Renewable and Sustainable Energy Reviews 186 (2023) 113653

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

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

Original research article

# A deep learning approach to predict and optimise energy in fish processing industries

Ali Ghoroghi [a,*], Ioan Petri [a], Yacine Rezgui [a], Ateyah Alzahrani [a,b]

[a] School of Engineering, Cardiff University, Cardiff, UK
[b] School of Engineering, Al-Qunfudah Umm Al-Qura University, Makkah 715, Saudi Arabia

ARTICLE INFO

*Keywords:*
Multilayer perceptron
Genetic algorithm
Simulated annealing
Recurrent neural network
Long short-term memory
Gated recurrent unit

ABSTRACT

The fish processing sector is experiencing increased pressure to reduce its energy consumption and carbon footprint as a response to (a) an increasingly stringent energy regulatory landscape, (b) rising fuel prices, and (c) the incentives to improve social and environmental performance. In this paper, a standalone forecasting computational platform is developed to optimise energy usage and reduce energy costs. This short-term forecasting model is achieved using an artificial neural network (ANN) to predict power and temperature at thirty-minute intervals in two cold rooms of a fish processing plant. The proposed ANN function is optimised by genetic algorithms (GA) with simulated annealing algorithms (SA) to model the relationships between future temperature and power and the system variables affecting them. To assess the accuracy of the proposed method, extensive experiments were conducted using real-world data sets. The results of the experiments indicate that the proposed ANN model performs with higher accuracy than (a) the long short-term memory (LSTM) as an artificial recurrent neural network (RNN) architecture, (b) peephole-LSTM, and (c) the gated recurrent unit (GRU). This paper finds that using GA & SA algorithms; ANN parameters can be optimised. The RMSE obtained by the ANN compared with the second-ranked method GRU was consequently 16% and 4% less for the predicted temperature and power. The results in one particular site demonstrate energy cost savings in the range of 15%–18% after applying the forecast-optimiser approach. The proposed prediction model is used in a fish processing plant for energy management and is scalable to other sites.

## 1. Introduction

Fish processing involves sizeable and increasing levels of energy consumption and carbon emissions due to continuous needs for refrigeration, air conditioning, and ice making, as well as the reliance on fossil fuels [1–3]. Recent technological advancements in areas such as the Internet of Things (IoT) and Artificial Intelligence have paved the way to a digital transformation of the industry reflected by the more informed use of resources with significant carbon and cost reductions [4]. The coordination of energy flows and optimisation of energy use through clean energy generation and storage can decarbonise fishery ports while stimulating the development of energy communities in the local ecosystem [5]. One key area of the fish processing plant is the cold room. Dedicated temperature-controlled cold rooms are commonplace as fish is very perishable. Maintaining a constant temperature in cold rooms is challenging due to the dynamic nature of the processes involved with the regular transport of products, equipment and people entering and leaving the port. Furthermore, other constraints such as the build-up of frost, the cold rooms' energy efficiency, and energy costs must be considered in an optimal energy management system of the cold rooms [6].

This study aims to research, design, and develop a set of prediction, simulation, and optimisation modules to deliver more informed energy management in the fish processing industry. Our primary research questions are:

(a) How hybridising an ANN with evolutionary algorithms such as genetic algorithm (GA) and simulated annealing algorithm (SA) can enhance the mapping capability of the ANN?
(b) How does modelling the usage profile of the fish processing industry and predicting and optimising real-time energy interaction with pricing schemes and power demand help reduce energy costs and carbon footprint?
(c) What are the benefits of the proposed ANN framework in implementing smart energy grids and energy transition for industries?

* Corresponding author.
*E-mail address:* Ghoroghi@cardiff.ac.uk (A. Ghoroghi).

https://doi.org/10.1016/j.rser.2023.113653
Received 21 January 2023; Received in revised form 12 August 2023; Accepted 14 August 2023
Available online 27 August 2023

**Nomenclature**

| | |
|---|---|
| $AI$ | Artificial Intelligence |
| $ALO$ | Ant Lion Optimizer |
| $ANN$ | Artificial Neural Network |
| $AR$ | Auto-Regressive |
| $ARIMA$ | Auto-Regressive Integrated Moving Average |
| $DNN$ | Deep Neural Network |
| $ELU$ | Exponential Linear Unit |
| $GA$ | Genetic Algorithms |
| $GHG$ | Green-House Gas |
| $GOA$ | Grasshopper Optimization Algorithm |
| $GRU$ | Gated Recurrent Unit |
| $IMPVP$ | International Performance Measurement & Verification Protocol |
| $IoT$ | Internet of Things |
| $LSTM$ | Long Short-Term Memory |
| $MA$ | Moving Average |
| $MAE$ | Man Absolute Error |
| $MAPE$ | Mean Absolute Percentage Error |
| $MLP$ | Multi-Layer Perceptron neural network |
| $MSE$ | Mean Square Error |
| $PSO$ | Particle Swarm Optimisation |
| $ReLU$ | Rectified Linear Unit |
| $RMSE$ | Root Mean Square Error |
| $RNN$ | Recurrent Neural Network |
| $SA$ | Simulated Annealing |

The above research questions involve a common objective to develop a data-driven predictive model that can predict the power and temperature of cold rooms in thirty-minute intervals. Power consumption and temperature can be considered as typically univariate time series. As such, general time series forecasting methods are commonplace for electric load forecasting. These include autoregressive (AR), moving average (MA), autoregressive integrated moving average (ARIMA) models and a number of their variants. However, the abovementioned techniques are applicable when the observed and future time series are assumed to be linearly related. These methods are inferior for time series, which possess significant non-linear characteristics [7]. Another forecasting method class is based on artificial neural networks (ANNs). ANNs mimic the human nervous system in their parallel distributed processing. They can be applied to solving complex problems, including approximating functions and pattern recognition. Forecasting variables are multifaceted and complicated so a wholly accurate forecast is impossible. Nevertheless, a suitable model with intermittent monitoring of results can improve periodic forecasting accuracy to identify whether any errors are concordant with the predetermined ranges. Contrasting the linear forecasting of the ARIMA-based methods, ANNs can use a set of non-linear methods which are both data-driven and self-adaptive. ANNs are considered capable of approximating any non-linear function, with sufficient efficacy, especially for complex models [8,9].

The significant contribution of this paper is the introduction a precise multilayer perceptron neural network (MLP) as a subset of a deep neural network (DNN) model for power and temperature forecasting. A comparison of the efficacies of several DNN forecasting methods is then made. This paper explores how a predictive model can be integrated with an optimiser to achieve optimal energy management to reduce carbon footprint and energy costs. This is achieved using an optimiser (prediction data) in situ at a fish processing site. The proposed approach is informed by an in-depth energy audit conducted on the pilot. The optimisation model is based on sensitivity analysis that determines the degree of importance of a variable concerning the stipulated objectives.

The proposed model can forecast the temperature and power for intervals of half an hour. The inputs to the MLP are set-point, the engaged capacity of cold rooms, the season of the year, current temperature and current power. The outputs are predicted temperature and power. The forecasting model is trained and validated with measured cold-room data. The model was prepared to forecast temperature and power in a real trial site at Atlantic Dawn in Killybegs, Ireland. The accuracy of the proposed MLP's model was subsequently compared with the accuracy of long short-term memory (LSTM), peephole-LSTM, and gated recurrent unit (GRU) forecasting techniques.

The remaining contents of this paper are organised as follows. Section 2 discusses related work and provides an overview of ANNs. The methodology is described in several steps in Section 3. The data from testing and validating the model is described in Section 4. This includes an evaluation of the proposed model against other DNN approaches. Section 5 discusses and contextualises the energy cost savings for a fishery site. The findings are evaluated and concluded in Section 6.

## 2. Related work

Forecasting energy consumption holds significant practical importance when addressing net-zero emissions, sustainability, and energy efficiency. Precise energy usage forecasts enable governments, policymakers, and organisations to plan for the future energy landscape, facilitating the creation of informed policies and strategies promoting renewable energy adoption, energy efficiency measures, and sustainability initiatives. Integrating renewable energy sources effectively is critical to achieving net-zero emissions and sustainability objectives. Energy usage forecasts facilitate the optimal deployment of renewable energy during peak demand, ensuring a stable and reliable energy supply from clean sources. Responsible resource management is essential for sustainable energy usage, and energy usage forecasts aid in identifying areas where energy-saving measures and energy-efficient technologies can be implemented to reduce consumption.

One potential solution to the infrastructural challenges in energy supply is the digitalisation of energy management systems. This can also help integrate renewable energy sources into the energy management system. To achieve digitalisation, smart technologies such as Artificial intelligence (AI) in forecasting and smart grids can be implemented to strive for environmental sustainability [10]. Intelligent software can optimise decision-making and operations within the energy management system by controlling the power demand relative to the system's energy supply. However, there are still some limitations to the implementation of smart technologies. Business owners' primary concern is the significant initial investment needed to achieve digitisation and incorporate a smart system into their facilities. Another consideration is the difficulty in finding causes for malfunctions in the system, which may rarely arise in complex automation controller algorithms. Finally, a greater demand for computational capacity is often needed for smart technologies to solve complex problems. Therefore, regulators and policymakers look for evidence of the impact of intelligent technologies on the energy system, with incentives to reduce greenhouse gas emissions and energy costs where possible [11,12].

The indoor temperature of buildings and the associated energy consumption can be estimated in two ways. These include white-box (or physical) versus black-box (or empirical) models. Both models use weather parameters and indoor actuators' variables as inputs and predict energy or temperature as outputs [13]. In white-box or physical modelling, heating and cooling demands can be estimated by an energy simulation program, from which the system capacity and zone temperatures are calculated [14]. Conversely, black-box modelling implements a data mining technique to extract information from models. These models, one example of which is neural networks, rely on experimental data.

With the developments in AI, novel machine-learning (ML) methods have been used for power forecasting in cases where energy data needs to be carefully managed. Compared to traditional statistical forecasting methods, AI technology displays clear advantages in its ability to analyse significant amounts of data quickly [15,16]. Some commonly used load forecast methods are linear regression [17,18], autoregressive methods [19], and artificial neural networks [15,20,21]. A cold room's power and temperature forecast is associated with the season (weather), capacity changes and set-point characteristics. Time series models use mathematical formulation and are often less suitable than AI methods in forecasting multi-variable scenarios [14].

ML is utilised for energy usage forecasting due to the complexity and non-linearity of energy data influenced by external factors like weather, economic conditions, and human behaviour. ML models, particularly neural networks such as ANN, LSTM, and GRU, are adept at capturing intricate patterns and dependencies present in such data, resulting in accurate energy usage predictions. ML algorithms efficiently handle large datasets, extracting valuable insights and generating precise forecasts from extensive historical energy consumption data. The adaptability of ML models is crucial as energy demand patterns may change due to emerging technologies, policy shifts, or societal changes, and ML can adapt and update predictions based on new data. Additionally, ML techniques automate feature identification, reducing the need for manual feature engineering and improving prediction accuracy by understanding the factors significantly impacting energy usage. As ML models continuously learn and improve with more energy usage data, they remain up-to-date with evolving consumption patterns [22–25]. In conclusion, the application of machine learning in forecasting energy usage is of utmost importance in addressing contemporary challenges related to net-zero emissions, sustainability, and energy efficiency. Accurate energy forecasts facilitated by ML aid in developing strategies to reduce carbon emissions, optimise energy utilisation, and contribute to achieving net-zero emissions goals. This paper addresses the theoretical and practical aspects of ML for smart technologies in the fishery industry.

By training through previous data sets, ANNs imitate the human nervous system and can identify non-linear, complex relationships between variables. ANNs successfully predicted energy consumption or production at time intervals after training on a historical data set. Several papers have considered using ANNs to predict room temperatures and power consumption [26,27]. Further work has proposed using these predictions in energy-saving control strategies when air-conditioning systems are in use [13]. D. Yang et al. presented ANN use in researching optimal start and endpoints in operation used to control cooling systems [28]. This was done to reduce energy consumption effectively. ANNs were also used to control air-conditioning systems by predicting indoor temperature data [29]. Zhao et al. reviewed many techniques which aimed to predict building energy consumption [30]. The idea of utilising ANNs for load and temperature prediction was introduced back in 1995, where an ANN model was shown short-term load for daily load predictions [31]. ANNs have become popular amongst applications controlling heating, ventilation, energy consumption and power in buildings or rooms [32]. Ekonomou et al. used an ANN-based method alongside the wavelet denoising algorithm [33]. This was used to establish electricity load data into different frequency signals. The wavelet denoising algorithm provides neural network training with valid electricity load data, thus improving load forecasting accuracy.

Combination models integrating energy simulation and GA can be used to select optimal design parameters and help conserve energy usage [34–37]. The most common applications of GAs are for designing neural network architectures or for improving their learning rate [38]. Several related studies make use of GA for various energy management applications. N.Kampelis et al. have used ANN and GA for an energy management problem focusing on optimising the ANN prediction at a building and district level [39]. Similar works, [40,41], demonstrate how energy systems using solar power are optimised using GA to achieve improved efficiency. O. Hazem Mohammed et al. provides an example of a standalone hybrid energy system optimising using GA [42]. H. Lu et al. analysed energy quality management for a micro-energy network and applied GA to optimise the energy distribution in a tourist area [43]. Some authors [44,45] showed that sorting GA could reduce water pumps' electricity requirements and pollution emissions. A double-injection diesel engine is optimised by using a hybrid model of ANN and GA [46]. The model was able to reduce fuel consumption and computational time deficiency. S. Nikbakht expanded on the ANN model by assigning weights to multiple trained models and subsequently coupled them with the non-dominated sorting GA (NSGA-II). This resulted in a significant increase in neural network prediction accuracy [47].

This study aims to research, design, model, and develop a multilayer perceptron neural network (MLP) as a subset of a deep neural network (DNN) to deliver a smarter and optimised energy management in the fish processing industry. We model, predict and optimise the operation profile in the cold rooms of a site and explore how the real-time energy optimisation interacts with pricing schemes and power demand.

## 3. Methodology

This study aims to reduce energy costs and carbon footprint by modelling the usage profile of particular selected sites and optimising their energy use. This is done by exploring the interaction of real-time energy optimisation with pricing schemes and power demand.

The motivation for creating models of the usage profiles of selected sites is to establish an approach that can be scaled to similar sites in the future with similar energy management requirements. We use an ANN model to determine a relationship between inputs and outputs to support further modelling, identification, and prediction analysis [48,49]. We developed an optimiser that uses tabular data sets with an ANN forecasting model selected over other deep learning methods. The ANN provides improved flexibility and the ability to learn a mapping from inputs to outputs. As such, an optimisation strategy considering energy management is proposed by combining hour-ahead and real-time scheduling (Fig. 1).

This section presents the prediction framework and associated required mechanisms for conducting near-real-time energy optimisation in the fish processing plant. The proposed prediction framework consists of four main stages:

(a) analysis of environmental variables;
(b) data collection and cleansing/filtering;
(c) ANN-based prediction; and
(d) model evaluation

In the stage of analysing environmental variables, data sensitivity analysis is implemented to determine the inputs of MLP. The data collection and cleansing/filtering stage covers gathering environmental variables and removing abnormal noise from the data. Data filtering is usually used as a time series modelling tool. The ANN-based prediction stage covers the development of a prediction engine. Finally, GA and SA are embedded in this MLP to enhance outputs, and the accuracy of the prediction models was validated by statistical measures, compared with other established forecasting models and tested against real-time data by an optimisation platform.

### 3.1. Analysis of environmental variables

The study aims to model the usage profile of sites and explore how real-time energy optimisation interacts with the price of power and power demand; therefore, the primary analysis's objective was to identify the data necessary for accurate temperature and power forecasts.

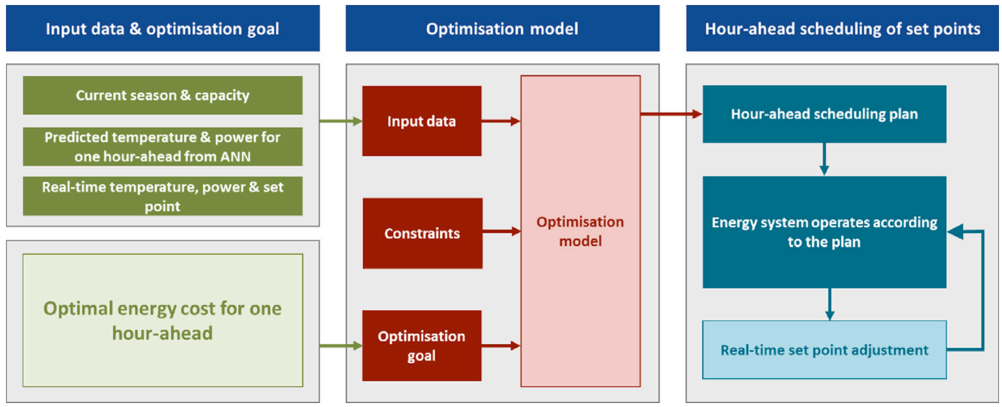



**Fig. 1.** The architecture of the optimisation strategy. The predicted temperature and power are obtained using the ANN, with inputs illustrated in Fig. 2.

In [50], a simulation model of the pilot project has been developed and calibrated using the open-source EnergyPlus simulation engine that considers a wide range of parameters linked to the building envelope and constituents materials, such as the "U" value, which factors in heat transfers and energy losses across the pilot building. The calibrated simulation model has informed the model's accuracy in the forecasting model's development.

The calibrated energy model that was developed factored in all the equipment and machinery for the fish processing, using both performance data gleaned from the manufacturer documentation (including power capacity) and on-site performance and operational data. The energy model did not factor in the relationship between the refrigeration power and the mechanical power of the compressor (e=Pf /pm) due to a lack of credible data. This is identified as a limitation and direction for follow-on research.

The other potential parameters influencing temperature and power prediction in cold rooms were identified. These included current power & temperature, the volume of products in the cold rooms (capacity states), set points and weather characteristics. Weather factors such as wind, ambient temperature and humidity were obtained using historical weather data.

In turn, the effect of adding each of the inputs was considered and assessed. This process was repeated for each permutation of input variables until the extra variables' addition did not equate to significant improvements in the model's performance. For example, ambient, humidity and wind speed were found to have no significant impact on the model's accuracy and were not considered. Our analysis showed that power use and temperature prediction are affected by the volume of products, season, set point, start temperature and power. The season of the year is used to consider ambient temperature changing across the year. Other parameters could be analysed in further investigations. These include the size of cold rooms, product surface temperature and layout of products in the rooms.

### 3.2. Data collecting and cleansing/filtering

The historical data sets comprised the sensory data collected by ten sensors and data produced using the calibrated simulation model. Around one-year historical data were consolidated (between 28 January till the end of November 2019) that informed the development of the forecasting model. Some of the historical data were sourced from the same sensor units as permitted by the project facilities or pilot facilities, whereas the second part was data from the calibrated simulation model. From these two sources, a total data set of around 95,000 samples were obtained after reprocessing. Based on the ambient characteristic of the particular site in Ireland, two seasons were considered from March to August for the summer trial and October to April for the winter trial. Extensive site data streams were cleaned and modelled to facilitate accurate profiles. This helps to deliver applicable prediction algorithms that reflect live scenarios. The data set for this study were disorganised and incomplete due to many factors. Consequently, a novel method was required to restore the data for use in ML models, so a deep learning-based model of missing data was applied.

**Table 1**
Statistical parameters of the inputs.

| | Minimum | Maximum | Mean | Deviation |
|---|---|---|---|---|
| Set-point | −25 | 0 | −17.77 | 9.45 |
| Season | 0 | 1 | N/A | N/A |
| Power | 1.89 | 138.46 | 69.19 | 35.21 |
| Temperature | −23.94 | −15.82 | −20.41 | 1.89 |
| Capacity | 0.5 | 1 | N/A | N/A |

This task was based on temperatures and power consumption in two cold rooms. The temperature and power consumption measurements were carried out through two defined seasons (generally labelled winter and summer). Temperature measurements are an average of temperatures obtained from ten sensors in two cold rooms at five-minute intervals.

Based on our literature survey [51,52], a common sampling interval for monitoring cold room temperatures in a fish processing factory is every 30 min to 1 h. This interval allows for frequent enough measurements to detect any temperature fluctuations or anomalies that may affect the quality and safety of the stored fish. Therefore, for the scenarios selected, the authors chose a conservative time interval of five minutes to allow sensing to be aligned with the actuation. For efficient energy management, the process needs to use sensor input for the latest configuration of the building and return adequate setpoints within the same interval. As such, an actuation performed at 30-minute intervals using five minutes intervals for sensing is satisfactory, given the nature of the activity in the room and possible changing scenarios [53].

The data set was cleaned and explored using visualisation tools. The data relating to two daily defrost, and some abnormal fluctuations were removed from the data set. The data set's basic statistics are presented in Table 1, which shows the minimum, maximum, mean and deviation of all the data set variables. Following the gathering of historical data, the next step in training was its normalisation.

### 3.3. ANN-based prediction

In an ANN, the processing units are layered neurones, with the input and output layers separated by hidden layers.

On each hidden layer, the $i$th neuron sums the weighted outputs $x_j$ of the previous layer and adds a bias $b_{ij}$ to the sum. The sums are then processed by the activation function $\varphi(\cdot)$ to generate outputs $y_i$, which

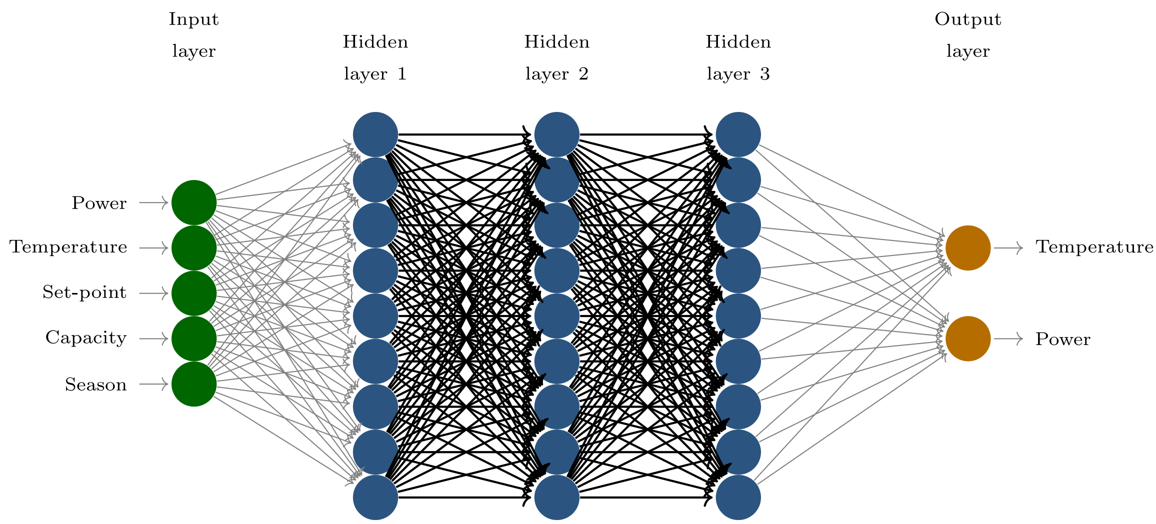



**Fig. 2.** The architecture of the proposed ANN.

work as the inputs for neurons on the following layer:

$$y_i = \varphi\left(\sum_{j=1}^{N} w_{ij}x_j + b_{ij}\right) \tag{1}$$

where $N$ is the total number of outputs received from the previous layer, $x_j$ is $j$th output from the previous layer, $y_i$ is the output of the $i$th neuron on the current layer, $w_{ij}$ and $b_{ij}$ are the weight and bias of the $j$th input on the $i$th neuron, and $\varphi(\cdot)$ is the activation function. ANN parameters (the weight $w_{ij}$ and bias $b_{ij}$) are determined from learning data in a supervised training process [54].

Power forecasts are usually performed a few steps ahead of the current time, mainly when the forecast's accuracy is essential for the economy and reliability of the power being provided [55,56]. In this task, accurate forecasting is performed for up to six steps ahead, with each step being in five-minute intervals. The optimum number of lags, neurons and hidden layers and the activation function had to be considered to produce an optimal model. A set of potential training methods were selected based on the nature of the data, and the optimum combination of the parameters above was implemented and used to generate a forecast. A neural network was trained to create a relationship between variables (environmental and control variables) and objectives based on historical data. A trial and error approach was employed to find the optimal number of hidden neurons, starting with just one hidden neuron in the hidden layer. The number of hidden neurons was incrementally increased until the MSE value was achieved. Finally, ANN is modelled, with five neurons in the first layer and three hidden layers, each containing nine neurons and two in the output layer.

The first layer (input layer) receives inputs as follows: Current temperature, Current power, Capacity, Set-point and Season. The last layer supplies the power and temperature assessed by the network and organises the responses obtained (Fig. 2).

After choosing relevant input variables, an appropriate number of lags for each variable had to be determined. This is fundamental for complex problems with numerous inputs, and no prior knowledge exists to identify possible lags. A set of different activation functions are compared in terms of forecast accuracy. A hyperbolic tangent activation function is used in the current model because of its squashing effect on minimal and massive values while maintaining near-linearity for mid-range values. Training a neural network involves setting the most suitable weights on the input of each unit. Here, the total number of instances is 95 000. The instance splitting method is random. The percentage of training instances is set to 60%, the number of selection instances set to 20%, and the number of testing instances set to 20%. Mean square error (MSE), (Eq. (2)) of the MLP model is used as criteria for the performance of the model. MSE is the standard error value used to measure fitness for predictive values. The neural network algorithm aims to minimise MSE [57].

$$\mathrm{MSE} = \frac{1}{N}\sum_{n=1}^{N}\left(y_n - \hat{y}_n\right)^2 \tag{2}$$

where $y_n$ denotes the measured value, $\hat{y}_n$ is the estimated value, and $N$ represents the sample size. The quasi-Newton method algorithm was used here for training as the optimisation algorithm stops when a specific condition is satisfied. The initial value of the training loss was 3.8265, and the final value after 509 iterations was 0.124476. The initial value of the selection loss was 3.84683, and the final value after 509 iterations was 0.118187.

### 3.4. Applying genetic algorithm & Simulated annealing in MLP

Despite the ability of the NN to have a feature selection function if given some regularisation penalty or sparsity to the connections between input nodes and the hidden layer, and despite advancements in learning theory which have proven many properties of gradient descent techniques, in this paper, we apply GA and SA to optimise the NN.

Feature selection algorithms are used to deal with high-dimensional data. These algorithms encompass identifying relevant features and eliminating inapplicable and repetitious features. Choosing a subset of data can reduce the data size and storage space needed, affecting the processing time and storage costs [58]. The weights in the MLP are updated periodically by the gradient descent method. However, errors in the initial weight selection can result in the MLP quickly falling into the local minimum or a slow convergence rate, affecting the system's accuracy.

When considering AI-based algorithms, widely used choices of algorithms for feature selection include; GA, particle swarm optimisation (PSO), Grasshopper Optimization Algorithm (GOA), Ant lion optimizer (ALO), and SA. At first, GOA and GA were considered. GOA is a "swarm intelligence algorithm" inspired by grasshoppers foraging and swarming behaviour. The literature contains examples of the hybridisation of GOA with other optimisation algorithms, such as metaheuristics and ML [59–63]. When comparing GA and GOA implementation, the results in this study show that both algorithms are similar in terms of accuracy. Therefore, we used GA as an optimiser of feature selection. GA is a metaheuristic algorithm used for solving constrained and unconstrained optimisation problems. It solves optimisation problems by emulating the natural selection process in organisms' biological evolution. A GA

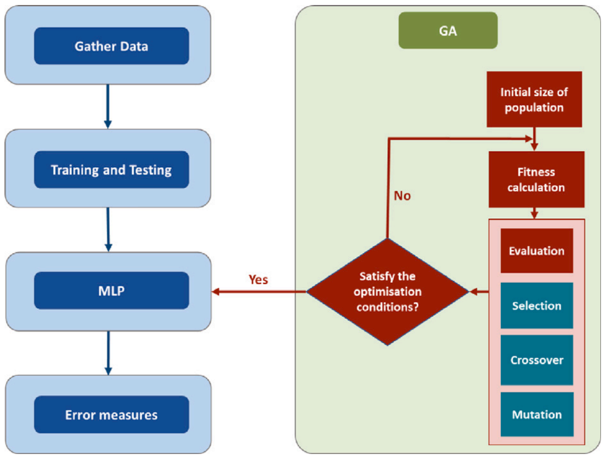



**Fig. 3.** Flowchart of GA tuned MLP.

is advantageous due to its global optimisation and fast convergence rate. Consequently, a GA can be incorporated into an MLP model to optimally find the proper weights and thus maximise the prediction capabilities of the model. The optimisation of a GA is dependent on the population size. As per the natural selection process, the GA operates because the fitter an individual in the population is, the better they will be represented in the population over time [64–66]. In this process, the individual's fitness consistently increases, and the error norm is reduced over time. The reciprocal of the error norm of the MLP is taken as the fitness function of the GA. The fitness calculation method is presented as follows:

$$F = \frac{1}{\frac{1}{N}\sum_{n=1}^{N}\left(y_n - \hat{y}_n\right)^2} \tag{3}$$

The superior operation is carried out in the GA, dependent on the fitness of the individuals within the population. A proportion of the population with higher fitness values is selected for the next generation. Variations within the individuals of the population are introduced through crossover and mutation to form a unique next generation. It is then assessed whether enough generations have been produced and the iterations completed or whether a certain performance by the population has been achieved, which is considered satisfactory. If these conditions are not met, the iterations continue, or the strategy is modified before the process starts again. Once the optimal termination conditions are achieved, the initial weightings for that process are taken, and the optimal weighting is known. Fig. 3 is a flowchart outlining the optimisation of a NN using the GA.

### 3.4.1. *Genetic algorithm parameters tuning*

The parameters the GA used in this study are presented: with a population size of 100, the error values of the NN model were at a minimum, which had the best prediction performance. Therefore, the GA here is used with a population size of 100 and a tolerance of 0.01 for selection loss in training. Fixed length chromosome representation is the most predominant in the field of GA [67]. The problem posed in this paper is suitable for a fixed-length solution. For the input variables to evolve, one individual in the population represents a set of possible inputs. A chromosome represents the features with a length equal to the total number of inputs features [68]. For the GA, binary encoding is applied to the chromosomes, with selected features holding one value and non-feature selected items being given zero. A uniform crossover method with a probability of 0.8 and a ranked-based fitness assignment method was used in the algorithm. The bit flip mutation allows for variations in the features included in the population. This involves flipping each bit and allows more exploration of the population. A low mutation rate of 0.02 was used to strike a balance between enough variation to explore the space of features but not an excess of variation.

Simulated Annealing (SA) is another metaheuristic algorithm that can approximate the global optimum of a given function. It is named after the metallurgy annealing process, which applies heating and cooling in a specific manner to affect a metal's behaviour and properties.

To prevent overfitting, the cross-validation method is commonly used [54,69]. Two algorithms, GA and SA, are used to approximate global optimisation in the search space to implement cross-validation. The input selection obtains the optimal subset of inputs for the model's best loss. After several analyses, GA is chosen to search for the subset of inputs, producing minimal selection error. To be consistent with the best prediction outcome, GA is used with a population size of 380 and a tolerance of 0.01 for selection loss in training. For the GA, selected features were given one value, and all others were zero. A balance of crossover and mutation had to be achieved to introduce variation of the desired amount. A uniform crossover of probability 0.8 and a lower mutation rate of 0.03 were used to produce an appropriate but not excessive variation.

### 3.4.2. *Simulated annealing algorithm parameters tuning*

Furthermore, SA is used for the selection loss in training to identify the optimal number of neurons and order selection. SA metaheuristic is inspired by annealing's physical process, where a piece of metal takes on a more stable state following heat treatment and subsequent cooling. To avoid falling into the local optimum, SA accepts a possible worsening move $m$ in its neighbourhood at a certain probability that is inversely affected by $\delta E$, the energy variation before and after the given move. Parameter $temp$ (temperature) is introduced, which controls how frequently worsening moves are accepted. Initially, $temp$ is set to a value that can be computed heuristically, and is then updated according to a cooling schedule with time. Given a $temp$, the probability of selecting a move $m$ is computed;

$$P(\delta E) = e^{-\left(\frac{\delta E}{temp}\right)} \tag{4}$$

Commonly, using a cooling rate parameter $\gamma$, the temperature can be decreased and updated as Eq. (5) [28,70].

$$temp = \gamma.temp \tag{5}$$

To be consistent with the best prediction outcome, the initial temperature used was 1000. The final temperature, in this case, was 0.1. A cooling rate of 0.5 and a tolerance of 0.01 were selected.

GA and SA parameters were made by trial and error. Parameters were identified that provided the best results in a reasonable amount of time. As per NN models (supervised)–MLP regressor–scikitlearn.org, the time complexity of backpropagation can be shown as $\mathcal{O}(i.t.m.h^k.l)$, where $i$ is the number of iterations, $m$ is the number of features, $h$ is the number of nodes per hidden layer, $k$ is the number of hidden layers and $l$ is the output layer neurons [71].

## 4. Model evaluation

For performance evaluation of the model, statistical parameters, including mean absolute percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE) were used. MAPE is used as a measure of prediction accuracy given by a percentage of deviation error. RMSE is a quadratic scoring rule that can also measure the average magnitude of the error. Similarly, MAE measures the average magnitude of the errors in a set of predictions without considering their direction. The MAPE and RMSE, and MAE are defined by Eqs. (6a), (6b) and (6c), respectively, where $y_n$ denotes the measured value, $\hat{y}_n$ is the estimated value, and $N$ represents the sample size.

$$\text{MAPE} = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{y_n - \hat{y}_n}{y_n}\right| \times 100\% \tag{6a}$$

**Table 2**
Errors for training, selection and testing.

| | Training | Selection | Testing |
|---|---|---|---|
| Sum squared error | 481 701 | 166 858 | 167 574 |
| Mean squared error | 223.943 | 233.694 | 237.356 |
| Root mean squared error | 14.9647 | 15.2871 | 15.4064 |

$$\text{RMSE} = \sqrt{MSE} \tag{6b}$$

$$\text{MAE} = \frac{1}{N}\sum_{n=1}^{N} |y_n - \hat{y}_n| \tag{6c}$$

Table 2 illustrates all errors measured for the training, selection and testing stages.

### 4.1. Comparison with other DNN approaches

LSTM and GRU networks are both types of NNs, but they are specifically designed for processing sequential data. LSTM and GRU are specialised variants of recurrent neural networks (RNNs) designed to overcome the limitations of traditional RNNs in processing sequential data. They both have internal mechanisms to capture long-term dependencies and handle vanishing/exploding gradient problems, making them practical for tasks involving sequences [72]. Ferreira et al. proposed a complete and self-contained presentation of the mathematical foundations of LSTM and GRU [73]. Their ability to handle sequential data and capture long-term dependencies makes them well-suited for various tasks. Therefore they are applied in Natural Language Processing (NLP), Time Series Forecasting, Healthcare, Weather Prediction, Robotics and Speech Recognition. Skrobek et al. [74] employed LSTM, Bidirectional Long Short-Term Memory (BiLSTM), and GRU to predict the mass of an adsorption bed in the fixed and fluidised bed. The GRU overcame the others in predicting the mass of both the fluidised and fixed beds. Also, LTSM and GRU are used in energy management, such as load forecasting, demand response, and fault detection. LSTMs have been applied successfully for building energy optimisation using deep reinforcement learning. GRUs have shown exemplary performance in energy load forecasting, which helps grid planning and management. Overall, LSTM and GRU models can potentially represent temporal dynamics in energy systems and improve energy management [75]. Sahhad et al. [76] performed experiments on LSTM, CNN-LSTM and CNN-GRU Short-Term Residential Load Forecasting. Phan et al. [77] proposed a model for forecasting solar power generation using GRU and LSTM. They showed that a GRU could deal with a larger dataset because it has fewer parameters and a shorter training time than LSTM.

The presented MLP was compared with the following methods: LSTM, peephole-LSTM and GRU. Most techniques for predicting power demand include recurrent neural network (RNN)-based LSTMs using time-series data and natural language processing [78–80]. RNN is used to assign missing values and to accommodate for the nonlinearity of meteorological time-series data [81]. RNNs, are a type of neural network that allows for a previously generated output to be used as a subsequent input while having hidden states. Given an input sequence $x = (x_1, \dots, x_T)$, a standard RNN computes the hidden vector sequence $h = (h_1, \dots, h_T)$ and output vector sequence $y = (y_1, \dots, y_T)$ by iterating the following equations from $t = 1$ to $T$:

$$h_t = \varphi(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{7a}$$

$$y_t = W_{hy}h_t + b_y \tag{7b}$$

where the $W$ terms denote weight matrices, the $b$ terms denote bias vectors and $\varphi$ is the hidden layer function [82]. Fig. 4 shows a simple unrolled RNN.

The recurrent layer of RNNs contains feedback loops so they can hold information in memory over time. However, the training of standard RNNs to successfully solve problems that require learning long-term temporal dependencies is challenging. The vanishing gradient problem can explain this. Hochreiter and Schmidhuber [83] introduced a type of RNN and a subset of DNN in 1997; the LSTM networks.

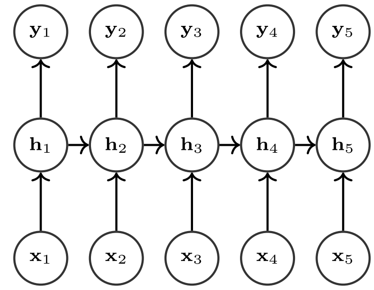



**Fig. 4.** A schematic of a simple unrolled RNN.

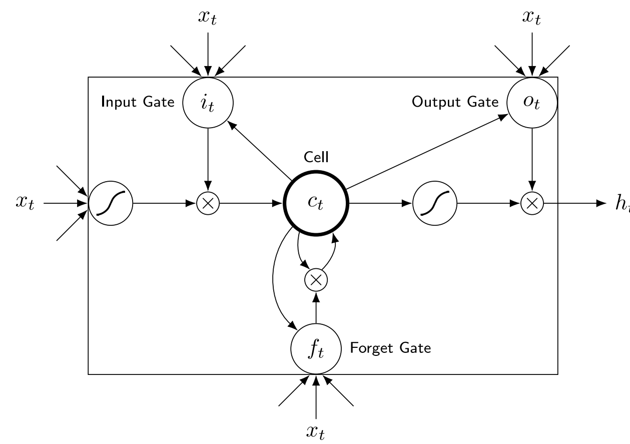



**Fig. 5.** A schematic of an LSTM memory cell.

LSTM implements gate and memory cells in each hidden layer. A memory block includes an input gate *i*, a forget gate *f*, an output gate *o*, and self-connected memory cells *C*. The input gate controls the entry of the activations to the memory cell; similarly, the output gate functions to filter cell activations appropriately and output to the successive network. The forget gate aims to help the network remove past input data and reset the memory cell. For LSTM, the following composite function is implemented:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \tag{8a}$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \tag{8b}$$

$$c_t = f_t x_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{8c}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \tag{8d}$$

$$h_t = o_t \tanh(c_t) \tag{8e}$$

where $\sigma$ is the sigmoid, the $W$ terms denote weight matrices and the $b$ terms denote bias vectors [82]. The access and storage of information for a long time period is made possible by the careful use of multiplicative gates. Consequently, this construct aims to alleviate the vanishing gradient problem [84], making LSTM suitable for problems with long-term dependencies. Fig. 5 shows a single LSTM memory cell [82].

One problem with the LSTM is that the gates cannot access information from the memory cell output when the output gate is closed. As such, the LSTM cannot know how long the memory should be held for the model.

To overcome this, peephole connections introduced by Gers and Schmidhuber [85] can be implemented in the LSTM memory cells, and the gate layers can then observe the cell state [86]. Also, a variation of the LTSM is GRU, introduced by Cho et al. [87]. GRU implements

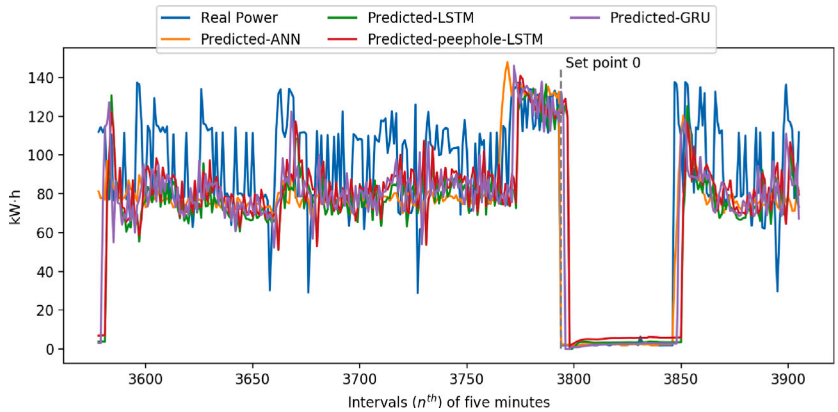



**Fig. 6.** Power forecasting performances.

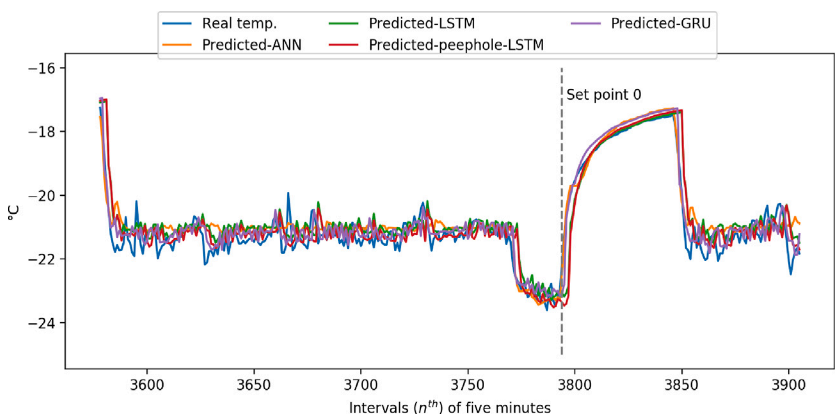



**Fig. 7.** Temperature forecasting performances.

**Table 3**
The accuracy of DNN approaches expressed in terms of the statistical measures.

| Algorithms | Temp. | | | Power | | |
|---|---|---|---|---|---|---|
| | MAPE (%) | RMSE | MAE | MAPE (%) | RMSE | MAE |
| **Proposed model** | **1.695** | **0.514** | **0.348** | **63.643** | **19.954** | **12.448** |
| LSTM | 1.857 | 0.623 | 0.386 | 66.065 | 25.459 | 15.634 |
| Peephole-LSTM | 1.835 | 0.633 | 0.380 | 65.959 | 24.991 | 15.519 |
| GRU | 1.827 | 0.612 | 0.376 | 64.680 | 20.766 | 13.046 |

a single update gate by combining the forget and input gates. GRU also merges cell states and hidden states. As such, the model becomes simpler than conventional LSTM models [86]. Cheng et al. [88] used an LSTM model to predict the power demand for small power systems with non-linear non-critical characteristics.

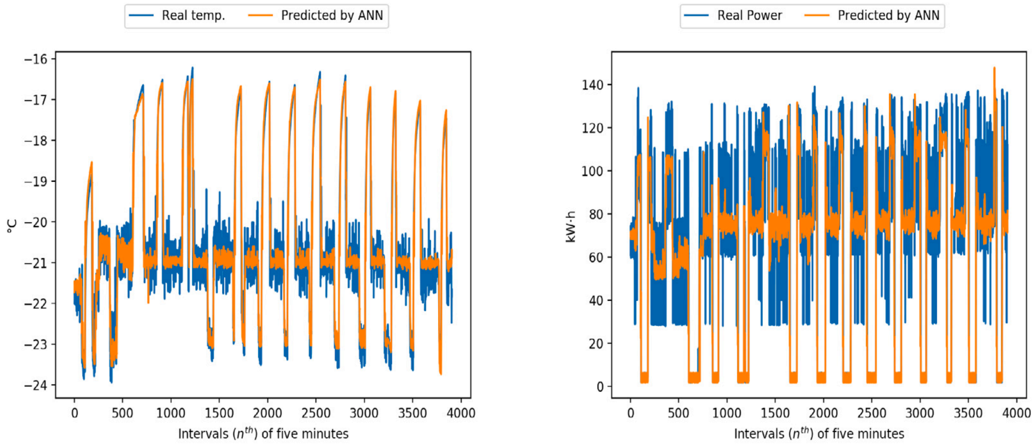
Power usage was considered in smaller groups and divided further to consider the power usage of individual households. Subsequently, a study using an LSTM model was conducted to forecast each household's power usage [88]. Future power demand was forecast using current power production in a solar power farm.

To produce an optimal model, the following had to be considered and optimised: the number of neurons, the number of hidden layers, the learning rate, batch size and the number of epochs. Exponential linear unit (ELU) was set as an activation function of the LSTM model, and leaky rectified linear unit (ReLU) as an activation function of peephole-LSTM and GRU models. While ELU smooths slowly until its output equals $-\alpha$, leaky ReLU sharply converges. ReLU does not allow for negative inputs, but leaky ReLU provides for a small, non-zero, constant gradient $\alpha$.

MAPE, RMSE and MAE were used to evaluate the performance of the forecasting models. The results of these are summarised in Table 3. These scores are reported as averages across different runs of the models.

Among the included DNN approaches, the proposed model and GRU showed significantly lower error measures than LSTM and peephole-LSTM. GRU performed better in temperature and power predictions than LSTM and peephole-LSTM; it showed notably higher accuracy in power predictions than these two approaches. For temperature, the proposed model had MAPE of 1.695% compared to the 1.827% of GRU. Similarly, scores of 0.514 and 0.348 in RMSE and MAE, respectively, showed the proposed model had, on average, the least error in temperature predictions. For power predictions, the proposed model had a MAPE of 63.643%, which was better than the GRU measure of 64.680%. The proposed model produced a minimum MAE of 12.448 and an RMSE of 19.954. While the error measures for GRU and the proposed model are slightly similar, the accuracy and the proposed model's ability to utilise tabular data sets make it the ideal approach for the application of this paper.

Figs. 6 and 7 represent the behaviour of ANN, LSTM, peephole-LSTM and GRU against real data (power and temperature). They show more details of the forecasting trends. Fig. 6 shows the fluctuations for one day of real power compared to the DNN approaches' predictions. When the set-point is set to zero (shown by a vertical drop in real power towards zero), power consumption is followed most closely by predicting the proposed model. The most significant forecasting errors of the other DNN approaches occur when the set-point changes to zero. Peephole-LSTM shows variation compared to that of real data. Fig. 7 shows the DNN models' behaviour against real temperature data for one day. At the point of the set-point changing to zero, the proposed model's prediction closely follows the actual temperature initially but overshoots its prediction and shows fluctuations in its behaviour. LSTM prediction follows the actual temperature more closely.




(a) Predicted and real temperatures shown for predictions of thirty-minute intervals for a fourteen-day period.

(b) Predicted and real power shown for predictions of thirty-minute intervals for a fourteen-day period.

**Fig. 8.** Model evaluation: Predicted vs. real temperature.

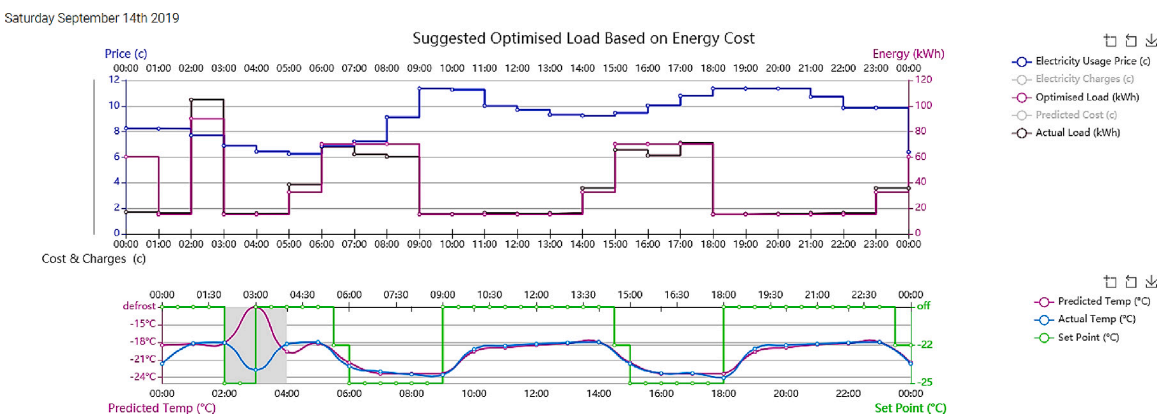



**Fig. 9.** Simulation results from an optimisation platform using one-day ahead prediction. The optimiser was developed by the telecommunications software and systems group (TSSG), Waterford, Ireland.

### *4.2. Evaluating the accuracy of the proposed model against actual data*

Real-world data was used to evaluate the developed model's standard performance indices and the relationship between variables. Fig. 8 shows the predicted temperature and power accuracy for half an hour intervals against fourteen days of historical trial data.

## 5. Energy cost savings

The proposed model was developed as an optimiser in an industry's energy management system. As forecasting is an ongoing goal, and MLP is used, the forecasting process is never-ending, constantly monitoring accuracy. There will be occasions where the forecasting models must be adapted to coincide with changes in data or the future goal. For model performance evaluation, observed data from 2020 was compared to the previous year's validation period. After slight calibration of the model using 2020 data, the model was validated by real-time data from 2020 for the same period as 2019. Almost all models need to be calibrated in this way using observed data. Year to year, models like the proposed MLP may encounter uncertainties from different sources. These include data, parameter, and model structure uncertainty [89].

Fig. 9 shows the application of this model by the optimiser in real-time. In this application, the optimiser makes a one-hour prediction using two predictions of thirty-minute intervals. The output of the first prediction is used as the second's input to produce a one-hour temperature prediction. The use of a multi-stage prediction is expected to make a cumulative error. However, the model performs within the 0.5 °C accuracy threshold required for this application. The optimiser then assesses the real-time energy cost and adjusts the combination of set points for the next hour. The prediction model's ability to stay within the 0.5 °C error limit while experiencing cumulative error reflects the accuracy of the proposed model.

In the next stage, using a homogeneous method for all pilots, energy efficiency and greenhouse gas (GHG) emissions were assessed, and economic efficiency factoring in this project was considered. The IMPVP (International Performance Measurement & Verification Protocol) was used to measure economic efficiency. Option B, the Retrofit Isolation option from IPMVP, is selected. It was essential to keep the facility's conditions as consistent as possible before and after implementation measurement periods. The savings calculation was made using the simplified equation as any calibration error would affect both the baseline and post-implementation period equally:

Expected savings (kWh) = Expected old energy use (kWh) - Expected new energy use (kWh).

Two baseline periods in a particular site in Ireland were considered. The first period was from the 1st of September to the 31st of October 2019. For one cold room with 75% of capacity engaged, the hourly average energy saving is 2.77 kWh, and the hourly average cost saving is €162.6. The daily average energy-saving and cost-saving are 66.4

**Table 4**
Measured consumption and savings obtained in the nominated pilot.

| Period | # of cold rooms - Capacity engaged | Before retrofit | After retrofit | Energy savings [kWh] | Cost avoided [€] | $CO_2$ emissions avoided [kg] |
|---|---|---|---|---|---|---|
| | | EC [kWh] | EC [kWh] | | | |
| First baseline | One- %75 | 67 291 | 62 253 | 4037 | 237 193 | 1716 |
| Second baseline | Two- %100 | 133 436 | 126 712 | 6224 | 320 081 | 2645 |

kWh and €3901.7. This equates the first baseline period as energy saving and cost saving of 4038 kWh and €237,193 respectively. The second period was from the 20th of April to the 30th of June 2019. For two cold rooms with 100% of capacity engaged, the hourly average energy saving is 4 kWh, and the hourly average cost saving is €187.4. The daily average energy-saving and cost-saving are 94.5 kWh and €4497.6. This equates the first baseline period as energy saving and cost saving of 4038 kWh and €237,193 respectively. The $CO_2$ emission intensity (kg $CO_2$/kWh) is calculated as the ratio of $CO_2$ emissions from electricity production. $CO_2$ emission avoided during the baseline periods is obtained using the conversion rate for Ireland (0.425 kg $CO_2$ per kWh, source: The European Environment Agency - 2016). Table 4 shows the average power usage for both periods and the savings calculated. The results in the site show energy cost savings in the range of 15%–18% after applying the forecast-optimiser platform.

## 6. Conclusion

This paper aims to develop an accurate forecasting model using cold rooms' short-term temperature and power to optimise energy across the fish processing chain. The resulting prediction model provides the optimised decision capability to respond to real-time changes in energy demand through monitored real-time data. The real-time generated decision can actuate on-site control facilities to take actions according to any pre-defined optimisation objectives, either single or multi-objectives. For its applicability to the optimiser, the prediction models' accuracy was validated by statistical measures and testing against real-time data by an optimisation platform. Furthermore, the proposed model was evaluated and compared to several forecasting models to predict power and temperature. These included LSTM as an artificial RNN architecture, peephole-LSTM and GRU. We can see that the proposed model and GRU had error measures, although Figs. 6 and 7 for power and temperature, respectively, show very similar behaviours in DNN models when the set-point is zero. We conclude that the proposed MLP is superior in its use for the application above due to its accuracy and flexibility. This MLP allows two-fold or multi-stage prediction required for one-hour ahead forecasting and uses tabular data. Furthermore, the findings show that the model can substantially improve with GA and SA methods. The deep learning approach for temperature and energy predictions is scalable for other similar industries using site-specific data predictions.

This model is then used as an input to the optimisation module, directing the system to follow the most economical path while maintaining site constraints. The real-time optimiser aimed to change the energy demand of the cold rooms by controlling the set points for the cooling system based on the predicted temperature behaviour. The economic approach is derived from live market data in electrical market pricing and available on-site renewable energy generation. Consequently, the optimiser could change the energy-management system to minimise energy usage with the most cost-efficient times relative to the grid's power pricing. The IPMVP protocols were used to assess the economic efficiency of the energy management system. The results show significant cost savings after applying the forecast-optimiser platform. Furthermore, there is a reduction in carbon dioxide emissions. These modules combine to deliver a scalable, intelligent solution for the fish processing industry, which relies on innovative systems to compete with other sectors reliant on the energy market. Thus, we conclude that hybridising an ANN with evolutionary algorithms can enhance its modelling capability. This modelling can predict the energy profile of a fishery site in situ and optimise real-time energy use with real-time energy cost.

We conclude that these approaches can be used with an economic mindset, using live market data in electrical market pricing and available on-site renewable energy generation. This can lead to a reduction in energy costs and carbon footprint. This study shows how digital technologies can support smart energy grids to improve the performance gap and ensure a successful transition towards sustainable industries.

Proposed future work includes implementing methodologies for parameter tuning, which can further enhance forecasting accuracy. It may be beneficial for this application to consider parameters such as the size of cold rooms, product surface temperature and layout of products in the rooms. The season categories used in the prediction model could be expanded by using a broad set of new data across different times in the year. Alternatively, real-time ambient temperature could be implemented as an input to replace season categories, which would accommodate unpredictable changes in weather during a season.

## CRediT authorship contribution statement

**Ali Ghoroghi:** Machine learning and simulation work and authoring of the paper, Verified the underlying data and validated the results. **Ioan Petri:** Conceived the study, designed, and led the research in their capacity of Principal Investigators on the EU INTERREG piSCES Project: "Smart Cluster Energy System for the Fish Processing Industry, grant number: 504460 ", Supervision, Review and editing. **Yacine Rezgui:** Conceived the study, designed, and led the research in their capacity of Principal Investigators on the EU INTERREG piSCES Project: "Smart Cluster Energy System for the Fish Processing Industry, grant number: 504460 ", Supervision, Review and editing. **Ateyah Alzahrani:** Carried out experiments and data analysis.

## 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.

## Data availability

The authors do not have permission to share data.

## Acknowledgements

This work is part of the EU INTERREG piSCES Project: "Smart Cluster Energy System for the Fish Processing Industry", grant number: 504460. The authors would like to thank the editors, reviewers, and Telecommunications Software and Systems Group, Waterford, Ireland. All authors approved the version of the manuscript to be published.

## References


[1] Jiang B, Sun Z, Liu M. China's energy development strategy under the low-carbon economy. Energy 2010;35(11):4257–64. http://dx.doi.org/10.1016/j.energy.2009.12.040.

[2] Pelletier N, Audsley E, Brodt S, Garnett T, Henriksson P, Kendall A, Kramer KJ, Murphy D, Nemecek T, Troell M. Energy intensity of agriculture and food systems. Annu Rev Environ Resour 2011;36(1):223–46. http://dx.doi.org/10.1146/annurev-environ-081710-161014.

[3] Kroodsma DA, Mayorga J, Hochberg T, Miller NA, Boerder K, Ferretti F, Wilson A, Bergman B, White TD, Block BA, Woods P, Sullivan B, Costello C, Worm B. Tracking the global footprint of fisheries. Science 2018;359(6378):904–8. http://dx.doi.org/10.1126/science.aao5646.

[4] Ghobakhloo M. Industry 4.0, digitization, and opportunities for sustainability. J Clean Prod 2020;252:119869.

[5] Acciaro M, Ghiara H, Cusano MI. Energy management in seaports: A new role for port authorities. Energy Policy 2014;71:4–12.

[6] Melis RB. New techniques and methods for cold chain monitoring and tracking in perishable products (Ph.D. thesis, Ph.D. thesis), Universidad Politécnica de Madrid; 2016.

[7] Zheng J, Xu C, Zhang Z, Li X. Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network. In: 2017 51st annual conference on information sciences and systems. CISS, 2017, p. 1–6.

[8] Li J, Wang J. The load forecasting model based on Bayes-GRNN. J Softw 2012;7. http://dx.doi.org/10.4304/jsw.7.6.1273-1280.

[9] Panigrahi S, Karali Y, Behera DH. Time series forecasting using evolutionary neural network. Int J Comput Appl 2013;75:13–7. http://dx.doi.org/10.5120/13146-0553.

[10] Petri I, Yuce B, Kwan A, Rezgui Y. An intelligent analytics system for real-time catchment regulation and water management. IEEE Trans Ind Inform 2018;14(9):3970–81. http://dx.doi.org/10.1109/TII.2017.2782338.

[11] Rhodes A. Digitalisation of energy: An energy futures lab briefing paper. Tech. rep., Imperial College London; 2020, URL http://hdl.handle.net/10044/1/78885.

[12] Ahmad T, Zhang D, Huang C, Zhang H, Dai N, Song Y, Chen H. Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities. J Clean Prod 2021;289:125834. http://dx.doi.org/10.1016/j.jclepro.2021.125834.

[13] Yang I-H, Kim K-W. Prediction of the time of room air temperature descending for heating systems in buildings. Build Environ 2004;39(1):19–29. http://dx.doi.org/10.1016/j.buildenv.2003.08.003.

[14] Lawan S, Abidin WAWZ, Chai W, Baharun A, Masri T. Different models of wind speed prediction; a comprehensive review. 2014.

[15] Metaxiotis K, Kagiannas A, Askounis D, Psarras J. Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. Energy Convers Manage 2003;44(9):1525–34. http://dx.doi.org/10.1016/S0196-8904(02)00148-6.

[16] Ghoroghi A, Rezgui Y, Petri I, Beach T. Advances in application of machine learning to life cycle assessment: a literature review. Int J Life Cycle Assess 2022;27(3):433–56. http://dx.doi.org/10.1007/s11367-022-02030-3.

[17] Javed F, Arshad N, Wallin F, Vassileva I, Dahlquist E. Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting. Appl Energy 2012;96:150–60. http://dx.doi.org/10.1016/j.apenergy.2012.02.027, Smart Grids.

[18] Iwafune Y, Yagita Y, Ikegami T, Ogimoto K. Short-term forecasting of residential building load for distributed energy management. In: 2014 IEEE international energy conference. ENERGYCON, 2014, p. 1197–204. http://dx.doi.org/10.1109/ENERGYCON.2014.6850575.

[19] Sevlian R, Rajagopal R. Short term electricity load forecasting on varying levels of aggregation. Trans Power Syst 2014.

[20] Hippert HS, Pedreira CE, Souza RC. Neural networks for short-term load forecasting: a review and evaluation. IEEE Trans Power Syst 2001;16(1):44–55. http://dx.doi.org/10.1109/59.910780.

[21] Tzafestas S, Tzafestas E. Computational intelligence techniques for short-term electric load forecasting. J Intell Robot Syst 2001;31(1):7–68. http://dx.doi.org/10.1023/A:1012402930055.

[22] Skrobek D, Krzywanski J, Sosnowski M, Uddin GM, Ashraf WM, Grabowska K, Zylka A, Kulakowska A, Nowak W. Artificial intelligence for energy processes and systems: Applications and perspectives. Energies 2023;16(8). http://dx.doi.org/10.3390/en16083441, URL https://www.mdpi.com/1996-1073/16/8/3441.

[23] Ashraf WM, Uddin GM, Arafat SM, Krzywanski J, Xiaonan W. Strategic-level performance enhancement of a 660 MWe supercritical power plant and emissions reduction by AI approach. Energy Convers Manage 2021;250:114913. http://dx.doi.org/10.1016/j.enconman.2021.114913, URL https://www.sciencedirect.com/science/article/pii/S019689042101089X.

[24] Amjad A, Ashraf WM, Uddin GM, Krzywanski J. Artificial intelligence model of fuel blendings as a step toward the zero emissions optimization of a 660 MWe supercritical power plant performance. Energy Sci Eng 2023.

[25] Krzywanski J. Advanced AI applications in energy and environmental engineering systems. Energies 2022;15(15). http://dx.doi.org/10.3390/en15155621, URL https://www.mdpi.com/1996-1073/15/15/5621.

[26] Mustafaraj G, Lowry G, Chen J. Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office. Lancet 2011;43:1452–60. http://dx.doi.org/10.1016/j.enbuild.2011.02.007.

[27] Underwood C. HVAC control systems: Modelling, analysis and design. Routledge; 2005, p. 1–340. http://dx.doi.org/10.4324/9780203028704.

[28] Yang D, Wang X, Tian X, Zhang Y. Improving monarch butterfly optimization through simulated annealing strategy. J Ambient Intell Humaniz Comput 2020. http://dx.doi.org/10.1007/s12652-020-01702-y.

[29] Ruano A, Crispim E, Conceição E, Lúcio M. Prediction of building's temperature using neural networks models. Energy Build 2006;38:682–94. http://dx.doi.org/10.1016/j.enbuild.2005.09.007.

[30] xiang Zhao H, Magoulès F. A review on the prediction of building energy consumption. Renew Sustain Energy Rev 2012;16(6):3586–92. http://dx.doi.org/10.1016/j.rser.2012.02.049.

[31] Kiartzis S, Bakirtzis A, Petridis V. Short-term load forecasting using neural networks. Electr Power Syst Res 1995;33(1):1–6. http://dx.doi.org/10.1016/0378-7796(95)00920-D.

[32] Neto A, Fiorelli F. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. Energy Build 2008;40:2169–76. http://dx.doi.org/10.1016/j.enbuild.2008.06.013.

[33] Ekonomou L, Christodoulou C, Mladenov V. A short-term load forecasting method using artificial neural networks and wavelet analysis. Int J Emerg Electr Power Syst 2016;1:64–8.

[34] Howell SK, Wicaksono H, Yuce B, McGlinn K, Rezgui Y. User centered neuro-fuzzy energy management through semantic-based optimization. IEEE Trans Cybern 2019;49(9):3278–92.

[35] Ferdyn-Grygierek J, Grygierek K. Multi-variable optimization of building thermal design using genetic algorithms. Energies 2017;10:1570. http://dx.doi.org/10.3390/en10101570.

[36] Yuce B, Rezgui Y. An ANN-GA semantic rule-based system to reduce the gap between predicted and actual energy consumption in buildings. IEEE Trans Autom Sci Eng 2017;14(3):1351–63.

[37] Reynolds J, Ahmad MW, Rezgui Y, Hippolyte J-L. Operational supply and demand optimisation of a multi-vector district energy system using artificial neural networks and a genetic algorithm. Appl Energy 2019;235:699–713. http://dx.doi.org/10.1016/j.apenergy.2018.11.001.

[38] Yalkin C, Korkmaz E. Neural network world: A neural network based selection method for genetic algorithms. Neural Netw World 2012;22:495–510. http://dx.doi.org/10.14311/NNW.2012.22.030.

[39] Kampelis N, Tsekeri E, Kolokotsa D, Kalaitzakis K, Isidori D, Cristalli C. Development of demand response energy management optimization at building and district levels using genetic algorithm and artificial neural network modelling power predictions. Energies 2018;11(11). http://dx.doi.org/10.3390/en11113012, URL https://www.mdpi.com/1996-1073/11/11/3012.

[40] Asgher U, Babar Rasheed M, Al-Sumaiti AS, Ur-Rahman A, Ali I, Alzaidi A, Alamri A. Smart energy optimization using heuristic algorithm in smart grid with integration of solar energy sources. Energies 2018;11(12). http://dx.doi.org/10.3390/en11123494, URL https://www.mdpi.com/1996-1073/11/12/3494.

[41] Hong Y-Y, Yo P-S. Novel genetic algorithm-based energy management in a factory power system considering uncertain photovoltaic energies. Appl Sci 2017;7(5). http://dx.doi.org/10.3390/app7050438, URL https://www.mdpi.com/2076-3417/7/5/438.

[42] Hazem Mohammed O, Amirat Y, Benbouzid M. Economical evaluation and optimal energy management of a stand-alone hybrid energy system handling in genetic algorithm strategies. Int J Electron 2018;7(10). http://dx.doi.org/10.3390/electronics7100233, URL https://www.mdpi.com/2079-9292/7/10/233.

[43] Lu H, Yang J, Alanne K. Energy quality management for a micro energy network integrated with renewables in a tourist area: A Chinese case study. Energies 2018;11(4). http://dx.doi.org/10.3390/en11041007, URL https://www.mdpi.com/1996-1073/11/4/1007.

[44] Lorencin I, Anđelić N, Mrzljak V, Car Z. Genetic algorithm approach to design of multi-layer perceptron for combined cycle power plant electrical power output estimation. Energies 2019;12:4352.

[45] Sadatiyan A. SM, Miller CJ. PEPSO: Reducing electricity usage and associated pollution emissions of water pumps. Water 2017;9(9). http://dx.doi.org/10.3390/w9090640, URL https://www.mdpi.com/2073-4441/9/9/640.

[46] Alonso JM, Alvarruiz F, Desantes JM, Hernandez L, Hernandez V, Molto G. Combining neural networks and genetic algorithms to predict and reduce diesel engine emissions. IEEE Trans Evol Comput 2007;11(1):46–55. http://dx.doi.org/10.1109/TEVC.2006.876364.

[47] Nikbakht S, Anitescu C, Rabczuk T. Optimizing the neural network hyperparameters utilizing genetic algorithm. J Zhejiang Univ Sci 2021;22(6):407–26. http://dx.doi.org/10.1631/jzus.A2000384.

[48] Tsakiri K, Marsellos A, Kapetanakis S. Artificial neural network and multiple linear regression for flood prediction in Mohawk River, New York. Water 2018;10(9). http://dx.doi.org/10.3390/w10091158, URL https://www.mdpi.com/2073-4441/10/9/1158.
[49] Mostafa N, Ramadan HSM, Elfarouk O. Renewable energy management in smart grids by using big data analytics and machine learning. Mach Learn Appl 2022;9:100363. http://dx.doi.org/10.1016/j.mlwa.2022.100363, URL https://www.sciencedirect.com/science/article/pii/S2666827022000597.
[50] Alzahrani A, Petri I, Rezgui Y, Ghoroghi A. Optimal control-based price strategies for smart fishery ports micro-grids. In: 2021 IEEE international conference on engineering, technology and innovation (ICE/ITMC). 2021, p. 1–8. http://dx.doi.org/10.1109/ICE/ITMC52061.2021.9570267.
[51] Alzahrani A. Decarbonising fishery ports through smart cluster energy systems (Ph.D. thesis), Cardiff university, Engineering school; 2021.
[52] Jacobsen S, Fossan K. Temporal variations in the glaze uptake on individually quick frozen prawns as monitored by the CODEX standard and the enthalpy method. J Food Eng 2001;48(3):227–33. http://dx.doi.org/10.1016/S0260-8774(00)00162-X, URL https://www.sciencedirect.com/science/article/pii/S026087740000162X.
[53] Petri I, Li H, Rezgui Y, Chunfeng Y, Yuce B, Jayan B. A HPC based cloud model for real-time energy optimisation. Enterp Inf Syst 2016;10(1):108–28. http://dx.doi.org/10.1080/17517575.2014.919053.
[54] Chu Y, Pedro H, Li M, Coimbra C. Real-time forecasting of solar irradiance ramps with smart image processing. J Sol Energy 2015;114. http://dx.doi.org/10.1016/j.solener.2015.01.024.
[55] Liu M, Quilumba FL, Lee W. Dispatch scheduling for a wind farm with hybrid energy storage based on wind and LMP forecasting. IEEE Trans Ind 2015;51(3):1970–7. http://dx.doi.org/10.1109/TIA.2014.2372043.
[56] Sideratos G, Hatziargyriou ND. An advanced statistical method for wind power forecasting. IEEE Trans Power Syst 2007;22:258–65.
[57] Ben-Nakhi A, Mahmoud M. Energy conservation in buildings through efficient A/C control using neural networks. Appl Energy 2002;73:5–23. http://dx.doi.org/10.1016/S0306-2619(02)00027-2.
[58] Zeinali Y, Niaki STA. Heart sound classification using signal processing and machine learning algorithms. Mach Learn Appl 2022;7:100206. http://dx.doi.org/10.1016/j.mlwa.2021.100206, URL https://www.sciencedirect.com/science/article/pii/S2666827021001031.
[59] Meraihi Y, Gabis AB, Mirjalili S, Ramdane-Cherif A. Grasshopper optimization algorithm: Theory, variants, and applications. IEEE Access 2021;9:50001–24. http://dx.doi.org/10.1109/ACCESS.2021.3067597.
[60] Renukadevi T, Karunakaran S. Optimizing deep belief network parameters using grasshopper algorithm for liver disease classification. Int J Imaging Syst Technol 2020;30(1):168–84. http://dx.doi.org/10.1002/ima.22375, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/ima.22375. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ima.22375.
[61] Jape VM, Suryawanshi HM, Modak JP. An efficient grasshopper optimization with recurrent neural network controller-based induction motor to replace flywheel of the process machine. Trans Inst Meas Control 2021;43(1):151–66. http://dx.doi.org/10.1177/0142331220938205.
[62] Moghanian S, Saravi FB, Javidi G, Sheybani EO. GOAMLP: Network intrusion detection with multilayer perceptron and grasshopper optimization algorithm. IEEE Access 2020;8:215202–13. http://dx.doi.org/10.1109/ACCESS.2020.3040740.
[63] Talaat M, Farahat M, Mansour N, Hatata A. Load forecasting based on grasshopper optimization and a multilayer feed-forward neural network using regressive approach. Energy 2020;196:117087. http://dx.doi.org/10.1016/j.energy.2020.117087, URL https://www.sciencedirect.com/science/article/pii/S0360544220301948.
[64] Sharma DK, Hota HS, Brown K, Handa R. Integration of genetic algorithm with artificial neural network for stock market forecasting. Int J Syst Assur Eng Manag 2021;1–14. http://dx.doi.org/10.1007/s13198-021-01209-5, PMC8367767[pmcid].
[65] Tang S-Z, Li M-J, Wang F-L, He Y-L, Tao W-Q. Fouling potential prediction and multi-objective optimization of a flue gas heat exchanger using neural networks and genetic algorithms. Int J Heat Mass Transfer 2020;152:119488. http://dx.doi.org/10.1016/j.ijheatmasstransfer.2020.119488, URL https://www.sciencedirect.com/science/article/pii/S0017931019351877.
[66] Shirzadi A, Asadi S, Shahabi H, Ronoud S, Clague JJ, Khosravi K, Pham BT, Ahmad BB, Bui DT. A novel ensemble learning based on Bayesian Belief Network coupled with an extreme learning machine for flash flood susceptibility mapping. Eng Appl Artif Intell 2020;96:103971. http://dx.doi.org/10.1016/j.engappai.2020.103971, URL https://www.sciencedirect.com/science/article/pii/S0952197620302797.
[67] Aliefa MH, Suyanto S. Variable-length chromosome for optimizing the structure of recurrent neural network. In: 2020 international conference on data science and its applications (ICoDSA). 2020, p. 1–5. http://dx.doi.org/10.1109/ICoDSA50139.2020.9213012.
[68] Yao X. Evolving artificial neural networks. Proc IEEE Inst Electr Electron Eng 1999;87(9):1423–47. http://dx.doi.org/10.1109/5.784219.
[69] Cook D, Ragsdale C, Major R. Combining a neural network with a genetic algorithm for process parameter optimization. Eng Appl Artif Intell 2000;13(4):391–6. http://dx.doi.org/10.1016/S0952-1976(00)00021-X, URL https://www.sciencedirect.com/science/article/pii/S095219760000021X.
[70] Urli T. Hybrid meta-heuristics for combinatorial optimization. Constraints 2015;20(4):473. http://dx.doi.org/10.1007/s10601-015-9209-7.
[71] Alla H, Moumoun L, Balouki Y. A multilayer perceptron neural network with selective-data training for flight arrival delay prediction. Sci Program 2021;2021:5558918. http://dx.doi.org/10.1155/2021/5558918.
[72] Sang S, Qu F, Nie P. Ensembles of gradient boosting recurrent neural network for time series data prediction. IEEE Access 2021;1. http://dx.doi.org/10.1109/ACCESS.2021.3082519.
[73] Ferreira GO, Ravazzi C, Dabbene F, Calafiore GC, Fiore M. Forecasting network traffic: A survey and tutorial with open-source comparative evaluation. IEEE Access 2023;11:6018–44. http://dx.doi.org/10.1109/ACCESS.2023.3236261.
[74] Skrobek D, Krzywanski J, Sosnowski M, Kulakowska A, Zylka A, Grabowska K, Ciesielska K, Nowak W. Implementation of deep learning methods in prediction of adsorption processes. Adv Eng Softw 2022;173(C). http://dx.doi.org/10.1016/j.advengsoft.2022.103190.
[75] Choi W, Lee S. Performance evaluation of deep learning architectures for load and temperature forecasting under dataset size constraints and seasonality. Energy Build 2023;288:113027. http://dx.doi.org/10.1016/j.enbuild.2023.113027, URL https://www.sciencedirect.com/science/article/pii/S0378778823002578.
[76] Sajjad M, Khan ZA, Ullah A, Hussain T, Ullah W, Lee MY, Baik SW. A novel CNN-GRU-based hybrid approach for short-term residential load forecasting. IEEE Access 2020;8:143759–68. http://dx.doi.org/10.1109/ACCESS.2020.3009537.
[77] Phan Q-T, Wu Y-K, Phan Q-D, Lo H-Y. A novel forecasting model for solar power generation by a deep learning framework with data preprocessing and postprocessing. IEEE Trans Ind Appl 2023;59(1):220–31. http://dx.doi.org/10.1109/TIA.2022.3212999.
[78] Wen T, Gasic M, Mrksic N, Su P, Vandyke D, Young SJ. Semantically conditioned LSTM-based natural language generation for spoken dialogue systems. 2015, CoRR abs/1508.01745.
[79] Sundermeyer M, Ney H, Schlüter R. From feedforward to recurrent LSTM neural networks for language modeling. IEEE/ACM Trans Audio Speech Lang Process 2015;23(3):517–29. http://dx.doi.org/10.1109/TASLP.2015.2400218.
[80] Sak H, Senior A, Beaufays F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In: Proc. annu. conf. int. speech commun. assoc. INTERSPEECH. 2014, p. 338–42.
[81] Alsaeegh M, Shukur O. Using multiple regression model and RNN for imputing the missing values of PM10 datasets. Int J Adv Sci Eng Inf Technol 2020;10:2582. http://dx.doi.org/10.18517/ijaseit.10.6.11236.
[82] Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks. In: 2013 IEEE international conference on acoustics, speech and signal processing. 2013, p. 6645–9.
[83] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9:1735–80. http://dx.doi.org/10.1162/neco.1997.9.8.1735.
[84] Graves A. Supervised sequence labelling with recurrent neural networks. Studies in computational intelligence, Berlin: Springer; 2012, http://dx.doi.org/10.1007/978-3-642-24797-2.
[85] Gers FA, Schmidhuber J. Recurrent nets that time and count. In: Proceedings of the IEEE-INNS-ENNS international joint conference on neural networks. IJCNN 2000. Neural computing: New challenges and perspectives for the new millennium, Vol. 3. 2000, p. 189–94.
[86] Olah C. Understanding LSTM networks. 2015, URL http://colah.github.io/posts/2015-08-Understanding-LSTMs/.
[87] Cho K, van Merrienboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. 2014, CoRR abs/1406.1078.
[88] Kong W, Dong ZY, Jia Y, Hill DJ, Xu Y, Zhang Y. Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Trans Smart Grid 2019;10(1):841–51.
[89] Cheng K-S, Lien Y-T, Wu Y-C, Su Y-F. On the criteria of model performance evaluation for real-time flood forecasting. Stoch Environ Res Risk Assess 2016;31:1123–46.