performs marginally better at several of the remaining setpoints. In the series configuration, by contrast, CEI shows a small but more consistent advantage, yielding a slightly higher average saving and outperforming UCB-LB at most of the reported operating points. This suggests that

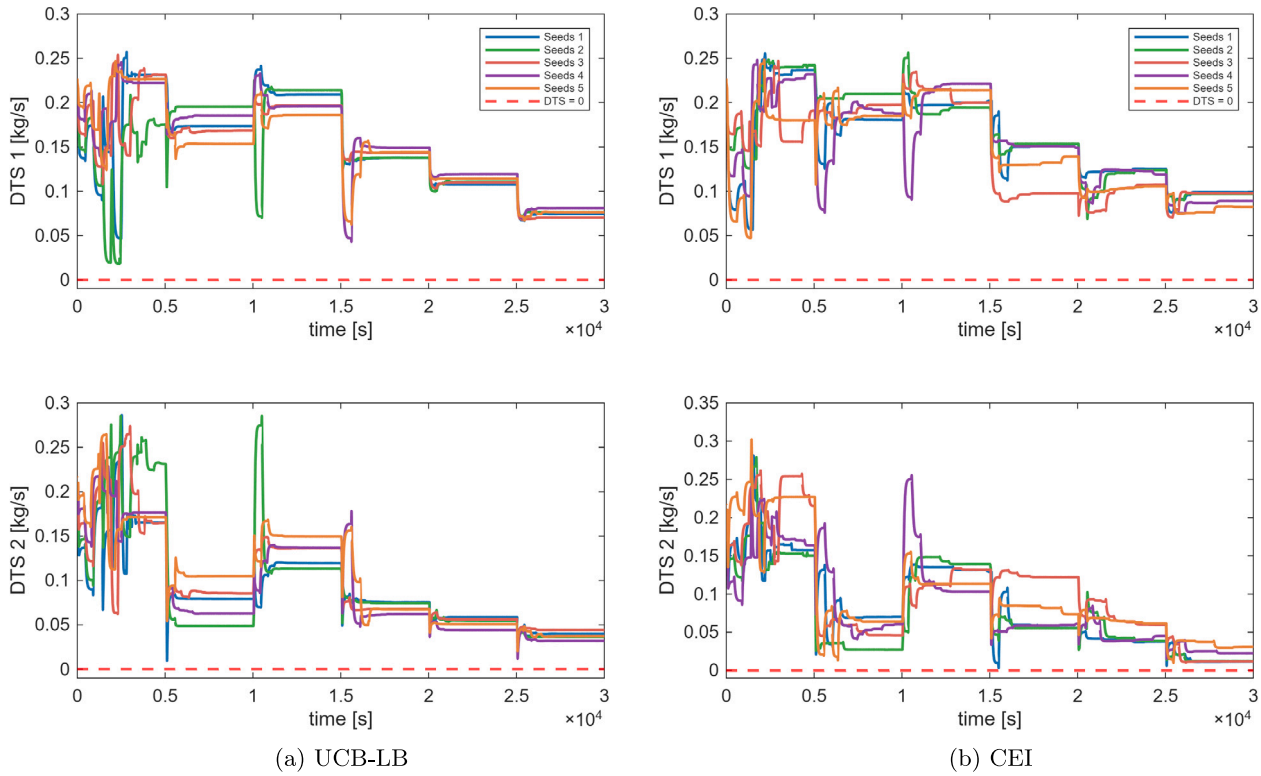


Fig. 6. Distance-to-surge trajectories of the two series compressors across five initialization runs: (a) UCB-LB and (b) CEI.

the more exploratory behavior of CEI can be beneficial when the load-sharing problem is more strongly coupled and the efficient feasible region is harder to identify. There is also no noticeable change in the computational effort when using the different strategies. Overall, the results indicate that the main distinction between the two acquisition functions is not a large systematic gap in attainable energy savings, but rather the trade-off between conservativeness and exploration: CEI offers simpler deployment and slightly stronger economic performance in the series case, whereas UCB-LB provides comparable efficiency with a more conservative operating profile, at the expense of additional tuning effort.

Calibration of the surge-margin surrogates. The conservative interpretation of the learned surge constraints depends on the validity of the GP confidence bounds. In practical deployment, their calibration could be assessed using independent or sequentially collected validation data by comparing measured DTS values with the corresponding predictive intervals and monitoring empirical coverage over the relevant operating region. Persistent undercoverage would indicate that the kernel, noise assumptions, or confidence multiplier require revision. More rigorous approaches include practically evaluable uncertainty bounds for GP regression (Fiedler et al., 2021) and data-driven estimation of the RKHS norm used in constructing GP confidence intervals (Tokmak et al., 2025). Such empirical checks are useful diagnostics, but they do not by themselves establish uniform high-probability coverage in regions with limited data. The present simulation study empirically evaluates the observed surge margins but does not provide a formal calibration guarantee for the GP confidence bounds.

Interpretability of the learned surrogate landscapes. Since the load-sharing optimization problem for the series configuration is two-dimensional ($\mathbf{T} = [T_1, T_2]^\top$), the predicted mean values of the total power consumption, the discharge pressure, and the surge stability boundaries for both compressors can be visualized directly, as shown in Fig. 10. The learned GPs, reconstructed using data collected from one of the simulation runs, reveal the complex inter-dependencies that govern station efficiency and stability.

1. **Nonlinear demand mapping:** The discharge pressure contours confirm a significant nonlinear dependence on the torque vector. While the maximum pressure is achieved near the upper bounds of the torque domain (T_1, T_2), the mid-range demand manifold exhibits a characteristic curvature. In this regime, the sensitivity of the pressure to individual torque adjustments is non-uniform, demonstrating that the framework successfully captures the aerodynamic coupling between the stages without relying on an explicit thermodynamic model.
2. **Antagonistic surge stability:** A key observation is the antagonistic relationship between the surge margins of the two compressors. The stability boundary for the upstream unit (C_1) is primarily reduced when T_1 is relatively high and T_2 is relatively low, whereas the downstream unit (C_2) approaches surge under the opposite torque allocation. This creates a competing constraint landscape in which moving one unit further from instability can bring the other closer to it. At low discharge pressure demands (below 244,000 Pa), the intersection of the safe operating regions constricts into a narrow feasible diagonal, indicating that the system must be governed with high precision to remain within this safety corridor.
3. **Efficiency-stability trade-off:** The power consumption surface P reveals that, for a given pressure setpoint, the global minimum is consistently shifted toward a higher torque allocation for the second stage ($T_2 > T_1$). Crucially, the most energy-efficient operating points in low-demand regimes are located immediately adjacent to the surge stability boundary of the downstream compressor. This near-boundary behavior helps explain why load-sharing optimization is difficult in practice: achieving peak efficiency requires operation close to the stability limit, a task that the proposed CBO performs autonomously through its safety-aware formulation.

Implications of the torque-based modeling approach. A practical advantage of the proposed framework is that the torque-based formulation

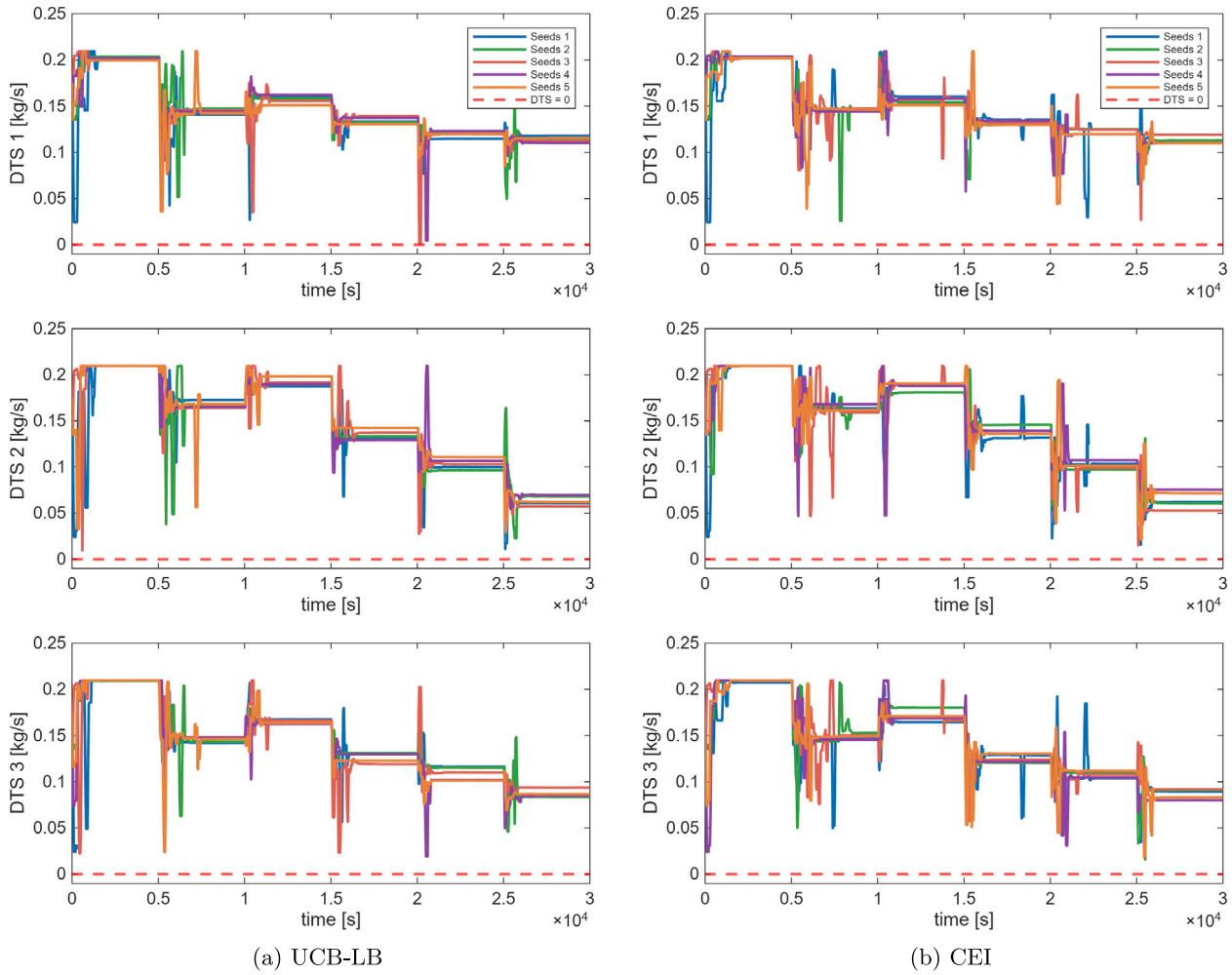
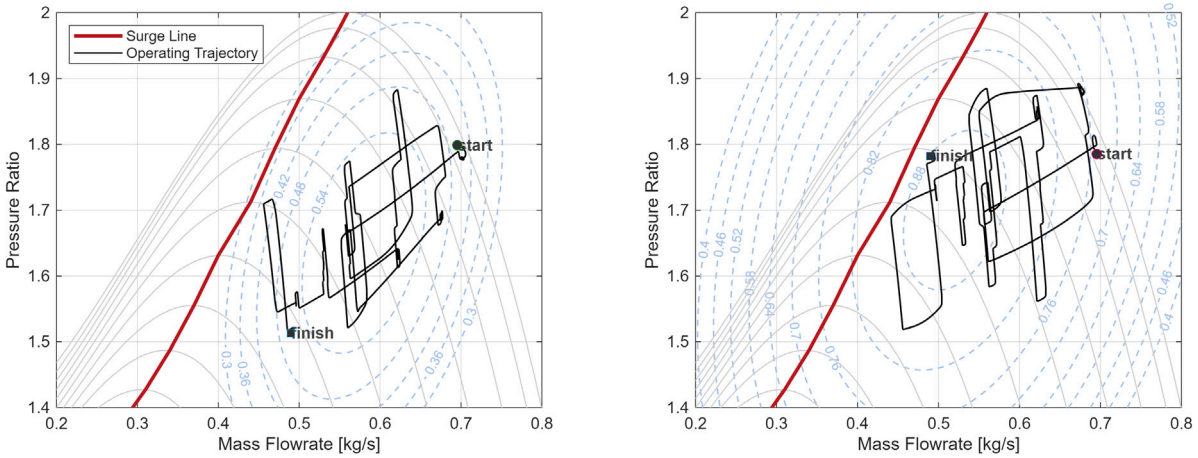


Fig. 7. Distance-to-surge trajectories of the three parallel compressors across five initialization runs: (a) UCB-LB and (b) CEI.

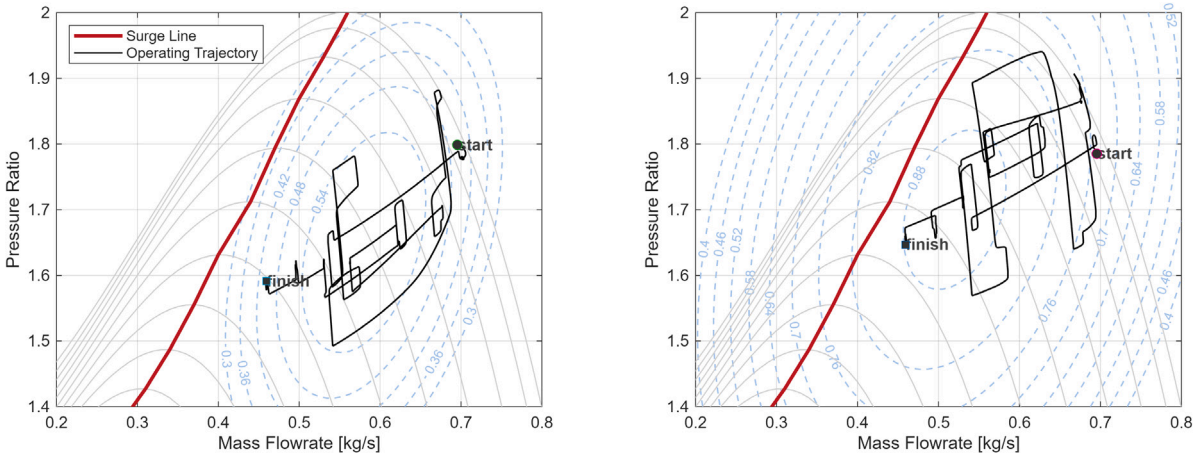
is particularly straightforward to implement in compressor stations equipped with variable-frequency electric drive systems. In such systems, the drive torque, or closely related drive commands, can be adjusted directly through the electrical drive architecture, making the load-sharing optimizer naturally compatible with the available actuation structure. The simulation studies and formulations presented in this work are aligned with this setting and therefore reflect electrically driven compressor stations. For compressor stations driven by gas turbines or steam turbines, the problem becomes more involved because the driver dynamics and performance characteristics introduce an additional layer between the supervisory command and the resulting shaft torque. In such cases, shaft torque is no longer the most direct manipulated variable from an implementation perspective. Nevertheless, the central philosophy of the proposed framework remains applicable. For turbine-driven compressors, variables such as fuel-flow valve position or steam-flow valve position constitute the fundamental degrees of freedom at the driver level. In principle, reformulating the optimization problem in terms of these driver-side inputs, rather than shaft torque directly, should allow a very similar data-driven and constraint-aware optimization architecture to be constructed. This suggests that the proposed framework can be generalized to a broader class of compressor stations provided that the selected decision variables reflect the true actuation interface of the driver system.

Computational load and feasibility of real-time deployment. From a computational standpoint, the proposed load-sharing optimizer is sufficiently lightweight for real-time deployment in the problem settings considered here. The computation time requirement is relative to the interval between successive supervisory updates, since the optimization layer is event-driven and is executed only when a new steady-state operating point is accepted. In the reported simulation scenarios, each 30,000 s closed-loop study, including solution of the dynamic compressor models (as the simulated plant), steady-state detection, GP updates, and repeated acquisition-function optimization, required approximately 60 s of wall-clock time on a 64-bit Windows 11 Enterprise workstation equipped with an Intel® Core™ Ultra 9 185H processor (2.30 GHz) and 32 GB RAM, corresponding to an aggregate execution speed of about 500 times faster than real time. The shortest observed interval between two successive supervisory actions was approximately 50 s. These results therefore indicate that the CBO framework can be executed comfortably within the available decision windows. More broadly, this level of computational demand is consistent with the type of local deployment already reported for advanced optimization and control applications on embedded and edge platforms, including Raspberry Pi hardware and industrial PC-based edge devices (Wei and Calautit, 2024).

Hierarchical control and disturbance rejection. The proposed framework optimizes steady-state efficiency while tracking the prescribed station demand. However, the steady-state detection required for reliable GP updates introduces latency and may limit the framework's response to high-frequency process disturbances. A fast local PID feedback loop could therefore be added to maintain the station demand variable near its reference between updates of the load-sharing optimizer. Such a loop would constitute an additional feedback-control component and was not included in the present CBO simulations. It would complement



(a) CEI



(b) UCB-LB

Fig. 8. Operating trajectories on the compressor maps for one representative series-configuration run: (a) CEI and (b) UCB-LB. In both cases, the compressors remain within the stable region, to the right of the surge control line, throughout the simulation.

the load-sharing optimizer operating at the supervisory layer and the existing local anti-surge protection layer.

Alternative constraint-control architectures. Alternative RTO architectures can place a dedicated constraint-feedback controller between the load-sharing optimizer and the plant and use an admissible constraint setpoint as an optimization variable (Skogestad, 2023). Restricting this setpoint to the safe range can prevent the optimizer from issuing an inadmissible constraint reference without requiring a GP constraint model. This setpoint-level guarantee nevertheless depends on a measurable constraint variable, sufficient actuation, and robust tracking by the constraint controller. In the compressor problem considered here, the torque inputs jointly affect station demand, power consumption, and multiple coupled DTS margins. Implementing active constraint control such as proposed by Krishnamoorthy (2025) would therefore require a different feedback structure and allocation of manipulated variables. The present work retains GP-based constraint modeling together with the separate local anti-surge protection layer, while recognizing active constraint control as a relevant alternative architecture.

Series and parallel compressor arrangements and the operating envelope. A fundamental distinction between the parallel and series configurations lies in the nature of the system resistance curve, that is, the relationship imposed by the surrounding process or pipe network between the delivered flowrate and the pressure rise required from the compressor system. This curve represents the external load against which the compressors must operate, and the feasible operating point is determined by its interaction with the compressor characteristics. In the parallel architecture, each unit faces a static resistance profile dictated by the common suction and discharge header pressures. Consequently, the load-sharing task is primarily a search for the most efficient distribution of mass flow across the individual efficiency maps, as the pressure ratio remains a shared and approximately fixed constraint for all participating units, as shown in Figs. 9(a) and 9(b). In contrast, the series configuration is characterized by dynamic inter-stage coupling, where the resistance curve of the upstream compressor (C_1) is directly modulated by the operating state of the downstream compressor (C_2) and vice versa. Because the discharge of C_1 feeds the suction of C_2, any adjustment in the second-stage torque setpoint shifts the intermediate pressure and thereby alters the back-pressure experienced by the first stage. This interaction creates a floating resistance landscape for both compressors that the CBO framework must navigate, as shown in Figs. 8(a) and 8(b). The optimizer must therefore account for the fact that an efficiency gain in one stage may inadvertently shift the other stage into a higher-resistance or lower-stability regime. The ability of the GP surrogates to learn these nonlinear cross-stage dependencies without an explicit model of the inter-stage volumes is an important advantage. This result indicates that the framework can manage the mutual influence inherent in series compression, a task that traditional equal-torque or equal-speed strategies may fail to optimize effectively.

Changing operating conditions and contextual extension. The present surrogate models use only the torque vector as input because we assume

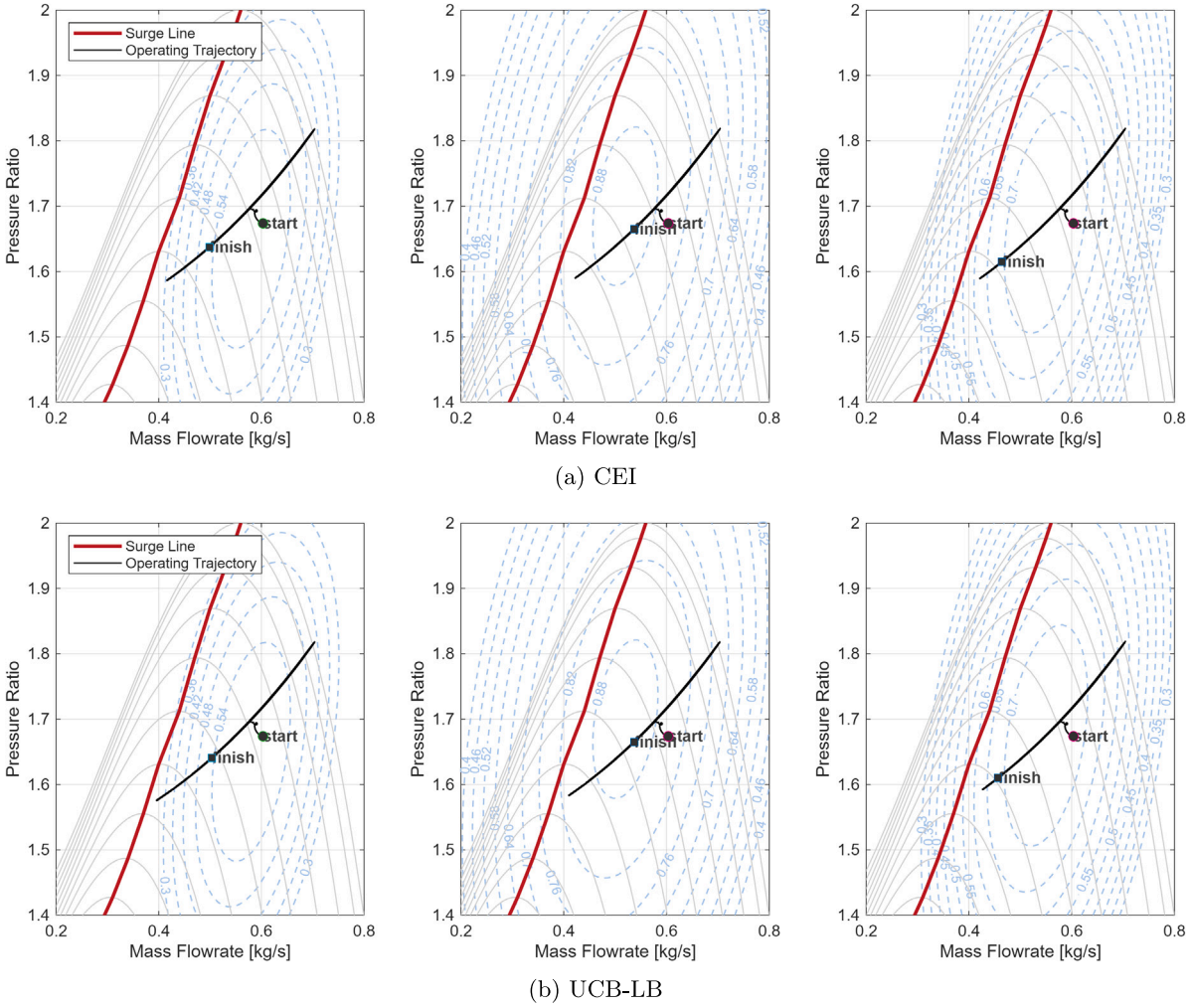


Fig. 9. Operating trajectories on the compressor maps for one representative parallel-configuration run: (a) CEI and (b) UCB-LB. In both cases, all three compressors remain within the stable region, to the right of the surge control line, throughout the simulation.

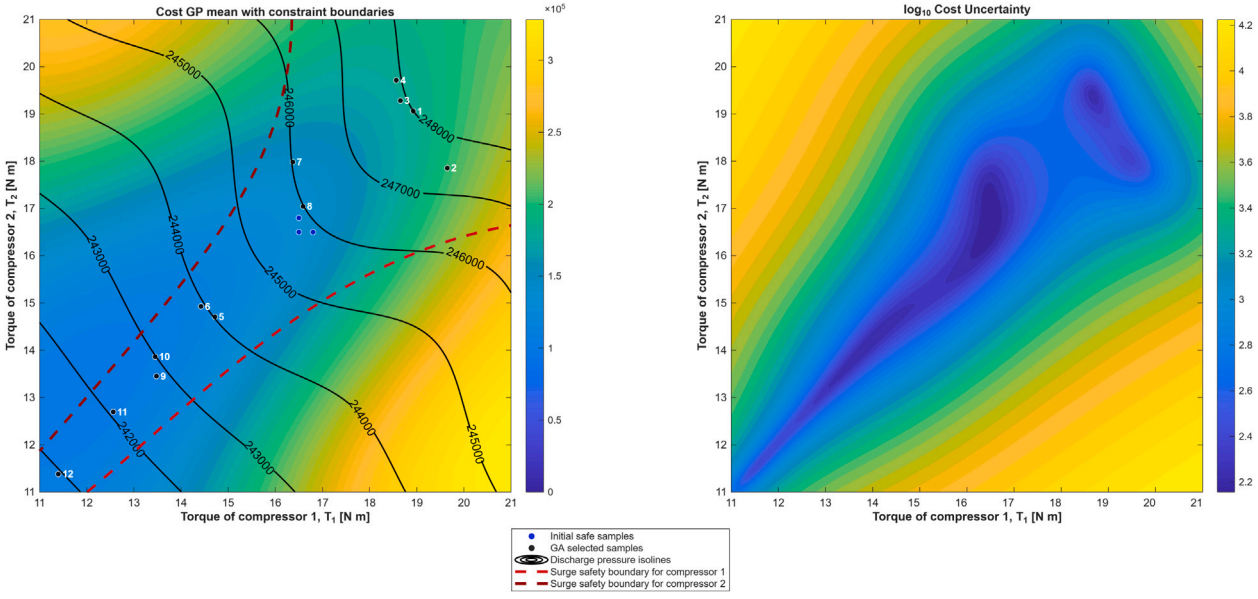


Fig. 10. Learned GP surrogate landscapes for the series configuration (Eq. (25)). Left: GP mean total power shown by the color field, final mean discharge pressure shown by black contour lines, and the mean surge-safety boundaries of the two compressors shown by red dashed lines. The initial safe samples (blue dots) and subsequent GA-selected samples (numbered black dots) are also indicated. Right: logarithm of the GP predictive uncertainty for total power.

fixed external operating conditions in this study. Changes in the station demand reference are accommodated directly through h_{ref} in the acquisition function. However, changes in other operating conditions, particularly the suction and discharge interface pressures, would alter the compressor operating points and the learned mappings for power, demand, and DTS. To use observations collected under different external operating conditions, the proposed methodology must be extended to represent those conditions explicitly and prevent degradation. This can be addressed using contextual BO, which has been formulated for online control problems with black-box objectives and constraints affected by observed exogenous disturbances (Xu et al., 2023).

Specifically, the GP inputs can be augmented with a vector of measured, non-manipulated context variables **c**, leading to surrogates of the form:

$$f(\mathbf{T}, \mathbf{c}), \qquad f \in \{P, h, S_1, \ldots, S_{n_c}\}.$$

The context could include interface pressures, inlet conditions, and other measured disturbances. The GP-UCB framework can be extended directly to incorporate such information by modeling the response over the joint decision–context space and evaluating the confidence bound conditional on the observed context (Krause and Ong, 2011). For the EI-based formulation, additional care is required because the incumbent must be defined for the current context rather than as a single global best observation. A contextual EI construction can instead measure improvement relative to the predicted optimum for the newly observed context (Park, 2020).

Gradual changes in compressor characteristics could also be addressed using a temporal context, sliding-window updates, or recency-weighted observations (Li et al., 2024). Changes in compressor availability are outside the continuous-operation scope considered here and would require reconfiguration of the active decision variables and surrogate models.

5. Conclusions

This work presented a unified data-driven framework for real-time load-sharing optimization of compressor stations in both parallel and series configurations. The load-sharing optimizer operates at the supervisory layer and treats the station as a black-box steady-state mapping from shaft torque inputs to measured performance variables, thereby avoiding explicit use of compressor maps within the optimization layer. Constrained Bayesian optimization was used to learn the economic and surge-related constraint structure directly from plant-response data and to identify improved load-sharing decisions under demand-tracking and operability requirements.

The simulation studies showed that the proposed framework can consistently improve load allocation relative to equal-load operation across different initialization scenarios and for both compressor-station architectures considered. In the reported cases, the method achieved energy savings of up to 2.6% for the parallel configuration and 3.1% for the series configuration, while maintaining safe steady-state surge margins and without observed surge control line crossings or anti-surge controller activation. These results indicate that measured DTS information can be used effectively as a safety-relevant feedback signal within a data-driven load-sharing optimization framework.

The findings of the study motivate several directions for future work. First, extending the framework to account more explicitly for transient behavior would strengthen its applicability in situations with frequent disturbances or rapid operating changes. Second, scaling the method to larger hybrid compressor networks will require more computationally efficient surrogate and acquisition-optimization strategies, with sparse GP formulations representing one promising direction (Liu et al., 2020). Third, future work should consider contextual BO formulations in which measured non-manipulated operating conditions, particularly the upstream and downstream station interface pressures, are included as GP inputs.

Finally, long-term deployment will additionally require adaptation mechanisms as compressor characteristics and operability boundaries evolve. Possible approaches include sliding-window updates, recency-weighted regression, and forgetting-factor-based sparse GP schemes (Gómez-Verdejo et al., 2024). These methods would reduce the influence of stale data and keep the surrogate models representative of current station behavior.

CRediT authorship contribution statement

Liqiu Dong: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mohamed Becha:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Johannes Köhler:** Writing – review & editing, Supervision, Methodology. **Marta Zagorowska:** Writing – review & editing, Methodology. **Mehmet Mercangöz:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used OpenAI's ChatGPT to assist with editorial checks of clarity, consistency, and completeness. OpenAI's ChatGPT was also used to generate some of the icons and illustrations used in the graphical abstract. All authors thoroughly reviewed and edited the manuscript and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Simulation models and compressor surge characterization

This appendix summarizes the dynamic compressor model and the surge-related definitions used to construct the simulation environments for the parallel and series case studies. These relations underpin the numerical experiments but are collected here to keep the main text focused on the load-sharing problem and its solution by the load-sharing optimizer operating at the supervisory layer.

A.1. Single compressor model

The study adopts the centrifugal compressor model established by Greitzer (1976) and subsequently refined by Gravdahl and Egeland (2012), which was previously validated by our research team in Cortinovis et al. (2015). As illustrated in Fig. 11, the system consists of a centrifugal compressor connected to suction, discharge, and recycle lines. The recycle line connects the outlet back to the inlet and provides the actuation path used by the anti-surge protection system.

The dynamic model is given by:

$$\dot{p}_1 = C_1\left(m_{\mathrm{in}} + m_{\mathrm{rec}} - m_{\mathrm{comp}}\right), \tag{31}$$

$$\dot{p}_2 = C_2\left(m_{\mathrm{comp}} - m_{\mathrm{rec}} - m_{\mathrm{out}}\right), \tag{32}$$

$$\dot{m}_{\mathrm{comp}} = C_3\left(\Pi\, p_1 - p_2\right), \tag{33}$$

$$\dot{\Pi} = C_4\left(\Pi_{\mathrm{ss}} - \Pi\right), \tag{34}$$

$$\dot{\omega} = C_5\left(\tau - \tau_{\mathrm{comp}}\right), \tag{35}$$

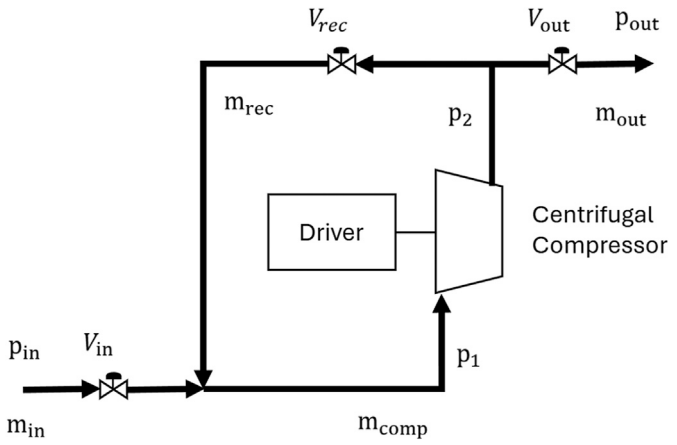


Fig. 11. Schematic of a single centrifugal compressor with suction, discharge, and recycle paths.

$$\dot{m}_{\mathrm{rec}} = C_6\left(m_{\mathrm{rec,ss}} - m_{\mathrm{rec}}\right), \tag{36}$$

where m_{in}, m_{out}, m_{comp}, and m_{rec} denote the inlet, outlet, compressor, and recycle mass flowrates, respectively; p_1 and p_2 are the suction and discharge pressures; Π is the compressor pressure ratio; Π_{ss} is its steady-state value; ω is the shaft speed; τ is the applied shaft torque; τ_{comp} is the torque required for compression; and C_i, $i = 1, \ldots, 6$, are constant model parameters.

The inlet, outlet, and recycle flows are computed from pressure-drop relationships:

$$m_{\mathrm{in}} = K_{\mathrm{in}} v_{\mathrm{in}} \sqrt{|p_{\mathrm{in}} - p_1|}, \tag{37}$$

$$m_{\mathrm{out}} = K_{\mathrm{out}} v_{\mathrm{out}} \sqrt{|p_2 - p_{\mathrm{out}}|}, \tag{38}$$

$$m_{\mathrm{rec,ss}} = K_{\mathrm{rec}} v_{\mathrm{rec}} \sqrt{|p_2 - p_1|}, \tag{39}$$

$$\tau_{\mathrm{comp}} = \sigma_{\mathrm{slip}} r^2 \omega \, |m_{\mathrm{comp}}|, \tag{40}$$

where K_{in}, K_{out}, and K_{rec} are the valve gains; v_{in}, v_{out}, and v_{rec} are the corresponding valve openings; and p_{in} and p_{out} are the boundary pressures.

The shaft power is computed as:

$$P = \frac{y_p}{\eta} m_{\mathrm{comp}}, \tag{41}$$

where y_p is the polytropic head:

$$y_p = \frac{Z_{\mathrm{in}} R T_{\mathrm{in}}}{M_{\mathrm{W}}} \frac{\kappa}{\kappa - 1} \left(\Pi^{\frac{\kappa-1}{\kappa}} - 1 \right), \tag{42}$$

and η is the compressor efficiency. In (42), Z_{in} is the inlet compressibility factor, R is the gas constant, T_{in} is the inlet temperature, M_{W} is the molecular weight of the gas mixture, and κ is the polytropic exponent.

The steady-state pressure ratio is obtained from a polynomial compressor map:

$$\Pi_{\mathrm{ss}} = \sum_{i=0}^{3} \sum_{j=0}^{3} \alpha_{i,j} \omega^i m_{\mathrm{comp}}^j. \tag{43}$$

The isentropic efficiency is approximated as a quadratic function of mass flowrate and pressure ratio around the reference point $(m_{\mathrm{comp},0}, \Pi_0)$:

$$\eta = \eta_0 - \left[\gamma_1 (m_{\mathrm{comp}} - m_{\mathrm{comp},0})^2 - \gamma_2 (m_{\mathrm{comp}} - m_{\mathrm{comp},0})(\Pi - \Pi_0) + \gamma_3 (\Pi - \Pi_0)^2 \right]. \tag{44}$$

These relations define the dynamic simulation model from which the steady-state operating points used in the optimization problems of Section 2.1 are obtained.

A.2. Intermediate pressure dynamics in the series configuration

For the series compressor configuration, the pressure between the two compressors is not imposed as a fixed boundary condition. Instead, it is determined dynamically from a mass balance over the intermediate control volume connecting the discharge of the first compressor to the suction of the second compressor. This intermediate section represents the gas holdup between the two stages, including the intercooler outlet and connecting pipework.

Let p_{int} denote the intermediate pressure, V_{int} the effective gas volume of the intermediate section, T_{int} the corresponding gas temperature, and m_{int} the gas mass contained in this control volume. Assuming ideal-gas behavior, the intermediate pressure is related to the contained gas mass through:

$$p_{\mathrm{int}} V_{\mathrm{int}} = m_{\mathrm{int}} \frac{R}{M_{\mathrm{W}}} T_{\mathrm{int}}, \tag{45}$$

where R is the universal gas constant and M_{W} is the molecular weight of the gas mixture.

The mass balance over the intermediate volume is written as:

$$\frac{d m_{\mathrm{int}}}{dt} = m_{\mathrm{out},1} - m_{\mathrm{in},2}, \tag{46}$$

where $m_{\mathrm{out},1}$ is the discharge mass flowrate of the first compressor and $m_{\mathrm{in},2}$ is the inlet mass flowrate of the second compressor.

Assuming that the effective intermediate volume V_{int} and temperature T_{int} remain constant (we considered an intercooler between the compressor stages which cools the gas from the first stage to a fixed temperature), differentiation of (45) together with (46) gives:

$$\dot{p}_{\mathrm{int}} = \frac{R T_{\mathrm{int}}}{M_{\mathrm{W}} V_{\mathrm{int}}} \left(m_{\mathrm{out},1} - m_{\mathrm{in},2} \right). \tag{47}$$

This dynamic relation is used to determine the floating intermediate pressure in the series configuration. Consequently, the suction pressure of the second compressor is not fixed externally, but evolves according to the imbalance between the upstream discharge flow and downstream intake flow. This coupling is the main source of interaction between the two compressors in the series case.

A.3. Compressor map and surge characterization

The operational envelope of a centrifugal compressor is commonly represented through its compressor map, which relates pressure ratio, mass flowrate, and rotational speed. In addition to efficiency contours, this map identifies the feasible operating region bounded by mechanical and aerodynamic limits. Of particular importance is the surge boundary, which defines the onset of unstable operation at low flowrates.

Surge is an aerodynamic instability that occurs when the compressor can no longer sustain forward flow against the imposed pressure rise. Under such conditions, the flow may decelerate, stall, and eventually reverse, producing oscillatory behavior that can evolve into a damaging limit cycle. Repeated surge events can cause severe mechanical vibration, thermal stress, seal degradation, and accelerated equipment wear (Du et al., 2025; Dixon and Hall, 2013).

To avoid operation too close to the physical surge boundary, compressor control practice employs a Surge Control Line (SCL), located to the safe side of the surge limit by means of an additional design margin. The operating margin relative to this boundary is often quantified through a DTS indicator. In the present work, the surge constraint for compressor i is represented generically through a signed safety function:

$$S_i(m_{\mathrm{comp},i}, \Pi_i, \omega_i) \geq 0, \tag{48}$$

where positive values indicate safe operation and negative values indicate incursion into the unsafe region.

For the simulations used in this study, the surge boundary can be expressed in reduced form as:

$$m_{\text{surge},i}(\Pi_i) = s_{0,i} + s_{1,i}\Pi_i, \tag{49}$$

and the corresponding surge margin is defined as:

$$S_i = m_{\text{comp},i} - m_{\text{surge},i}(\Pi_i). \tag{50}$$

Hence, the constraint $S_i \geq 0$ used in the main-body optimization problems enforces operation to the safe side of the prescribed surge control line. In the dynamic implementation, this quantity is monitored continuously by the anti-surge protection layer and is also provided to the load-sharing optimizer as an operability indicator.

A.4. Anti-surge controller

Each compressor is equipped with an anti-surge controller (ASC) acting on the recycle valve. The controlled variable is defined as the signed distance between the current operating point and the surge control line (SCL). For compressor i, this quantity is written as:

$$e_{\text{ASC},i} = S_i(m_{\text{comp},i}, \Pi_i, \omega_i), \tag{51}$$

where $S_i(\cdot)$ denotes the DTS metric. By construction, $e_{\text{ASC},i} > 0$ indicates that the compressor is operating to the safe side of the SCL, $e_{\text{ASC},i} = 0$ corresponds to operation on the SCL, and $e_{\text{ASC},i} < 0$ indicates that the operating point has entered the anti-surge protection region.

The recycle valve opening fraction, denoted by $u_{\text{ASV},i} \in [0, 1]$, is manipulated so as to increase recycle flow whenever the compressor approaches surge. The controller is activated only when the surge margin becomes negative. In normal operation, when $e_{\text{ASC},i} > 0$, the anti-surge valve remains fully shut and the controller integral state is held in anti-windup mode in order to prevent integrator accumulation at the lower saturation bound. When $e_{\text{ASC},i} < 0$, the controller opens the recycle valve, thereby increasing the compressor flow and driving the operating point back toward the SCL. A convenient implementation is:

$$u_{\text{ASV},i} = \text{sat}_{[0,1]}\left(K_{p,i}(-e_{\text{ASC},i}) + K_{I,i}z_i\right), \tag{52}$$

$$\dot{z}_i = \begin{cases} 0, & e_{\text{ASC},i} > 0 \text{ and } u_{\text{ASV},i} = 0, \\ -e_{\text{ASC},i}, & \text{otherwise}, \end{cases} \tag{53}$$

where $K_{p,i}$ and $K_{I,i}$ are the proportional and integral gains, respectively, z_i is the integral state, and $\text{sat}_{[0,1]}(\cdot)$ denotes saturation to the admissible valve range.

Once the operating point returns to the safe side of the SCL, the valve is closed gradually rather than instantaneously. Depending on the operating conditions, the closed-loop system may either settle on the SCL with a partially open recycle valve, corresponding to $e_{\text{ASC},i} \approx 0$, or recover to a point further to the right of the SCL, where sufficient positive surge margin has been re-established and the anti-surge valve can return to its fully closed position. In practice, the latter behavior is generally preferred because it minimizes unnecessary recycle and therefore reduces the efficiency penalty associated with prolonged anti-surge action. Additional feed-forward control formulations to force the recycle valve to quickly open when the operating point is approaching the surge or the SCL are also used in practice but are not implemented in our simulation environment (Amin et al., 2021).

Appendix B. Offline model-based reference optimum as an oracle

To benchmark the proposed CBO framework against a model-based steady-state optimum, an offline reference solution was computed using the full mechanistic simulation model given in Appendices A.1 and A.2. This benchmark is referred to as the *oracle* solution because it assumes access to the true steady-state model, including the underlying compressor characteristics and feasibility constraints, in contrast to the proposed load-sharing optimizer, which relies only on measured plant-response data.

For each prescribed station demand target, the oracle solution was defined as the feasible steady-state operating point that minimizes the total compressor power while satisfying the corresponding demand requirement and all imposed operability constraints. In the parallel configuration, the demand variable is the total outlet mass flowrate, whereas in the series configuration it is the final discharge pressure.

The oracle benchmark was obtained numerically by parameterizing the admissible steady-state operating space with a set of candidate operating-point variables and evaluating a dense collection of such points using the full mechanistic model. For each demand target, the search used a uniform 500 × 500 grid, corresponding to 250,000 candidate combinations before feasibility filtering. For the series configuration, the pressure ratios of the two stages, Π_1 and Π_2, were independently discretized over $[1.0, 1.8]$ using 500 values each, giving an interval of approximately 1.60×10^{-3}. For the parallel configuration, the branch flowrates m_1 and m_2 were independently discretized over $[0.01, 2.2]$ kg s^{-1} using 500 values each, giving an interval of approximately 4.39×10^{-3} kg s^{-1}. The third branch flowrate was determined from the demand equality, $m_3 = M_{\text{ref}} - m_1 - m_2$. For each grid point, the associated steady-state station variables were recovered by solving the corresponding steady-state closure relations of the compressor network. The resulting operating point was then evaluated in terms of total power consumption, demand satisfaction, surge-margin indicators, and actuator-related feasibility conditions. Candidates that violated the demand specification or any feasibility constraint were discarded. Among the remaining feasible candidates, the one with minimum total power was retained as the oracle solution for that demand level.

In addition to this oracle benchmark, an equal-load baseline was also computed using the same mechanistic model and feasibility checks, but with a symmetric load-sharing structure imposed across the compressors. This provides a consistent comparison between the proposed data-driven optimizer, a conventional equal-load policy, and a model-based reference optimum.

It is important to emphasize that the oracle solution is an *offline* benchmark and not an implementable real-time strategy within the assumptions of the proposed framework. It relies on direct access to the true mechanistic steady-state model, which is intentionally not supplied to the CBO-based load-sharing optimizer. The oracle therefore serves as a best-case model-based reference for assessing the performance of the proposed data-driven approach.

Data availability

The MATLAB/Simulink simulation models used for this study as well as the MATLAB scripts for the real-time constrained Bayesian optimization implementations are publicly available at https://github.com/AISL-at-Imperial-College-London/CBO_comp.

References

Ahmed, Akhil, Zagorowska, Marta, Del Rio-Chanona, Ehecatl Antonio, Mercangöz, Mehmet, 2022. Application of Gaussian processes to online approximation of compressor maps for load-sharing in a compressor station. In: 2022 European Control Conference. ECC, pp. 205–212. http://dx.doi.org/10.23919/ECC55457.2022.9838042.

Amin, Arslan Ahmed, Maqsood, Muhammad Taimoor, ul Hasan, Khalid Mahmood, 2021. Surge protection of centrifugal compressors using advanced anti-surge control system. Meas. Control. 54 (5–6), 967–982. http://dx.doi.org/10.1177/0020294020983372.

Amini, Sasan, Vannieuwenhuyse, Inneke, Morales-Hernández, Alejandro, 2025. Constrained Bayesian optimization: A review. IEEE Access 13, 1581–1593. http://dx.doi.org/10.1109/ACCESS.2024.3522876.

Berkenkamp, Felix, Krause, Andreas, Schoellig, Angela P., 2023. Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics. Mach. Learn. 112 (10), 3713–3747. http://dx.doi.org/10.1007/s10994-021-06019-1.

Borraz-Sánchez, C., Haugland, D., 2013. Optimization methods for pipeline transportation of natural gas with variable specific gravity and compressibility. Trans. Oper. Res. 21, 524–541. http://dx.doi.org/10.1007/s11750-011-0210-z.

Boyce, Meherwan P., 2003. Centrifugal Compressors: a Basic Guide. PennWell Books.

Chaudhari, Subodh, Abbas, Ahmad, Botts, Alex, Sundaramoorthy, Senthil, Wenning, Thomas, 2023. Load sharing energy savings methodology for systems with multiple centrifugal compressors. J. Clean. Prod. 433, 139630. http://dx.doi.org/10.1016/j.jclepro.2023.139630, URL https://www.sciencedirect.com/science/article/pii/S0959652623037885.

Chowdhury, Sayak Ray, Gopalan, Aditya, 2017. On kernelized multi-armed bandits. In: Precup, Doina, Teh, Yee Whye (Eds.), Proceedings of the 34th International Conference on Machine Learning. In: Proceedings of Machine Learning Research, vol. 70, PMLR, pp. 844–853, URL https://proceedings.mlr.press/v70/chowdhury17a.html.

Cortinovis, A., Ferreau, H.J., Lewandowski, D., Mercangöz, M., 2015. Experimental evaluation of MPC-based anti-surge and process control for electric driven centrifugal gas compressors. J. Process Control 34, 13–25. http://dx.doi.org/10.1016/j.jprocont.2015.07.001, URL https://www.sciencedirect.com/science/article/pii/S0959152415001390.

Cortinovis, Andrea, Mercangöz, Mehmet, Zovadelli, Matteo, Pareschi, Diego, De Marco, Antonio, Bittanti, Sergio, 2016. Online performance tracking and load sharing optimization for parallel operation of gas compressors. Comput. Chem. Eng. 88, 145–156. http://dx.doi.org/10.1016/j.compchemeng.2016.01.012, URL https://www.sciencedirect.com/science/article/pii/S0098135416300035.

de Marco, Fabio Capelassi Gavazzi, Elias, Gustavo Passos, 2011. Fuel consumption model on natural gas compression stations driven by two-shaft gas turbine. In: Proceedings of the PSIG Annual Meeting. In: PSIG Annual Meeting, pp. PSIG–1107, arXiv:https://onepetro.org/PSIGAM/proceedings-pdf/PSIG11/PSIG11/PSIG-1107/1663531/psig-1107.pdf.

Desimpel, Stefan, Dorbec, Matthieu, Van Geem, Kevin M., Stevens, Christian V., 2026. Bayesian optimization for chemical reactions. Chem. Soc. Rev. 55, 2731–2775. http://dx.doi.org/10.1039/D5CS00962F.

Dixon, S. Larry, Hall, Cesare, 2013. Fluid Mechanics and Thermodynamics of Turbomachinery. Butterworth-Heinemann.

Du, Jinshi, Zhang, Yu, Martínez García, Miguel, Spencer, Adrian, 2025. Review on compressor surge monitoring, modeling and anti-surge control approaches. J. Dyn. Monit. Diagn. 4 (4), 292–307. http://dx.doi.org/10.37965/jdmd.2025.889, URL https://ojs.istp-press.com/dmd/article/view/889.

Fiedler, Christian, Scherer, Carsten W., Trimpe, Sebastian, 2021. Practical and rigorous uncertainty bounds for Gaussian process regression. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 35, pp. 7439–7447.

Frazier, Peter I., 2018. A tutorial on Bayesian optimization. URL https://arxiv.org/abs/1807.02811.

Gardner, Jacob, Kusner, Matt, Xu, Zhixiang, Weinberger, Kilian, Cunningham, John, 2014. Bayesian optimization with inequality constraints. In: Xing, Eric P., Jebara, Tony (Eds.), Proceedings of the 31st International Conference on Machine Learning. In: Proceedings of Machine Learning Research, vol. 32, PMLR, Bejing, China, pp. 937–945, URL https://proceedings.mlr.press/v32/gardner14.html.

Gómez-Verdejo, Vanessa, Parrado-Hernández, Emilio, Martínez-Ramón, Manel, 2024. Adaptive sparse Gaussian process. IEEE Trans. Neural Networks Learn. Syst. 35 (11), 16383–16395. http://dx.doi.org/10.1109/TNNLS.2023.3294089.

Gravdahl, Jan Tommy, Egeland, Olav, 2012. Compressor Surge and Rotating Stall: Modeling and Control. Springer Science & Business Media.

Greitzer, E.M., 1976. Surge and rotating stall in axial flow compressors—Part I: Theoretical compression system model. J. Eng. Power 98 (2), 190–198. http://dx.doi.org/10.1115/1.3446138, arXiv:https://asmedigitalcollection.asme.org/gasturbinespower/article-pdf/98/2/190/5886245/190_1.pdf.

Jones, Donald R., Schonlau, Matthias, Welch, William J., 1998. Efficient global optimization of expensive black-box functions. J. Global Optim. 13 (4), 455–492.

Kapp, Sean, Wang, Chunxu, McNelly, Megan, Romeiko, Xiabo, Choi, Jun-Ki, 2024. A comprehensive analysis of the energy, economic, and environmental impacts of industrial variable frequency drives. J. Clean. Prod. 434, 140474. http://dx.doi.org/10.1016/j.jclepro.2023.140474, URL https://www.sciencedirect.com/science/article/pii/S0959652623046322.

Krause, Andreas, Hübotter, Jonas, 2025. Probabilistic artificial intelligence. URL https://arxiv.org/abs/2502.05244.

Krause, Andreas, Ong, Cheng, 2011. Contextual Gaussian process bandit optimization. In: Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K. (Eds.), In: Advances in Neural Information Processing Systems, vol. 24, Curran Associates, Inc., URL https://proceedings.neurips.cc/paper_files/paper/2011/file/f3f1b7fc5a8779a9e618e1f23a7b7860-Paper.pdf.

Krishnamoorthy, Dinesh, 2025. ECCBO: An inherently safe Bayesian optimization with embedded constraint control for real-time process optimization. J. Process Control 152, 103467. http://dx.doi.org/10.1016/j.jprocont.2025.103467, URL https://www.sciencedirect.com/science/article/pii/S0959152425000952.

Krishnamoorthy, Dinesh, Doyle III, Francis J., 2023. Model-free real-time optimization of process systems using safe Bayesian optimization. AIChE J. 69 (4), e17993.

Kudva, Akshay, Sorourifar, Farshud, Paulson, Joel A., 2022. Constrained robust Bayesian optimization of expensive noisy black-box functions with guaranteed regret bounds. AIChE J. 68 (12), e17857. http://dx.doi.org/10.1002/aic.17857.

Kumar, Sant, Cortinovis, Andrea, 2017. Load sharing optimization for parallel and serial compressor stations. In: 2017 IEEE Conference on Control Technology and Applications. CCTA, pp. 499–504. http://dx.doi.org/10.1109/CCTA.2017.8062511.

Kurz, Rainer, Brun, Klaus, 2010. Assessment of compressors in gas storage applications. J. Eng. Gas Turbines Power 132 (6), 062402. http://dx.doi.org/10.1115/1.4000147.

Kurz, Rainer, Brun, Klaus, 2012. Fouling mechanisms in axial compressors. J. Eng. Gas Turbines Power 134 (3), 032401. http://dx.doi.org/10.1115/1.4004403.

Kurz, Rainer, Lubomirsky, Matt, Brun, Klaus, 2012. Gas compressor station economic optimization. Int. J. Rotating Mach. 2012 (1), 715017.

Kurz, Rainer, Mistry, Jay, McBain, Marybeth, Lubomirsky, Matt, 2023. Compressor applications in the decarbonization discussion. IEEE Trans. Ind. Appl. 59 (6), 7929–7939. http://dx.doi.org/10.1109/TIA.2023.3299890.

Lassahn, Gerrit, Bok, Felix, Soelch, Maximilian, Bayer, Justin, Knorn, Steffi, 2026. Two-stage optimal control of large-scale industrial pressure network. Control Eng. Pract. 172, 106925. http://dx.doi.org/10.1016/j.conengprac.2026.106925, URL https://www.sciencedirect.com/science/article/pii/S0967066126001693.

Li, Xia, Cui, Tao, Huang, Kun, Ma, Xin, 2021. Optimization of load sharing for parallel compressors using a novel hybrid intelligent algorithm. Energy Sci. Eng. 9 (3), 330–342. http://dx.doi.org/10.1002/ese3.821, URL https://scijournals.onlinelibrary.wiley.com/doi/abs/10.1002/ese3.821.

Li, Jialin, Zagorowska, Marta, De Pasquale, Giulia, Rupenyan, Alisa, Lygeros, John, 2024. Safe time-varying optimization based on Gaussian processes with spatio-temporal kernel. In: Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., Zhang, C. (Eds.), In: Advances in Neural Information Processing Systems, vol. 37, Curran Associates, Inc., pp. 95326–95355. http://dx.doi.org/10.52202/079017-3020, URL https://proceedings.neurips.cc/paper_files/paper/2024/file/ad30b5f8274286c4ee12de843f03a07f-Paper-Conference.pdf.

Liu, Haitao, Ong, Yew-Soon, Shen, Xiaobo, Cai, Jianfei, 2020. When Gaussian process meets big data: A review of scalable GPs. IEEE Trans. Neural Networks Learn. Syst. 31 (11), 4405–4423. http://dx.doi.org/10.1109/TNNLS.2019.2957109.

Liu, Shengchun, Wu, Sicheng, Hu, Yukun, Li, Hailong, 2019. Comparative analysis of air and CO2 as working fluids for compressed and liquefied gas energy storage technologies. Energy Convers. Manage. 181, 608–620. http://dx.doi.org/10.1016/j.enconman.2018.12.031, URL https://www.sciencedirect.com/science/article/pii/S0196890418313669.

Maathuis, Hauke F., De Breuker, Roeland, Castro, Saullo G.P., 2025. Scaling Bayesian optimization for high-dimensional and large-scale constrained spaces. AIAA J. 63 (11), 4947–4957. http://dx.doi.org/10.2514/1.J065252.

Marler, R. Timothy, Arora, Jasbir S., 2004. Survey of multi-objective optimization methods for engineering. Struct. Multidiscip. Optim. 26 (6), 369–395.

Milosavljevic, Predrag, Marchetti, Alejandro G., Cortinovis, Andrea, Faulwasser, Timm, Mercangöz, Mehmet, Bonvin, Dominique, 2020. Real-time optimization of load sharing for gas compressors in the presence of uncertainty. Appl. Energy 272, 114883. http://dx.doi.org/10.1016/j.apenergy.2020.114883, URL https://www.sciencedirect.com/science/article/pii/S0306261920303950.

Mitchell, Melanie, 1999. A Introduction to Genetic Algorithms. The MIT Press.

Mokhatab, Saeid, Mak, John Y., Valappil, Jaleel V., Wood, David A., 2013. Handbook of Liquefied Natural Gas. Gulf Professional Publishing.

National Gas Transmission, 2025. Annual Environmental Report 2024/25. Regulatory Report, National Gas, URL https://www.nationalgas.com/responsibility/environment.

Park, Jinkyoo, 2020. Contextual Bayesian optimization with trust region (CBOTR) and its application to cooperative wind farm control in region 2. Sustain. Energy Technol. Assessments 38, 100679. http://dx.doi.org/10.1016/j.seta.2020.100679, URL https://www.sciencedirect.com/science/article/pii/S2213138819303108.

Rasmussen, Carl Edward, Williams, Christopher K.I., 2005. Gaussian Processes for Machine Learning. The MIT Press, http://dx.doi.org/10.7551/mitpress/3206.001.0001, arXiv:https://direct.mit.edu/book-pdf/2514321/book_9780262256834.pdf.

del Rio Chanona, E.A., Petsagkourakis, P., Bradford, E., Graciano, J.E. Alves, Chachuat, B., 2021. Real-time optimization meets Bayesian optimization and derivative-free optimization: A tale of modifier adaptation. Comput. Chem. Eng. 147, 107249. http://dx.doi.org/10.1016/j.compchemeng.2021.107249, URL https://www.sciencedirect.com/science/article/pii/S0098135421000272.

Shahriari, Bobak, Swersky, Kevin, Wang, Ziyu, Adams, Ryan P., de Freitas, Nando, 2016. Taking the human out of the loop: A review of Bayesian optimization. Proc. IEEE 104 (1), 148–175. http://dx.doi.org/10.1109/JPROC.2015.2494218.

Shaw, D.C., 1994. Pipeline system optimization: a tutorial. In: Pipeline Simulation Interest Group, 26th Annual Meeting, San Diego, CA.

Silva, Thiago Lima, Camponogara, Eduardo, 2014. A computational analysis of multidimensional piecewise-linear models with applications to oil production optimization. European J. Oper. Res. 232 (3), 630–642. http://dx.doi.org/10.1016/j.ejor.2013.07.040, URL https://www.sciencedirect.com/science/article/pii/S0377221713006425.

Skogestad, Sigurd, 2023. Advanced control using decomposition and simple elements. Annu. Rev. Control. 56, 100903. http://dx.doi.org/10.1016/j.arcontrol.2023.100903, URL https://www.sciencedirect.com/science/article/pii/S1367578823000676.

Sui, Yanan, Gotovos, Alkis, Burdick, Joel, Krause, Andreas, 2015. Safe exploration for optimization with Gaussian processes. In: Bach, Francis, Blei, David (Eds.), Proceedings of the 32nd International Conference on Machine Learning. In: Proceedings of Machine Learning Research, vol. 37, PMLR, Lille, France, pp. 997–1005, URL https://proceedings.mlr.press/v37/sui15.html.

Tirnovan, R., Giurgea, S., Miraoui, A., Cirrincione, M., 2008. Surrogate modelling of compressor characteristics for fuel-cell applications. Appl. Energy 85 (5), 394–403. http://dx.doi.org/10.1016/j.apenergy.2007.07.003, URL https://www.sciencedirect.com/science/article/pii/S0306261907001158.

Tokmak, Abdullah, Krishnan, Kiran G., Schön, Thomas B., Baumann, Dominik, 2025. Safe exploration in reproducing kernel Hilbert spaces. URL https://arxiv.org/abs/2503.10352.

Ungredda, Juan, Branke, Juergen, 2024. Bayesian optimisation for constrained problems. ACM Trans. Model. Comput. Simul. 34 (2), http://dx.doi.org/10.1145/3641544.

Variny, Miroslav, Jediná, Dominika, Rimár, Miroslav, Kizek, Ján, Kšiňanová, Marianna, 2021. Cutting oxygen production-related greenhouse gas emissions by improved compression heat management in a cryogenic air separation unit. Int. J. Environ. Res. Public Health 18 (19), http://dx.doi.org/10.3390/ijerph181910370, URL https://www.mdpi.com/1660-4601/18/19/10370.

Viholainen, J., Grönman, K., Jaatinen-Värri, A., Grönman, A., Ukkonen, P., Luoranen, M., 2015. Centrifugal compressor efficiency improvement and its environmental impact in waste water treatment. Energy Convers. Manage. 101, 336–342. http://dx.doi.org/10.1016/j.enconman.2015.05.054, URL https://www.sciencedirect.com/science/article/pii/S0196890415005099.

Wang, Xilu, Jin, Yaochu, Schmitt, Sebastian, Olhofer, Markus, 2023. Recent advances in Bayesian optimization. ACM Comput. Surv. 55 (13s), http://dx.doi.org/10.1145/3582078.

Wei, Zhichen, Calautit, John Kaiser, 2024. Field experiment testing of a low-cost model predictive controller (MPC) for building heating systems and analysis of phase change material (PCM) integration. Appl. Energy 360, 122750. http://dx.doi.org/10.1016/j.apenergy.2024.122750, URL https://www.sciencedirect.com/science/article/pii/S0306261924001338.

Widell, K.N., Eikevik, T., 2010. Reducing power consumption in multi-compressor refrigeration systems. Int. J. Refrig. 33 (1), 88–94. http://dx.doi.org/10.1016/j.ijrefrig.2009.08.006, URL https://www.sciencedirect.com/science/article/pii/S0140700709001911.

Xenos, Dionysios P., Cicciotti, Matteo, Kopanos, Georgios M., Bouaswaig, Ala E.F., Kahrs, Olaf, Martinez-Botas, Ricardo, Thornhill, Nina F., 2015. Optimization of a network of compressors in parallel: Real time optimization (RTO) of compressors in chemical plants – An industrial case study. Appl. Energy 144, 51–63. http://dx.doi.org/10.1016/j.apenergy.2015.01.010, URL https://www.sciencedirect.com/science/article/pii/S0306261915000161.

Xu, Wenjie, Jiang, Yuning, Svetozarevic, Bratislav, Jones, Colin N., 2023. Primal-dual contextual Bayesian optimization for control system online optimization with time-average constraints. In: 2023 62nd IEEE Conference on Decision and Control. CDC, pp. 4112–4117. http://dx.doi.org/10.1109/CDC49753.2023.10383653.

Zagorowska, M., Degner, M., Ortmann, L., Ahmed, A., Bolognani, S., del Rio Chanona, E.A., Mercangöz, M., 2023. Online feedback optimization of compressor stations with model adaptation using Gaussian process regression. J. Process Control 121, 119–133. http://dx.doi.org/10.1016/j.jprocont.2022.12.001, URL https://www.sciencedirect.com/science/article/pii/S0959152422002220.

Zagorowska, Marta, Skourup, Charlotte, Thornhill, Nina F., 2020. Influence of compressor degradation on optimal operation of a compressor station. Comput. Chem. Eng. 143, 107104. http://dx.doi.org/10.1016/j.compchemeng.2020.107104, URL https://www.sciencedirect.com/science/article/pii/S0098135420305627.

Zhao, Yongning, Xu, Xiandong, Qadrdan, Meysam, Wu, Jianzhong, 2021. Optimal operation of compressor units in gas networks to provide flexibility to power systems. Appl. Energy 290, 116740. http://dx.doi.org/10.1016/j.apenergy.2021.116740, URL https://www.sciencedirect.com/science/article/pii/S030626192100252X.