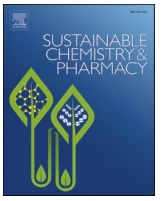
Sustainable Chemistry and Pharmacy 43 (2025) 101874

Contents lists available at ScienceDirect

Sustainable Chemistry and Pharmacy

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



State-enhanced attention network for optimisation of energy and yield in gas atomised metal powder production

Abiodun Ayodeji [a,*], Evelyne El Masri [a], Tom Williamson [b], Mohmmad Ali Asgar Abbas [a], Tat-Hean Gan [a]

[a] Brunel Innovation Centre, Brunel University London, Uxbridge, UB8 3PH, UK
[b] Atomising Systems Limited, 371 Coleford Rd, Darnall, Sheffield, S9 5NF, UK

ARTICLE INFO

Handling Editor: Borhane Mahjoub

Keywords:
Gas atomisation
Metal powder production
Artificial intelligence
Attention network
Process optimisation
Additive manufacturing

ABSTRACT

Gas atomisation is a widely used technique for producing spherical metal powder feedstock for additive manufacturing. However, the process parameters suffer from variability and inefficiency in balancing powder yield, energy consumption, and particle size distribution. Optimising these complex, interdependent parameters pose a significant challenge. This work proposes a novel State-Enhanced Attention Network architecture in a framework that simultaneously optimises yield and energy consumption during nitrogen gas atomisation for sustainable metal powder production. The novelty lies in integrating processed long-term memory states with the attention mechanism, enabling nuanced attention weighting. This allows the model to leverage global sequence context and recent state information for improved yield and energy predictions. The proposed network is trained and integrated into a non-dominated sorting genetic algorithm to enable multi-objective optimisation. This framework evolves a set of Pareto-optimal solutions that balance trade-offs between maximising yield and minimising energy consumption. The approach is evaluated using augmented real-world data from an industrial gas atomisation plant. The proposed model demonstrates significantly improved predictive accuracy on real-world datasets, compared with baseline deep learning models. Results highlight the capabilities of the proposed technique for automated, data-driven optimisation of gas atomisation, simultaneously improving yield, energy efficiency, quality control, and sustainability. The integrated deep learning and evolutionary optimisation framework also provides an innovative solution for enhanced control of additive manufacturing powder production processes.

1. Introduction

Gas atomisation is a widely used method for producing metal powder for additive manufacturing (AM). This process involves the conversion of molten metal into a spray of droplets which solidify into powder. In this process, molten metal alloys are first refined and degassed, then poured into a gas nozzle (Singh and Verma, 2018). A high-pressure gas stream passing through the gas atomisation nozzle impinges upon a molten metal stream, breaking it up and quenching the molten stream to form metal powder (Munyaka and Yadavalli, 2022). This process is ideal for AM due to the geometrical properties of the powder particles that can be achieved (Singh and Verma, 2018).

* Corresponding author.
E-mail address: ayod_abe@yahoo.com (A. Ayodeji).

https://doi.org/10.1016/j.scp.2024.101874
Received 22 July 2024; Received in revised form 23 September 2024; Accepted 28 November 2024
Available online 5 December 2024

One of the key characteristics of powder suitable for AM is their sphericity. During the free fall inside a tower in the gas atomisation process, the surface tension of the metal pulls the drop into a sphere (Munyaka and Yadavalli, 2022). To protect the metal from oxidation, the atomising gas - usually nitrogen or argon - is often used as the atomising medium due to its high stability and non-reactivity with most metals.

As opposed to water atomisation, plasma or induction-based process, the gas atomising process is scalable and applicable to a wide range of alloys including reactive materials. However, the gas atomising process has low process efficiency in terms of powder yield, requiring screening to remove oversized particles. Moreover, gas atomisation has process parameters which affect the deposition quality, powder properties, particle size distribution, powder yield and energy consumption, and the process offers less control over particle morphology compared to other processes. This necessitates careful control of process parameters to ensure optimal and sustainable additive manufacturing powder production.

To produce high-quality gas-atomised powder at low cost often involves complex experimental trial and error. The optimal process parameters vary based on the type of superalloy used, the specific manufacturing equipment, and environmental conditions (Tamura et al., 2021). Consequently, significant resources are required to fine-tune the process for a specific application. Furthermore, there can be significant variability in the physical properties of metal powder from different suppliers, even when the chemical composition is the same, (Dawes et al., 2015). Hence there is a need for a tailored tool that enables process optimisation and quality control during gas atomisation to achieve consistent powder properties.

Machine learning presents a promising approach, as it can rapidly analyse large multidimensional datasets from sensors and instrumentation to develop predictive models for powder characteristics. Implementing machine learning-based process monitoring and control can minimize the variability in metal powder properties regardless of production scale. Previously, several attempts have been made to optimise the gas atomising process using data-driven tools. Tamura et al. (2021) present the Bayesian optimisation technique to determine the melt temperature and gas pressure that improves the additive manufacturing powder yield. Focusing on powder sizes that are less than 53 μm, the method improves the yield of molten Ni–Co-based super-alloy from 30% to 70%. However, the work did not discuss the energy utilised to achieve such an increase in yield, which questions the sustainability of the process. Optimising energy consumption contributes to more sustainable production practices, reducing the environmental footprint of the gas atomising process.

Shang et al. (2023) propose a multi-objective optimisation strategy to refine the process parameters in the vertical zone refining ultra-high purity indium, with machine learning. The authors utilise the Synthetic Minority Over-sampling Technique (SMOTE) technique to augment the dataset and overcome the sample imbalance problem of the experimental dataset. Then a K-nearest neighbor model is used to predict the total impurity content in the product.

Machine learning (ML) techniques have also been used in other additive manufacturing applications. To evaluate the overall quality of the AM building part (Gor et al., 2022), applied the support vector machine to predict the density in Powder Bed Fusion Additive Manufactured components. Several machine learning algorithms were examined by Eshabilov et al. (Eshkabilov et al., 2022) to characterize the effects of the printing process parameters on relative density, hardness, yield strength, and tensile strength in manufactured parts. Although the work presents a useful application of ML models, the goal is for part quality prediction. An application of machine learning to aid in the design of high-entropy alloys, and common ML workflow for alloy production optimisation is discussed in Liu et al. (2022).

The reviewed literature on ML application in additive manufacturing presents innovative approaches for process optimisation. However, most of the proposed methods are for optimising a single objective. For instance, most of the literature optimises for yield without considering the effect on energy. In a gas atomising process where yield and energy are largely dependent, it is crucial to simultaneously capture the trade-offs and relationship between the two to find a balanced solution that optimises both. In addition, the literature surveyed shows that linear models are the conventional tool for most of the optimisation tasks. However, while linear models provide a simple starting point, they may fail to capture nonlinear relationships and interactions between process variables. The gas atomisation process likely involves complex physics with nonlinear effects that cannot be adequately modelled by linear regression. To enable more flexible modelling and capture intricate parameter relationships, advanced nonlinear techniques like deep learning (DL) models need to be explored. DL models can learn complex mapping from data without relying on rigid linear assumptions.

Further, the literature lacks a robust framework for data augmentation that allows exploring a wider range of process parameters beyond the original limited data. By generating synthetic samples across expanded ranges, the model could be used to predict outcomes for novel parameter settings not present in the original data. This enables more comprehensive process optimisation while reducing the cost of running experiments. This would also mitigate overfitting, as ML models are prone to overfitting with limited original samples.

Moreover, the atomisation efficiency also relies heavily on nozzle design, as nozzle setup plays a pivotal role in determining the quality and yield of the final product in the gas atomising process of metal powder production (Anderson and Terpstra, 2002). Balancing the environmental impact with the industrial necessity of gas atomisation presents a complex challenge. This underscores the need for innovative approaches to reduce energy use and CO_2 emissions in this process, without compromising the quality of the atomised metal powder. Hence, there is a need for studies that move beyond generic process optimisation to targeted enhancements for specialised powder-process applications (Lewis, 2022).

Towards sustainable production of additive manufacturing powder, this work proposes a novel deep learning framework to address key limitations of prior optimisation attempts for the gas atomisation process. Specifically, this work develops a tailored deep learning architecture called State-Enhanced Attention Network (SEAN). This hybrid model combines Convolutional Neural Networks (CNNs) to extract informative spatial features, Long-short Term Memory (LSTM) layers to learn temporal dynamics, and attention mechanisms to focus on relevant inputs. The key innovation is the injection of LSTM state outputs into the attention module. This state enhancement

provides the attention mechanism with both local sequential context from the LSTM outputs as well as global dependencies from the full cell state. The integrated state representations also enable more nuanced attention weighting to improve predictive performance.

The SEAN architecture was selected as it combines complementary techniques that are well-suited for modelling the gas atomisation process. SEAN is trained on augmented real-world data from a gas atomising process. To obtain optimal parameter settings for the atomising process, the Non-dominated Sorting Genetic Algorithm II (NSGA-II) is utilised to simultaneously optimise for the required yield and energy by selecting optimal parameter settings. This work contributes to knowledge by.

1. Proposing a novel deep learning architecture called State-Enhanced Attention Network (SEAN)
2. Using the SEAN architecture in a framework that simultaneously optimises the yield and energy consumption in gas-atomised metal powder production. The framework is also tailored for a particular nozzle setting.
3. Employing an innovative data augmentation technique to expand the dataset and train models on a wider parameter space. This resolves data quantity and variability issues common in real-world industrial processes.
4. Implementing a deep learning-based multi-objective optimisation with NSGA-II to find optimal trade-offs between competing objectives of maximising yield and minimising energy.

The goal is to develop an accurate deep-learning-based optimisation tool for the gas atomisation process by leveraging advanced machine learning, to enable robust optimisation and control of the process. To be deployed in a data analytics platform, the proposed technique is an innovative framework useful for process understanding and optimisation in gas-based metal powder production with fewer real-world datasets.

2. Background

2.1. 1-D convolutional units

A one-dimensional convolutional neural network (1D CNN) is designed to process one-dimensional sequential data such as time series or text. It applies convolutional filters to the input sequence to extract informative features. 1D CNNs are a specialised tool for extracting informative features from sequential data (Ayodeji et al., 2022). In contrast to 2D convolutions commonly used for images, 1D CNNs are designed for efficient processing of one-dimensional input signals. 1D CNNs apply convolutional filters to the input sequence, learning to detect patterns and meaningful features. The filters slide across the sequence, computing dot products between the filter values and local regions of the input. Multiple filters are applied to extract multiple feature maps, each detecting a different pattern.

Compared to 2D CNNs, 1D CNNs have reduced computational complexity and are more suitable for limited, low-dimensional sequential data. They train faster and generalise better given the constrained 1D input space. Moreover, key advantages of 1D CNNs include efficient feature extraction from 1D sequences capturing local dependencies and orderings, layerwise abstraction learning both low and high-level features, high performance on tasks with limited training data, ability to handle high signal variation and noise robustly, and low computational requirements allowing real-time and low-cost applications (Kiranyaz et al., 2021). Mathematically, a 1D convolution operation involves applying a filter $w \in R^K$ of length K to a 1D input $x \in R^N$ to produce a feature map $z \in R^{N-K+1}$. This is formulated as:

$$z[i] = b + \sum_{\{j=0\}}^{\{K-1\}} x(i+j)w(j) \tag{1}$$

Where b is a bias term. The filter w slides across the input x to compute the dot product between w and each *K-gram* from x. The convolutional filters extract local patterns from the sequence, acting as feature detectors. Multiple filters can be learned to obtain multiple feature maps. After the convolutional layer, additional operations are commonly applied, such as padding the input, applying rectified linear activation or down sampling with pooling to reduce dimension. In addition, stacking convolutional layers allows the network to learn hierarchical feature representations. The convolutional encoding extracts spatial features that retain sequence order and local context.

2.2. Long-short term memory network

Long-short-term memory networks (LSTMs) are a type of recurrent neural network (RNN) that are well-suited for learning long-term dependencies in sequential data. The key component of an LSTM is the cell state Ct, which acts as a memory to store relevant information from previous timesteps. This allows the network to preserve information over long sequences. The LSTM also contains input, output and forget gates, which regulate the flow of information into and out of the cell state. The input gate controls what new information is stored. The output gate controls what is outputted. The forget gate adapts by removing or retaining information from the cell state. These gating mechanisms allow LSTMs to overcome the vanishing gradient problem faced by standard RNNs. By learning to open and close access to the cell state, LSTMs can model longer context and handle exploding or vanishing gradients. The LSTM is mathematically defined as (Wang et al., 2021):

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{2}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\widetilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \tag{5}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t * \tanh(C_t) \tag{7}$$

Where f_t is the forget gate's activation, i_t is the input gate's activation , $\widetilde{C}_t$ is the candidate cell state, C_t is the new cell state, o_t is the output gate's activation, h_t is the new hidden state, and σ denotes the sigmoid function. W *and* b represent the weights and biases for their respective gates. The combination of cell state and gating mechanisms allow LSTMs to learn complex sequential dependencies and patterns that would be inaccessible to standard RNNs. This makes them ideal for processing long sequences of data such as text, audio, and time series.

2.3. Attention mechanism

The Attention layer allows models to focus selectively on parts of the input that are most relevant to the task or prediction. Since its introduction for neural machine translation, it has been widely used in tasks such as image segmentation, object detection and natural language processing. The key motivation is that when processing long sequences, not all inputs may be equally important. Attention provides a way to extract the most salient information and give appropriate weight to the inputs. Mathematically, attention computes alignment scores between the target output and each input element (e.g. encoder hidden states). Given a target output y_t, the alignment score $\alpha_{\{ti\}}$ for each input element h_i is computed as:

$$\alpha_{ti} = \text{align}(y_t, h_i) \tag{8}$$

The alignment scores are normalized using the softmax function to obtain attention weights:

$$a_{ti} = \frac{\exp(\alpha_{ti})}{\sum_k \exp(\alpha_{tk})} \tag{9}$$

The context vector c_t is then computed as the weighted sum of the input elements based on these alignments:

$$c_t = \sum_i a_{ti} * h_i \tag{10}$$

The context vector summarizes the relevant source elements for predicting the target. This context vector is concatenated or combined with the target output y_t to get the attentional representation used for prediction. Different alignment functions align() can be used, including dot product, bilinear, additive and multi-layer perceptron functions. The choice depends on computational efficiency and performance.

Attention thus allows dynamic focusing on relevant parts of potentially long input sequences when generating each output. It has become a very effective technique in RNNs/LSTMs for sequence modelling and transduction problems like machine translation, speech recognition, and text summarization. Variants of attention networks, including multi-head attention (Ayodeji et al.), self-attention (Ayodeji et al.), and hierarchical attention (Bi et al., 2023) have been used to learn important information in time series, and sequential datasets. This implementation was found to provide optimal results, which serve as a motivation for the current work.

3. Method

3.1. SEAN model architecture

The state-enhanced attention network (SEAN) model is a hybrid neural network architecture that combines convolutional neural networks (CNNs), long short-term memory networks (LSTMs), and attention mechanisms with an innovative approach to integrating LSTM states in the attention module. This model is designed to capture both spatial and temporal relationships in the data, as well as to highlight the importance of certain time steps within a sequence. This section discusses the fundamental components and the mathematical formulation of the model.

3.1.1. Convolutional layers

The model begins with two one-dimensional convolutional layers, each followed by a rectified linear unit (*ReLU*) activation function. The input data is passed through two 1D convolutional layers with 128 filters, kernel size 2, and *ReLU* activation. Convolutional layers extract local patterns and motifs from the input time series. The convolutions also help detect low-level features like trends, seasonality, etc. Using small kernel size and multiple layers allows learning hierarchical representations. *ReLU* activation introduces non-linearity.

Let $X \in R^{T\times F}$ denote the input matrix for a single sample, where T is the number of time steps and F is the number of features. The convolutional layers transform the input as follows:

$$X_{conv} = \mathrm{ReLU}(W_{conv} * X + b_{conv}) \tag{11}$$

where * denotes the convolution operation , W_{conv} represents the convolutional filters, and b_{conv} is the bias term.

3.1.2. The LSTM layers

The output of the convolutional layers is reshaped and fed into a stack of LSTM layers. Each LSTM layer processes the data sequentially and captures long-term dependencies. The fourth LSTM layer is special as it returns both the sequence output and the last time step's hidden and cell states, denoted by $\mathrm{lstm\_out}, \mathrm{lstm\_h}, and\ \mathrm{lstm\_c},$ respectively. The LSTM's cell state carries long-term context information about the entire input sequence. This provide useful global sequence dependencies to the attention layer. LSTM's cell state and hidden state act as memory to capture contextual information. The cell state flows through the entire sequence while the hidden state captures the output. The model uses LSTMs gates (input, output, forget) to modulate information flow and mitigate vanishing gradients. Multiple layers allow learning more complex temporal dynamics.

3.1.3. State enhancement and attention mechanism

The model further processes the hidden and cell states through dense layers to obtain refined state representations *state_h and state_c*:.

$$state\_h = Dense_{64}(lstm\_h) \tag{12}$$

$$state\_c = Dense_{64}(lstm\_c) \tag{13}$$

The hidden state captures the LSTM's output for the last time step, providing insight into the most recent state. These states are processed through a Dense layer to reduce their dimensionality to 64 units - the same as the attention vector size. These processed states are then dynamically repeated to match the sequence length of the LSTM output using Lambda layers, facilitating their integration into the attention mechanism. An attention layer is added after the LSTMs to focus on relevant parts of the input sequence for predicting the outputs. It computes a query, key, and value vector using dense layers on the LSTM outputs. The dot product of query and key determines the attention weights, which are applied to the values to get the attention vector. This acts as a weighting mechanism to amplify important parts of the sequence. The attention mechanism is applied to the LSTM sequence output:

$$attn = Attention(Dense_{64}(lstm\_out), Dense_{64}(lstm\_out), Dense_{64}(lstm\_out)) \tag{14}$$

The Attention function computes a weighted sum of the values $((v)),$ where the weight assigned to each value is computed by a compatibility function of the query $((q))$ with the corresponding key $((k))$. The model concatenates the LSTM output, attention output, and repeated processed states, flattens the combined features, and passes them through separate dense layers to produce two outputs: yield and energy.

$$\mathrm{output\_yield} = \mathrm{Dense}_1(\mathrm{Flatten}(\mathrm{Concatenate}([\mathrm{lstm\_out}, \mathrm{attn}, \mathrm{state\_h}, \mathrm{state\_c}]))) \tag{15}$$

$$\mathrm{output\_energy} = \mathrm{Dense}_1(\mathrm{Flatten}(\mathrm{Concatenate}([\mathrm{lstm\_out}, \mathrm{attn}, \mathrm{state\_h}, \mathrm{state\_c}]))) \tag{16}$$

To align the state vectors with the attention output, they are tiled/repeated to match the sequence length using a Lambda layer. The key novelty in this algorithm lies in the integration of the processed LSTM states with the attention mechanism. By dynamically repeating the processed states to match the sequence length and concatenating them with the LSTM output and attention output, the model can leverage both the global sequence context and the recent state information to make more accurate predictions of yield and energy.

3.1.4. Loss function and optimisation

The model is compiled with Mean Squared Error (MSE) loss functions for both outputs and optimised using the Adam optimiser. The model also tracks Mean Absolute Error (MAE) as an additional performance metric during training.

$$\mathrm{Loss} = \mathrm{MSE}\left(\mathrm{output\_yield}, y_{\mathrm{true\_yield}}\right) + \mathrm{MSE}\left(\mathrm{output\_energy}, y_{\mathrm{true\_energy}}\right) \tag{17}$$

The proposed *SEAN* model captures complex relationships in the data by effectively combining spatial feature extraction, temporal dynamics learning, and state-enhanced attentional focusing, aiming to improve predictive performance on tasks with both spatial and temporal dimensions. The model concatenates the LSTM output, attention output, and repeated processed states, flattens the combined features, and passes them through separate dense layers to produce two outputs: yield and energy. By concatenating the processed states with the attention output, the model can utilise both focused local information from attention, and holistic sequence/state context. Overall, this state injection technique allows the attention mechanism to be enhanced with relevant state information spanning long-range dependencies as well as the most recent output. This adds a useful signal for the final prediction task.

3.2. Non-dominated sorting genetic algorithm II (NSGA-II)

The SEAN model is integrated into the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to significantly augment the gas atomising process parameter optimisation. NSGA-II is a benchmarked multi-objective optimisation algorithm that is well-suited for finding a set of solutions that balance between different trade-offs—in this case, yield and energy. NSGA-II is an evolutionary algorithm that uses principles of selection, crossover, and mutation inspired by natural evolution to evolve a population of solutions towards the Pareto front over generations. Specifically, the objective function is a tuple of two elements, each representing one of the objectives to be optimised.

1. The absolute difference between the predicted yield and the required yield.
2. The absolute difference between the predicted energy consumption and the required energy consumption.

Mathematically, the objective function can be expressed as follows:

$$\text{Objective Function} = \left(\left|Y_{\text{predicted}} - Y_{\text{required}}\right|, \left|E_{\text{predicted}} - E_{\text{required}}\right|\right) \tag{18}$$

Where $Y_{\text{predicted}}$ is the model's predicted yield based on the input parameters, Y_{required} is the user-specified required yield, $E_{\text{predicted}}$ is the model's predicted energy consumption based on the input parameters, and E_{required} is the user-specified required energy consumption. The NSGA-II algorithm seeks to minimize both elements of this tuple, striving for a solution that closely matches the user's specified yield and energy requirements. The weights in the fitness function indicate that both are to be minimized in the multi-objective optimisation context.

The output of the optimisation process is a Pareto front of solutions that represent the best trade-offs between the yield and energy objectives. Users can review these solutions and select the most appropriate one according to their specific criteria or constraints. This approach allows for a sophisticated decision-making process that is informed by advanced modelling and optimisation techniques, demonstrating a practical application of NSGA-II in enhancing manufacturing processes.

In the optimisation framework, the NSGA-II is used to tune the spray nozzle parameters (temperature, gas flow, pressure, metal flow rate) to meet the required yield and energy constraints. The NSGA-II algorithm evolves the population over generations to approach the Pareto front of optimal combinations of parameters that satisfy the targets. The top solutions provide tunings for the process variables that can achieve the desired yield and energy requirements. Fig. 1 below shows the full graph of the proposed SEAN and Fig. 2 shows the development flowchart of the optimisation framework. Algorithmically, the NSGA-II with the proposed SEAN

Fig. 1. Full graph of the proposed SEAN.

works as described in Table 1.

3.3. Dataset description

The data used in evaluating the proposed SEAN is obtained from Atomising Systems Limited's industrial gas atomising process plant. At the metal powder processing plant, metal is melted in coreless induction furnaces in an air atmosphere and then poured into the atomiser. This uses high-pressure nitrogen gas to form a spray of molten metal, which freezes and is collected via a cyclone. The cyclone separates the gas and solid particles, and nitrogen is discharged through a reverse pulse jet filter.

The energy consumed by atomiser components is captured by power meters, and this data is fed into the PLC where it is retrieved for analysis. The gas consumption is captured by the PLC whilst atomising, and from this volume, an energy consumption figure is obtained by taking the integral of the power over time for each component in the process. Using a particular nozzle design (referred to as Nozzle 1 for confidentiality reasons), 10x 316L runs were completed on the large gas atomiser, using typical settings for the AM powder size ($-53 + 20$ μm). The gas pressure was set to 28 bar, and the gas temperature was set to 600 °C (Williamson and Mellor, 2023). Each run followed the same material charging pattern and had the same target temperature to pour at. The data for each run is subsequently collected in a CSV file, with the yield and other relevant parameters. A total of 58 data points is obtained through the process.

3.4. Data processing and analysis

The original dataset consists of 58 data points representing process parameter settings, power consumption metrics, and powder yield for the nozzle 1 setup. To compute the total energy consumption per kilogram of powder (Total AM kWh/kg), the energy usage is divided into 3 main parts (Williamson and Mellor, 2023).

1. Melting energy: This represents the initial period starting from when power is applied to the furnace until the atomiser begins operating. During this stage, the metal charge is melted and heated to the pouring temperature. Approximately 80% of the power consumed in this period stems from the induction furnace, while the remainder is from ancillary equipment like cooling circuits and idle gas heaters.
2. Atomising energy: This encompasses the period from atomiser start-up until shutdown of the furnace and atomiser, marking the end of one cycle. The power consumption includes the furnace holding power, recirculation fan, active gas heaters, cooling circuits, and other ancillary equipment.
3. Nitrogen energy: The energy required to produce the nitrogen gas supply is calculated separately from atomising energy to better understand its impact on efficiency. A conversion factor of 0.69 kWh/m^3 is utilised to estimate nitrogen energy based on flow rates.

To calculate the metered energy for melting and atomising periods, a specialised preprocessing script was developed. The script

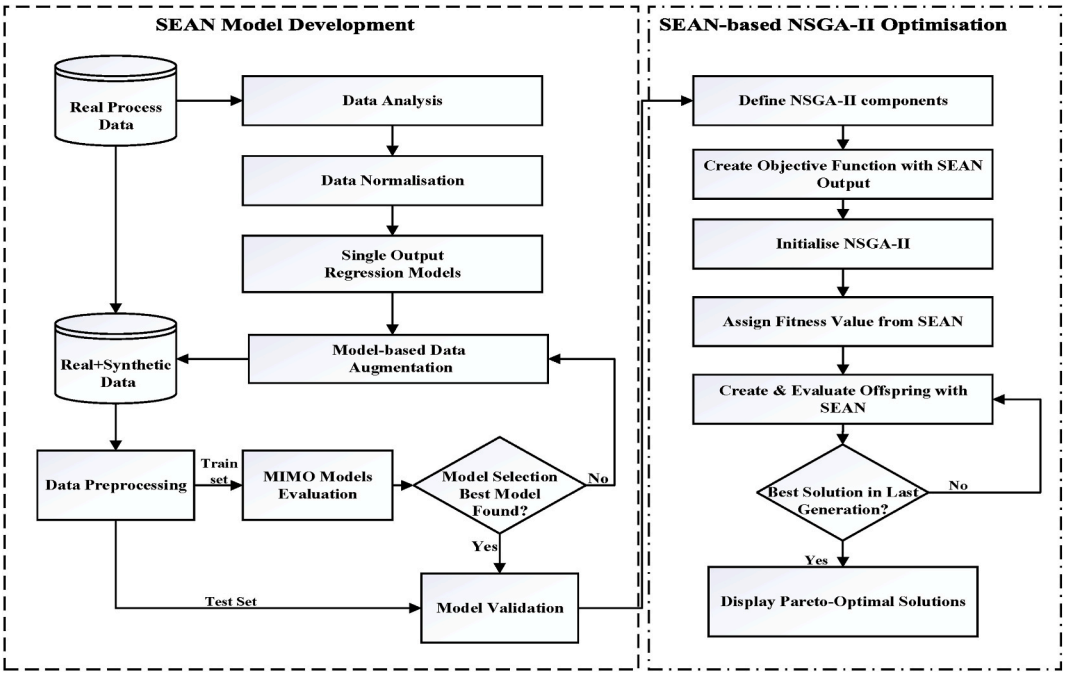


Fig. 2. The flowchart of the optimisation framework.

Table 1
The SEAN with NSGA-II algorithm.

```
Input: pre-trained SEAN, scale, CalcucluateObjectives; ErrorHandling
Parameters: NSGA2AttrsDefined; MaxGenerations; bounds
Output: population; topSolutions; metrics; obj1, obj2
If !NSGA2AttrsDefined:Define NSGA2Attrs
Create NSGA2Toolbox()
population < - CreateInitialPopulation()
For x in population:
  xScaled = Scale(x)
  yPred = Predict(SEANModel, xScaled)
  obj1, obj2 = CalculateObjectives(yPred)
  Evaluate(x, obj1, obj2)
For i ← 1 to MaxGenerations:
  parents < - SelectParents(population)
  offspring < - Crossover(parents) + Mutate(offspring)
  For x in offspring:
    xScaled = Scale(x)
    yPred = Predict(SEANModel, xScaled)
    obj1, obj2 = CalculateObjectives(yPred)
    Evaluate(x, obj1, obj2)
If offspring > bounds:
  offspring < - Clip(offspring, bounds)
  population < - Select(offspring + population)
  hallOfFame.Update(population)
  topSolutions < - GetTop(hallOfFame, n = 3)
  For x in topSolutions:
    xScaled = Scale(x)
    yPred = Predict(SEA Model, xScaled)
    Display(yPred)
  Display(topSolutions, metrics)
Else:
  NoOp()
```

parses time stamps in the data to compute the duration between rows. It then calculates total energy as the product of power and elapsed time for all power-related variables. Statistical measures like mean and standard deviation are also computed for non-power variables. The consolidated energy consumption data is combined into a final normalised data frame. The first 10 representative data points and associated variables used for SEAN training are presented in Table 2.

3.5. Data augmentation

This section describes the process undertaken for model training, synthetic data generation, and subsequent prediction of yield and energy consumption using machine learning models. The primary objective was to develop predictive models capable of estimating yield (20–53 μm yield, %) and energy consumption (Total AM kWh/kg) based on key input parameters shown in Table 2 above.

Due to the high cost of additional experiments, only 58 data points collected from the gas atomisation process were initially available. To ensure consistency across input features, the selected data was scaled using the *StandardScaler*, function in *ScikitLearn* library, which normalized the features to have zero mean and unit variance. This pre-processing step is essential to improve model performance, particularly for algorithms that are sensitive to feature scaling. The datapoint is then used to train and evaluate five conventional machine learning models. The performance of five machine learning models was evaluated for the simultaneous prediction of two target variables: yield (20–53 μm yield, %) and energy consumption (Total AM kWh/kg). The models tested include K-Nearest Neighbors (KNN), Linear Regression, Random Forest, XGBoost, and Support Vector Regression (SVR). Each model's performance was assessed based on Mean Squared Error (MSE) and the coefficient of determination (R^2), with train-test split being 70/30.

Table 2
Parameters used for model evaluation.

Run No	Temperature (^{0}C)	Gas Flow (m^3/min)	Pressure (Bar)	Metal Flow Rate (kg/min)	20–53 μm Yield (%)	Total AM (kWh/kg)
S235420	551.65	17.82	25.74	15.32	40.19	5.46
S235421	553.22	17.40	25.44	16.99	39.54	4.91
S235422	547.47	17.80	25.70	14.04	38.69	5.63
S235423	550.77	17.82	25.80	15.46	34.81	5.68
S235424	553.65	17.90	25.76	14.82	35.99	5.70
S235425	544.77	18.48	25.58	11.87	38.09	6.32
S235426	550.71	18.66	25.81	11.38	40.36	6.12
S235427	552.33	18.46	25.61	11.67	40.20	6.21
S235428	525.57	18.31	25.45	12.50	39.24	6.45
S235429	534.09	18.32	25.40	12.50	40.60	5.63

The evaluation result of the model is as shown in Table 3 below.

As shown in Table 3 above, for yield prediction, the models demonstrated generally poor performance, with all models yielding near-zero MSE values, suggesting minimal variance in the predicted outcomes. However, the R^2 values ranged from −0.24 to 0.11, indicating the models explained little to none of the variance in yield, with Linear Regression being the most effective ($R^2 = 0.11$) and KNN and XGBoost showing negative R^2 values, reflecting poor generalisation.

For energy prediction, the models exhibited improved performance. The Random Forest model performed the best, achieving an MSE of 0.07 and an R^2 of 0.41, indicating moderate predictive power. SVR and XGBoost also performed reasonably well with R^2 values of 0.40 and 0.38, respectively. In contrast, Linear Regression demonstrated the weakest performance for energy prediction, with an R^2 of −0.97, suggesting a significant inability to capture the relationship between the input features and energy consumption.

Overall, the results indicate that the models struggled to predict yield accurately, while energy prediction showed more promise, particularly with non-linear models like SVR and Random Forest. Direct regression modelling on this limited real data resulted in poor performance for predicting yield and energy consumption because the real data was too small. Also, the real data exhibited low variability in the parameter ranges of pressure, temperature, flow rate, and metal flow rate, as shown in Table 4.

Consequently, to address the paucity of data, synthetic data generation was employed to expand the dataset. First, the five conventional machine learning models are retrained and re-evaluated to predict one output (yield or energy) separately, as opposed to predicting the two outputs simultaneously. This results in a significantly improved performance for Linear Regression and Random Forest models with the Linear Regression model having MSE and R^2 values of 0.03 and 0.75 respectively for yield prediction, and the Random Forest having MSE and R^2 values of 0.01 and 0.81 respectively.

To generate the synthetic dataset, the parameter range for each input is first expanded as shown in Table 4, based on domain knowledge and practical considerations. This expanded sampling allowed exploration of a wider parameter space. Then Random values were generated for each of the selected features within the predefined ranges. Specifically, the nozzle temperature was randomly sampled between 500 °C and 600 °C, flow rate between 15 and 24 m^3/min, nozzle pressure between 22 and 28 bar, and metal flow rate between 8 and 18 kg/min as shown in Table 4. These ranges were selected to simulate realistic conditions that might be observed in the manufacturing environment. A total of 10,000 synthetic data points is generated.

The synthetic dataset is then scaled using the same *StandardScaler* applied to the training data to maintain consistency. Once scaled, the best performing machine learning model for each target output was utilised to predict additional synthetic data points. The trained models were used to predict either yield or energy consumption for each synthetic data point. The Linear Regression model predicted the yield (20–53 μm yield, %), while the Random Forest Regressor provided predictions for energy consumption (Total AM kWh/kg). This process enabled the generation of synthetic outputs, which is then analysed to assess potential trends and behaviours in the absence of additional real-world data. The synthetic dataset, now containing both the predicted yield and energy values, was exported to a CSV file for further analysis. The synthetic dataset was also verified by a domain expert to ensure consistency with the real-world metal powder production data.

In total, 10,000 synthetic data points were generated through this approach. These were combined with the original 58 real data points, resulting in an augmented dataset of 10058 points. This dataset exhibited greater variability in the input parameters and output targets. The integration of real and synthetically generated data was performed to enable more robust modelling and optimisation while avoiding the need for additional expensive physical experiments.

The synthetic with original samples act to regularise the model, reducing overfitting and improving robustness. This allowed exploring a greater parameter space beyond the narrowly constrained original data. The synthetic data also allows advanced models to be trained on a larger and more varied data distribution, facilitating robust multi-objective optimisation across a wider process range. The resulting augmented dataset was used to evaluate the proposed SEAN model for predicting yield and energy consumption simultaneously. To evaluate the performance of the proposed model and test for overfitting, the SEAN model was compared to other deep learning architectures. Since traditional regression models assume linearity and independent residual, the SEAN model is compared with other conventional deep learning models.

4. Result and discussion

This study employs the Keras API on Tensorflow to develop the SEAN model. Experiments are conducted on *Intel Core i9-12900HK*, with 64G RAM. Additionally, comparison experiments are carried out on the cloud-based platform, Google Colab. The real-world gas atomisation process data and synthetic data generated through regression models are concatenated to create an augmented dataset with greater variability. Pre-processing steps of standardization and reshaping are applied to prepare the data for modeling. The

Table 3
Conventional ML model performance on multi-output prediction.

Model	Yield Prediction		Energy Prediction	
	MSE	R2	MSE	R2
KNN	0.00	−0.24	0.08	0.29
Linear Regression	0.00	0.11	0.22	−0.97
Random Forest	0.00	0.05	0.07	0.41
XGBoost	0.00	−0.17	0.07	0.38
SVR	0.00	−0.02	0.06	0.40

Table 4
Real and synthetic data parameter range.

Data source	Pressure (bar)		Temperature (°C)		Flow (m^3/min)		Metal flow rate (kg/min)	
	Min	Max	Min	Max	Min	Max	Min	Max
Real	24.20	25.83	511.73	553.65	17.34	22.16	8.7	16.99
Synthetic	22	28	500	600	15	24	8	18

combined dataset is split into training and test sets (70/30). Multiple deep learning models are defined and trained, including vanilla convolution neural networks (CNN), long-short-term memory network (LSTM), attention-based long-short-term memory and the proposed SEAN architecture, as described in Section 3. To ensure fairness, all models are trained using a similar optimiser.

4.1. Evaluation metrics

The performance of predictive models in the gas atomising process was quantitatively assessed using three widely recognized statistical metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination, R-squared (R^2). These metrics provide a comprehensive evaluation of model accuracy, consistency, and predictive power.

4.1.1. Root mean square error (RMSE)

RMSE is a quadratic scoring rule that measures the average magnitude of the error. It's the square root of the average of squared differences between prediction and actual observation. Mathematically, it is represented as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2} \tag{19}$$

where y_i is the actual value, $\widehat{y}_i$ is the predicted value from the model, and n is the number of observation. The squaring of errors before averaging ensures larger errors are emphasized and provides a metric that is in the same units as the predicted value.

4.1.2. Mean absolute error (MAE)

MAE measures the average absolute difference between predicted values and actual values, offering a linear score that represents the average magnitude of the errors in a set of predictions, without considering their direction:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y}_i| \tag{20}$$

Where $|y_i - \widehat{y}_i|$ denotes the absolute value of the difference between the actual and the predicted values. Unlike RMSE, MAE gives an equal weight to all errors, which provides a direct indication of the average prediction error.

4.1.3. The coefficient of determination (R^2)

R^2, also known as the coefficient of determination, indicates the proportion of the variance in the dependent variable that is predictable from the independent variables:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{21}$$

where $\overline{y}$ is the mean of the actual values y_i. R^2 values range from *-inf* to 1, where 1 indicates perfect prediction, 0 indicates that the model predicts as well as the mean of the actual values, and negative values means the model is worse than the mean of the actual values. A higher R^2 score represents a model that more closely fits the observed data. These metrics were computed to evaluate the SEAN model's predictive performance on a test set. RMSE and MAE offer a direct interpretation of model error in terms of the unit of measurement, providing insights into the average prediction error and variability. Meanwhile, R^2 contextualizes the model's predictive power by comparing its performance to a baseline mean model. Together, these metrics facilitate a comprehensive assessment of the model's predictive accuracy and its suitability for optimising the gas atomising process parameters.

Each model is fitted on the appropriate training data. All models evaluated are trained for 700 epochs with a batch size of 32 and 20% validation split. Model evaluation is performed by making predictions on the test set and calculating performance metrics such as mean squared error, R^2, root mean squared error, and mean absolute error for both the yield and energy consumption targets. To avoid overfitting, the *early_stopping* and *checkpoint* callback functions are implemented. The *early_stopping* function monitors the validation loss and stops model training once the loss stops improving after 5 epochs, and the *checkpoint* saves the best model at the epoch with the validation accuracy. The *Adam* optimiser and a fixed learning rate of 3e-5 were used for all experiments. To show the performance of the proposed SEAN, Fig. 3 shows the training and validation loss of the model, and Fig. 4 shows the prediction vs yield

of the SEAN model.

It is observed that the model achieved 98% accuracy during training. It is also seen that the accuracy is maintained on its prediction on the test set as well. The predicted vs actual yield and energy curve shown in Figs. 4 and 5 also shows that the models performs significantly better on the majority of the data points in the test set.

4.2. SEAN ablation study and comparison with conventional models

To properly account for the impressive performance achieved by the SEAN model, Table 5 presents the proposed model's comparison with other MIMO deep learning models that have been found to perform well on numerical datasets. This comparison is structured as an ablation study, systematically deconstructing the SEAN model by incrementally adding layers to reveal the impact of each component.

The intent is to distil the unique contributions of the attention module and state enhancements. The models compared are the Convolution Neural Network (CNN), the Long Short-Term Memory Network (LSTM), hierarchical CNN-LSTM, LSTM-Attention and the vanilla CLAN model without the state-enhanced attention module. All models are evaluated for 700 epochs, batch size of 32, and validation split of 20%. All models also use the same number of neurons, kernel size, optimiser, and call-back functions as specified for SEAN. Table 5 shows the predictive performance comparison of all evaluated models and the proposed SEAN on an out-of-sample test set.

4.3. Yield prediction

The SEAN model achieved a perfect RMSE and MAE of 0.00 and an R^2 of 0.97 for Yield prediction, outperforming all other models. The CNN, LSTM-Attention, and vCLAN models all had RMSE and MAE values that were higher by 100%, indicating that the SEAN model halved the error in prediction compared to these models. The LSTM and CNNLSTM models had even greater RMSE and MAE values, with SEAN improving upon these by 100%. In terms of R^2, which indicates the proportion of variance explained by the model, SEAN outperformed the CNN, LSTM-Attention, and vCLAN models by 14.12%, 4.30%, and 10.23%, respectively. The LSTM and CNNLSTM models had negative R^2 scores, rendering the comparison inapplicable as they failed to outperform the baseline model of predicting the mean.

4.4. Energy prediction

For energy prediction, the SEAN model again demonstrated superior performance with an RMSE of 0.10, an R^2 of 0.98, and a MAE of 0.05. The CNN model's error metrics were higher by 44.44% for RMSE and 58.33% for MAE, with a slight decrease in R^2 performance by 3.16%. This indicates that the SEAN model was significantly more accurate in predicting energy output. The LSTM and CNNLSTM models exhibited inferior performance with RMSE and MAE values that were more than 6000% higher than those of the SEAN model, along with negative R^2 scores, showing that these models were unsuitable for the energy prediction task in this context.

The LSTM-attention and vCLAN models showed an increase in RMSE by 28.57% and 16.67%, respectively, compared to the SEAN model. The MAE was also higher by 37.50% and 16.67% for these two models, respectively. The R^2 scores for the LSTM-Attention and vCLAN models were close to that of the SEAN model, with only slight decreases of 1.03% and no change, respectively. The SEAN model consistently outperformed other architectures across all metrics and outputs, demonstrating the effectiveness of state-enhanced attention mechanisms in capturing both spatial and temporal dependencies within the data. The model's integration of LSTM states into the attention mechanism has provided a significant improvement in predictive accuracy, particularly notable in the complex task of energy prediction in gas atomising process. The negative R^2 scores observed in LSTM and CNNLSTM models for energy highlight the importance of model choice in the context of specific prediction tasks. The SEAN model's robustness across different metrics establishes its superiority and potential as a reliable tool for predictive analysis in multidimensional time-series data.

The remarkable improvements observed with the SEAN model are directly traceable to the integrated LSTM states—both hidden

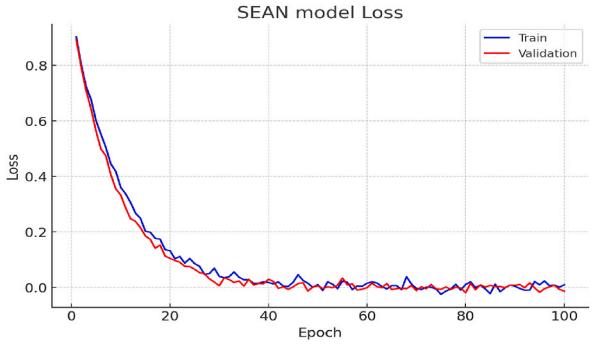


Fig. 3. SEAN training and validation loss curve.

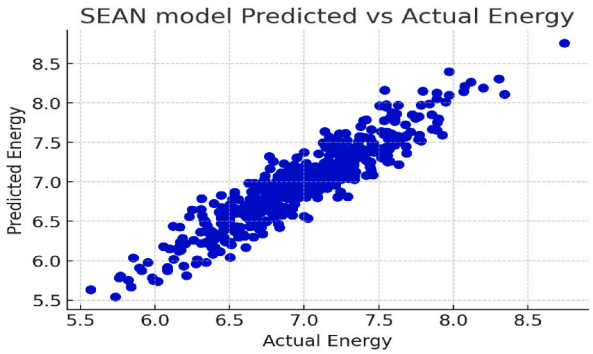


Fig. 4. SEAN predicted vs actual yield in test data.

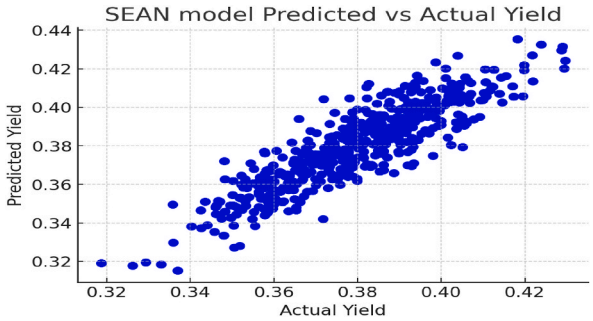


Fig. 5. SEAN predicted vs actual energy in test data.

Table 5
Comparison results of baseline models with the proposed SEAN.

Models	RMSE		R^2		MAE	
	Yield	Energy	Yield	Energy	Yield	Energy
CNN	0.01	0.18	0.85	0.95	0.01	0.12
LSTM	0.03	6.03	−0.04	−54.40	0.02	5.98
CNNLSTM	0.03	6.04	−0.01	−54.54	0.02	5.98
LSTM-Attention	0.01	0.14	0.93	0.97	0.01	0.08
vCLAN	0.01	0.12	0.88	0.98	0.01	0.06
SEAN	0.00	0.10	0.97	0.98	0.00	0.05

and cell states—into its attention mechanism. A major feature of the SEAN model is the enhanced feature representation achieved by processing LSTM states through dense layers and subsequently concatenating them with the LSTM and attention outputs. This state integration permits a nuanced focus on salient features across the temporal domain, thereby enriching the model's predictive precision. The attention mechanism further refines the model's performance by dynamically weighting input sequence components, proving especially beneficial in a process where certain temporal segments yield more informative insights than others. Such a comprehensive feature set is pivotal in capturing a broad spectrum of data dependencies, which is indispensable in the complex gas atomising process.

4.5. Evaluation of the SEAN model on the hold-out test set

To address concerns regarding potential bias introduced by using an augmented dataset with synthetic data generated from a limited original dataset of 58 points, several changes were implemented to ensure a more unbiased evaluation of the model's performance. Firstly, before any data augmentation or model training, the original dataset of 58 points was split into training and test sets. Specifically, 80% of the data was used for training, while 20% was reserved as a hold-out test set. This reserved test set was not used in any training or data augmentation processes.

The models were trained and validated using the training portion of the original data combined with synthetic data generated exclusively from this training set. This approach ensured that the hold-out test set remained untouched and was used solely for evaluating the model's generalisation performance. After training, the SEAN model was evaluated on the reserved 20% real-world data

to assess its performance on data not influenced by synthetic augmentation. This approach provided a more accurate measure of the model's ability to generalise to unseen real-world conditions. The results are presented in Table 6 below.

The evaluation of the SEAN model on the reserved 20% real-world test set showed strong predictive performance for 20–53 μm Yield (%). The model achieved an R-squared (R^2) of 0.82, indicating it captures the variance in the yield data well. The Mean Squared Error (MSE) of 0.00 and Mean Absolute Error (MAE) of 0.05 reflect the model's accuracy, with minimal deviation from actual values, suggesting SEAN is well-suited to predicting yield due to the more straightforward relationships in the input features.

However, the model's performance was weaker for Total AM kWh/kg, with an MSE of 0.24, RMSE of 0.49, and MAE of 0.40, indicating less precise predictions. The R-squared (R^2) of 0.75 suggests a reasonable fit, but lower than for yield. This reduction in performance may be due to the complexity of energy consumption data, influenced by multiple factors and operational conditions that the model struggles to capture fully. The lower performance on energy predictions could be due to the complex nature of the metric, which involves interactions and variables not fully represented in the training data. SEAN may be sensitive to noise and variability, leading to less accurate predictions. Although the evaluation did not directly use synthetic data, any biases introduced during training could affect its ability to generalise to complex real-world scenarios involving energy use. Overall, SEAN performs well for simpler metrics like yield but will require further tuning, additional real-world data, or model refinement to improve its accuracy for predicting more complex outcomes such as energy consumption.

4.6. Integration of SEAN model with NSGA-II

The research leverages the predictive power of the SEAN model within an NSGA-II framework to optimise process parameters for gas atomisation used in metal powder production. By using the SEAN model in the NSGA-II algorithm, the optimisation process capitalizes on the SEAN model's superior ability to capture the complex relationships within the gas atomising process parameters. This integration allows the NSGA-II algorithm to more effectively navigate the solution space and identify parameter configurations that yield the desired outcomes with higher precision.

Upon execution, the NSGA-II algorithm, powered by the SEAN model, generates a diverse set of Pareto-optimal solutions. Each solution represents a unique set of process parameters that balance the trade-off between yield and energy. The optimisation process is constrained by setting bounds for temperature, gas flow, pressure, and metal flow rate, ensuring that the solutions are within feasible and realistic operational limits.

The top solutions extracted from the NSGA-II algorithm reveal the optimal combinations of process parameters that reflect the predefined yield and energy requirements. To make the tool interactive, the *Streamlit* is used for the user interface design. The *Streamlit* interface presents an interactive exploration of these solutions, allowing stakeholders to visualise the impacts of various parameter adjustments on the desired outputs. The yield versus energy prediction scatter plot for the top solutions provides a visual representation of the multi-objective optimisation results. This interactive plot illustrates the trade-offs and relationships between yield and energy, offering insights into how varying one parameter affects the other. Figs. 6 and 7 below show the UI for the optimisation tool.

The implications of the SEAN model's predictive capabilities are transformative for the domain of metal powder production. Accurate yield predictions pave the way for improved process efficiency and quality control, ensuring that the metal powder adhere to stringent quality standards. Moreover, the model's energy output predictions present a new era of cost savings by optimising energy consumption—an important factor towards sustainable and economical metal powder production.

Moreover, the integration of the SEAN model into the NSGA-II optimisation algorithm has profound implications for the gas atomising process. The approach facilitates nuanced control over the atomisation process, allowing for the fine-tuning of parameters to achieve optimal yield and energy consumption simultaneously. Moreover, by identifying parameter settings that correlate with higher yield and lower energy requirements, the process becomes more efficient, reducing operational costs and waste. In addition, the model's predictive accuracy ensures that the powder produced consistently meets quality specifications, leading to fewer defects and higher customer satisfaction.

5. Limitations

The analysis above shows the significant improvement enabled by the state-enhanced attention module of the proposed SEAN in optimising the gas atomising process. However, it is also pertinent to mention a few limitations in this work. Firstly. this study utilised machine learning models for yield and energy consumption, based on a relatively small real-world data. While the models performed adequately within the range of the training data, their ability to extrapolate beyond this range is inherently limited.

As such, the generation of synthetic data from extended parameter ranges may introduce uncertainties, especially when predictions fall outside the original feature space. Although the parameter ranges for synthetic data generation were expanded cautiously, informed by domain knowledge to avoid unrealistic extremes, readers should consider this approach to synthetic data generation exploratory rather than definitive, and future work would benefit from broader datasets that reduce the reliance on extrapolation. An alternative would be active learning approach, where the model identifies the most informative data points for labelling through targeted experiments. This strategy can significantly reduce the number of costly experiments required while ensuring that the data collected is highly valuable for model training.

Secondly, the computational efficiency of the SEAN, when integrated with the NSGA-II algorithm, is impacted by slower execution times owing to the complex nature of the deep learning architecture (optimisation runs take approximately 195s on Intel Core i9-12900HK, with 64G RAM). Performing the optimisation in a cloud platform would significantly improve the speed and enhance the tool's applicability in real-time scenarios or for use with large-scale datasets.

Table 6
SEAN performance on hold-out real-world dataset.

Metric	20–53 μm Yield (%)	Total AM kWh/kg
Mean Squared Error (MSE)	0.00	0.24
R-squared (R^2)	0.82	0.75
Root Mean Squared Error (RMSE)	0.05	0.49
Mean Absolute Error (MAE)	0.05	0.40

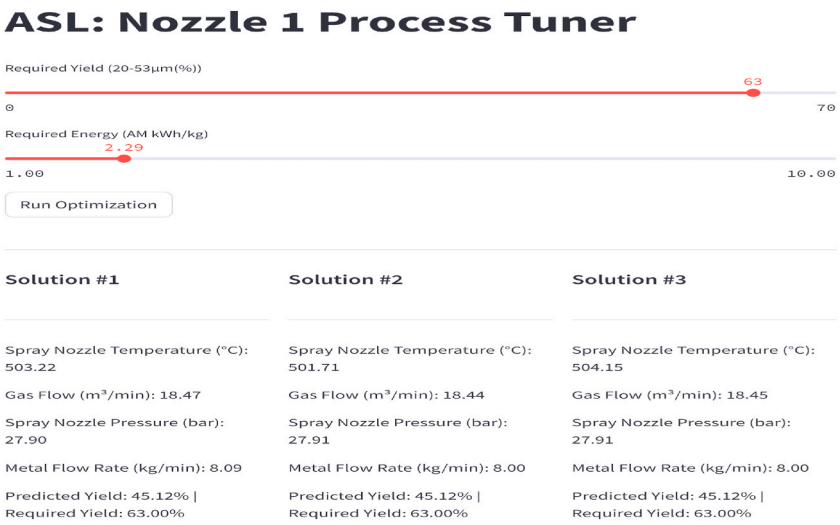


Fig. 6. Streamlit input and output pane for the optimisation tool.

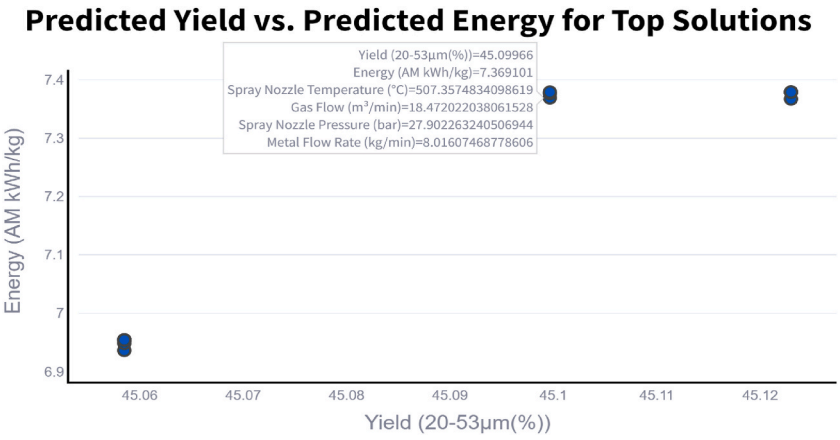


Fig. 7. Streamlit visualisation of the Pareto front.

Thirdly, the optimisation framework's reliance on a static dataset, sourced and stored in a CSV file. The absence of real-time data integration means that the model does not learn from continuously evolving process conditions. Lastly, the current framework does not provide functionality for saving optimisation runs, which is a crucial feature for conducting comparative analyses over time. The ability to record and revisit optimisation runs would enable a more thorough examination of the model's performance and the efficacy of the optimisation outcomes.

6. Conclusion

In this study, an innovative integration of the State-Enhanced Attention Network (SEAN) within the Non-dominated Sorting Genetic Algorithm II (NSGA-II) is presented. This represents a significant leap in the automation of metal powder production processes. This paper's deep learning architecture stands out for its dynamic adaptability, handling varying sequence lengths with precision,

which makes it a formidable tool for real-time process monitoring and control.

The synergy between SEAN model's deep learning capabilities and the evolutionary computation power of NSGA-II has been demonstrated to enhance both the efficiency and sustainability of the gas atomisation process. This is pivotal for the industry's ongoing shift towards more sustainable manufacturing practices. The practical application of Streamlit visualisation tools further enables stakeholders to glean actionable insights and make data-driven adjustments to the manufacturing parameters, thereby reinforcing the model's practical utility in real-world scenarios.

This research serves as a substantive contribution to the field of process optimisation using artificial intelligence, setting a benchmark for subsequent improvements. Future enhancements of this framework will focus on integrating real-time data, improving computational speed, and developing features to track optimisation performance over time. These advancements would improve the model's applicability and ensure its relevance in the fast-evolving landscape of industrial manufacturing and artificial intelligence.

CRediT authorship contribution statement

Abiodun Ayodeji: Writing – original draft, Software, Formal analysis, Data curation, Conceptualization. **Evelyne El Masri:** Writing – review & editing, Investigation, Funding acquisition, Formal analysis. **Tom Williamson:** Resources, Formal analysis, Data curation. **Mohmmad Ali Asgar Abbas:** Writing – review & editing, Software. **Tat-Hean Gan:** Writing – review & editing, Supervision, Project administration.

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.

Acknowledgement

This work was funded by the UK's Department of Energy, Security and Net Zero, under the Industrial Energy Transformation Fund (IETF22034).

Data availability

Data will be made available on request.

References

Anderson, I.E., Terpstra, R.L., 2002. Progress toward Gas Atomization Processing with Increased Uniformity and Control. Elsevier BV.

Ayodeji, A., Wang, W., Su, J., Yuan, J. and Liu, X. An Empirical Evaluation of Attention-Based Multihead Deep Learning Models for Improved Remaining Useful Life Prediction.

Ayodeji, A., Wang, Z., Wang, W., Qin, W., Yang, C., Xu, S., Liu, X., 2022. Causal augmented ConvNet: a temporal memory dilated convolution model for long-sequence time series prediction. ISA Trans. 123, 200–217. https://doi.org/10.1016/j.isatra.2021.05.026. Available at:

Bi, H., Lu, L., Meng, Y., 2023. Hierarchical attention network for multivariate time series long-term forecasting. Appl. Intell. 53 (5), 5060–5071. https://doi.org/10.1007/s10489-022-03825-5. Available at:

Dawes, J., Bowerman, R., Trepleton, R., 2015. Introduction to the additive manufacturing powder metallurgy supply chain. Johnson Matthey Technology Review 59 (3), 243–256. https://doi.org/10.1595/205651315X688686. Available at:

Eshkabilov, S., Ara, I., Azarmi, F., 2022. *A Comprehensive Investigation on Application of Machine Learning for Optimization of Process Parameters of Laser Powder Bed Fusion Processed 316L Stainless Steel* Research Square Platform LLC.

Gor, M., Dobriyal, A., Wankhede, V., Sahlot, P., Grzelak, K., Kluczyński, J., Łuszczek, J., 2022. Density prediction in powder bed fusion additive manufacturing: machine learning-based techniques. Appl. Sci. 12 (14), 7271. https://doi.org/10.3390/app12147271. Available at:

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D.J., 2021. 1D convolutional neural networks and applications: a survey. Mech. Syst. Signal Process. 151, 107398. https://doi.org/10.1016/j.ymssp.2020.107398. Available at:

Lewis, G., 2022. *Aspects of the Powder in Metal Additive Manufacturing: A Review* Scientific. Research Publishing, Inc.

Liu, X., Xu, P., Zhao, J., Lu, W., Li, M., Wang, G., 2022. Material machine learning for alloys: applications, challenges and perspectives. J. Alloys Compd. 921, 165984. https://doi.org/10.1016/j.jallcom.2022.165984. Available at:

Munyaka, J.B., Yadavalli, S.V., 2022. Inventory management concepts and implementations: a systematic review. S. Afr. J. Ind. Eng. 32 (2), 15–36. https://doi.org/10.7166/33-2-2527. Available at:

Shang, Z., Lian, Z., Li, M., Han, K., Zheng, H., 2023. Machine-learning-assisted multi-objective optimization in vertical zone refining of ultra-high purity indium. Sep. Purif. Technol. 305, 122430. https://doi.org/10.1016/j.seppur.2022.122430. Available at:

Singh, D., Verma, A., 2018. Inventory management in supply chain. Mater. Today Proc. 5 (2), 3867–3872. https://doi.org/10.1016/j.matpr.2017.11.641. Available at:

Tamura, R., Osada, T., Minagawa, K., Kohata, T., Hirosawa, M., Tsuda, K., Kawagishi, K., 2021. Machine learning-driven optimization in powder manufacturing of Ni-Co based superalloy. Mater. Des. 198, 109290. https://doi.org/10.1016/j.matdes.2020.109290. Available at:

Wang, H., Peng, M., Miao, Z., Liu, Y., Ayodeji, A., Hao, C., 2021. Remaining useful life prediction techniques for electric valves based on convolution auto encoder and long short term memory. ISA Trans. 108, 333–342. https://doi.org/10.1016/j.isatra.2020.08.031. Available at:

Williamson, T., Mellor, A., 2023. Specification of Industrial Process Energy Consumption. Available at: (Accessed 23 October 2023).