Findings Surrogate Modeling to Quantify the Resilience of an Energy System Design

1 Introduction

This research project aims to create AI-driven surrogate models to predict the cost and curtailment of a given energy system design in the Netherlands. This has been done by training several machine learning models, both simple traditional ones(ML Models) and more advanced neural networks (Neural Network), on labeled training data. The training data was generated through the original Calliope model(Calliope), where several input variables- the gas price and the wind capacity- were varied and Monte Carlo sampling techniques were applied(Input & Monte Carlo). The output of this data consisted of monthly predictions of both the cost and curtailment(Neural Network). Costs refer to the operational costs of the given design for a specific month. Curtailment represents the deliberate reduction of the output of renewable energy below what could have been produced in order to balance energy supply and demand or due to transmission constraints. Furthermore, the wind capacity was encoded for the surrogate models(Data preprocessing), such that it was represented by merely one feature. Moreover, months were circularly encoded using sine and cosine techniques to reflect seasonality in the models. Lastly, unseen data was created, again using Monte Carlo sampling(Input & Monte Carlo), to test the overall resilience of the energy system design. In this document, the findings of all the modeling are being discussed, by presenting the results and discussing the performance of the models for different cases.

2 Results

2.1 Evaluation of the Surrogate Models

The following table 1, shows the performance of the best surrogate models per case. Case refers here to the combination of inputs that have been varied for that specific model together with the output variable. The performance has been tested by looking at the MAE (Mean Absolute Error) of each model. Subsequently, the model with the lowest MAE was chosen as the final model for that case. To give a better understanding of the relative size of the error (MAE value), the standard deviation of the training data set is also stated in table 1. For most cases the 3 layer neural network has the best performance (Results Training Neural Networks) ; these neural networks consist of 100-500 units per layer.

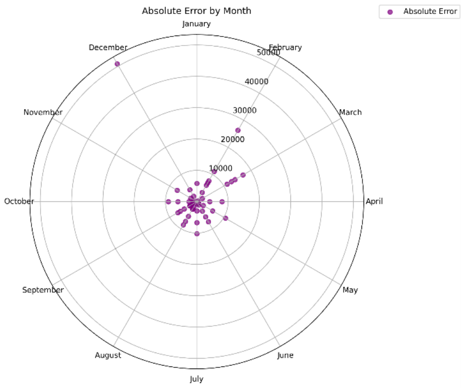
Additionally, for evaluating the performance of the surrogate models, a weight metric for different seasons during the year has been applied. Winter months were discarded as more important and therefore weighted more heavily. For policymakers that prioritize costs or curtailment reduction during winter months, this could be a meaningful metric. The final MAE resulted in about a 10% increase compared to the unweighted MAE. However, since this holds for all cases, the weighted MAE has been left out of table 1, to ensure clarity of the results.

Inputs	**Output**	**Best Model**	**Final MAE**	**Std**
Gas Price	Cost	Linear Regression	149.8	392145
Gas Price, Month	Cost	3L Neural Network	6224	412106
Wind, Month	Cost	3L Neural Network	18001	75282
Gas Price, Month	Curtailment	Decision Tree	6114	678186
Wind, Month	Curtailment	3L Neural Network	177152	496279

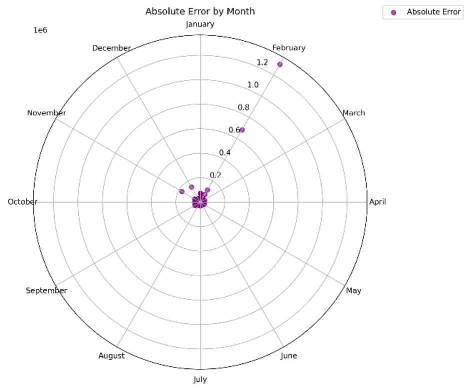
Table 1: Performance of the Best Surrogate Models.

2.2 Performance of the Surrogate Models

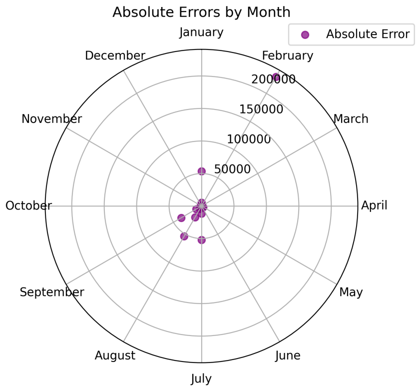
In figure 1, plots of the best models per case, see table 1, display the performance in terms of absolute errors for all predicted values per month. For three cases, the models perform generally well except for the month February, see figures 1a, 1b and 1c. As for the fourth case, the surrogate models performed worse on the autumn months, see 1d. Moreover, the errors are higher compared to the other cases. However, after training the surrogate models with just the monthly variable as input, it followed that the resulting errors were almost twice as high. This indicates that, even though the performance of the case with wind and month as input and curtailment as output (figure 1d), is not as good as for the other cases, the encoded wind variable still contributes to a better outcome of the predicted curtailment.



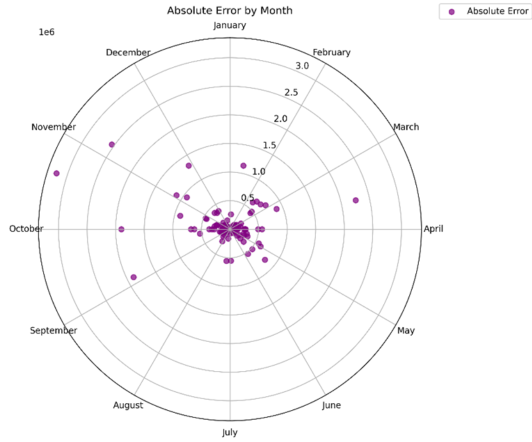
(a) Gas Price, Month; Cost



(b) Wind, Month; Cost



(c) Gas Price, Month; Curtailment



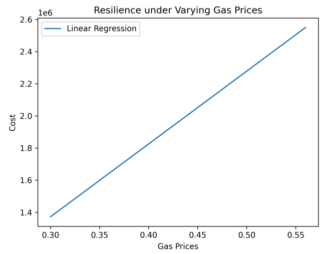
(d) Wind, Month; Curtailment

Figure 1: Absolute Errors per Month for Different Input Variables and Outputs.

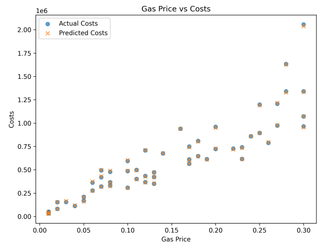
2.3 Interpretation of the Predictions

From figure 2, relationships between the input variables and the predicted output- either cost or curtailment- can be seen. It follows that the gas price has a linear correlation with the predicted costs, see figures 2a and 2b. For the curtailment, however, different gas prices result in similar predictions. Therefore, these two variables appear to be relatively independent of each other.

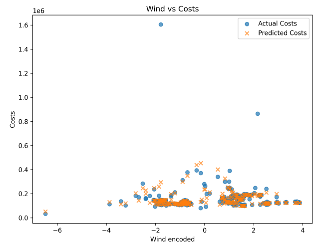
As for the encoded wind variable, figure 2c shows that the costs are, except for two outliers, mostly of the same order. Curtailment shows a negative trend, where a larger value of the encoded wind results in a lower curtailment. Figure 2e also shows that outliers are not captured by the surrogate model, but the negative trend is.



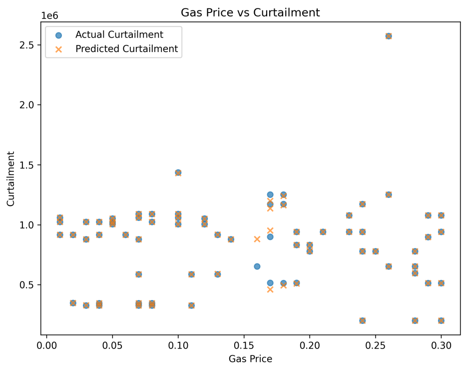
(a) Gas Price; Cost



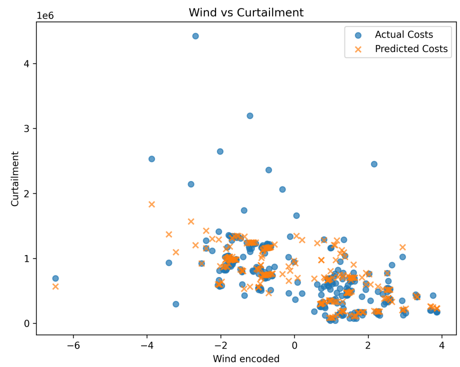
(b) Gas Price, Month; Cost



(c) Wind, Month; Cost



(d) Gas Price, Month; Curtailment



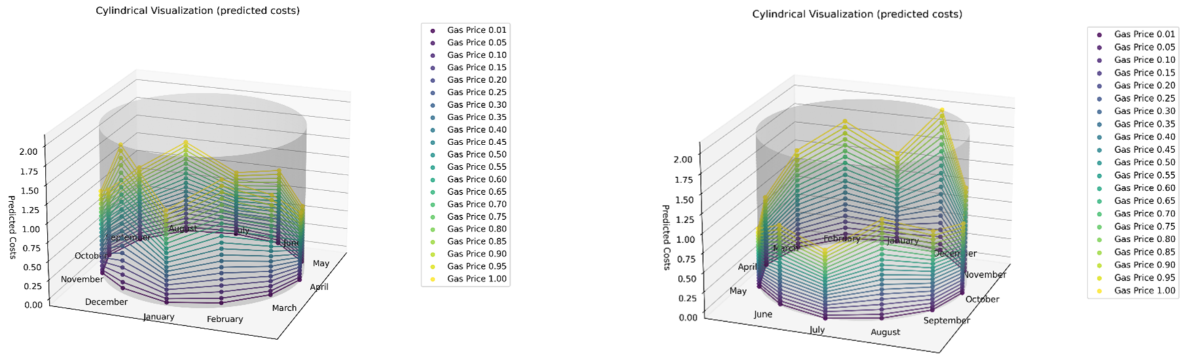
(e) Wind, Month; Curtailment

Figure 2: Actual and Predicted Values for Different Input Variables and Outputs.

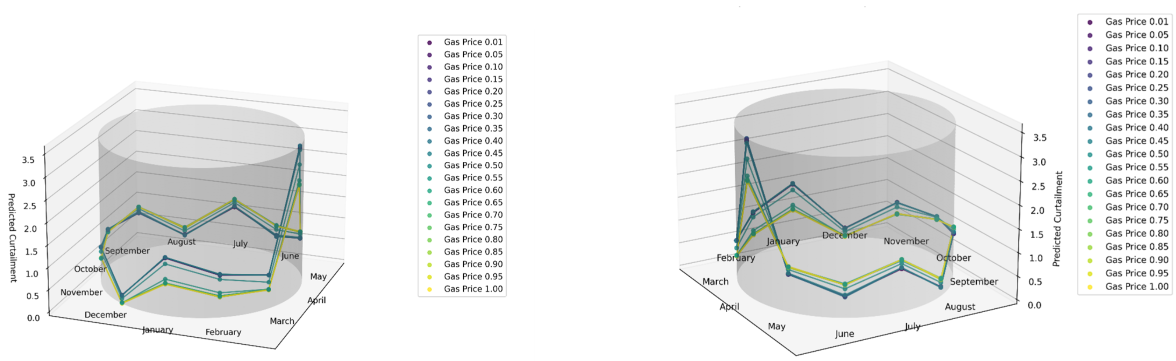
2.4 Application Surrogate Models

After having evaluated the performance of the surrogate models for different cases, new unseen data- generated using Monte Carlo sampling- was used to make predictions with these models. Because the surrogate models have not captured a clear relationship between the encoded wind variable and the output variables (costs and curtailment), visualizations of these predictions are hard to interpret. As for the gas price variable, however, relationships have been found and are displayed in figure 3. Figure 3a shows the months on the x and y axis, circular encoded, and predicted costs for different gas prices on the z axis. It follows that, just like in figure 2b, a higher gas price linearly results in a higher predicted cost.

Figure 3b again shows the circularly encoded months on the x and y axes, but this time the predicted curtailment is displayed on the z axis. It follows that, in spring, the predicted curtailment is higher compared to other months. This could be explained by the fact that, during these months, the weather conditions result in a relatively high amount of renewable energy, whereas the total energy demand is generally lower than the supply. This spike in curtailment holds for all different gas prices and is of the same order.



(a) Inputs: Gas Price, Month - Output: Cost.



(b) Inputs: Gas Price, Month - Output: Curtailment.

Figure 3: 3D Cylindrical Plots of the Predicted Values of Unseen Data using Gas Price and Month as Input Variables.

3 Conclusion

The results demonstrate that out of all the models tested, three-layer neural networks consistently delivered the best overall performance for most cases. A variety of machine learning models and neural networks with 1 to 4 layers were repeatedly trained and tested on the dataset, but the three-layer neural networks consistently stood out. Particularly with 100 until 500 units/layer as configuration had the best performance. It can be concluded that surrogate modeling for this amount of variables works the best using neural network architectures of this size and depth.

Furthermore, we can conclude that the surrogate model has a strong performance in capturing the relationship between input variables gas price, month, and output variables curtailment or costs. Specifically, the model performs well in predicting costs based on gas price and month, as this relationship is mostly linear for the given energy system design. Furthermore, curtailment can also be accurately predicted based on gas price and month, with higher curtailment values typically occurring in the spring months. The model also captures the relationship between wind, month, and costs effectively, showing good performance in this area. However, when it comes to predicting curtailment based on wind, the model struggles to find the relationship, leading to poorer performance. That said, the results are still better than when only the month is considered. The decline in performance with the encoded wind variable partly stems from the encoding method used for the wind variable, as well as the relatively small training data set. Therefore, future research on encoding strategies is necessary to improve performance for this specific case. Additionally, a weighted mean absolute error (MAE) metric was introduced, transforming the traditional MAE by applying weights to each month. With a change of +- 10% This allows the surrogate model to be evaluated with greater flexibility, enabling users to prioritize specific months when assessing performance.

4 Discussion

The results indicate several points for discussion. Firstly, the monthly variable was encoded using sine and cosine values. This feature engineering step substantially improved the performance of the surrogate model and is worth revisiting in future research. Furthermore, the time series of the wind were encoded into one value per month, reducing over 700 hourly values across 12 locations into a single feature. This encoded wind feature resulted in mixed performance. The model was able to capture the relationship between the encoded wind value, month and costs, although it lead to a larger error than for other cases. On the contrary, the curtailment was hard to predict when varying the wind and using this encoded value. In future research, this problem can be improved by in-depth exploration of the feature engineering. Encoding both the temporal and the regional parts of the wind time series is challenging and therefore meaningful to research.

To expand the model and make it more applicable for real-world scenarios, more variables could be taken into account. Variables such as photovoltaic capacity, electricity, and wind capacity offshore are all time series and therefore should be encoded into meaningful features as well. Moreover, to make it more applicable for real-world scenarios, the energy system design option should also be less simplified, to better simulate the effect of varying these input variables. This can be investigated to expand the surrogate model to eventually capture importance between all the variables. In this way, the surrogate models predictions for costs and curtailment are based on all variables the Calliope model also uses as input, therefore making the surrogate model a more useful tool.

Lastly, the performance of the surrogate model can always be improved by increasing the number of data points. Gathering the data points through running Calliope is time-costly because of Calliope's running time and the formatting of the calliope data into data the surrogate model can be trained on. That is why the datasets used for training and evaluation were relatively small (from 100 to 1500 data points). Running Calliope more extensively in the future and finding an efficient way of processing this data into useful training and test formats will increase the performance of the surrogate model.

In summary, future research for surrogate modeling based on Calliope energy optimization involves three key areas: (1) feature engineering for time series variables with temporal and locational components, (2) incorporating additional variables such as photovoltaic and offshore wind capacity, and (3) increasing the dataset size by optimizing Calliope runs and data formatting processes. By addressing these challenges, surrogate models can become more comprehensive, accurate, and broadly applicable.