PISTON COMPRESSOR PERFORMANCE WITH TERNARY MIXTURES: MACHINE LEARNING SURROGATE MODELS

Alexandra Welp[1]*, Ilyas Adaköy[1], Burak Atakan[1]

[1]Thermodynamics, University of Duisburg-Essen, 47057 Duisburg, Germany

*Corresponding Author: alexandra.welp@uni-due.de

ABSTRACT

Mixtures with a large temperature glide can help reducing the exergy losses in the heat exchangers of heat pumps, for the charging step of Carnot Batteries, but also in other thermodynamic cycles. Although the heat exchanger performance may be well evaluated in such calculations, the overall picture might be wrong, because the efficiency of compressors significantly impacts the overall efficiency of heat pumps. The isentropic efficiency depends on different parameters, like the compressor speed, pressure level, pressure ratio or the working fluid composition. Despite this, many scientists use constant efficiencies within thermodynamic cycle calculations for evaluating different working fluids. Especially when assessing the influence of fluids and fluid mixtures on the compressor performance, detailed cycle-resolved physical models are needed to obtain good predictions, leading to time consuming calculations. When optimizing fluid composition within thermodynamic cycles frequently, time-intensive calls to the compressor model are necessary, thus faster alternatives are being sought.

This work explores the use of surrogate models based on physical models of compressors in heat pumps. The surrogate models are created through machine learning algorithms to reduce computation time and to access the importance of operation parameters. The investigated system consists of a ternary mixture of pentane, isobutane and propane of variable composition that is compressed at different pressure ratios. A fluid screening is established to verify the suitability of a certain fluid mixture and inlet condition prior to training, which already reduces the area of interest by 68 %. The resulting (non-) suitability was systematically analysed with regard to resulting available temperature-glide and isentropic efficiency, volumetric efficiency, and outlet temperature. Based on this, a database is created to provide the data set for machine learning. Multilayer perceptrons (MLP) with hyperparameter optimization are selected as regressors based on their promising performance in previous work. To assess the influence of the different parameters, we trained several MLPs with different feature combinations and performed permutation feature importance analysis for the targets isentropic efficiency, volumetric efficiency, and outlet temperature. Using superheating temperature difference, inlet temperature, pressure ratio, and mole fractions as features, we found that the pentane mole fraction had the highest permutation feature importance for the isentropic efficiency. Replacing the mole fractions with the inlet density reduced the relevant features to pressure ratio, inlet temperature and density while still achieving excellent prediction accuracies, with an average relative deviation of 1.5 % for the isentropic efficiency, 0.4 % for the volumetric efficiency and 0.56 % for the outlet temperature. The surrogate model reduces computational time by 99.4% for 100 000 data points compared to the semi-empirical model, including the database creation. This offers a significant advantage for computational optimization within the trained parameter space.

1 INTRODUCTION

The avoidance of entropy production is of particular importance for Carnot Batteries which combines the strengths but also weaknesses of heat pumps and power cycles. Therefore any avoidable loss must be minimized to achieve a sufficiently good round-trip-efficiency (Atakan, 2024). The utilisation of zeotropic mixtures in heat pumps and other thermodynamic cycle processes including Carnot Batteries

37th INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

335

(e. g. in Bernehed (2021) or Koen *et al.* (2021)) offers an opportunity to reduce the exergy loss in heat exchangers indicated by temperature differences. For a heat pump, the coefficient of performance (COP) is usually used for assessment. Zühlsdorf *et al.* (2018) investigated binary mixtures of natural refrigerants and found an increased performance by up to 27 %. Guo *et al.* (2019) also studied subcritical binary mixtures, but extended the study to a cost estimate and concluded that off-the-shelf components can be used. Huang *et al.* (2023) extended the idea to use a novel configuration to actively adjust the mixture composition depending on the operating condition. In these studies, cited as examples, the focus was only on the advantage of temperature glide in the heat exchanger, but the variation of the compressor performance with fluid composition was not addressed. As Roskosch *et al.* (2021) have shown, the compressor efficiencies are fluid- and composition-specific, which has a significant influence on the COP. Neumaier *et al.* (2023) have investigated this further by comparing COPs for different fluids with individually scaled and optimised piston compressors and found that the best-performing refrigerant is not even identified among the best ten refrigerants when constant isentropic efficiencies were assumed. Reliable models for determining fluid specific efficiencies can be computationally intensive. This is particularly problematic when computational optimisations are carried out with numerous calls. In addition to computationally efficient implementation in alternative programming languages (as carried out by Neumaier *et al.* (2023)), a physical model can be replaced by a surrogate model. Various machine learning (ML) methods are available for this purpose, which are used in literature and practice, e.g. for failure diagnosis and monitoring over time (e.g. for LNG compressors in Hidalgo-Mompeán *et al.* (2021) and for gas-turbines in Liu and Karimi (2020)). Also the performance prediction using ML (e.g. in Vering *et al.* (2024) and Wu *et al.* (2020)) and optimization procedures (e. g. Joly *et al.* (2019), Masood *et al.* (2021) and Zhu *et al.* (2020)) were discussed recently in several publications. What distinguishes these approaches is the training data (experimental or simulated) and the use of regressors or classifiers or their combination. A fast prediction of the fluid-dependent compressor performance, as needed in optimization and fluid screening is not available in the literature and is addressed here. Also, the sensitivity of process and fluid parameters on the compressor performance it is not clear from the literature and is analysed here. In this work, we aim to analyse the suitability and limitations of using multilayer perceptrons (MLP) for a specific use case. A semi-empirical compressor model from literature is approximated by a surrogate model using an MLP regressor for the prediction of the isentropic efficiency η_{is}, volumetric efficiency η_{vol}, and the outlet temperature T_o. Inlet parameters are selected appropriately for a heat pump process, which is also the charging process in Carnot Batteries, in order to limit the total optimisation effort to plausible parameter ranges. The resulting input parameter range is analysed with regard to the efficiency (isentropic and volumetric) and outlet temperature ranges that can actually be achieved. A permutation feature importance analysis is carried out to assess the relevance of different parameters and adapt the MLP. The predictions of the regressor are discussed and analysed regarding consistency and computing time savings.

2   STRATEGY AND MODELS

The general modelling sequence is shown in Figure 1. The procedures and models will be explained in this order. The strategy here is to screen fluid mixture compositions within the interesting and accessible temperature and pressure bounds, evaluate the predictions of the physical model for exemplary points within the feasible range, and to use these points to construct surrogate models from ML. The models were implemented as Python scripts.

The investigated parameter range is shown in Table 1. The superheating temperature difference ΔT is chosen to ensure a save compressor operation without condensation. The pressure ratio Π is limited to 8 to ensure high volumetric efficiencies. The inlet temperature T_{in} was determined regarding our use case that is explained in the next section. We investigate a ternary mixture of pentane (subscript C5), isobutane (C4) and propane. The mole fractions x_{C5} and x_{C4} are only limited between 0 and 1 prior to the analysis, together with the condition that all mole fractions are positive and sum up to 1. The data points are uniformly distributed in the parameter range using Saltelli's extension of the Sobol' sequence (Herman and Usher, 2017; Iwanaga *et al.*, 2022).
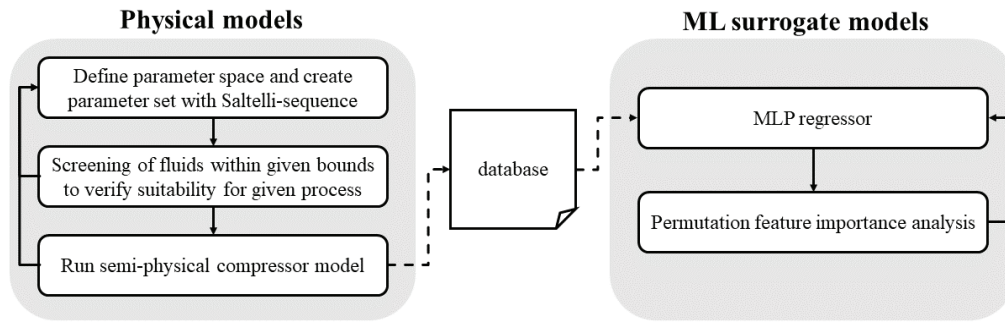
37th INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

https://doi.org/10.52202/077185-0029   336

Figure 1: Overview of the procedure and interaction of the individual modelling steps.

Table 1: Chosen parameter range (x_{C5} limited after fluid screening).

ΔT	Π	T_{in}	x_{C5}	x_{C4}
2-15 K	2-8	273-308 K	0-1 (0-0.65)	0-1

2.1 Fluid screening

The analysed piston compressor is planned to be used in the charging heat pump cycle of a Carnot Battery, which is operated with pairs of high-temperature (s2, s3) and low-temperature (s1, s4) storage tanks and charging/discharging can therefore be regarded as quasi steady-state. It is operated with zeotropic fluids with a temperature glide, to reduce the temperature difference between the working fluid and the sensible storage fluid. The process is shown in the T-Ḣ diagram in Figure 2. The state-change of the secondary fluids is also shown in the diagram. Four isobars for the pressures p_{max}, p_l, p_h and p_{min} are shown as dashed lines. The state points of the heat pump process are numbered from 1 to 4 and will be used in the further explanation. State points 3 and 4 are further subdivided into 3a and 3b, 4a and 4b respectively, depending on whether they are cooled down to the saturated vapor state with quality $x = 0$ or to environmental temperature T_{env}. This results in three requirements for the fluid to
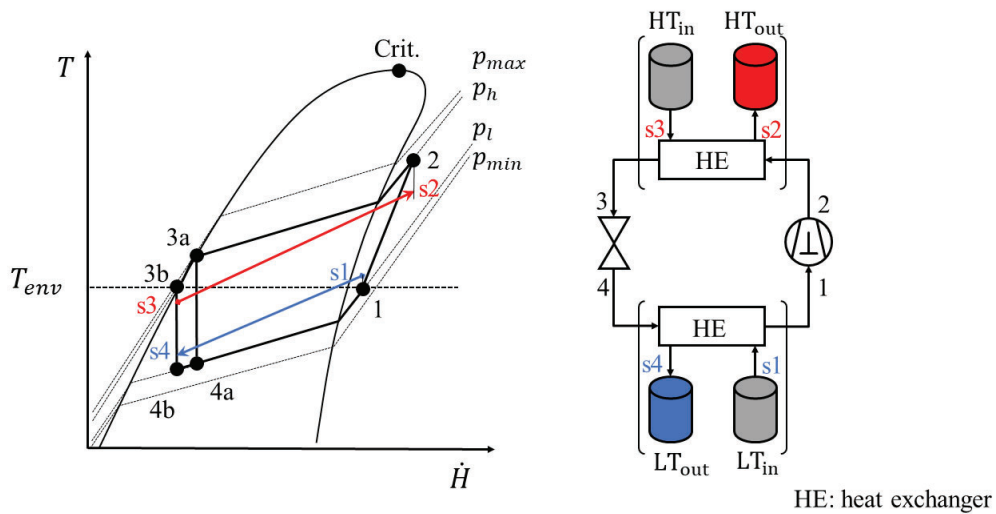

Figure 2: Qualitative representation of a heat pump cycle with zeotropic mixture and sensible storage in quasi-stationary operation in T-Ḣ-diagram and flow chart.

37ᵗʰ INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

337

https://doi.org/10.52202/077185-0029

ensure a safe and stable operation of the cycle. Firstly, the cycle is designed to be subcritical, so the higher pressure p_h must not exceed the critical pressure p_{crit}.

$$p_h < p_{crit} \tag{1}$$

Secondly, the minimum allowable pressure is set to $p_{min} = 1$ bar to prevent the entrance of air in case of a leakage. The maximum pressure depends on the chosen components and is $p_{max} = 22$ bar for the given set-up. Equation (2) summarizes the requirement.

$$p_{min} < p_{l,h} < p_{max} \tag{2}$$

Thirdly, the lower storage temperature of the high-temperature storage T_{s3} and the higher storage temperature of the low temperature storage T_{s1} should be close to environmental temperature $T_{env} = 298$ K to allow a simple start-up of the system and avoid unnecessary losses during standstill. As both are directly related to the corresponding temperatures of the working fluid at saturation state, the latter is used to express the third requirement in equation (3).

$$T_1 \leq T_{env} \leq T_3 \ \rightarrow T(p_l, x = 1) \leq T_{env} \leq T(p_h, x = 0) \tag{3}$$

After the fluid screening, operation points are labelled as either *suitable*, or according to the reason of the infeasibility, by *fail: p_l*, *fail: p_h*, *fail: T_{env}*, and *fail: other* (includes p_{crit}).

2.2 Semi-physical compressor model

The semi-physical piston-compressor model of Roskosch *et al.* (2017) is used with the (constant) parameters as given in the reference to investigate the influence of different fluids and operation parameters and to generate a database. The cycle-resolved model for reciprocating compressors uses fluid properties from REFPROP (Huber *et al.*, 2018) and semi-physical correlations to calculate valve losses. One value of an experimental operating point is needed to fit two parameters. The time dependent energy and mass balances are solved. The heat transfer to and from the walls is evaluated, and finally the output state, the integral work, and the mass flow rate are calculated. The model was compared to 63 measured points of different compressors and fluids, including isobutane and propane, and showed mean prediction errors of 3.0 % for the isentropic efficiency and 2.3 % for the volumetric efficiency, as described in the reference.

The input state and the pressure ratio Π are input variables, together with the fluid composition. Here, a compression expansion cycle is resolved in 3600 steps. The calculation is repeated for several cycles, until the difference of p, T and the temperature of the thermal mass of the compressor T_{th} between the first and last cycle step is below a prescribed deviation (see equation (4)). In some cases, the model does not converge due to oscillations of the system pressure around the discharge pressure. 4 % of the operating points did not converge after 100 iterations; these values were not considered for further evaluations. Typical calculation times per condition are in the range of 7 to 8 minutes.

$$\Delta T_{0,3600} + \Delta p_{0,3600} + \Delta T_{th_{0,3600}} < 0.01 \tag{4}$$

2.3 Surrogate Models from machine learning models

Because the computation times for the physical model are considerable, surrogate models, also called meta-models, were developed based on the physical model. The ML models implemented in the open source library *scikit-learn* (Pedregosa *et al.*, 2011) for Python are used here.

An MLP is a fully interconnected neural network. For each neuron, the inputs i_i are multiplied by the weights w_i and added to the bias vector b. The result is then processed in an activation function like the rectified linear unit function (ReLU), the logistic sigmoid function, or the hyperbolic tangent function. (Geron, 2022)

For each of the three target variables a separate MLP was trained. The features were scaled to a feature range between 0 and 1 before training and transformed back for evaluation. Hyperparameters were optimized using a randomized search, sampling 100 combinations each and using 3-fold cross-

37th INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

https://doi.org/10.52202/077185-0029   338

validation. The maximum number of iterations is limited to 2000. 20 % of the data is withheld for testing. The coefficient of determination R^2 and the root-mean-square error $RMSE$ are used for scoring, whereby only the latter is used to refit the model. The hyperparameter range, with the best values highlighted, is shown in Table 2. Along this work different input parameters were used, as indicated in the top row, leading to the models MLP 1-3.

Table 2: Parameter selection and results of MLP hyperparameter optimization.
η_{is} in bold, η_{vol} in italics, T_{out} underlined.

	MLP 1 $\Delta T, \Pi, T_{in}, x_{C5}, x_{C4}$	MLP 2 $\Delta T, \Pi, T_{in}, \rho_{in}$	MLP 3 Π, T_{in}, ρ_{in}
Hidden layers	**2**, 3, *4*	2, **3**, *4*	2, 3, *4*
Neurons	75, *100*	75, **100**	75,100
Activation function	**ReLU**, Tanh, logistic	**ReLU**, *Tanh*, logistic	**ReLU**, Tanh, logistic
Alpha	**0.001**, 0.01	0.001, **0.01**	**0.001**, 0.01
Solver	*Adam*, **lbfgs**, sgd	*Adam*, **lbfgs**, sgd	*Adam*, **lbfgs**, sgd
Batch size	20, **50**, 100, 200	20, **50**, 100, 200	20, *50*, 100, **200**
Learning rate	**Adaptive**, *constant*	**Adaptive**, constant	**Adaptive**, constant
Initial learning rate	*0.001*, **0.01**	*0.001*, **0.01**	**0.001**, 0.01

2.4 Feature importance as measure of sensitivity

The feature importance of complex computational models can be used to calculate sensitive measures as an alternative to a global sensitivity analysis (like Saltelli *et al.* (2008)). It can be assessed with permutation based variable importance indices, as presented for RF by Antoniadis *et al.* (2021).

The importance i for feature j is calculated $k = 1 \dots K$ times, whereby the columns of the data set get randomly shuffled, and is then compared to the reference value s of the model. In this way the contribution of each feature to a model's performance is assessed. The permutation feature importance can be determined for untrained data and surpass impurity-based importance for strongly biased and high cardinality features. (Pedregosa *et al.*, 2011)

$$i_j = s - \frac{1}{K} \cdot \sum_{k=1}^{K} s_{k,j} \tag{5}$$

Here, the MLPs described in the previous section are used and the feature importance is assessed for evaluating the $RMSE$. The investigated input parameters $\Delta T, \Pi, T_{in}, x_{C5}, x_{C4}$ were chosen because they are independent from each other, which is not the case for x_{C3} and p_{in}, which are not included. The function *permutation_importance* implemented in *scikit-learn* was used with 100 permutation per feature.

3 RESULTS

The results are presented in the following section, starting with the screening of fluid mixtures with respect to the prescribed pressures and temperatures. For the fluids with the wanted thermal properties, the physical compressor model is run to create the database for ML. This is followed by creating ML models. The regressor is trained as a surrogate for the physical model. In the permutation feature importance analysis, the most important parameters for ML are analysed. The total data set prior to screening with limited x_{C5} consists of 1729 points, among which 556 are classified as suitable and used for the MLP regressor.
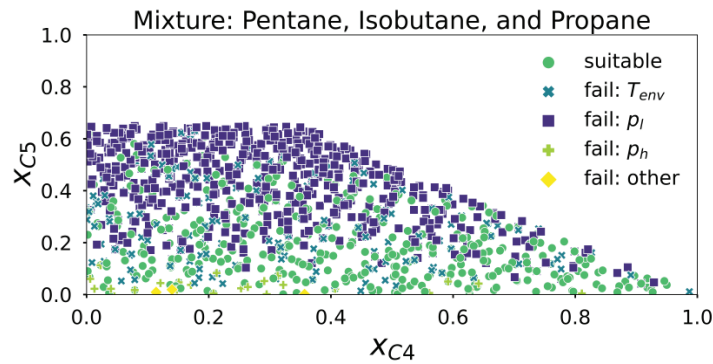
37th INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

339

Figure 3: Results of the fluid screening with suitability of the mixture composition for present problem definition, x_{C5} limited.

3.1 Fluid Screening

The three components pentane, isobutane and propane were selected for the ternary mixture based on their saturation pressures for the desired compressor inlet temperature range. Figure 3 shows the results of the fluid screening for the range of parameters in Table 1. The pentane mole fraction x_{C5} is plotted versus the isobutane mole fraction x_{C4}, while the propane mole fraction can be derived from the others. In the legend, the subdivision is defined according to suitability or non-suitability. It can clearly be seen that mole fractions of pentane above 65 % lead to evaporation pressures below the minimum pressure, which was expected from the saturation vapor pressure curve. The parameter range of pentane is therefore limited to $x_{C5} = 0.65$ in the following tasks, which already significantly reduces the calculation time. Still, 52 % of the points analysed are not suitable due to the lower pressure criterion. Only 32 % of the uniformly distributed parameter combinations are suitable for operation under the given boundary conditions. The other requirements lead to the exclusion of 16 % of the data points, but no clear boundaries can be determined that would allow a further refinement of the parameter range. However, a fluid screening prior to the optimization process can significantly reduce the computational effort by avoiding calculations for inaccessible operating points.

The main motivation for using zeotropic mixtures instead of pure substances is the temperature glide during phase change. Figure 4 visualizes the dependence of the temperature glide at p_h on the pressure
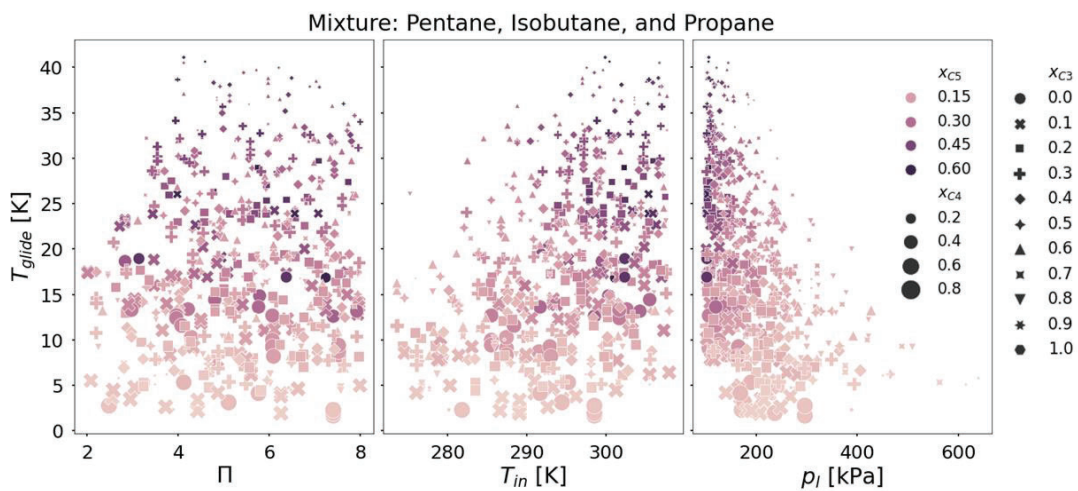

Figure 4: Dependance of temperature glide during phase change at p_h from pressure ratio (left), inlet temperature (centre) and lower pressure (right), only suitable-classified data points shown.

37th INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

https://doi.org/10.52202/077185-0029   340

ratio, the inlet temperature, and the inlet pressure for the suitable mixtures. The mixture composition is indicated by the colour, the size, and the shape of the data points. The temperature glide increases with the pentane mole fraction, while large isobutane mole fractions result in lower temperature glides, with the combination of pentane and propane appearing to be advantageous. The pressure ratio has no strong effect on the temperature glide under the given boundary conditions. However, increasing compressor inlet temperatures tend to lead to higher temperature glides. The inlet pressure has the greatest influence: only at pressures below 300 kPa the temperature glides exceed 30 K. What was not taken into account in the screening was the uniformity of the temperature glide, which is however an important measure to assess the temperature difference between working and secondary fluid in a heat exchanger. But this will not be addressed here, where we focus on the compressor performance.

3.2 Results of the physical model

The different parameter combinations are now used as inputs of the semi-physical compressor model, which ran in parallel on different cores. To analyse the impacts and interdependencies in detail, the entire parameter range whether suitable or not is calculated and shown first. Figure 5 shows the isentropic efficiency η_{is}, the volumetric efficiency η_{vol}, and the outlet temperature T_{out} as a function of the pressure ratio. The color shows the dependence on inlet temperature and density ρ_{in} for the suitable data points. Unsuitable data points are shown in grey and their shape indicates the respective
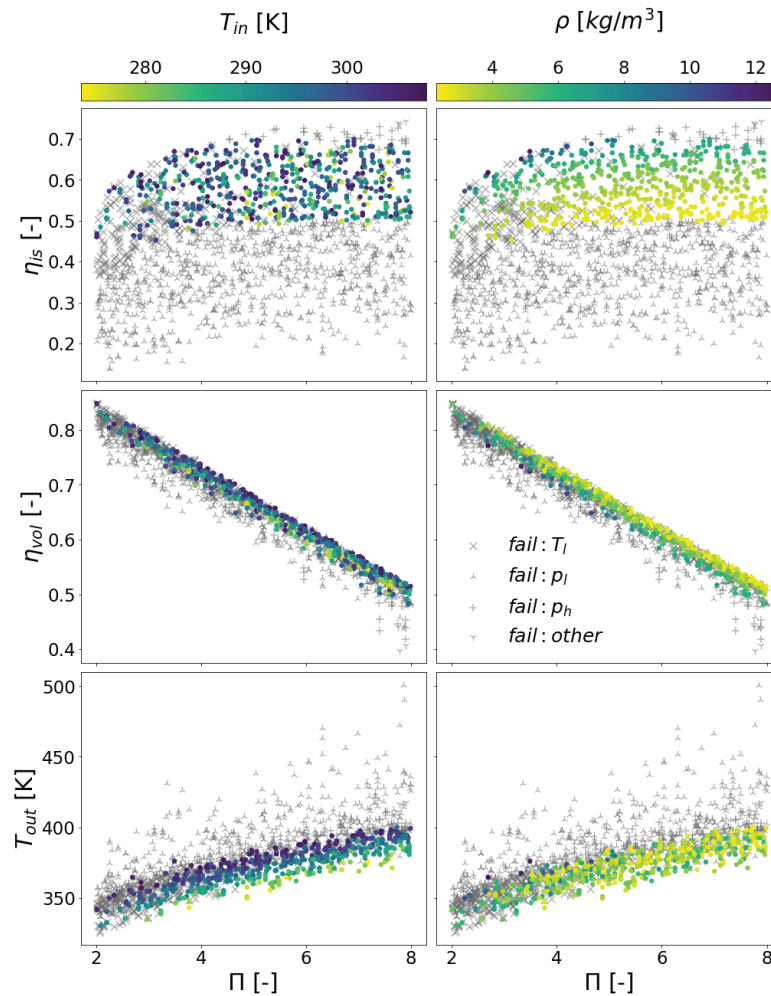

Figure 5: η_{is}, η_{vol} and T_{out} over Π. Colour shows inlet temperature (left) or density (right) for suitable data points. Unsuitable points in grey, the shapes indicate the failing criterion.

37th INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

341

criteria. The isentropic efficiency varies between 13.8 % and 74.3 %, showing a wide variety depending on the pressure ratio, the inlet pressure, and the inlet density. The density dependency has been explained in literature, leading to the conclusion that the use of constant isentropic efficiencies without taking the actual fluid properties into account leads to significant errors (Roskosch *et al.* (2021)). The isentropic efficiency strongly increases with the inlet density. Almost all suitable parameter combinations show an isentropic efficiency above 45.2 %. Conditions leading to lower efficiencies are eliminated mainly because of the minimum pressure criterion. The fluids leading to the highest efficiencies above 69.9 % violate the maximum pressure criterion. Low pressure ratios can lead to too low outlet temperatures of the condenser and are therefore not suitable. It should be noted that these conclusions depend on the used compressor and might be shifted to other levels and fluids, depending on its design points. The centre figures show the volumetric efficiency, which strongly falls with the pressure ratio and varies between 39.7 % and 84.9 %. The inlet temperature and the inlet density only have a minor impact, as they influence the value by about 5 % for a fixed pressure ratio. In contrast to the isentropic efficiency, high densities have a negative effect on the volumetric efficiency. A direct dependency between the suitability and the volumetric efficiencies is not seen, although the unsuitable data points show the lowest volumetric efficiencies. In the plots on the bottom, the outlet temperatures are plotted, which increase with the pressure ratio and depend on the inlet temperature. The outlet temperature varies by about 25 K for a fixed pressure ratio. Data points that violate the minimum pressure criterion exit at the highest outlet temperatures. No dependency on the density is found.

3.3 Accuracy and feature importance of MLP

The first surrogate model (MLP 1) is trained with the features from Table 1. The targets were the three from the previous section for all three surrogate models discussed. First, the accuracy of the prediction is analysed. Then, the influence of each parameter is identified using the permutation feature analysis and the MLP is adapted. Table 3 shows the metrics R^2, the root-mean-square error $RMSE$, and additionally the average relative error for all MLPs analysed during this section. All MLPs were trained by optimizing R^2 and $RMSE$, while only the latter was used for refitting. The first structure with $\Delta T, \Pi, T_{in}, x_{C5}$ and x_{C4} as features shows a very high R^2 between 0.980 and 1.000. The $RMSE$ shows excellent agreement with deviations of 0.8 % for η_{is}, 0.5 % for η_{vol}, and 0.2 K for T_{out}. The average relative error is between 0.03 % and 1.1 %. The hyperparameter optimisation takes 459 s, while the

Table 3: R^2, $RMSE$, and average relative error for MLP with different features.

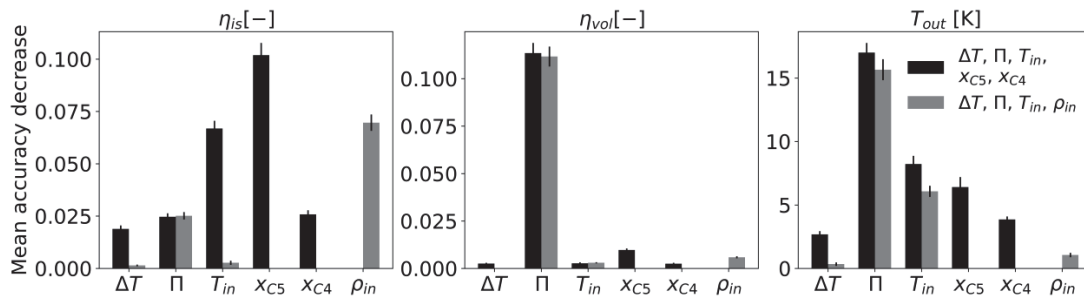|		**R^2**	**$RMSE$**	**Av. rel. error**
η_{is}	MLP 1	0.980	0.8 %	1.1 %
	MLP 2	0.962	1.2 %	1.7 %
	MLP 3	0.972	1.0 %	1.5 %
η_{vol}	MLP 1	0.997	0.5 %	0.6 %
	MLP 2	0.993	0.7 %	0.9 %
	MLP 3	0.999	0.3 %	0.4 %
T_{out}	MLP 1	1.000	0.2 K	0.03 %
	MLP 2	0.965	2.6 K	0.54 %
	MLP 3	0.965	2.5 K	0.56 %

Figure 6: Results of permutation feature importance, accuracy measure negative $RMSE$. Black: with mole fractions, grey: with inlet density.

calculation of one point with the surrogate model is reduced to ca. 9 µs. The results of the permutation feature importance analysis are shown in black in Figure 6 for MLP 1. The analysis was performed on the test set for each of the target parameters, using the negative $RMSE$ as the accuracy metric. The largest increase in $RMSE$ and therefore the largest influence on the isentropic efficiency is the pentane mole fraction with 10.2 % ± 0.6 %, followed by the inlet temperature with 6.7 % ± 0.4 %, and the isobutane mole fraction with 2.6 % ± 0.2 %, and the pressure ratio with 2.5 % ± 0.2 %. The volumetric efficiency mainly depends on the pressure ratio with a mean accuracy decrease of 11.4 % ± 0.5 %, which has already been deducted from Figure 5. The outlet temperature is strongly dependant on the pressure ratio, which increases the $RMSE$ by 17.0 K ± 0.8 K, and to some extend on the inlet temperature with 8.2 K ± 0.7 K and the mole fractions. Surprisingly, the isobutane mole fraction has a minor impact throughout. Therefore, and because of the significant influence of the inlet density reported in the literature, the second analysis was performed with an MLP regressor, where the mole fractions were replaced by the inlet density as input (MLP 2). $RMSE$ is increased to 1.2 % for η_{is}, to 0.7 % for η_{vol}, and to 2.6 K for T_{out}. The average relative error increases slightly by 0.6 % for η_{is}, by 0.3 % for the volumetric efficiency, and by 0.5 % for T_{out}, as expected. The calculation time is reduced to 399 s. Looking at the permutation feature importance of MLP 2 (shown in grey in Figure 6), the significant influence of ρ_{in} on the isentropic efficiency is demonstrated with a mean increase in RMSE of 7.0 % ± 0.4 %, but this is below the values of the mole fractions. Interestingly, the inlet temperature becomes almost insignificant, while the pressure ratio feature importance remains almost constant. For the volumetric efficiency, feature relevance remains unchanged. The influence of the inlet density on the outlet temperature is with 1.1 K ± 0.2 K significantly smaller than the influence of the mole fractions in MLP 1. The smaller decrease in prediction accuracy when using ρ_{in} instead of the mole fractions explains why the prediction accuracy of MLP 2 decreases. The permutation feature importance analysis for the features $\Delta T, \Pi, T_{in}, \rho_{in}$ revealed that the superheating temperature difference does not have an impact on the target values once the mole fractions are replaced. A possible explanation might be that the features are not independent from each other. By performing the fluid-screening, the former uniformly distributed parameter space is restricted (see also Figure 4). Therefore, a possible correlation
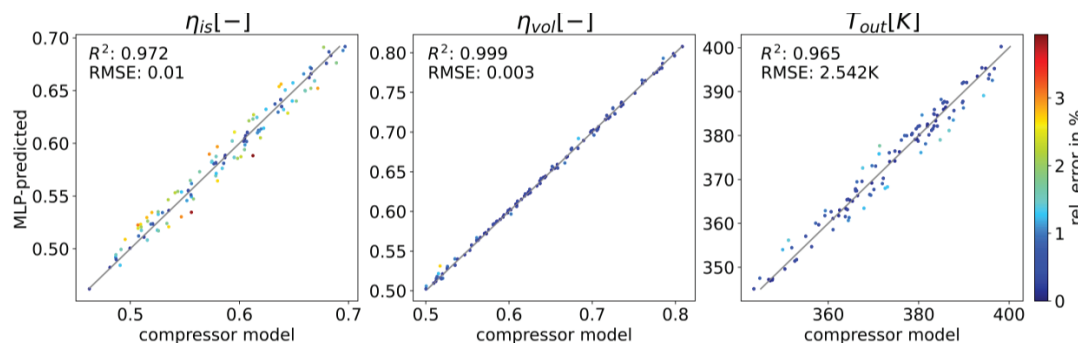

Figure 7: Results of MLP over result from physical model, colour indicates relative error.

cannot be entirely excluded. Thus, we did not expect any loss of accuracy by omitting the feature and trained the final MLP 3 with Π, T_{in}, ρ_{in} as only features. The results confirm this hypothesis and MLP 3 even outdoes the previous MLP 2, proving that the use of irrelevant features can even deteriorate a models prediction accuracy.

The accuracy of MLP 3 with Π, T_{in}, ρ_{in} as features is visualised in Figure 7. The predicted result of MLP 3 on the test set is plotted versus the results from the semi-physical model. The good agreement between the ML model and the physical model can also be seen visually. For the isentropic efficiency, 95 % of the data points show a relative deviation below 2.9 %; for the volumetric efficiency below 0.98 % and for the outlet temperature even below 1.3 %. The agreement justifies the use of an MLP with reduced features to predict the results of the semi-physical compressor model in principle. Limitations and transferability are analysed and explained in more detail in the following section. Overall, all models show excellent forecasting accuracy and thus meet our basic requirement for a surrogate model. The average relative error of the physical model is given as 3.0 % for the isentropic efficiency and 2.3 % for the volumetric efficiency. By using the MLP, the uncertainties are increased by 1.5 % for η_{is} and by 0.4 % for η_{vol}, which seems well acceptable.

4   DISCUSSION

The surrogate model is capable to predict the target values with good accuracy. It is also of interest whether the three target values are consistent with each other and can be used to calculate other values for which the MLP 3 has not been trained. As an example, the consistency of the outlet entropy and enthalpy is checked. They are calculated using REFPROP from T_{out} and p_h.

$$s_{out}, h_{out} = f(T_{out}, p_h) \tag{6}$$

The results are shown in Figure 8. The derived values are well predicted, although the MLP was not explicitly trained for them. The mean relative errors are with 0.7 % for h_{out} and 0.6 % for s_{out} within the same order of magnitude as for T_{out} with 0.6 %. Thus, the ML results can be used for further evaluations, without adding much uncertainty.
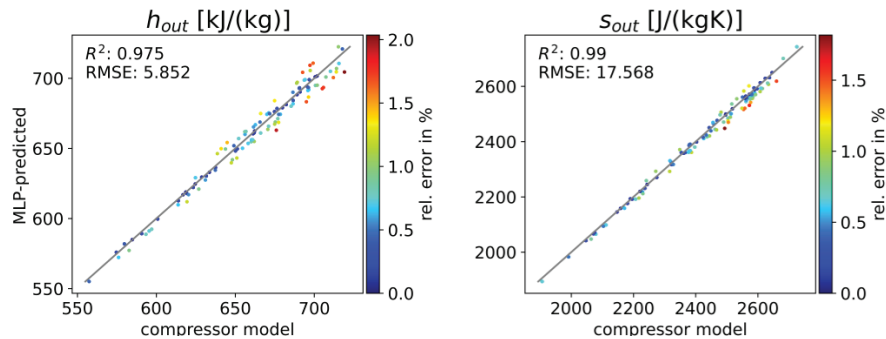

Figure 8: Outlet enthalpy and entropy calculated in postprocessing with the MLP-results.

Reducing computation time was the second objective of this study, alongside accuracy. The computation of the 556 data points with the semi-physical model took 1.53 h. Training of the MLP regressor including hyperparameter optimization took only 383 s and predicting took less than 10.0 ms. The computation time for one data point was thus reduced by a factor of $9.1 \cdot 10^{-6}$. Of course, including the time needed to generate the database, the procedure is only worthwhile if more points are included in the analysis than are needed for the training. For example, when 100 000 data points are to be evaluated along an optimization, the computational time can be reduced by 99.4 %, allowing optimization over a large parameter space. Therefore, the examined procedure promises significant time savings.

5 CONCLUSIONS

The compressor performance is crucial in heat pumps and Carnot Batteries, while detailed compressor models need relatively long computation times. Accurate surrogate models can on the one hand be used to reduce the computation time and on the other hand to determine the key features influencing the performance. This is especially interesting for fluid and fluid mixture selection and ranking. In this work, we investigated the use of ML-based surrogate models for regression and for interpretation of the feature sensitivity for a fixed compressor geometry but varying inlet and outlet conditions for different ternary alkane mixtures. Because several mixtures are not of interest for heat pump or Carnot Battery applications, due too high or too low pressures at some prescribed temperatures, we propose to exclude such mixtures in advance and to train the surrogate model only with appropriate mixtures. This systematic fluid screening prior to surrogate model development can already significantly reduce the computational effort. The permutation feature importance can be evaluated with the MLP for different input variables, showing for the specific case that the mole fraction of the compound with the lowest vapor pressure has the highest sensitivity on the isentropic efficiency, which in turn can be converted to the influence of the fluid density on isentropic efficiency. Thus, physical interpretations are possible in these cases. For higher accuracy, the use of a surrogate model, as applied here using the example of a multilayer perceptron regressor, provides reliable results within the set parameter range with significantly reduced computing time, with a very small reduction of accuracy compared to the physical compressor model. ML models moreover offer a time-saving alternative to other methods of sensitivity analysis and are not subject to any restrictions regarding the exact number of parameter combinations. As is always the case with regression, one should not use the model beyond the trained regime.

DECLARATION OF AI-ASSISTED TECHNOLOGIES

During the preparation of this work the authors used DeepL in order to improve readability and language. The authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

REFERENCES

Antoniadis, A., Lambert-Lacroix, S. and Poggi, J.-M. 2021. Random forests for global sensitivity analysis: A selective review. *Reliability Engineering & System Safety*, vol. 206: p. 107312.

Atakan, B. 2024. Thermodynamic Analysis of Carnot Batteries: Storage Temperatures, Sensitivities, and Optimal Parameters.

Bernehed, A. 2021. ZeoPTES: Zeotropic Pumped Thermal Energy Storage with an Ammonia–Water Mixture as Working Fluid. *Energy Tech*, vol. 9, no. 11.

Geron, A. 2022. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd Edition,* O'Reilly Media.

Guo, H., Gong, M. and Qin, X. 2019. Performance analysis of a modified subcritical zeotropic mixture recuperative high-temperature heat pump. *Applied Energy*, vol. 237: pp. 338–52.

Herman, J. and Usher, W. 2017. SALib: An open-source Python library for Sensitivity Analysis. *JOSS*, vol. 2, no. 9: p. 97.

Hidalgo-Mompeán, F., Gómez Fernández, J. F., Cerruela-García, G. and Crespo Márquez, A. 2021. Dimensionality analysis in machine learning failure detection models. A case study with LNG compressors. *Computers in Industry*, vol. 128: p. 103434.

Huang, Y., Chen, J., Chen, Y., Luo, X., Liang, Y., He, J. and Yang, Z. 2023. Performance explorations of a novel high temperature heat pump with multi-adjusted compositions of zeotropic mixture. *Applied Thermal Engineering*, vol. 235: p. 121409.

Huber, M., Harvey, A., Lemmon, E., Hardin, G., Bell, I. and McLinden, M. 2018. *NIST Reference Fluid Thermodynamic and Transport Properties Database (REFPROP) Version 10 - SRD 23,* National Institute of Standards and Technology. (In en).

Iwanaga, T., Usher, W. and Herman, J. 2022. Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses. *SESMO*, vol. 4: p. 18155.

37th INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

345

Joly, M., Sarkar, S. and Mehta, D. 2019. Machine Learning Enabled Adaptive Optimization of a Transonic Compressor Rotor With Precompression. *Journal of Turbomachinery*, vol. 141, no. 5.

Koen, A., Farres-Antunez, P., Macnaghten, J. and White, A. 2021. A low-temperature glide cycle for pumped thermal energy storage. *Journal of Energy Storage*, vol. 42: p. 103038.

Liu, Z. and Karimi, I. A. 2020. Gas turbine performance prediction via machine learning. *Energy*, vol. 192: p. 116627.

Masood, Z., Khan, S. and Qian, L. 2021. Machine learning-based surrogate model for accelerating simulation-driven optimisation of hydropower Kaplan turbine. *Renewable Energy*, vol. 173: pp. 827–48.

Neumaier, L., Roskosch, D., Schilling, J., Bauer, G., Gross, J. and Bardow, A. 2023. Refrigerant Selection for Heat Pumps: The Compressor Makes the Difference. *Energy Tech*, vol. 11, no. 4.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, vol. 12: pp. 2825–30.

Roskosch, D., Venzik, V. and Atakan, B. 2017. Thermodynamic model for reciprocating compressors with the focus on fluid dependent efficiencies. *International Journal of Refrigeration*, vol. 84: pp. 104–16.

Roskosch, D., Venzik, V., Schilling, J., Bardow, A. and Atakan, B. 2021. Beyond Temperature Glide: The Compressor is Key to Realizing Benefits of Zeotropic Mixtures in Heat Pumps. *Energy Tech*, vol. 9, no. 4.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. 2008. *Global sensitivity analysis. The primer.* Chichester, Wiley. (In en).

Vering, C., Möntenich, J., Rätz, M., Klebig, T., Streblow, R. and Müller, D. 2024. Process model design for positive displacement compressors and their experimental validation: Comparison of Optimal Experimental Design and Machine Learning. *International Journal of Refrigeration*, vol. 157: pp. 1–11.

Wu, X., Liu, B., Ricks, N. and Ghorbaniasl, G. 2020. Surrogate Models for Performance Prediction of Axial Compressors Using through-Flow Approach. *Energies*, vol. 13, no. 1: p. 169.

Zhu, W., Chebeir, J. and Romagnoli, J. A. 2020. Operation optimization of a cryogenic NGL recovery unit using deep learning based surrogate modeling. *Computers & Chemical Engineering*, vol. 137: p. 106815.

Zühlsdorf, B., Jensen, J. K., Cignitti, S., Madsen, C. and Elmegaard, B. 2018. Analysis of temperature glide matching of heat pumps with zeotropic working fluid mixtures for different temperature glides. *Energy*, vol. 153: pp. 650–60.

ACKNOWLEDGEMENT

37th INTERNATIONAL CONFERENCE ON EFFICIENCY, COST, OPTIMIZATION, SIMULATION AND ENVIRONMENTAL IMPACT OF ENERGY SYSTEMS, 30 JUNE - 4 JULY, 2024, RHODES, GREECE

https://doi.org/10.52202/077185-0029   346