euspen's 25th International Conference & Exhibition, Zaragoza, ES, June 2025
www.euspen.eu

Manufacturing scheduling for sustainability

Jon Hall[1], Zebokhon Abduraimova[1], Marco Franchino[1], Matteo Zamarco[1], Clementine Tace Morgan[1]

[1]*University of Sheffield, Advanced Manufacturing Research Centre, Sheffield, S9 1ZA, UK*

c.t.morgan@sheffield.ac.uk, jon.hall@sheffield.ac.uk

Abstract
As most of the world strives for improved sustainability, and the UK works towards its Net Zero 2050 goals, emissions from manufacturing have a significant part to play. One potential method of improving the sustainability of manufacturing processes is to reduce the CO_2 equivalent (CO_2E) of the process. The CO_2E of a process can be reduced by using energy from renewable sources as these have a lower carbon intensity.
The carbon intensity data of the UK electricity supply is published live, and a forecast is provided to predict the carbon intensity for the next 48 hrs. The carbon intensity score is based on weather prediction data. Historical and forecast weather data could therefore be used for predictive energy optimisation in manufacturing facilities.
This paper looks to investigate if the carbon intensity forecast and weather data can be used to schedule manufacturing, so electricity is used when CO_2E is at its lowest. A methodology for capturing future predictive data was developed, using it to generate a predictive CO_2E for a given process. Based on this an algorithm was developed that suggests manufacturing time from carbon intensity forecast data and enables the forecast to be compared directly with the actual CO_2E data when the process is complete. Scenario testing was carried out using simulation to investigate when this would be applicable for manufacturers and when it would provide the most benefit.
This work gives manufacturers the tools to examine the viability of scheduling manufacturing to minimise CO_2E of their processes and reduce energy costs.

decision-making, manufacturing, scheduling, sustainable development

1. Introduction

In the UK energy production is split between renewable energy sources such as wind and solar power, and fossil fuel sources such as coal and gas. The power output of the renewable energy sources fluctuates based on weather conditions, so easily controlled energy sources such as gas-fired power stations are used to meet demand when renewable energy production is insufficient. As a result of this, the amount of Carbon dioxide equivalent (CO_2E) produced by an energy-intensive manufacturing process varies depending on weather conditions. CO_2E is a standardised measurement to understand the 'Global Warming Potential' of greenhouse gases, as defined by the Intergovernmental Panel on Climate Change (IPCC) of the United Nations [1]. The carbon intensity of energy produced from a mix of renewable and fossil fuel sources can be calculated to represent how many grams of CO_2E are released to produce a kilowatt hour of electricity. This can then be used to determine how much CO_2E is attributed to a manufacturing process which uses that electricity. Project Butterfly, funded by Made Smarter in 2022, developed a carbon intensity calculator which used material consumption and energy usage to determine the CO_2E of manufacturing processes [2]. From this work, it was determined that the carbon intensity of a manufacturing process could be predicted based on weather forecast data.

This paper details the development and deployment of a machine learning model capable of predicting the carbon intensity of a manufacturing process based on weather forecast data. The model uses publicly available weather forecast data to determine how much power renewable energy sources will contribute to the grid over the next 48 hours and generates a time dependant value for carbon intensity during that period. The program can be used to schedule a manufacturing process (defined by its power consumption and duration) within the next 48 hours to minimise the CO_2E produced by the execution of the process.

Section 2 describes the software development process for the carbon intensity prediction. Details are given on how the data used for training and testing the program was cleaned and formatted prior to the training of the model, and how additional features were generated from the data to improve the quality of the prediction. The process of exploratory data analysis which was used to gain insights into overall trends in the data is also described. Some representative predictions of the model are analysed and used to compare the prediction quality from different data sources. Section 3 describes how results from the prediction model can be used in discrete event simulations (DES) to schedule more complex manufacturing processes for minimal CO_2E production across multiple machines and processes. Section 4 discusses the results of the discrete event simulations and compares simulations which prioritise sustainability with those that prioritise production rate.

2. Methodology

This section provides an overview of the methodology used to build a carbon intensity prediction model based on weather and irradiance data. It describes the development of two prediction models using different data sources with varying sizes and qualities. After evaluating their performance, the most stable model was selected as the best choice for further use.

2.1. Data collection and processing

The data collection step involved the acquisition of both historical and forecast data, ensuring that they contain similar, aligned variables. Historical data was targeted for acquisition over a period of at least two years to ensure that the prediction model could capture the seasonality of the data. The national historical carbon intensity dataset was available for free download from the National Energy System Operator (NESO) covering the period from January 1, 2009, to the present [3].

Midas Open is the open data version of the Met Office Integrated Data Archive System (MIDAS), containing weather and global irradiance observations dating back to 1853 [4]. The dataset has inconsistencies due to the historical nature of the data. For this project, the latest version available at the time—v202308—was used, with weather data from Heathrow, London from 2019 to the end of 2022. Forecast weather data was requested daily from the Met Office DataHub Atmospheric Model for Heathrow, London, to align with the historical Midas Open weather data [5]. Accurate and comprehensive historical irradiance and weather data were downloaded from SolCast for the available station in Birmingham, England, covering the extended period from 2007 to 2024 [6]. This dataset offers greater consistency and accuracy compared to Midas Open, as it spans a longer, uninterrupted timeframe. To align with the SolCast historical data, forecast irradiance data and forecast weather data were requested daily from SolCast and Met Office Site Specific APIs respectively.

Renewable quarterly electricity capacity and generation statistics were acquired from the Department for Energy Security and Net Zero, covering the period from 2010 to the second quarter of 2024. Estimates for the third and fourth quarters of 2024 were calculated using the percentage change from the same quarters in 2023 [7].

The units for historical weather and radiation observations were converted to ensure consistency with the forecast data units; for example, air temperature was converted from Celsius to Kelvin. The units of the historical weather observations were aligned with the Met Office Site-Specific forecast weather data, while the units of the historical irradiance data were matched with those of the SolCast irradiance forecast data.

2.2. Feature Engineering

The weather, energy capacity, and carbon intensity datasets were merged based on datetime, resulting in two datasets for prediction model training:

Dataset 1 – Met office: Midas Open historical weather and radiation data, quarterly energy trends statistics, and carbon intensity datasets for the period from 2019 to 2023.

Dataset 2 – SolCast: SolCast historical weather and radiation data, quarterly energy trends statistics, and carbon intensity datasets for the period from 2010 to the second quarter of 2024.

For the first set, the available raw weather features that aligned with the corresponding forecast data included wind speed, wind direction, gust speed, total global radiation, air temperature, total cloud cover, visibility, mean sea level pressure, dew point temperature, relative humidity, and station pressure.

For the second set, the available raw weather features that aligned with the corresponding forecast data comprised wind speed, wind direction, air temperature, azimuth, cloud opacity, dew point temperature, diffuse horizontal irradiance (dhi), direct normal irradiance (dni), global horizontal irradiance (ghi), precipitation rate, relative humidity, and zenith.

These native features were engineered to create additional variables capturing interactions between wind speed and air temperature, temperature differences between air and dew point, and the squared wind speed to help the model better understand the relationships between weather observations and carbon intensity.

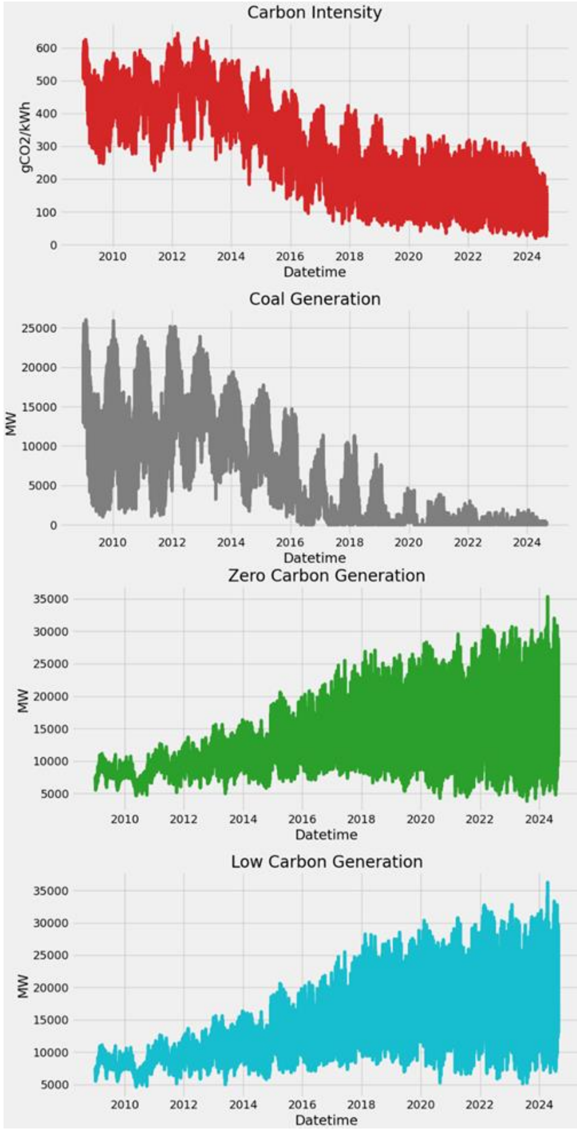

Figure 1. Carbon intensity and energy generation trends in the UK between 2010 and 2024.

The cyclical nature of features such as wind direction and time, was represented by calculating their sine and cosine values, which were added as new features to assist the model in capturing cyclical trends in the data. This is supported by cyclical behaviours in the historical data, which includes annual, monthly, weekly, and daily fluctuations in carbon intensity.

The quarterly energy trends statistics, which capture the increasing trend of renewable energy sources, were incorporated as features. This includes installed capacities for wind and solar energy generation, as well as the percentage of their load factors and shares of generated electricity.

Some exploratory data analysis was conducted to gain an in-depth understanding of the data, its structure, patterns, and the relationships between variables through visualisation. Overall, there is a downward trend in carbon intensity over time, driven by a reduction in coal usage and the increasing adoption of low and zero-carbon energy generation sources, as shown in **Error! Reference source not found.**. This trend is further reinforced by the growth in installed wind and solar capacities and their increasing share in energy generation over time.

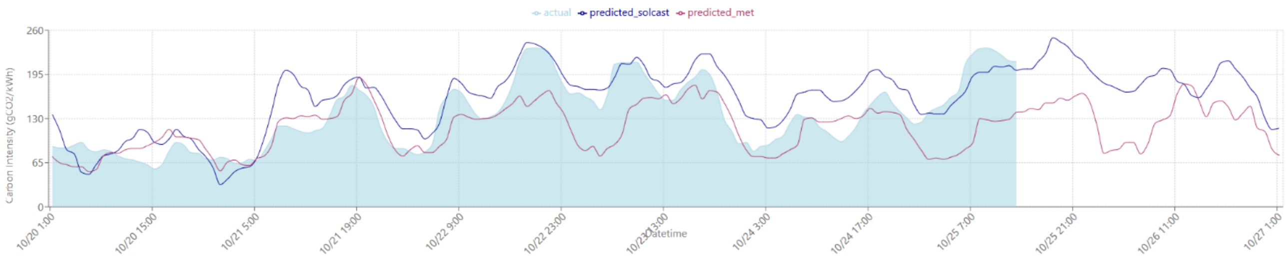

Figure 2. Carbon intensity over time as predicted using the SolCast model (blue), and the Met Office model (red). Actual carbon intensity values are overlaid for comparison.

2.3. Building the Prediction Model

A regression algorithm from the open-source XGBoost library was selected to build the prediction model for both datasets: historical data from SolCast and Met Office [8]. The Gradient Boosting Tree Regressor was chosen for this project due to its effectiveness in handling continuous prediction tasks, its ability to manage complex relationships in the data, and its high performance in various regression scenarios, including time-series forecasting of carbon intensity. XGBoost also automatically handles missing values without requiring explicit imputation during the data preprocessing stages.

For the SolCast dataset, the training set includes records from 2010 to 2021, while the testing set consists of data from 2022 to 2024, representing approximately 17% of the data for testing and 83% for training. The Met Office dataset was split similarly, with the training set covering the period from 2019 to 2021 (75%) and the testing set covering 2022 (25%), ensuring that the testing part includes at least one full year.

To avoid overfitting and to ensure the model generalises well and performs accurately on new data, a technique called time series cross-validation was used. At each step, the model's performance was evaluated using mean squared error to provide insights into how well the model generalises across different time periods and performs on new, unseen data.

For this project, an automated hyperparameter optimisation framework—Optuna, which is based on Bayesian optimisation—was used across the time series cross-validation folds.

Hyperparameter tuning for the XGBoost Regressor model was conducted using time series cross-validation iterations on both datasets: SolCast and Met office. At the end of this process, models with optimal hyperparameters that performed best according to the chosen objective—mean squared error—were selected. The SolCast dataset spans a longer period, allowing the model to capture yearly trends, and the increasing energy statistics of wind and solar installed capacities over time, which balance out the overall decreasing carbon intensity trend.

The best models from both datasets were evaluated on their respective test sets. Predictions based on the SolCast data showed a stronger correlation with actual carbon intensity values, with the model achieving a higher R-squared score of 0.14. This indicates that the model explains 14% of the variance in the data. In contrast, the Met model's R-squared score is negative, suggesting it performs worse than a simple mean-based model. Both models exhibit relatively similar results in terms of root mean squared error (RMSE). The model built on Met Office's data exhibits instability, with performance degrading over time. In contrast, the SolCast model's performance measures, while showing slight variation, remain generally stable. Consequently, the model trained on SolCast historical data was selected for scheduling manufacturing processes based on carbon predictions from weather forecasts.

The resulting model was capable of predicting the carbon intensity of energy production within the next 48 hours. The correlation between the predicted carbon intensity trend and the actual values was 0.79, showing a strong correlation between the two. Figure 2 shows a representative seven day period which demonstrates the accuracy of the carbon intensity prediction. Particularly the models capacity to predict the maxima and minima in carbon intensity as a function of time. This is of particular interest when scheduling manufacturing processes, as it means that the model can identify the best and worst times to use a machine with regards to sustainability. Given the power consumption of the machine, this information can then be used to calculate the potential CO_2E reduction if the manufacturing process were scheduled to prioritise sustainability.

3. Discrete Event Simulation

A Discrete Event Simulation (DES) model was developed using AnyLogic software to examine the relationship between makespan (the time taken to complete a set of jobs in a manufacturing process) and carbon emissions in manufacturing scheduling. The primary objective was to evaluate the differences between the scheduling goals of minimising either parameter.

The DES model, shown in Figure 3**,** represents a sequential three-stage manufacturing process. Jobs may queue in unlimited capacity buffers before each stage. By varying the release times of jobs at each stage, different scheduling scenarios can be simulated. The simulation was run over 48 hours, corresponding to the prediction timeframe for carbon intensity (as detailed in section 2), which is a necessary input to estimate the carbon emissions of the process.

Synthetic data generation was employed initially, with variations in machine power consumption and carbon intensity sequences. This approach aimed to determine the influence of these inputs on the relationship between the two outputs. Job cycle times across the three processes were fixed at three hours and three jobs were to be scheduled. The research questions addressed in this phase were:

R1.1 What is the distribution of makespan and carbon emissions under varying power consumption and carbon intensity predictions?

R1.2 Are there specific trends or relationships between the two?

The second phase of the work focused on a real-world use case involving melding, cutting, and machining processes. Machine power consumption and job cycle times were derived from collected machine data. As in phase 1, three jobs were to be scheduled in a 48 hour period. The research questions were:

R2.1 What are the distributions of makespan and carbon emissions, randomising job scheduling for each simulation run?

R2.2 What is the minimal carbon emission achievable, and what is the corresponding makespan?

R2.3 What is the minimal makespan achievable, and what are the associated carbon emissions?

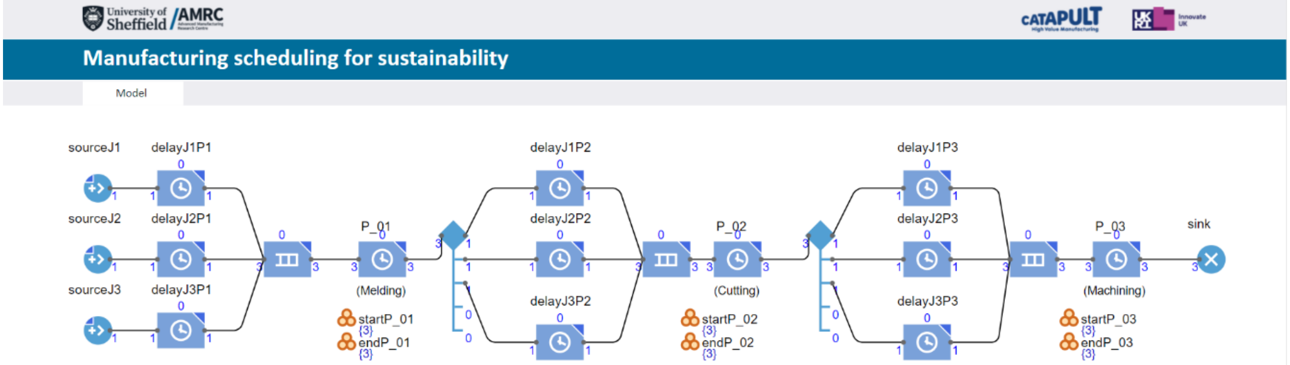

Figure 3. DES model for a three-stage manufacturing process.

4. Results

Results for the first phase of DES modelling work looked at the relationship between makespan and carbon emissions for an array of simulated manufacturing processes, varying the relative power consumption of the three processes for each simulation. Makespan distributions remained consistent across scenarios, while carbon emissions converged toward the average as the standard deviation of machine power consumption increased. For example, lower standard deviations result in evenly distributed emissions. In contrast, higher standard deviations yield narrower emission ranges. Despite these distinct patterns, no significant relationship emerges between the two output variables.

Results for the second phase of DES modelling demonstrated similar behaviours when using real-world machine data. Using collected machine data to simulate a manufacturing process involving 3 machines, 7.2 hours of machine time, and consuming a total of 578.7kWh of electricity. The analysis revealed a slightly positive correlation between makespan and carbon emissions, with reductions in makespan generally accompanied by lower carbon emission across most carbon intensity sequences. A visual example is provided in Figure 4.

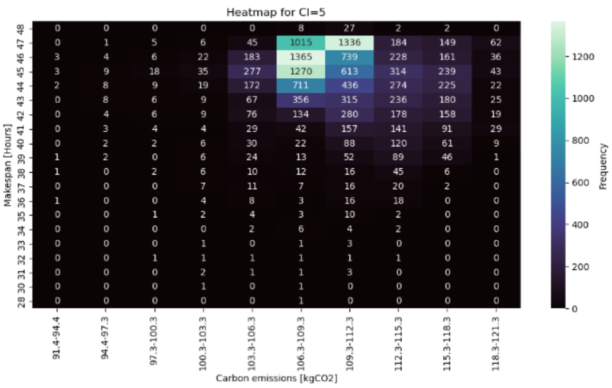

Figure 4. Frequency of simulation runs by makespan and carbon emissions, randomising job scheduling.

Regarding the two remaining research questions, Table 1 summarises the minimum makespan and carbon emission achieved under different optimisation objectives.

Table 1. DES results minimising both carbon emissions and makespan.

Objective	Emissions (kg CO_2E)	Makespan (hours)
Min (CO_2E)	81.10	20.0
Min (Makespan)	85.84	17.2

These results demonstrate that scheduling this manufacturing process to minimise CO_2E emissions reduced them by 4.74kg. Prioritising the sustainability of the process also created a production schedule which was close to the minimum makespan for this process.

5. Conclusions

The work described in this article demonstrates the development and application of a manufacturing scheduling model which minimises CO_2E emissions for a given process based on weather forecast data. The development and testing of the prediction model has been described, along with results of data analysis which showed trends in historical carbon intensity. This was factored into the model, along with the observed periodicity of the data. The resulting prediction model proved capable of predicting the optimal time within the next 48 hours to execute a process in order to minimise CO_2E emissions.

The prediction model was then used as part of a DES model to investigate how it could be used to schedule a more complex manufacturing process involving multiple machines. These simulations revealed positive relationship between the makespan and the carbon emissions of the manufacturing process, indicating that minimising carbon emissions would also create a process schedule with a short makespan. This has shown that the use of manufacturing scheduling based on Carbon intensity has the potential to reduce CO_2E of manufacturing processes. However, further validation on real world manufacturing datasets is required to fully understand the impact.

References

[1] Intergovernmental Panel on Climate Change (IPCC). (2001). Climate Change 2001: Synthesis Report. Watson, R.T et al Cambridge University Press

[2] Jadhav M, Morgan C AMRC Journal **18** 36-39 Available at: https://www.amrc.co.uk/files/document/515/1726047916_AMRC_Journal_18_AW-Digital.pdf [Accessed: Oct. 1, 2024]

[3] NESO. (2024). Historic generation mix and carbon intensity. Available at: https://www.neso.energy/data-portal/historic-generation-mix [Accessed: Oct. 1, 2024]

[4] CEDA Archive. (2024). Met Office MIDAS Open: UK Land Surface Stations Data (1853-current). Available at: https://catalogue.ceda.ac.uk/uuid/dbd451271eb04662beade68da43546e1/

[5] DataHub Met Office. (2024). Site-specific forecast data. Available at: https://datahub.metoffice.gov.uk/docs/f/category/site-specific/overview [Accessed: Sep. 1, 2024]

[6] SolCast. (2024). Historical Time Series. Global irradiance and weather data records designed and built for solar and cleantech. Available at: https://www.solcast.com/time-series [Accessed: Sep. 1, 2024]

[7] Department for Energy Security and Net Zero. "Energy Trends. Quarterly publication which presents data on the supply and demand of all the major fuels in the United Kingdom.". Gov.uk. Sep. 26, 2024. [Online]. Available: https://www.gov.uk/government/collections/energy-trends. [Accessed: Oct. 1, 2024]