European Journal of Pharmaceutical Sciences 210 (2025) 107102

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

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

Testing on continuous production of mefenamic acids—Design of experiment through simulation and process optimisation

Kai Eivind Wu [a,*], Cameron J. Brown [b], Murray Robertson [b], Blair F. Johnston [b], Rhys Lloyd [b], George Panoutsos [a,*]

[a] School of Electrical and Electronic Engineering, University of Sheffield, Sheffield, United Kingdom
[b] EPSRC Future Manufacturing Hub in Continuous Manufacturing and Advanced Crystallisation, University of Strathclyde, Glasgow, G1 1RD, United Kingdom

ARTICLE INFO

Keywords:
Continuous manufacturing
MSMPR
Many-objective optimisation
Machine learning

ABSTRACT

In the pharmaceutical manufacturing industry, continuous production methods have been recognised as providing several benefits compared to traditional batch production. These benefits include increased flexibility, higher product output, enhanced quality assurance through better monitoring techniques, and more consistent distribution of Active Pharmaceutical Ingredients (APIs). Despite these clear advantages, there is a lack of research focused on the simultaneous optimisation of multiple sub-processes in continuous manufacturing. This study explores the optimisation processes of continuous pharmaceutical production, explicitly targeting the production of mefenamic acid using wet milling (WM) and mixed-suspension mixed-product removal (MSMPR). We employ data-driven evolutionary optimisation algorithms to address these many-objective optimisation problems (MaOPs). High-fidelity model-generated data generated via the General Process Modelling System (gPROMS) is subsequently utilised to develop simpler surrogate models based on the Radial Basis Function Neural Network (RBFNN). This enables very fast simulations, suitable for use with computationally intensive machine learning algorithms. Utilising evolutionary optimisation algorithms, these models are used for model-based process optimisation. The efficacy of the MaOP approach is evaluated using a range of numeric and visual optimisation performance indicators. Our findings underscore the viability of integrating high-fidelity and surrogate models to discern functional relationships between dependent variables (objective functions) and independent variables (decision variables), providing a robust framework for process optimisation within the pharmaceutical domain. The approximated solutions are, on average, 58% better than the solutions obtained from Latin hypercube sampling. The chosen optimal solutions can form the basis of parameter setting in upcoming experimental campaigns. The significance of this work is in the demonstration, for the first time, of a many-objective optimisation framework for continuous pharmaceuticals production using simple surrogate models derived from high fidelity simulations using Machine Learning.

1. Introduction

In pharmaceutical production, the traditional approach has been batch-based manufacturing. This method entails processing a set amount of input materials in stages to produce the final drug product. Quality controls are implemented at each stage of processing, as well as at the end of production. This mode of production strictly adheres to predefined processing parameters, making mid-process adjustments difficult. As a result, if a batch fails to meet quality standards, it can lead to production delays and increased costs (Andersson and Zacché, 2019).

Continuous Manufacturing (CM) has emerged to address these challenges. CM, as opposed to its batch-based counterpart, possesses a more agile supply chain with lower residence time in the manufacturing process, which offers numerous advantages, including faster, more cost-effective scaling from lab scale to commercial production. This approach also facilitates rapid response to drug shortages and can support personalised medicine (Badman et al., 2019; Ierapetritou et al., 2016; Lee et al., 2015).

Still, CM is not yet widely used, due to strict process rules and challenges with the necessary real-time quality checks in ongoing processes (Hyer et al., 2024; Markarian, 2022). The ICH Quality Guidelines provide a series of standards developed by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), focusing on the quality aspects of pharmaceutical

* Corresponding Author
E-mail addresses: k.e.wu@sheffield.ac.uk (K.E. Wu), g.panoutsos@sheffield.ac.uk (G. Panoutsos).

https://doi.org/10.1016/j.ejps.2025.107102
Received 29 September 2024; Received in revised form 6 March 2025; Accepted 11 April 2025
Available online 29 April 2025
0928-0987/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

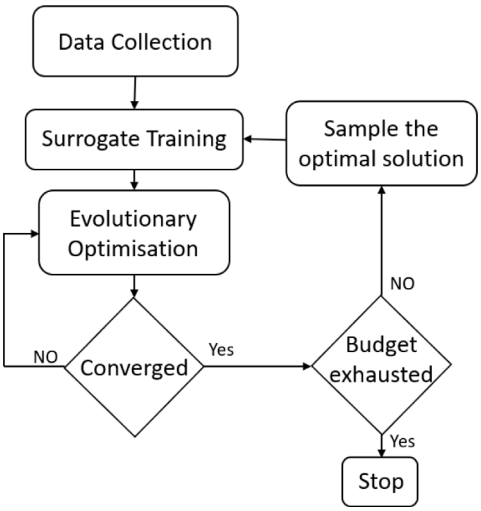


Fig. 1. Generation-based surrogate model management.

product development, manufacturing, and control and ensuring that medicines are safe, effective, and of high quality through aligning technical standards across countries (Anon, 2023).

Underpinning the transition to CM is the Quality by Design (QbD) paradigm. This approach emphasises understanding and controlling critical factors, such as raw material properties and production facilities, that influence product quality. Embracing QbD means employing a systematic methodology that incorporates analytical and risk-management techniques during the drug design, development, and manufacturing phases. Within this paradigm, the input determinants (independent variables) define the process outcomes (dependent variables) in functional terms.

In the proposed work, data-driven process modelling and optimisation is used to systematically address these challenges. This typically includes multivariate modelling methods, multi-objective and many-objective optimisation, quality assessment, as well as support for decision-making (visualisation, performance analyses) (Kim et al., 2024, 2023; Sagmeister et al., 2023). Fig. 1 displays the overall process flow of generation-based surrogate model management.

Data-driven optimisation presents a multidisciplinary approach to eliciting meaningful insights from complex, often noisy, and unstructured data. It encompasses the development of models and constraint functions derived from empirical data acquired through physical experiments and real-world observations, as well as computational and numerical simulations. A primary challenge in data-driven optimisation is addressing the inherent complexities of the data itself. For example, such data may exhibit heterogeneity, be affected by noise, undergo dynamic changes, or be either overly sparse or voluminous. Certain optimisation problems further add to this challenge, when the assessment of objectives or constraints requires resource-intensive procedures, such as costly physical experiments or intensive computational simulations. Notably, simulation-based optimisation, which might rely on techniques like computational fluid dynamics (CFD) and General Process Modelling System (gPROMS), can be time-consuming, with evaluations spanning from minutes to hours. Surrogate models can be established to mirror process knowledge, as simpler versions of high fidelity models. This can be achieved with Machine Learning, via the use of appropriate data to *train*. Surrogate models can then enable the optimisation of pharmaceutical processes, validate optimal solutions, and guide consequential operational decisions—instead of using high-fidelity models, which would deem the use of computationally expensive algorithms infeasible or impractical.

Surrogate models are simplified approximations of more complex processes. Knowing the extrapolation and interpolation properties of these models is critical to their effective application. Interpolation refers to the prediction of outputs within the range of observed data used to train the surrogate model. Extrapolation involves making predictions beyond the range of the training data. While interpolation offers more reliability, extrapolation should be approached cautiously due to inherent uncertainties and limitations (He et al., 2023; Kim et al., 2023).

Several very effective data-driven surrogate models have gained widespread recognition in model-based optimisation. For example, Polynomial Regression or Polynomial Response Surface (Queipo et al., 2005), Multi-Layer Perceptron (Gardner and Dorling, 1998), Support Vector Machines (Steinwart and Christmann, 2008), Decision Trees (Rokach and Maimon, 2007), Fuzzy Rule-Based Systems (Jin, 2000), Gaussian Process Regression Model (Emmerich et al., 2006), and Radial-Basis-Function Neural Networks (Koziel et al., 2011).

RBFNN and other ANN models have found substantial applications in the pharmaceutical and biomedical domains. Alshafiee et al. employed RBFNN to model the flowability of select pharmaceutical powders, capturing relationships with the flow function coefficient and the bulk density (Alshafiee et al., 2019). This model proved effective, especially for smaller datasets, due to its efficient training, interpretable nodes, and structural adaptability to data (Ding et al., 2022). Fragopoulos et al. (2020) utilised RBFNN (Radial Basis Function Neural Network) to assess thyroid lesions, while Wang and Chen (Wang and Chen, 2020) employed a similar approach to investigate the mechanism of acute toxicity. Both studies achieved commendable levels of accuracy in their respective evaluations. Furthermore, in the realm of chemical reactions and processes, Zhou et al. and Gbadago et al. utilised these networks for two-component reactions and butadiene synthesis, leveraging high-fidelity CFD models (Gbadago et al., 2021; Zhou et al., 2017). Velásco-Mejía et al. and Öner et al. expanded the applications to crystallisation processes (Öner et al., 2020; Velásco-Mejía et al., 2016). The former predicted crystallisation density outcomes and identified key parameters, while the latter innovated with real-time training on an automated laboratory crystallisation system using in-line data.

Evolutionary optimisation techniques stand distinct as gradient-free approaches. They utilise stochastic searches to conduct functional evaluations and implement random modifications on superior solutions, resulting in improved candidates through systematic selection of solutions. As iterations progress, the expectation is to identify the best solutions. These are often sub-optimal, as optimality is very challenging to be guaranteed. In contrast to traditional optimisation strategies, evolutionary methodologies excel at navigating intricate problems, offering a valuable ability to sidestep local optima in pursuit of the overarching global optimum.

Prominent many-objective evolutionary optimisation algorithms include Non-dominate Sorting Genetic Algorithm III (Deb and Jain, 2014), Reference Vector-Based Evolutionary Algorithm (RVEA) (Cheng et al., 2016), Knee-Driven Evolutionary Algorithm (Zhang et al., 2015), Two-Archive Evolutionary Algorithm (Elshamy et al., 2007), Preference-Based Algorithms (Reddy et al., 2019), Grid-Based Evolutionary Algorithm (Yang et al., 2013), Corner Sort Algorithm (Wang and Yao, 2014), and Push and Pull Search (Fan et al., 2019).

Performance metrics or indicators are vital tools for evaluating the quality of solutions in many-objective optimisation problems (MaOPs). They aid decision-makers in assessing the effectiveness of different optimisation algorithms. As the number of objectives grows, these indicators succinctly encapsulate essential information about the optimisation problems, ensuring accuracy. Detailed reviews on various quality indicators can be found in established studies (Wang et al., 2017; Audet et al., 2018). The most commonly employed performance metric is Hypervolume (HV) (Zitzler and Thiele, 1999). HV measures the volume of the objective space that is bounded by the set of calculated solutions and a set of reference points. A higher values (volume) indicates a better result.

(a) The wet milling (b) MSMPR setup

Fig. 2. Experiment setup for wet milling (a) and MSMPR (b).

Visualisation plays a pivotal role in decision-making for Pareto Front (PF) approximations. The Pareto Front represents a set of optimal solutions in multi-objective optimisation where improving one objective worsens another. A conflict of interest arises when competing objectives, such as cost vs. quality, require trade-offs. Decision-makers must choose a balanced solution based on priorities. It not only aids decision-makers in their interactive search for optimised solutions but also offers a means to compare the efficiencies of different optimisation algorithms. Factors such as dominance relationships, convergence, and diversity of approximations from different algorithms can be visually analysed (Deb and Jain, 2014; Filipic and Tusar, 2016). Filipič and Tušar (Filipic and Tusar, 2016) have summarised nine desired properties for an effective high-dimensional visualisation method. The method should show dominant relations between PF approximation sets, PF shape, and objective range, PF distribution, to maintain robustness, scalability, simplicity, and uniqueness and handle large and multiple data sets.

In continuous pharmaceutical manufacturing optimisation, the challenge often lies in discerning the relationships between a plethora of objectives and decision variables. Notably, this task becomes more arduous when the process involves multiple outputs. High-dimensional many-objective optimisation techniques must be used in these situations, which leads to the creation of a Pareto front (Wu and Panoutsos, 2021b,a).

In this paper, the focus is on Wet milling (WM) and mixed-suspension mixed-product removal (MSMPR). A typical experimental setup for these two processes is shown in Fig. 2; an example detailed description for these processes can be found in Urwin (2023). In this paper, the research work focuses on part of a preparation study for an upcoming experimental campaign. The hypothesis, and scope, of this research is to test if the proposed framework is suitable to capture the process behaviour sufficiently (in terms of modelling) to enable many-objective optimisation using machine learning. The main contributions of this paper are:

- Development of a surrogate data-driven model of the nonlinear continuous WM and MSMPR processes using machine learning (RBFNN), using process data generated from high-fidelity numerical simulations for the continuous production of mefenamic acid.
- Development of a many-objective optimisation framework for the WM and MSMPR processes, to identify optimal production conditions based on several objectives.
- Demonstration of a decision-making framework based on high-dimensional visualisation methods for optimal solutions.

2. Material and methods

The section discusses in more detail the methodologies utilised in the current work, namely the data generation using the high-fidelity simulations, the creation of the surrogate model, the many-objective optimisation algorithm and the optimisation performance indicators.

2.1. High fidelity simulations - gPROMS

gPROMS as a simulation platform for processes can be very computationally expensive, primarily due to the intensive computation required for solving systems of differential equations (Wu and Nauta, 2022).

Simpler surrogate models established based on data obtained from high-fidelity models, such as gPROMS, may effectively reduce the computational cost by replacing the high-fidelity but computationally expensive simulations of gPROMS so that the iterative process of design optimisation is significantly sped up. The feasibility of optimisation is thus improved in that a large number of evaluations to converge to a solution is made possible - a crucial bottleneck for many computationally intensive optimisation algorithms that rely on multiple iterations of model simulations.

The data for the mefenamic acid case study was generated using gPROMS Formulated Products, a powerful simulation tool renowned for its high fidelity in representing complex processes. This software serves as a convenient platform for data generation, offering a more efficient alternative to conducting numerous physical experiments. This

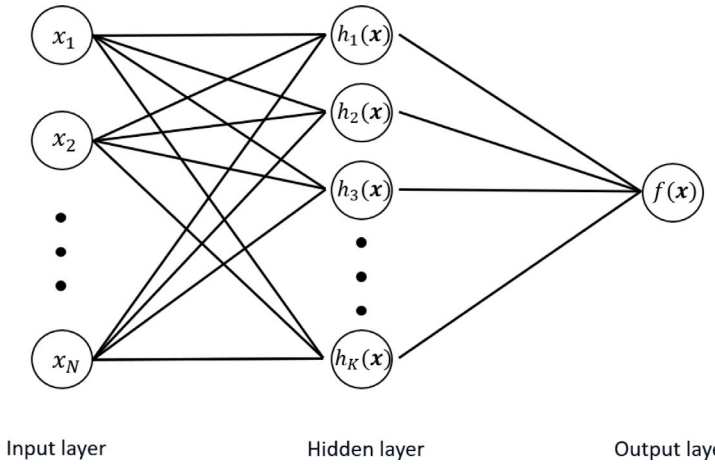


Fig. 3. A schematic view of the principal layout of an RBFNN.

approach substantially reduces the costs associated with data collection and significantly accelerates the process.

However, employing the high-fidelity model directly in optimisation poses practical challenges. This is primarily due to the considerable time required for functional evaluations and the need for tens of thousands of such evaluations (such as, in evolutionary optimisation). To overcome this challenge, we utilise the high-fidelity model to generate precise input and output data, only for inputs and outputs of interest. We then develop (train, using machine learning) surrogate models, which require significantly less time to evaluate.

2.2. Radial basis function neural network

In this research, we use the Radial Basis Function Neural Network (RBFNN) to capture the continuous manufacturing process. While many different machine learning structures could be used here, the RBFNN has universal approximation capability, as well as offers fast learning and forward prediction. Notably, it also offers the advantage of expedited training due to its inherent avoidance of local minima challenges (Lin et al., 2020). The RBFNN is a feed-forward neural network structure characterised by three layers: the input layer, a hidden layer, and an output layer. The hidden and output layers include the network's trainable parameters. The activation functions (hidden layer) are defined by a multi-dimensional Gaussian function (Radial Basis), where the standard deviation and the centre of the function serve as a trainable parameter. Back-propagation is used to find the best weights for the final model, with the loss function comprising of the model error(between the predicted output and simulated target values) (AlAlaween et al., 2017).

Fig. 3 illustrates the network structure for an RBFNN, where N represents the input vector's dimension (training data) and K signifies the dimension of the hidden layer.

To obtain the model parameters of an RBFNN, a typical training process consists of two distinct steps. In the initial step, unsupervised learning is employed, often utilising the K-means algorithm, to ascertain the critical parameters associated with the radial basis functions. These parameters encompass the determination of the number of hidden layers denoted as 'K' and the calculation of the spread denoted as σ. Subsequently, in the second step, a supervised approach is implemented, focusing on the minimisation of modelling errors. This stage is instrumental in learning the weights that connect the hidden layers with the output layers. It is important to note that the neuron functions within the output layer operate in a linear fashion.

In this study, we employ a three-layer RBF neural network to construct surrogate models for various output functions. The RBFNN, characterised by a Gaussian activation function, is mathematically formulated as follows (Rubio-Solis and Panoutsos, 2015):

$$f(\boldsymbol{x}) = \sum_{i=1}^{K} \omega_i \sum_{j=1}^{J} exp(-\frac{1}{2\sigma^2} \left\| x_j - c_i \right\|_2^2) + w_0 \tag{1}$$

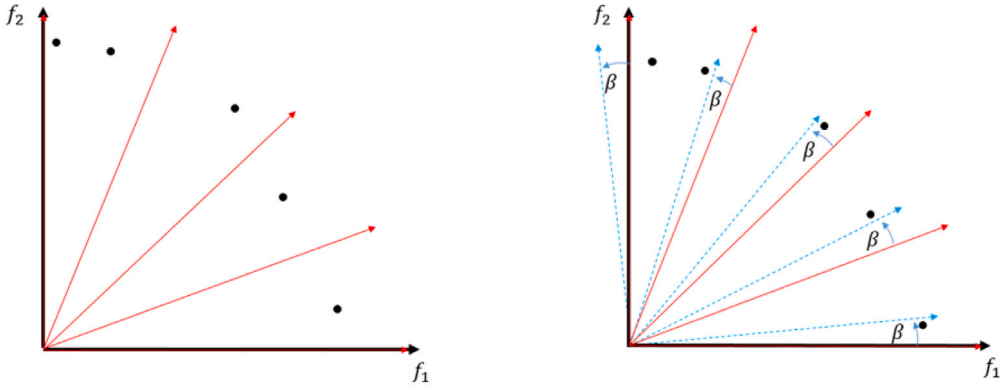
in this expression $\boldsymbol{x}$ is the input vector and f the output of the RBF neural network. K is the number of hidden neurons and w_i is the output weight relating to the i^th hidden neuron and the output neuron; w_0 is the constant bias for the output layer; J denotes the number of training data points; c_i is the centre of the Gaussian function for the ith hidden neuron; σ is the width parameter for each radial basis function; $\left\| x_j - c_i \right\|_2^2$ is the Euclidean distance of an input vector to its RBF centre.

$$\left\| x_j - c_i \right\|_2^2 = \sum_{n-1}^{N} \sqrt{(x_{n,j} - c_i)^2} \tag{2}$$

In Eq. (2), the parameters c_i and σ representing the centre and shared width of the Gaussian functions are established as initial conditions using simulated training data before training the weights ω_i in the RBF Neural Network. This initial phase involves grouping training data into clusters, from which the values for c_i and σ are determined for each cluster i within the range of '$i \in (i, \ldots, K)$'. It is worth noting that the number of clusters equals the number of RBF functions. In the context of this study, clustering and model training is performed using the computational framework described in Rubio-Solis and Panoutsos (2016) and Rubio-Solis et al. (2018).

2.3. Many-objective optimisation

This section outlines the Non-dominated Sorting Genetic Algorithm III (NSGA-III), one of the most popular gradient-free multiobjective optimisation algorithms. The NSGA-III algorithm in our case study serves as a key tool for conducting many-objective optimisation analyses, which aligns with the requirements of the mefenamic acid case study (*many-objective*, as opposed to multiobjective, refers to cases where greater than three objectives need to addressed simultaneously). The NSGA-III (Deb and Jain, 2014) algorithm follows a similar evolutionary computation framework to that of NSGA-II, encompassing initialisation, crossover, and mutation phases to generate an offspring population from the existing population. Where it distinguishes itself from NSGA-II is in the environmental selection process, as NSGA-III introduces the utilisation of a predefined set of widely and uniformly distributed reference points within the objective space. These reference points play a pivotal role in the selection process, thus enhancing diversity among the solutions.



(a) Locations of candidate solutions and reference vectors to start with.

(b) Rotated system of reference vectors with angle β.

Fig. 4. A 2D schematic view shows the rotation of reference vectors by β, where the sum of included angles is minimised to attain the minimal sum of net avertence angles.

2.4. Optimisation performance evaluation

A shortfall of heuristic and meta-heuristic optimisation algorithms is that while they can be very efficient and effective, understanding their performance can pose a significant challenge. This is because convergence analyses, and optimality guarantees cannot be derived in a straightforward fashion, as in the case for gradient-based methods. Therefore, there are numerous computational techniques developed for the post-hoc analysis of the performance of such algorithms and the quality of solutions (Riquelme et al., 2015). These techniques include numerical metrics of performance, as well as visualisation methods. Such methods are not trivial to develop, when considering the number of dimensions involved (e.g. visualisation of solution grouping in n dimensions, when n is high (which is the case for many-objective optimisation).

Hypervolume

Hypervolume (HV) (Zitzler and Thiele, 1999) serves as a metric for quantifying the volume within the objective space defined by the candidate solutions to a reference point. Typically, the Nadir point is selected as the reference, which is determined by combining the worst values across all objective functions present within the Pareto front. In the context of this metric, a higher value is indicative of a more desirable outcome (higher coverage of the solution space). While simple, this is one of the most commonly used measures of performance.

Inverse Ratio of Net Avertence angle

The inverse Ratio of Net Avertence angle (IRNA) (Wu and Panoutsos, 2021a) is classified as a unary diversity metric. The construction of this metric involves the use of reference vectors, with the aim of minimising the sum of the included angles. This minimisation is accomplished by systematically rotating the reference vector system in all dimensions, optimising the spatial angle for each rotation, which is shown in Fig. 4. This process removes possible systematic bias in assessing solutions, giving the solution set the highest possible diversity score. The process of evaluating diversity involves first identifying candidate solutions within an approximation set.

2.5. Visualisation methods

Visualisation methods (Filipič and Tušar, 2018), typically aim to provide an effective visual method for portraying solutions – the Pareto front – in multiple dimensions.

2.5.1. ProD

ProD (Projections on Reference Vector versus Distances) method, as described in Wu and Panoutsos (2021b), is designed to present Pareto front (PF) approximations through a combination of projections (Pro) onto reference vectors and distances (D) to those same reference vectors. In this context, a vector connects a nominal Ideal point and a nominal Nadir point based on the available non-dominated PF approximations. Fig. 5 shows a schematic view of the principle of ProD. The projection $r_{\parallel}$ is given as:

$$r_{\parallel} = [f_1, f_2, \dots, f_m] * \frac{RV}{|RV|} \tag{3}$$

in which m is the number of objective functions, and RV is the reference vector that is defined as:

$$RV = F_N - F_I \tag{4}$$

in which F_N is the vector of the nadir point and F_I is the vector of the ideal point. The ideal point is an auxiliary point with the least values of each objective function among PF approximations as coordinates, while the nadir point, on the contrary, consists of coordinates of the largest of each objective.

The following expression represents the distance to the reference vector:

$$r_{\perp} = \sin\theta * \sqrt{f_1^2 + f_2^2 + \dots + f_m^2} \tag{5}$$

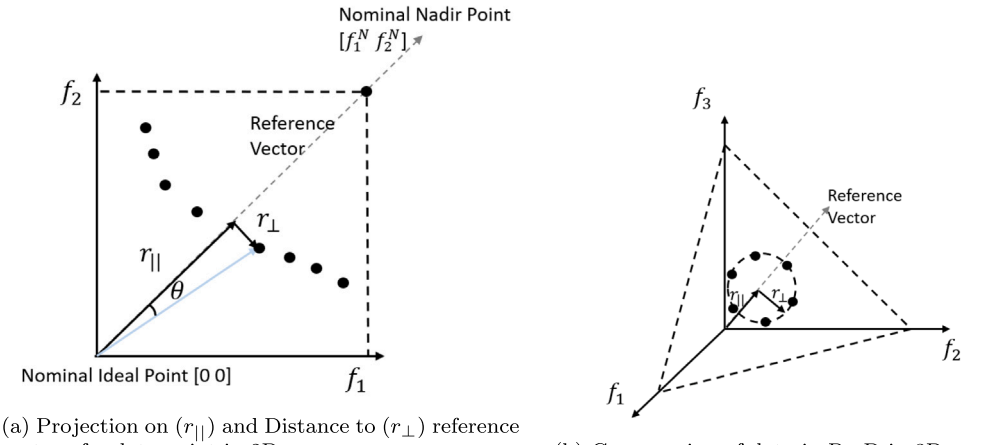
where angle θ between the vector of candidate solution and reference vector is calculated by

$$\theta = \cos^{-1} \frac{[f_1, f_2, \dots, f_m] * RV}{\sqrt{f_1^2 + f_2^2 + \dots + f_m^2} * |RV|} \tag{6}$$

ProD possesses a number of preferred properties of a visualisation method, i.e., display dominance relation, PF shape, objective range, PF distribution, robustness, handling large sets, handling multiple sets, scalability, simplicity, and uniqueness.

2.5.2. Parallel coordinates

Parallel Coordinates (Inselberg, 2009) is a visualisation method that projects high-dimensional vectors onto a two-dimensional space. It achieves this by using m equally spaced parallel axes, where m corresponds to the number of objective functions. Vectors are represented as poly-lines, connecting points on each axis according to their component values. This approach offers a simple, scalable solution for visualising multi-dimensional data, with no loss of information during the mapping process, however, it can be challenging to interpret for high complexity problems.



(a) Projection on ($r_{||}$) and Distance to ($r_\perp$) reference vector of a data point in 2D space.

(b) Compression of data in ProD in 3D space

Fig. 5. A schematic view of ProD in 2D (a) spaces and view in 3D (b) spaces.

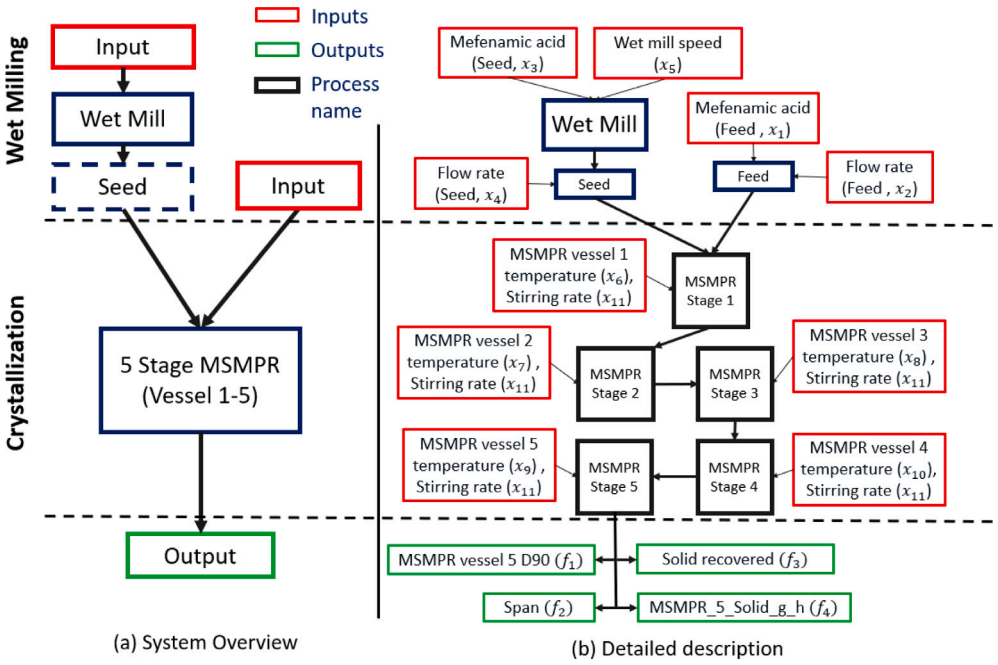


Fig. 6. The main portion of the mefenamic acid manufacturing process is shown, where the solid square boxes are the two major processing steps decision-makers are interested in optimisation.

3. Mefenamic acid production

Mefenamic acid, a commonly used non-steroidal anti-inflammatory drug for alleviating moderate pain, is the focus of this study. This entails the wet milling and crystallisation of mefenamic acid, as illustrated in Fig. 6(a).

The process commences with the wet milling of raw material, in the form of a seed, to obtain a specific particle size. Subsequently, the concentrated slurry derived from the wet milling step undergoes further processing. This involves dilution, achieved by blending the slurry with known compositions to attain a specified concentration, within the framework of mixed-suspension mixed-product removal (MSMPR) as illustrated in Fig. 6(b).

The continuous crystallisation test bed is employed to crystallise various active pharmaceutical ingredients (APIs) as a pivotal aspect of the manufacturing process.

Fig. 6(b) illustrates the specific segment of the mefenamic acid manufacturing process that serves as the primary focus of this study. Within this context, the primary objective of the decision-makers is to identify and optimise the operational conditions specifically related to the wet milling process and the crystallisation (MSMPR) that maximise a number of production/product objectives.

3.1. Decision variables, objective functions and constraints

Fig. 6(b) demonstrates the presence of four key outputs as of interest in our case study, as outlined in Table 1. This table also provides a comprehensive list of the inputs that play a pivotal role in the wet milling and crystallisation processes.

In Table 1, we present the primary outputs prioritised by decision-makers and their corresponding input parameters.

Table 1
List of inputs and output for the mefenamic process.

Inputs	Description
- Mefenamic acid (Feed) (x_1)	- Amount of acid added to the Feed.
- Pump flow rate (Feed) (x_2)	- Pump flow rate in Feed.
- Mefenamic acid (Seed) (x_3)	- Amount of acid added to the Seed.
- Pump flow rate (Seed) (x_4)	- Pump flow rate in the Seed.
- Wet mill speed (x_5)	- Rotating speed of the wet mill.
- MSMPR vessel temperature (x_6, x_7, x_8, x_9, x_{10})	- Temperature in 5 vessels during the continuous manufacturing process
- Stirring rate for all stages (x_{11})	- Speed of stirring in 5-stage MSMPR
Outputs	
- MSMPR_5_D90 (f_1)	- A size metric which represents the particle size 90% of particles are smaller than.
- Span (f_2)	- The width of the particle size distribution.
- Solid recovered (f_3)	- Percentage of the input mefenamic acid produced as solid crystals.
- MSMPR_5_Solid_g_h (f_4)	- The flow rate of mefenamic acid as solid crystals at steady state.

Table 2
Number of solutions generated and max. number of iterations.

Number of objectives	4
Number of decision variables	11
Number of repeated tests	24
Number of optimised solutions	4,000
Maximum number of evaluations	100,000

Table 3
The linear correlation coefficients between the inputs and the outputs for the mefenamic acid case study.

	f_1	f_2	f_3	f_4
x_1	0.011	−0.071	0.131	−0.016
x_2	0.216	0.083	−0.037	1
x_3	−0.469	0.166	0.276	0
x_4	−0.288	0.312	−0.162	0
x_5	−0.592	0.009	0.004	0
x_6	−0.22	−0.302	−0.157	0
x_7	−0.326	−0.523	−0.288	0
x_8	−0.236	−0.423	−0.354	0
x_9	−0.184	−0.349	−0.468	0
x_{10}	−0.071	−0.027	−0.887	0
x_{11}	−0.151	−0.081	0.029	0

4. Data acquisition, surrogate modelling and parameter settings for optimisation

This section encompasses the data acquisition, including the methodology for sampling the synthetic experimental data, and the analysis of the dependencies between the input variables and output functions (feature selection), towards the selection of the inputs for the surrogate model and the subsequent process optimisation (see Table 2).

4.1. Simulations and data sampling

A full factorial (symmetric design of experiments) high-fidelity simulation is run, yielding 100,000 samples. A smaller set, consisting of 4000 samples is randomly chosen, for the purpose of training the surrogate model. 50% of the samples are selected randomly for the surrogate model training and the rest 50% samples are used as the hold-out set for the purpose of model cross-validation. This process is repeated several times (24) to help establishing statistical measures of robustness for the model training.

4.2. Feature selection

In our study, the inclusion of an input variable in the formulation of a surrogate model for the output is confirmed based on its linear correlation coefficient with the output variable, relying on the assumption that the input–output relationships resemble linear behaviours; the Pearson correlation coefficient is used for this purpose.

Table 3 presents the correlation coefficients between the inputs and outputs. As an initial criterion, we employ a threshold of ±0.1 to assess the level of correlation between variables. Inputs that exhibit correlations meeting or exceeding this threshold are considered for inclusion in the surrogate modelling of the target functions. To account for strong nonlinear input–output relationships (Pearson correlation coefficient may fail to identify these), a *wrapper* feature selection approach (Jović et al., 2015) is used, whereby the Machine Learning model itself is used to interrogate the relevance of the inputs. The details of these dependencies are outlined in Table 3.

The correlated inputs and outputs are:

$f_1(x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{11})$
$f_2(x_3, x_4, x_5, x_6, x_7, x_8, x_9)$
$f_3(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})$
$f_4(x_2)$

Note: x_5 is found to be important for f_2 in the later modelling work.

4.3. Surrogate model structure and training

In the context of the RBFNN model structure, there is one tunable (learnable) hyperparameter, relating to the number of neurons in the hidden layer—denoted as M. Furthermore, there are two additional hyperparameters relating to the RBF activation functions; these are the widths and centres of the radial basis functions, represented as $\boldsymbol{\sigma}$ and $\boldsymbol{c}$; these are vectors with the same dimension as the input layer (one width and one centre for each input feature). To start the model training process, M, $\boldsymbol{\sigma}$ and $\boldsymbol{c}$ are established via data clustering. $\boldsymbol{\sigma}$ and $\boldsymbol{c}$ are further tuned, as part of the main model training part (also referred to in the literature as parametric model optimisation). The loss function in the Machine Learning framework is the *Root Mean Square Error*, using the error calculation between the model predicted and actual (simulated in our case) outputs. For model validation purposes, other statistical measures are used too, to ensure good model training and avoidance of overfitting; R^2 as well as *Mean Absolute Error* are used in the cross-validation process. Cross-validation includes *holding out* of the training process a portion of the data, and monitoring the performance in the hold-out sample set. Model training results are shown in Section 5.1.

4.4. Many-objective process optimisation framework

The MaOP algorithm used in this work is NSGA-III (Deb and Jain, 2014). The aim is to identify a Pareto Front which includes all non-dominated solutions for the optimisation task. The specific algorithm in our work is adopted from PlatEMO (Tian et al., 2017) with some modifications. NSGA-III can be used to solve many-objective optimisation problems, as in the mefenamic acid case study (a four-objective function optimisation problem). The quality metrics described in Section 2.4, Hypervolume (HV) and Inverse Ratio of Net Avertence Angle (IRNA), are utilised to support the analyses of the optimisation results.

The optimisation algorithm was set up as shown in Table 3; for the rest of the parameters set in the algorithm default values were used as in PlatEMO version 4.2 (Tian et al., 2017).

Box constraints were established, i.e., minimum and maximum allowed values for the inputs and outputs; these are listed in Table 4.

Table 4
The box constraints of the inputs and the outputs for the mefenamic acid.

Input/constraints	Max	Min	Interval
Mefenamic acid (Feed, x_1)	45	35	1
Pump flow rate (Feed, x_2)	20	10	1
Mefenamic acid (Seed, x_3)	6	2	0.1
Pump flow rate (Seed, x_4)	1.5	0.5	0.1
Wet mill speed (x_5)	11,000	5,500	100
MSMPR vessel 1 temperature (x_6)	55	35	1
MSMPR vessel 2 temperature (x_7)	55	35	1
MSMPR vessel 3 temperature (x_8)	50	30	1
MSMPR vessel 4 temperature (x_9)	45	25	1
MSMPR vessel 5 temperature (x_{10})	35	15	1
Stirring rate for all stages (x_{11})	600	200	1
Output/constraints	**Max**	**Min**	**Goal**
MSMPR_5_D90 (f_1)	326	10	=120
Span (f_2)	2.1	0.285	Minimise
Solid recovered (f_3)	56 (%)	27(%)	Maximise
MSMPR_5_Solid_g_h (f_4)	19.3	9.55	Maximise

Table 5
Hidden layer structure for each objective function.

Objective function	Number of neurons, M
f_1	400
f_2	800
f_3	300
f_4	8

5. Results

In this section, we present the outcomes of the surrogate modelling and optimisation results. These results are conveyed through various forms, including model fitness plots, Pareto Fronts plots shown in both the objective space and the decision space, as scatter plots, parallel coordinates plots, and Prod plots. Furthermore, we numerically assess the coverage of the Pareto Front approximations within the objective space, using HV and IRNA.

5.1. RBFNN surrogate model training results

The structural optimisation of the hidden layer, as described in Section 2.2, yields a number of hidden neurons, M, that is different for each objective function (output); this is shown in Table 5. For the parametric optimisation of the surrogate model (model training) the two hyperparameters of the RBFNN, $\boldsymbol{\sigma}$, and $\boldsymbol{c}$ are *learned* from the data using back-propagation. The cross-validation process, helps establishing hyperparameters while at the same time avoiding overfitting. This is achieved by monitoring the performance on the hold-out dataset, and selecting a model that offers a balanced performance between *training* and *validation*. For example, evidence of overfitting, would be a very good training performance with a poor validation performance (i.e. poor generalisation). The measures of model performance (accuracy) in terms of *RMSE*, *MAE* and R^2 are calculated for both the training and validation data to enable comparisons and analyses.

To assess the efficacy of the modelling, we compare between the simulated outputs and the predicted ones. These are outlined in Fig. 7, which includes fitness plots for both training and validation sets.

From visual observation (which is later on confirmed numerically, see Table 6) there overall fit appears to be satisfactory and there is no evidence of overfitting (significant differences between training and validation). One can observe that the data for f_4 is very sparse compared to the other objective functions and are locally concentrated. f_4 is a much simpler objective compared to the other objective functions and it also varies linearly with x_2.

The root mean square error (RMSE), the mean absolute error (MAE) and R^2 for the surrogate models are calculated and shown in Table 6. The RMSE and MAE values of the modelled objective functions are generally within 3% compared to the targeted values from that predicted by gPROMS, while R^2 values are generally higher than 0.94. The model performance is thus satisfactory.

Table 6
RMSE, MAE and R^2 for the objective function of the mefenamic acid case study.

	f_1	f_2	f_3	f_4
RMSE	0.0214	0.0184	0.0241	0.0066
MAE	0.0167	0.0129	0.0182	0.0052
R^2	0.9696	0.9405	0.9820	0.9997

5.2. Pareto front in parallel coordinates

It is challenging to visualise the Pareto Front for more than three objectives (dimensions). Possible methods that can be used are parallel coordinates and ProD. Fig. 8 displays the Pareto front of the mefenamic acid optimisation study, depicted in parallel coordinates. All the objective functions have been displayed with non-normalised and normalised values, which, in this case, use their respective upper- and lower-level limits so that the terms are displayed in their relative amounts (a quantity varies between 0 and 1).

All four objective functions are conflicting; see Figs. 8(a) and 8(b). The objective functions of the MSMPR_5_D90 (f_1), the Span (f_2, minimised) and the Solid Recovered (f_3, maximised) vary moderately under optimised operational conditions. The MSMPR_5_solid_g_h (f_4) changes significantly. It is recommended to the decision-makers that the balanced operational parameters of the MSMPR_5_D90 (f_1), the Span (f_2, minimised) and the Solid Recovered (f_3, maximised) should be chosen since the choice will, to a limited degree, influence the performance of MSMPR_5_solid_g_h (F4). The selected alternatives of sets of operational parameters should be verified with the high-fidelity model.

All the optimised solutions are *good choices* under specific situations. Prioritising f_4 will decrease the quality of the drug and increase the material waste. The decision-maker should only prioritise f_4 if maximising production is overwhelmingly more important than other factors.

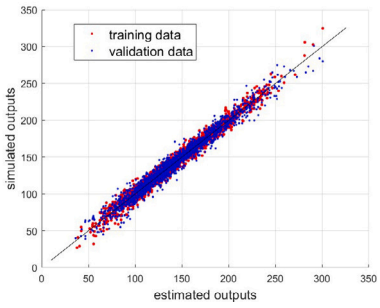
5.3. Pareto front in ProD plots

As before, before plotting in ProD, the objective functions have been normalised with their respective lower and upper limits. Fig. 9 displays the PF approximations of the mefenamic acid optimisation study in the ProD plot. The PF approximations are represented by the red dots. At the same time, the simulated outputs are shown in blue dots. It clearly shows that the PF approximations are located nearer to the origin than the simulated outputs, which means that the PF approximations represent better solutions compared to the simulated outputs.

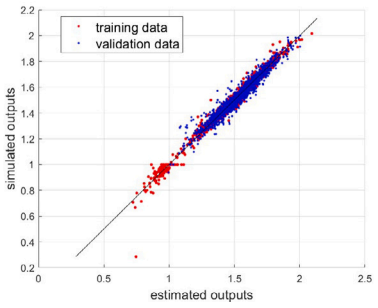
Comparing the optimality of the PF approximations versus the simulated data, the two data sets are subject to non-dominated sorting operations. The result is displayed in Fig. 10. It shows that only a few of the simulated outputs *survived* the evolutionary optimisation process, indicating that the optimisation process has been successfully conducted. Still, none of the simulated output should ideally survive after such a non-dominated sorting manoeuvre. When it does, it usually indicates that the PF approximations have not been found in certain areas in the objective space.

5.4. Pareto optimal sets (results in decision space) in parallel coordinates

The optimal solutions in the decision space (optimal solution set) are displayed in parallel coordinates in Fig. 11(a), together with the simulated input data in Fig. 11(b). The simulated input data shows a regular pattern since a full factorial design has been adopted in the simulations. The optimal solution set data have been normalised. Different decision variables have unique characteristics. Some are densely

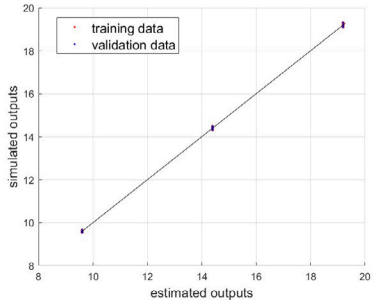


(a) The comparison between the simulated and predicted values of f_1.



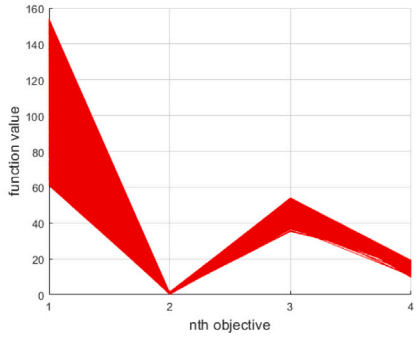
(b) The Comparison between the simulated and estimated function values of f_2

(c) The Comparison between the simulated and estimated function values of f_3

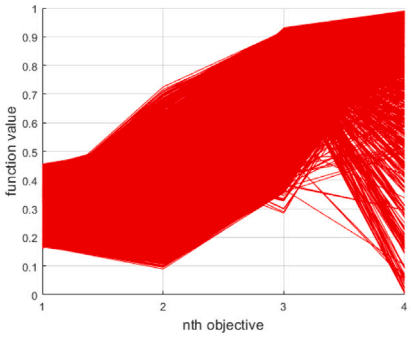


(d) The Comparison between the simulated and estimated function values of f_4

Fig. 7. Comparisons between the simulated and estimated objective function values. The solid line represents an exact match between the simulated and estimated output.



(a) Unnormalised



(b) Normalised

Fig. 8. The Pareto front of the mefenamic acid process: (a) Unnormalised and (b) normalised.

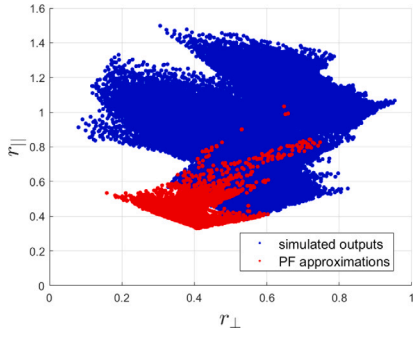


Fig. 9. The Pareto front of the mefenamic acid optimisation study in the ProD plot.

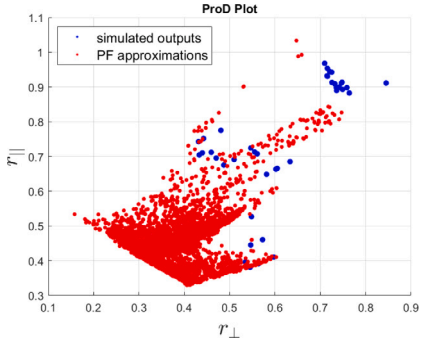
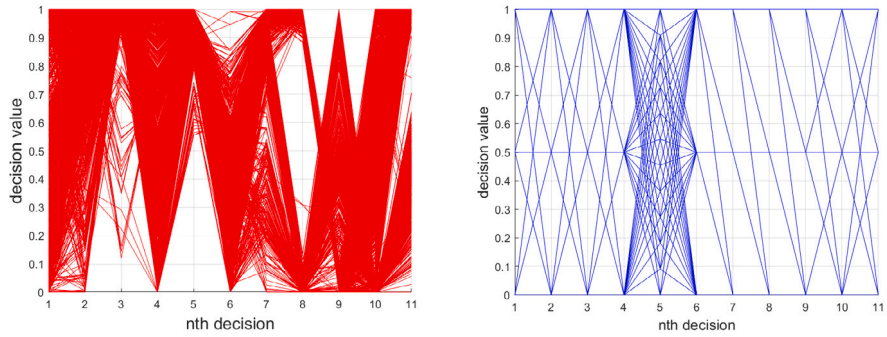
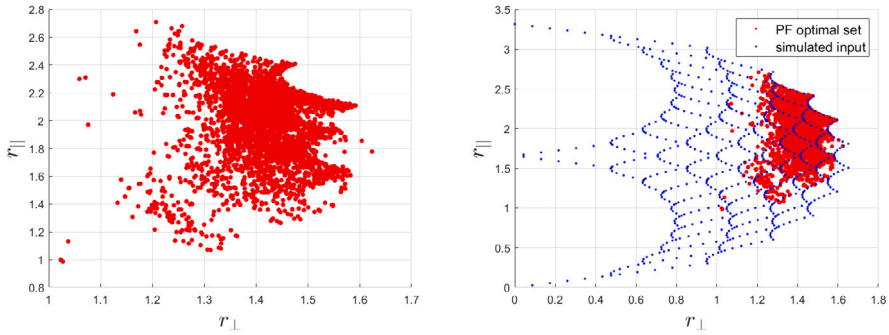


Fig. 10. The PF approximations and the non-dominated portion of simulated output.



(a) The Pareto optimal sets (b) The input data

Fig. 11. The Pareto optimal sets of the Mefenamic acid Case and compared with the raw input data.



(a) Pareto optimal sets (b) Pareto optimal sets and the simulated input

Fig. 12. The Pareto optimal sets (decision variables at optimised operational conditions) of the Mefenamic acid Case after an optimisation analysis (a) and compared with the raw input data (b).

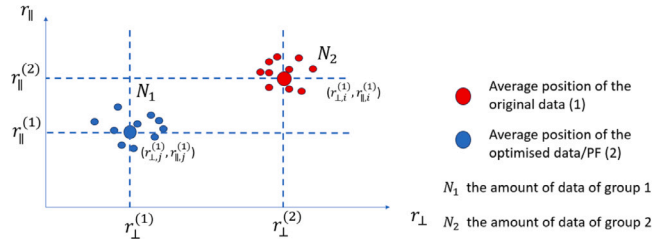


Fig. 13. The schematic view of evaluating the efficiency of an optimisation outcome.

populated (x_1, x_2, x_4, x_6, x_7, x_9, x_{10}) and extend to the whole scale. It is easier to select an optimal solution set for those decision variables. Others occupy partially (x_3, x_5, x_8, x_{11}) the range. The final choice of optimal process parameters should be made in areas where the solutions are densely populated. The solutions are more robust, meaning that any small fluctuations around the quantities would not jeopardise the optimality of the operations.

5.5. Pareto optimal sets (results in decision space) in ProD

The corresponding decision variable values at optimised operational conditions are displayed in a ProD plot in Fig. 12. Fig. 12(a) shows Pareto optimal sets (red dots). In contrast, they are plotted together with the simulated input data in Fig. 12(b), where the relative locations of the simulated inputs versus the Pareto optimal sets are indicated.

The optimal solutions in the decision space are relatively densely populated in a cluster, making choosing an ideal operational condition easier. A good choice might be the centre region of this cluster, providing a robust solution. Any minor discrepancies or changes in the operational parameters would not stop the system from running under optimal conditions.

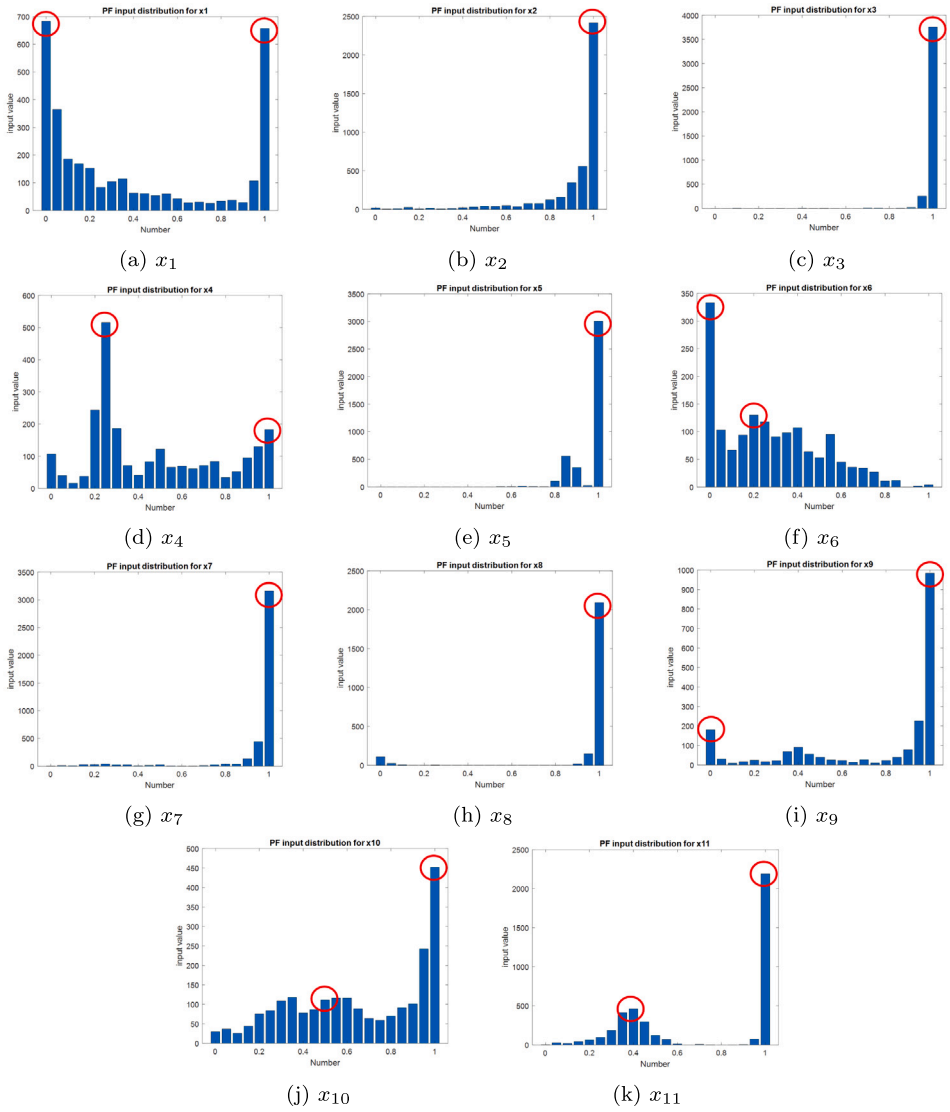
5.6. Quality of Pareto front

The evaluation of the quality of Pareto Front approximations revolves around two essential aspects. Firstly, we assess how closely these approximations align with the true Pareto Front, a criterion referred to as convergence. This is not possible here, as the true Pareto Front is not known. Secondly, we gauge the extent to which the Pareto Front is covered within the objective space, a measure termed diversity.

As the true Pareto Front is not known, to promote convergence, this study employs a rigorous approach involving multiple independent optimisation runs, totalling 24 runs. From these runs, the Pareto Front approximations are collected and subjected to non-dominated sorting, culminating in the formation of the final Pareto Front. This avoids *outlier* runs where the algorithm fails to converge, and convergence is poor.

Diversity is evaluated using numeric performance indicators, notably the Hyper Volume (HV) (Shang et al., 2021) and the Inverse Ratio of Net Avertence angle (IRNA) (Wu and Panoutsos, 2021a). These metrics provide insights into the distribution and coverage of the Pareto Front within the objective space. You can refer to Table 7 for specific HV and IRNA values associated with the Pareto Front approximations.

The maximum achievable values for both HV and IRNA metrics are 1.0. A HV value of 0.6732 indicates relatively low diversity, and the same holds true for an IRNA value of 0.1303. Notably, HV is nonlinear as a diversity metric, meaning that its value varies nonlinearly with



(a) x_1 (b) x_2 (c) x_3

(d) x_4 (e) x_5 (f) x_6

(g) x_7 (h) x_8 (i) x_9

(j) x_{10} (k) x_{11}

Fig. 14. Cardinality distributions of PF optimal sets.

Table 7
The HV and IRNA values of the PF approximations and the simulated data.

	PF approximations	The simulated data
HV	0.6732	0.3412
IRNA	0.1303	0.0033

increasing diversity. Consequently, it can be challenging to assess the coverage of the Pareto front in the objective space using HV alone.

However, IRNA serves as a linear diversity metric, and its value increases linearly with expanding Pareto front coverage. An IRNA value of 0.1303 signifies that the obtained Pareto front approximation encompasses only around 13% of the entire four-dimensional objective space. This limited coverage may have two possible explanations. Firstly, there could be additional Pareto front approximations that the optimisation process failed to discover. Secondly, it is possible that most of the undetected Pareto front approximations have been dominated by the existing ones.

Table 7 also presents the HV and IRNA values for the simulated data, which demonstrate significantly lower optimisation performance values. These values underscore the limited coverage of the simulated data within the objective space, forming the foundation for surrogate modelling. Investigating the impact of this sparse coverage of simulated data on the accuracy of surrogate models could be an interesting subject for future research.

5.7. Evaluation of the effect of optimisation - ProD

The vertical axis of ProD, see for instance Fig. 12, indicates the convergence of the optimisation, i.e., how far the optimisation process has reached. The effect of the optimisation process can thus be assessed by comparing the averaged $r_{\|}$ value of the obtained PF with that of the raw data; this principle is shown in Fig. 13.

The efficiency in an optimisation outcome or improvement in convergence (IC) compared with the simulated data before optimisation:

$$IC = \frac{r_{\|}^{(2)} - r_{\|}^{(1)}}{r_{\|}^{(1)}} = \frac{0.74 - 0.47}{0.47} = 0.58 \tag{7}$$

where: $r_{\|}^{(1)} = \frac{\sum_{j=1}^{N_1} r_{\|,j}^{(1)}}{N_1} = 0.47$ $r_{\|}^{(2)} = \frac{\sum_{j=1}^{N_1} r_{\|,j}^{(1)}}{N_1} = 0.74$

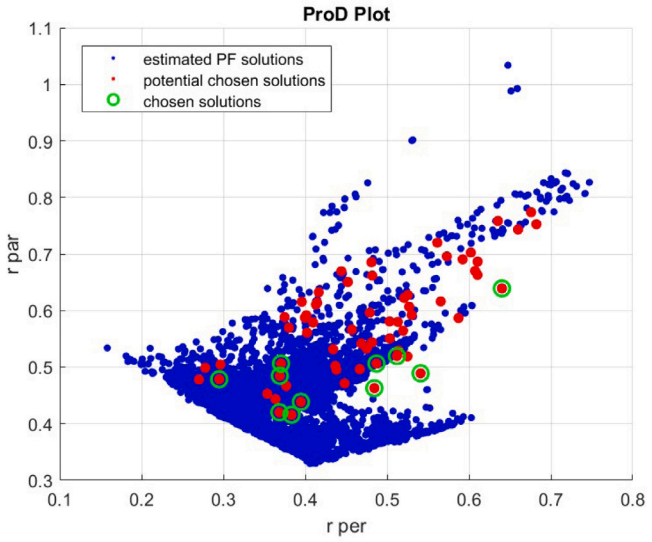


Fig. 15. The recommended chosen solutions for further experimental testing.

Table 8
The chosen decision/objective function values obtained for further testing while x_2=20 x_3=6 x_5=11000 x_7=55 x_8=50 x_{11}=600 are common.

Solution No.	x_1	x_4	x_6	x_9	x_{10}	f_1	f_2	f_3	f_4
1	35	0.75	35	25	25	111	0.813	44.4	19.2
2	45	0.75	35	25	25	111	0.997	47.2	19.2
3	35	1.50	35	25	25	64	0.529	43.8	19.2
4	35	0.75	39	45	25	120	1.171	42.4	19.2
5	45	0.75	39	45	25	120	1.392	47.2	19.2
6	35	0.75	35	25	35	111	0.535	38.1	19.2
7	45	0.75	35	25	35	111	0.748	42.4	19.2
8	35	1.50	35	25	35	63	0.185	37.4	19.2
9	45	1.50	35	25	35	63	0.353	41.0	19.2
10	45	1.50	39	25	35	57	0.373	41.0	19.2
11	35	0.75	39	45	35	120	1.144	35.1	19.2

5.8. Choice of solutions for further experimental testing

After determining the PF optimal sets, their cardinality distributions are calculated and arranged by dividing an input variable into 20 equal parts and counting the number of solutions that occurred in each part. (A similar technique to Kernel Density Estimation (KDE)). The result of totalling 11 scatter plots is depicted in Fig. 14. Then, the most frequent occurring decision variable values are identified and marked in red circles. By randomly combining these solutions, 64 potentially PF optimal sets are identified. They are then used in the surrogate models yielding multiple objective function sets. The final optimal solutions are found using non-dominant sorting on these sets with all previously found PF approximations. The outcomes are plotted in ProD in Fig. 15. Here, the red dots signify the 64 generated solutions, and the green dots show the targeting PF solutions among them, while the blue dots indicate the PF solutions from previous optimisation efforts. The corresponding PF optimal sets (corresponding to the green dots) are listed in Table 8 and can be targeted in a new experimental campaign.

6. Summary and conclusions

This study focuses on the process optimisation of mefenamic acid production based on simulated data from a high-fidelity model, gPROMS software. For the first time, a many-objective framework is considered, and to enable this a computationally fast surrogate model of the process has been developed with via the use of Machine Learning. The following conclusions are made based on the analysis performed:

- Using the process simulations via high-fidelity models (gPROMS) and simpler surrogate models based on Machine Learning to estimate the functional relationships of wet milling and MSMPR processes is feasible for process optimisation purposes;
- The Radial Basis Function Neural Network is an effective and modelling tool for the current study, for regression—this is used to establish a number of surrogate models of high accuracy in terms of low root mean square error, mean absolute error, and R^2 value.
- The model parameters are found at which the differences between modelling errors ($RMSE$, MAE or R^2) of training and validation offer a good balance, hence no overfitting and good generalisation properties.
- A clear Pareto Front surface is obtained. The corresponding decision variables under optimal conditions are located in clusters, making choosing the optimal operational conditions more accessible.
- The current approach identifies the critical factors in the continuous manufacturing process of mefenamic acids, in the form of raw material properties and production governing parameters, that influence the efficiency of production and product quality. Within this paradigm, the input determinants (independent variables) define the process outcomes (dependent variables) in functional terms. A systematic methodology is developed that incorporates analytical and risk-management considerations during the drug design, development, and manufacturing phases.

CRediT authorship contribution statement

Kai Eivind Wu: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Cameron J. Brown:** Writing – review & editing, Supervision, Resources, Data curation. **Murray Robertson:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization. **Blair F. Johnston:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Rhys Lloyd:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Conceptualization. **George Panoutsos:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Formal analysis, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT-4o in order to improve the readability and language of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Kai Eivind Wu reports financial support was provided by EPSRC Future Continuous Manufacturing and Advanced Crystallisation Research Hub. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is an UK EPSRC Future Manufacturing Feasibility Study for advanced pharmaceutical manufacturing, funded by the EPSRC, United Kingdom Future Continuous Manufacturing and Advanced Crystallisation Research Hub [EPSRC grant number: EP/P006965/1].

Data availability

Data will be made available on request.

References

AlAlaween, W.H., Mahfouf, M., Salman, A.D., 2017. Integrating the physics with data analytics for the hybrid modeling of the granulation process. AIChE J. 63 (11), 4761–4773. http://dx.doi.org/10.1002/aic.15831.

Alshafiee, M., AlAlaween, W.H., Markl, D., Soundaranathan, M., Almajaan, A., Walton, K., Blunt, L., Asare-Addo, K., 2019. A predictive integrated framework based on the radial basis function for the modelling of the flow of pharmaceutical powders. Int. J. Pharm. 568, 118542. http://dx.doi.org/10.1016/j.ijpharm.2019.118542.

Andersson, M., Zacché, M., 2019. The advantages of a 'Quality by Design' approach in pharma drug development. Pharma Manuf. http://dx.doi.org/10.13140/RG.2.2.21090.04802, Publisher: Recipharm.

Anon, 2023. International council for harmonisation of technical requirements for pharmaceuticals for human use (ICH). URL https://www.ich.org/.

Audet, C., Bigeon, J., Cartier, D., Le Digabel, S., Salomon, L., 2018. Performance indicators in multiobjective optimization. European J. Oper. Res. 292 (2), 397–422. http://dx.doi.org/10.1016/j.ejor.2020.11.016.

Badman, C., Cooney, C.L., Florence, A., Konstantinov, K., Krumme, M., Mascia, S., Nasr, M., Trout, B.L., 2019. Why we need continuous pharmaceutical manufacturing and how to make it happen. J. Pharm. Sci. 108 (11), 3521–3523. http://dx.doi.org/10.1016/j.xphs.2019.07.016.

Cheng, R., Jin, Y., Olhofer, M., Sendhoff, B., 2016. A reference vector guided evolutionary algorithm for many-objective optimization. IEEE Trans. Evol. Comput. 20 (5), 773–791. http://dx.doi.org/10.1109/TEVC.2016.2519378.

Deb, K., Jain, H., 2014. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints. IEEE Trans. Evol. Comput. 18 (4), 577–601. http://dx.doi.org/10.1109/TEVC.2013.2281535.

Ding, Y., Tiwari, P., Guo, F., Zou, Q., 2022. Shared subspace-based radial basis function neural network for identifying ncRNAs subcellular localization. Neural Netw. 156, 170–178. http://dx.doi.org/10.1016/j.neunet.2022.09.026.

Elshamy, W., Emara, H.M., Bahgat, A., 2007. Clubs-based particle swarm optimization. In: 2007 IEEE Swarm Intelligence Symposium. IEEE, Honolulu, HI, USA, pp. 289–296. http://dx.doi.org/10.1109/SIS.2007.367950.

Emmerich, M., Giannakoglou, K., Naujoks, B., 2006. Single- and multiobjective evolutionary optimization assisted by Gaussian random field metamodels. IEEE Trans. Evol. Comput. 10 (4), 421–439. http://dx.doi.org/10.1109/TEVC.2005.859463.

Fan, Z., Li, W., Cai, X., Li, H., Wei, C., Zhang, Q., Deb, K., Goodman, E., 2019. Push and pull search for solving constrained multi-objective optimization problems. Swarm Evol. Comput. 44, 665–679. http://dx.doi.org/10.1016/j.swevo.2018.08.017.

Filipic, B., Tusar, T., 2016. Visualization in multiobjective optimization. In: Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion. ACM, Denver Colorado USA, pp. 735–751. http://dx.doi.org/10.1145/2908961.2926994.

Filipič, B., Tušar, T., 2018. Visualization in multiobjective optimization. In: Proceedings of the Genetic and Evolutionary Computation Conference Companion. ACM, Kyoto Japan, pp. 858–879. http://dx.doi.org/10.1145/3205651.3207891, URL https://dl.acm.org/doi/10.1145/3205651.3207891.

Fragopoulos, C., Pouliakis, A., Meristoudis, C., Mastorakis, E., Margari, N., Chroniaris, N., Koufopoulos, N., Delides, A.G., Machairas, N., Ntomi, V., Nastos, K., Panayiotides, I.G., Pikoulis, E., Misiakos, E.P., 2020. Radial basis function artificial neural network for the investigation of thyroid cytological lesions. In: Stack, B.C. (Ed.), J. Thyroid. Res. 2020, 1–14. http://dx.doi.org/10.1155/2020/5464787.

Gardner, M.W., Dorling, S., 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmos. Environ. 32 (14–15), 2627–2636, Publisher: Elsevier.

Gbadago, D.Q., Moon, J., Kim, M., Hwang, S., 2021. A unified framework for the mathematical modelling, predictive analysis, and optimization of reaction systems using computational fluid dynamics, deep neural network and genetic algorithm: A case of butadiene synthesis. Chem. Eng. J. 409, 128163. http://dx.doi.org/10.1016/j.cej.2020.128163.

He, C., Zhang, Y., Gong, D., Ji, X., 2023. A review of surrogate-assisted evolutionary algorithms for expensive optimization problems. Expert Syst. Appl. 217, 119495. http://dx.doi.org/10.1016/j.eswa.2022.119495, URL https://linkinghub.elsevier.com/retrieve/pii/S0957417422025143.

Hyer, A., Gregory, D., Kay, K., Le, Q., Turnage, J., Gupton, F., Ferri, J.K., 2024. Continuous manufacturing of active pharmaceutical ingredients: Current trends and perspectives. Adv. Synth. Catal. 366 (3), 357–389. http://dx.doi.org/10.1002/adsc.202301137, URL https://onlinelibrary.wiley.com/doi/10.1002/adsc.202301137.

Ierapetritou, M., Muzzio, F., Reklaitis, G., 2016. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. AIChE J. 62 (6), 1846–1862. http://dx.doi.org/10.1002/aic.15210, URL https://aiche.onlinelibrary.wiley.com/doi/10.1002/aic.15210.

Inselberg, A., 2009. Parallel Coordinates. Springer New York, New York, NY, http://dx.doi.org/10.1007/978-0-387-68628-8.

Jin, Y., 2000. Fuzzy modeling of high-dimensional systems: complexity reduction and interpretability improvement. IEEE Trans. Fuzzy Syst. 8 (2), 212–221. http://dx.doi.org/10.1109/91.842154.

Jović, A., Brkić, K., Bogunović, N., 2015. A review of feature selection methods with applications. In: 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics. MIPRO, pp. 1200–1205. http://dx.doi.org/10.1109/MIPRO.2015.7160458.

Kim, J., Hayashi, Y., Badr, S., Okamoto, K., Hakogi, T., Furukawa, H., Yoshikawa, S., Nakanishi, H., Sugiyama, H., 2023. Hybrid modeling of an active pharmaceutical ingredient flow synthesis in a ring-opening reaction of an epoxide with a grignard reagent. Ind. Eng. Chem. Res. 62 (43), 17824–17834. http://dx.doi.org/10.1021/acs.iecr.3c02137, URL https://pubs.acs.org/doi/10.1021/acs.iecr.3c02137.

Kim, J., Hayashi, Y., Badr, S., Okamoto, K., Hakogi, T., Furukawa, H., Yoshikawa, S., Nakanishi, H., Sugiyama, H., 2024. Kinetic study and model-based design space determination for a drug substance flow synthesis using an amination reaction via nucleophilic aromatic substitution. Org. Process. Res. Dev. 28 (5), 1793–1805. http://dx.doi.org/10.1021/acs.oprd.3c00380, URL https://pubs.acs.org/doi/10.1021/acs.oprd.3c00380.

Koziel, S., Yang, X.-S., Kacprzyk, J. (Eds.), 2011. Computational Optimization, Methods and Algorithms. In: Studies in Computational Intelligence, vol. 356, Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-642-20859-1.

Lee, S.L., O'Connor, T.F., Yang, X., Cruz, C.N., Chatterjee, S., Madurawe, R.D., Moore, C.M.V., Yu, L.X., Woodcock, J., 2015. Modernizing pharmaceutical manufacturing: from batch to continuous production. J. Pharm. Innov. 10 (3), 191–199. http://dx.doi.org/10.1007/s12247-015-9215-8, URL http://link.springer.com/10.1007/s12247-015-9215-8.

Lin, H., Dai, Q., Zheng, L., Hong, H., Deng, W., Wu, F., 2020. Radial basis function artificial neural network able to accurately predict disinfection by-product levels in tap water: Taking haloacetic acids as a case study. Chemosphere 248, 125999. http://dx.doi.org/10.1016/j.chemosphere.2020.125999.

Markarian, J., 2022. Breaking through barriers to continuous manufacturing. Pharm. Technol. 46 (4).

Öner, M., Montes, F.C., Ståhlberg, T., Stocks, S.M., Bajtner, J.E., Sin, G., 2020. Comprehensive evaluation of a data driven control strategy: Experimental application to a pharmaceutical crystallization process. Chem. Eng. Res. Des. 163, 248–261. http://dx.doi.org/10.1016/j.cherd.2020.08.032.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Kevin Tucker, P., 2005. Surrogate-based analysis and optimization. Prog. Aerosp. Sci. 41 (1), 1–28. http://dx.doi.org/10.1016/j.paerosci.2005.02.001.

Reddy, J.P., Phanse, R., Nesarikar, V., 2019. Parameter estimation for roller compaction process using an instrumented vector TF mini roller compactor. Pharm. Dev. Technol. 24 (10), 1250–1257. http://dx.doi.org/10.1080/10837450.2019.1658775.

Riquelme, N., Von Lucken, C., Baran, B., 2015. Performance metrics in multi-objective optimization. In: 2015 Latin American Computing Conference. CLEI, IEEE, Arequipa, Peru, pp. 1–11. http://dx.doi.org/10.1109/CLEI.2015.7360024, URL http://ieeexplore.ieee.org/document/7360024/.

Rokach, L., Maimon, O.Z., 2007. Data Mining with Decision Trees: Theory and Applications, vol. 69, World scientific.

Rubio-Solis, A., Baraka, A., Panoutsos, G., Thornton, S., 2018. Data-driven interval type-2 fuzzy modelling for the classification of imbalanced data. In: Practical Issues of Intelligent Innovations. Springer, pp. 37–51.

Rubio-Solis, A., Panoutsos, G., 2015. Interval type-2 radial basis function neural network: A modeling framework. IEEE Trans. Fuzzy Syst. 23 (2), 457–473. http://dx.doi.org/10.1109/TFUZZ.2014.2315656.

Rubio-Solis, A., Panoutsos, G., 2016. Iterative information granulation for novelty detection in complex datasets. In: 2016 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE, IEEE, Vancouver, BC, Canada, pp. 953–960. http://dx.doi.org/10.1109/FUZZ-IEEE.2016.7737791.

Sagmeister, P., Schiller, C., Weiss, P., Silber, K., Knoll, S., Horn, M., Hone, C.A., Williams, J.D., Kappe, C.O., 2023. Accelerating reaction modeling using dynamic flow experiments, part 1: design space exploration. React. Chem. Eng. 8 (11), 2818–2825. http://dx.doi.org/10.1039/D3RE00243H, URL https://xlink.rsc.org/?DOI=D3RE00243H.

Shang, K., Ishibuchi, H., He, L., Pang, L.M., 2021. A survey on the hypervolume indicator in evolutionary multiobjective optimization. IEEE Trans. Evol. Comput. 25 (1), 1–20. http://dx.doi.org/10.1109/TEVC.2020.3013290, URL https://ieeexplore.ieee.org/document/9153850/.

Steinwart, I., Christmann, A., 2008. Support Vector Machines. In: Information Science and Statistics, Springer New York, New York, NY, http://dx.doi.org/10.1007/978-0-387-77242-4.

Tian, Y., Cheng, R., Zhang, X., Jin, Y., 2017. PlatEMO: A MATLAB platform for evolutionary multi-objective optimization [educational forum]. IEEE Comput. Intell. Mag. 12 (4), 73–87. http://dx.doi.org/10.1109/MCI.2017.2742868.

Urwin, S.J., 2023. Digital process design to define and deliver pharmaceutical particle attributes. Chem. Eng. Res. Des. 196, 726–749. http://dx.doi.org/10.1016/j.cherd.2023.07.003.

Velásco-Mejía, A., Vallejo-Becerra, V., Chávez-Ramírez, A., Torres-González, J., Reyes-Vidal, Y., Castañeda-Zaldivar, F., 2016. Modeling and optimization of a pharmaceutical crystallization process by using neural networks and genetic algorithms. Powder Technol. 292, 122–128. http://dx.doi.org/10.1016/j.powtec.2016.01.028.

Wang, Y., Chen, X., 2020. A joint optimization QSAR model of fathead minnow acute toxicity based on a radial basis function neural network and its consensus modeling. RSC Adv. 10 (36), 21292–21308. http://dx.doi.org/10.1039/D0RA02701D.

Wang, H., Jin, Y., Yao, X., 2017. Diversity assessment in many-objective optimization. IEEE Trans. Cybern. 47 (6), 1510–1522. http://dx.doi.org/10.1109/TCYB.2016.2550502.

Wang, H., Yao, X., 2014. Corner sort for Pareto-based many-objective optimization. IEEE Trans. Cybern. 44 (1), 92–102. http://dx.doi.org/10.1109/TCYB.2013.2247594.

Wu, L., Nauta, M., 2022. A rapid-prototype MPC tool based on gPROMS platform. http://dx.doi.org/10.48550/arXiv.2209.00092, arXiv:2209.00092 [eess].

Wu, K.E., Panoutsos, G., 2021a. A new diversity performance indicator for many-objective optimisation problems. In: 2021 IEEE Congress on Evolutionary Computation. CEC, pp. 144–152. http://dx.doi.org/10.1109/CEC45853.2021.9504903, Journal Abbreviation: 2021 IEEE Congress on Evolutionary Computation (CEC).

Wu, K.E., Panoutsos, G., 2021b. A visualisation method for Pareto front approximations in many-objective optimisation. In: 2021 IEEE Congress on Evolutionary Computation. CEC, IEEE, Kraków, Poland, pp. 1929–1937. http://dx.doi.org/10.1109/CEC45853.2021.9504904.

Yang, S., Li, M., Liu, X., Zheng, J., 2013. A grid-based evolutionary algorithm for many-objective optimization. IEEE Trans. Evol. Comput. 17 (5), 721–736. http://dx.doi.org/10.1109/TEVC.2012.2227145.

Zhang, X., Tian, Y., Jin, Y., 2015. A knee point-driven evolutionary algorithm for many-objective optimization. IEEE Trans. Evol. Comput. 19 (6), 761–776. http://dx.doi.org/10.1109/TEVC.2014.2378512.

Zhou, Z., Li, X., Zare, R.N., 2017. Optimizing chemical reactions with deep reinforcement learning. ACS Central Sci. 3 (12), 1337–1344. http://dx.doi.org/10.1021/acscentsci.7b00492.

Zitzler, E., Thiele, L., 1999. Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. IEEE Trans. Evol. Comput. 3 (4), 257–271. http://dx.doi.org/10.1109/4235.797969.