

BANDA, T.M., ZĂVOIANU, A.-C., PETROVSKI, A., WÖCKINGER, D. and BRAMERDORFER, G. 2023. Optimising linear regression for modelling the dynamic thermal behaviour of electrical machines using NSGA-II, NSGA-III and MOEA/D. In Stratulat, S., Marin, M., Negru, V. and Zaharie, D. (eds.) *Proceedings of the 25th International symposium on symbolic and numeric algorithms for scientific computing (SYNASC 2023)*, 11-14 September 2023, Nancy, France. Los Alamitos: IEEE Computer Society [online], pages 186-193. Available from: <https://doi.org/10.1109/SYNASC61333.2023.00032>

Optimising linear regression for modelling the dynamic thermal behaviour of electrical machines using NSGA-II, NSGA-III and MOEA/D.

BANDA, T.M., ZĂVOIANU, A.-C., PETROVSKI, A., WÖCKINGER, D. and BRAMERDORFER, G.

2023

This is the accepted manuscript version of the above paper. The version of record is available to purchase from the publisher's website: <https://doi.org/10.1109/SYNASC61333.2023.00032>

Optimising Linear Regression for Modelling the Dynamic Thermal Behaviour of Electrical Machines using NSGA-II, NSGA-III and MOEA/D

Tiwonge Msulira Banda
*National Subsea Centre
School of Computing
Robert Gordon University
Aberdeen, Scotland, UK
t.banda@rgu.ac.uk*

Alexandru-Ciprian Zăvoianu
*National Subsea Centre
School of Computing
Robert Gordon University
Aberdeen, Scotland, UK
c.zavoianu@rgu.ac.uk*

Andrei Petrovski
*National Subsea Centre
School of Computing
Robert Gordon University
Aberdeen, Scotland, UK
a.petrovski@rgu.ac.uk*

Daniel Wöckinger
*Institute for Electrical Drives and Power Electronics
Johannes Kepler University
Linz, Austria
daniel.woeckinger@jku.at*

Gerd Bramerdorfer
*Institute for Electrical Drives and Power Electronics
Johannes Kepler University
Linz, Austria
gerd.bramerdorfer@jku.at*

Abstract—For engineers to create durable and effective electrical assemblies, modelling and controlling heat transfer in rotating electrical machines (such as motors) is crucial. In this paper, we compare the performance of three multi-objective evolutionary algorithms, namely NSGA-II, NSGA-III, and MOEA/D in finding the best trade-offs between data collection costs/effort and expected modelling errors when creating low-complexity Linear Regression (LR) models that can accurately estimate key motor component temperatures under various operational scenarios. The algorithms are integrated into a multi-objective thermal modelling strategy that aims to guide the discovery of models that are suitable for microcontroller deployment. Our findings show that while NSGA-II and NSGA-III yield comparably good optimisation outcomes, with a slight, but statistically significant edge for NSGA-II, the results achieved by MOEA/D for this use case are below par.

Index Terms—data-driven thermal models, electrical machines, linear regression, cost vs accuracy, NSGA-II, NSGA-III, MOEA/D

I. INTRODUCTION AND MOTIVATION

As energy is transformed from electrical to mechanical form during operation, electrical machines (such as motors) generate heat. Engineers who design electrical machines are concerned about this heat because it lowers the machine's performance, shortens its lifespan, and occasionally even causes total machine failure. Therefore, modelling of heat in electrical machines (also known as thermal modelling), is crucial in assisting electrical engineers to better understand how heat affects the performance of the machines, and in designing effective and long-lasting electrical machines [1]. Thermal models can also be used to monitor and regulate operating electrical machines in real time [2]. The models for these applications must be as simple as possible in order to be installed and operated on (inexpensive) microcontrollers.

Our general research efforts concern the construction of low-complexity and explainable data-driven thermal models that can characterise the dynamic thermal behaviour of electrical machines. In a recent publication [3], we proposed a multi-objective optimisation strategy for discovering accurate, low-complexity and explainable linear regression (LR) models that can be used for estimating temperatures of key motor components under various operational scenarios. The aim of this strategy is to provide decision makers with a clear overview of optimal trade-offs between data collection costs, the expected modelling errors and the overall explainability of the generated thermal models. The strategy in [3] builds on the earlier works of Wöckinger et al. [4] and Zăvoianu et al. [5] in modelling the thermal behaviour of electrical machines using linear regression and synthetic features and is governed by a multi-objective evolutionary solver (i.e. NSGA-II).

In this present work, we evaluate the multi-objective optimisation thermal modelling strategy [3] by integrating three different state-of-the-art evolutionary solvers and comparing their performance and optimisation behavior. The numerical experiments indicate that any effective evolutionary solver can be integrated into the strategy, but caution the effectiveness of a purely decomposition-based solver on our type of problem formulation. Furthermore, the run-time variable frequency heatmaps we constructed in an effort to understand differences in solver performance, have also proven helpful in enhancing the interpretability of the overall thermal modelling strategy.

The rest of the paper is structured as follows: Section II describes the targeted modelling scenario and its requirements. In Section III we provide a description of the overall multi-objective thermal modelling strategy. Section IV covers a short background to multi-objective optimisation. In Section

V we describe the problem formulation and experimental setup. Section VI demonstrates the results and provides their interpretation, and finally, Section VII contains conclusions and an outlook on future work.

II. THERMAL MODELLING SCENARIO

Our thermal modelling scenario entails the creation of low-complexity regression models that can be applied to accurately estimate operating temperatures of the key components of a 3-phase brushless outer rotor permanent magnet synchronous motor under various operational scenarios. The motor is used in low-cost fans where higher levels of utilisation are expected. Creating a thermal model for this kind of motor using the traditional lumped parameter thermal network (LPTN) [6] is known to be challenging because the rotating bell leads to a complex flow of heat; and the air gap between the rotor bell and the mounting flange allows an exchange of air and heat, which further complicates the modelling [4]. That being the case, modelling the motor using data driven-techniques is ideal as long as relevant data is available. As with most modelling exercises, accuracy is critically important, but in this particular case, equally important is interpretability, microcontroller compatibility and the minimisation of modelling costs.

The motor has six key components that are of interest when wishing to monitor or manage heat. Two of these components are categorised by domain experts depending on their modelling importance as, high priority, and these are, the winding wire, T_w , and the static ring of the inner ball bearing, T_{bi} . The other components with priorities ranging from medium to low are the mounting flange, T_f ; the rotor, T_r ; the outer ring of the outer ball bearing, T_{bo} ; and the steel stator yoke, T_s . Domain experts have also identified two input variables as being highly relevant for modelling the thermal characteristics of the motor components. The two inputs are rotor speed, v , and electric current, I , and their importance is linked to their low cost and the ease with which they can be measured during operation. Conversely, other relevant inputs such as torque, ambient temperature and electric power input are rarely available or not available at all across most operational scenarios.

Given data availability restrictions within low-cost applications, our aim is to use the two available inputs, rotor speed (v), and electric current (I) to construct simple and accurate thermal models that can be used to estimate (the individual) temperatures of the two high priority motor components: the winding wire, T_w , and the static ring of the inner ball bearing, T_{bi} . Another requirement imposed by electrical engineers is that for the models to be effective, they need to have an average temperature error of less than ± 2 °C when the motor is used within its safe operating area (SOA). Fig. 1 provides a summary of the modelling requirements.

Domain specialists contributed 20 datasets of time series data, each representing a different operational situation of the motor under study, measured with a sample time of 2 seconds, in order to enable the development of dynamic thermal models. The 20 datasets (designated $DS_{01} \dots DS_{20}$)

have a total of 240,200 samples, with individual dataset sample sizes ranging from 571 to 16,201. The data was collected over the course of roughly 133.5 hours using an experimental testbench described in [4] and [7]. An illustration of how the two original modelling inputs (v and I) were varied across all 20 operational scenarios is shown in Fig. 2.

III. MULTI-OBJECTIVE THERMAL MODELLING STRATEGY

Fig. 3 provides a summary of the multi-objective thermal modelling strategy [3], which is aimed at providing decision makers (i.e., electrical engineers) with a complete overview of the optimal trade-offs between modelling costs and accuracy. Given that in our modelling scenario described in Section II there are only two input features to be used for modelling, the first task is to create additional synthetic features. This is done in two stages:

- Using domain knowledge in electrical engineering, it is known that in a typical electrical machine, torque (τ) is directly proportional to current (I) [8] and the total power losses are directly proportional to rotor speed (v) and I [9]. Thus, from a physical point of view, input variables based on several multiplicative combinations of v and I are considered suitable for thermal modelling resulting in the creation of 4 expert-suggested additional features: v^2 , v^3 , I^2 and $v \cdot I$.
- Using the Exponentially Weighted Moving Averages (EMAs) [10], as proposed by Kirchgässner, Wallscheid, and Böcker [11], to all the 6 features (2 original + 4 synthetic) based on v and I in an effort to smooth random fluctuations in the time series data and complement data samples with information regarding trends. EMA features were calculated using the formula in Equation 1:

$$\text{EMA}_{\alpha,t}(r) = \alpha \times r_t + (1 - \alpha) \times \text{EMA}_{\alpha,t-1}(r) \quad (1)$$

where, α is the weight, t is the current period, and r_t is the value of the time series r in the current period. A key aspect of using EMA is deciding how much weight to give to older observations. To capture a wide range of trends in the data, we generated 10 EMA weights using the formula $\alpha_i = 0.001 \cdot 2^i$, $i \in \{0, 1, 2, \dots, 9\}$ and used them to create 60 synthetic EMA features, one for each of the 6 features based on speed (v) and current (I).

The next step is the creation of the actual regression models. To discover the optimal combination of datasets and EMA weights that minimise the size of training samples and maximises model accuracy, we use a multi-objective-evolutionary algorithm (MOEA). NSGA-II [12] was used in [3], but in principle, any effective MOEA can be used. In this paper we will compare the performance of NSGA-II with two other state-of-the-art MOEAs, namely NSGA-III [13], and MOEA/D [14]. Finally, as excessive use of EMA-based synthetic features increases the complexity of LR models, thereby making them less understandable, a step-wise regularisation technique can also be employed as part of the modelling approach in order to remove features with low coefficients. While it was demonstrated in [3] that up to 50%

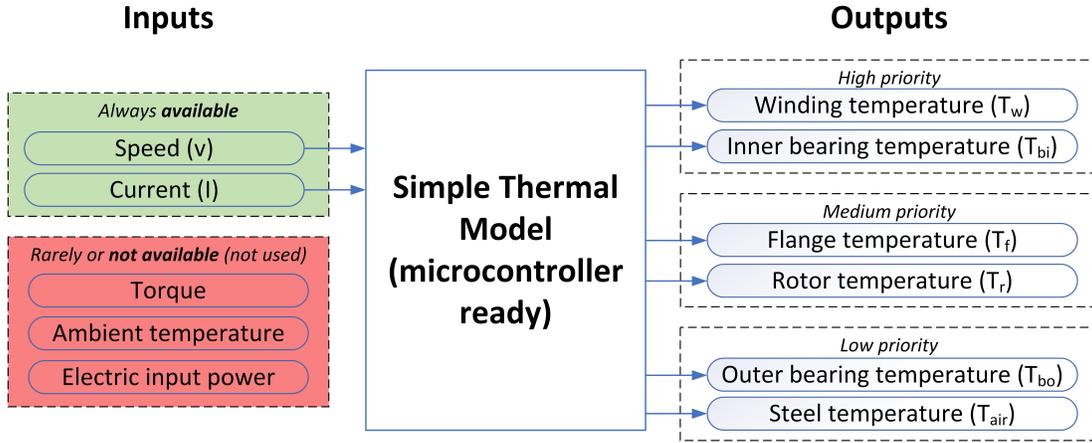


Fig. 1: Summary of modelling requirements with input and output variables. Only the 2 input variables marked as “always available” are used for modelling

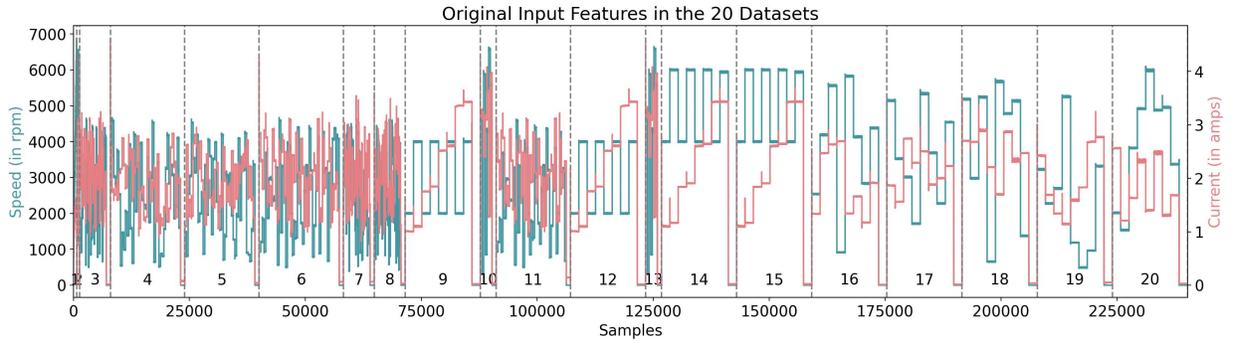


Fig. 2: Variation of the two original modelling inputs across the 20 datasets.

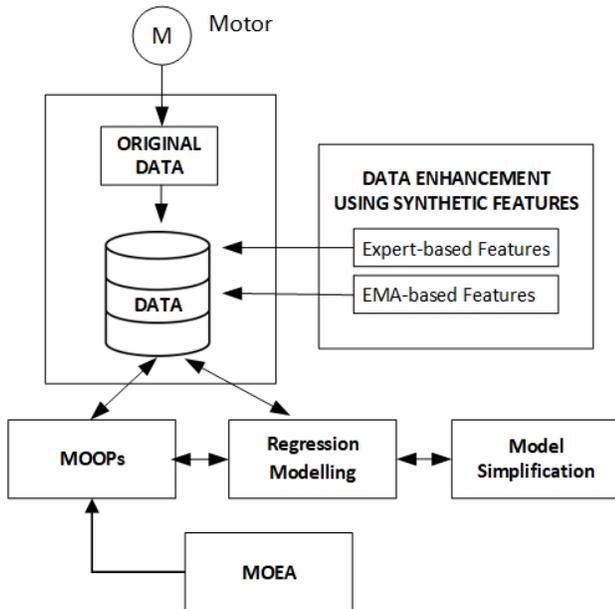


Fig. 3: Schematic of the multi-objective thermal modelling strategy (adapted from [3])

of the original features could be removed without significantly compromising the quality of the LR models, the numerical experiments described in this work do not make use of the model simplification step as our goal is to analyse the intrinsic ability of the tested multi-objective solvers to explore the main data collection effort vs model accuracy trade-off proposed by the considered use case.

IV. MULTI-OBJECTIVE OPTIMISATION

Multi-objective optimisation problems (MOOPs) are a class of optimisation problems where the goal is to maximise or minimise two or three conflicting objectives. If unconstrained, this can be formulated as follows:

$$\text{Minimise } F(x) = (f_1(x), \dots, f_m(x)), \quad m \leq 3 \quad (2)$$

where x is a solution candidate, and f_1 to f_m represent individual objectives.

In our real-life modelling problems, we want to minimise modelling costs while at the same time maximise model accuracy. Unsurprisingly, reducing modelling costs (i.e., the size of available training data) results in the degradation of the accuracy of the elicited dynamic thermal models. Optimal

solution candidates to our problems illustrate the best compromise / trade-off between the two objectives. By definition, the set of all optimal trade-off solutions is called the Pareto optimal set (PS) and its projection in objective space is called the (true) Pareto Front (PF_t). Akin to most real-life problems, the PS of our thermal modelling MOOPs is unknown.

Multi-objective evolutionary algorithms (MOEAs) denote a class of population-based nature-inspired metaheuristics that have over the years proved to be very effective at solving MOOPs. The working principle of MOEAs is to iteratively evolve a population of solution candidates that display an increased fitness by using specialised selection, crossover, mutation and survival operators. One of the main reasons MOEAs have become popular in both industry and research is their general ability to discover, after a single optimisation run, a set of Pareto non-dominated solutions (PN) that provides a good approximation of the PF . This provides decision makers with a set of optimal trade-offs to choose from [15]. Three of the most widely used state-of-the-art MOEAs are: the second version of the Non-dominated Sorting Genetic Algorithm (NSGA-II) [12], NSGA-III [13], and the Multi-objective Evolutionary Algorithm based on Decomposition (MOEA/D [14]). In this research we integrate these three solvers into the multi-objective thermal modelling strategy described in Section III and compare their performance. These three MOEAs have been chosen in light of their common usage and good performance in the optimisation of electrical machine designs. We continue by providing a brief description of the three solvers:

- **NSGA-II** is an improved version of the original NSGA [16], which was published in 1994. It solves MOOPs by initialising a population of randomly selection solutions and after evaluating their fitness, ranks them according to non-dominated fronts, with the best taking the top ranks. Diversity on the fronts is maintained using a method called crowding distance that aims to penalise solutions that are too close in objective space. Offspring solutions are created by applying the simulated binary crossover (SBX) and polynomial mutation (PM) genetic operators on parent solutions selected using tournament selection. NSGA-II is a very robust solver and has been used in numerous industrial applications [17]. An example of the application of NSGA-II in the optimisation of the design of five-phase induction machines is given in Pereira et al. [18].
- **NSGA-III** is a modified version of NSGA-II, and was proposed to address the challenges faced by NSGA-II when solving many-objective optimisation problems – i.e., problem with more than three objectives. The main difference with NSGA-II is that NSGA-III requires a widely distributed set of predefined reference points to guide it. The same reference points are used to achieve diversity and therefore NSGA-III does not use the crowding distance. The rest of the operators are the same as in NSGA-II. Examples of NSGA-III in electric machine

design optimisation are found in Sun, Xu, and Yao [19] and Hernandez et al. [20].

- **MOEA/D** solves an MOOP by decomposing it into multiple single objective sub-problems, each encoding a predefined trade-off. It then seeks to optimise each sub-problem using information from its neighbouring sub-problems. The relationship between neighbouring sub-problems is defined by the distance between them. In each generation, the population of solutions is composed of the best solutions that the algorithm has found for each sub-problem. In the original paper [14], MOEA/D outperformed NSGA-II on several benchmark MOOPs. Some examples of MOEA/D applied in electric machine design are given in Silva et al. [21] and Chen and Severson [22].

V. EXPERIMENTS

A. Problem Formulation

We formulated our MOOP grounded on the efficient usage of the 20 datasets ($DS_{01} \dots DS_{20}$) representing relevant operational scenarios (i.e., containing multiple load points) and EMA weights for generation of synthetic features capturing trends in the data. Given the cost and complexity of collecting data, it would be important to know how different combinations of load points are likely to help characterise the thermal behaviour of a particular motor component and the accuracy trade-offs related to their usage during modelling. We defined our MOOP using Equation 2, setting x as a n -dimensional vector of real-valued variables – i.e., $x_i \in D \subset \mathbb{R}, \forall 1 \leq i \leq n$; and $f_1 \in \mathbb{R}$ and $f_2 \in \mathbb{R}$ represent individual objectives:

- $f_1(x)$ = the total number of data samples in the training set encoded by x that are used for creating the LR model;
- $f_2(x)$ = the MAE or the MSE obtained by the trained LR model on the test set encoded by x .

In order to enable x to easily encode the training-test data split across our 20 datasets and create the necessary synthetic features based on the 10 EMA weights, we formulated the MOOPs as a typical 0,1 Knapsack problem, codified with real values [23] as illustrated in Fig. 4. A candidate solution is a decision vector of 30 real-values between 0 and 1 (i.e., $x \in [0, 1]^{30}$). For the first 20 variables, each variable x_i represents its associated dataset DS_i . If $x_i \geq 0.5$, then DS_i is selected and added to the training set of the modelling experiment. On the other hand, if $x_i < 0.5$, DS_i is added to the test set of the modelling experiment. The next 10 variables of the decision vector represent 10 predefined EMA weights. If a given weight is to be used (i.e., $x_i \geq 0.5, 21 \leq i \leq 30$), all the associated synthetic features (i.e., all 6 EMA features created with α_{i-21}) are used for training and testing the LR model that informs the accuracy of f_2 . In order to evaluate $F(x)$, a counting of the total number of samples in the training set is performed (i.e., f_1) and an LR model is first trained on the training set and then tested on the test set to inform f_2 . If a solution places all or no datasets in the training subset, the solution is heavily penalised across both fitness functions, thus making it

non-viable. It is noteworthy that since we are interested in the independent modelling of the 2 high priority motor component temperatures, we are considering two problems: MOOP- T_w and MOOP- T_{bi} .

B. Experimental Setup

We parameterised the three MOEAs using literature recommended parameters. We used standard genetic operators, i.e., Simulated Binary Crossover (SBX) [24] and polynomial mutation (PM) [25] for both NSGA-II and NSGA-III with crossover probability rate of 0.8, crossover distribution index of 20, mutation probability of $1/n$ and a mutation distribution index of 20. Population sizes for both solvers were set to 200. In NSGA-II we set offspring sizes to 200. NSGA-III does not require defining the offspring size. The number of reference points for NSGA-III, which is a key feature in the algorithm was set to a quarter of the population size.

For MOEA/D, we used the MOEADIEpsilon [26] variant with the Differential Evolution Crossover (with a crossover rate of 0.2, $F=0.5$), and Polynomial Mutation (with probability rate of $1/n$ and distribution index of 20) and the Tschebycheff aggregative function with the dimension set equal to the number of objectives). The neighbour size was set to 20, the neighbourhood selection probability to 0.9, and the maximum number of replaced solutions was set to 2.

We gave each of the three solvers a computational budget of 50,000 fitness evaluations, thereby evolving 250 generations. Given the stochastic nature of MOEAs, we carried 30 independent repeats of each optimisation run in order to enable a statistical comparison of the performance of the solvers. The experiments were conducted using jMetalPy, a Python-based framework for multi-objective optimization with metaheuristics [27], and Scikit-learn, a library for machine learning in Python [28], was used for the actual modelling.

VI. RESULTS AND INTERPRETATION

We first compare the performance of the three evolutionary solvers, NSGA-II, NSGA-III and MOEA/D by looking at the quality of the PN solutions they were able to generate in their 30 respective independent runs. For each solver, we combined all PN solutions generated by the repeated runs and selected the best among them using non-dominated sorting [29]. The resulting overall Pareto fronts (PF) for the three solvers are illustrated in Fig. 5 when modelling the temperature of the winding (a) and the inner bearing (b) using the MSE as the metric for model accuracy.

The results in Fig. 5 show a general increase in test errors when the size of the training set decreases. This is expected as limited training sets are likely to not contain all critical load points. Nevertheless, our approach can identify high-quality LR models that require a limited number of training samples. For instance, in Fig. 6 we present the performance of the LR T_{bi} model trained using only DS_3 and DS_{14} (i.e., 9.5% of all available samples) across all 20 datasets. The importance of also including the EMA weights in the optimisation is also highlighted in Fig. 6 by the observable

difference in prediction accuracy when contrasting with an LR model based on features selected using Principal Component Analysis instead (6 features explain over 92% of the observed T_{bi} variance in $DS_3 + DS_{14}$).

The plots in Fig. 5 indicate that (with a few exceptions) NSGA-II and NSGA-III identified almost the same optimal trade-off solutions. MOEA/D results are generally of lower quality than those obtained by the solvers implementing non-dominated sorting. The PN solutions generated by NSGA-II and NSGA-III are also widely spread across the PF whilst MOEA/D failed to identify solutions in the key high accuracy (i.e., low MSE) part of the Pareto front.

We proceeded to quantitatively measure the quality of the PFs generated by the three solvers using the hypervolume indicator [30]. Besides the hypervolume indicator, other indicators that are used for this task are: the generational distance [31], the inverse generational distance [32], and the epsilon indicator [33]. We opted to use the hypervolume indicator (Hv) as our unary PF quality measure because it is widely accepted in the MOEA community, has a theoretical proof of a monotonic convergence behaviour and can be easily used on problems with an unknown PF_t . This is because, $Hv(PF_c)$ measures the size of the objective space that PF_c dominates when considering an anti-optimal reference point [30]. Based on this, larger Hv values are preferred, but in order to make the numerical values more meaningful, computing the relative hypervolume as $Hr(PF_c) = \frac{Hv(PF_c)}{Hv(PF_t)}$ is advisable. In our case, as PF_t is unknown, we decided to assume it only contains the ideal point (0,0) that would denote an LR model that requires 0 training data and yields 0 errors. Conversely, the anti-optimal reference point was set at (5, 240200), denoting a hypothetical LR model that is trained using 100% of the data but falls well out of acceptable accuracy thresholds.

To obtain a clear overview of the progression of the three evolutionary solvers during the optimisation, we calculated the relative hypervolumes of their PFs at the end each generation and averaged them across the 30 independent optimisation runs. In Fig. 7 we show the comparative results of the three solvers for all the 250 generations when modelling T_w . The results show that NSGA-II and NSGA-III converge relatively faster, achieving average Hr of 84.06% and 83.94% respectively by the 25th generation, whereas MOEA/D only reaches Hr of 81.92% at that time. At the end of the 250 generations, NSGA-II reaches an average Hr of 84.99%, NSGA-III 84.92% and MOEA/D 82.23%.

While the Hr values for for NSGA-II are slightly higher (i.e., better) than those of NSGA-III, given the stochastic nature of the solvers, it is necessary to check if the observed differences in the end-of-the-run PN quality are statistically significant. We performed a one-sided Mann-Whitney U test [34] with a confidence interval of 0.025 that confirmed the superior average performance of NSGA-II (p-value=0.0010). A p-value of 3.0199^{-11} was obtained when comparing NSGA-III and MOEA/D, confirming that the former outperformed the latter.

Lastly, in order to gain insights regarding the behaviour

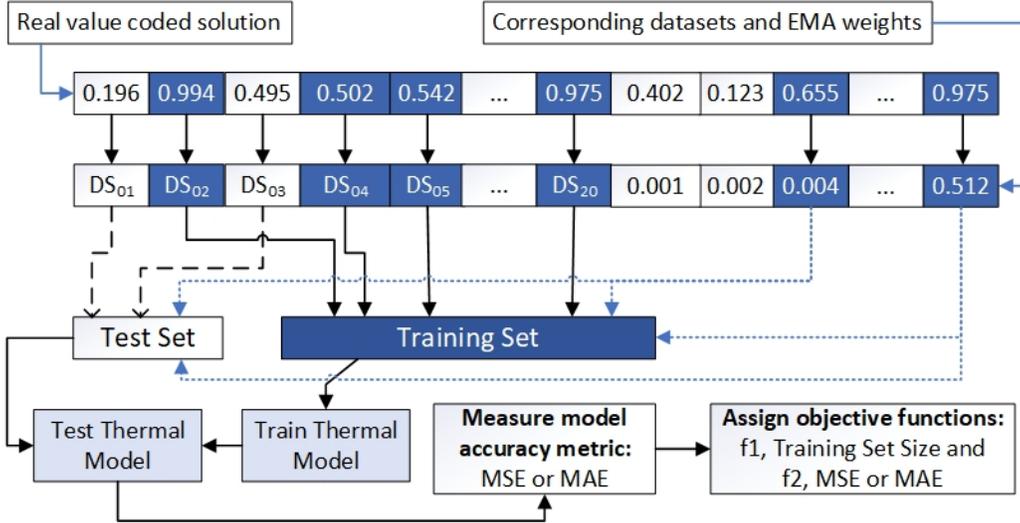


Fig. 4: Schematic of the problem formulation and codification

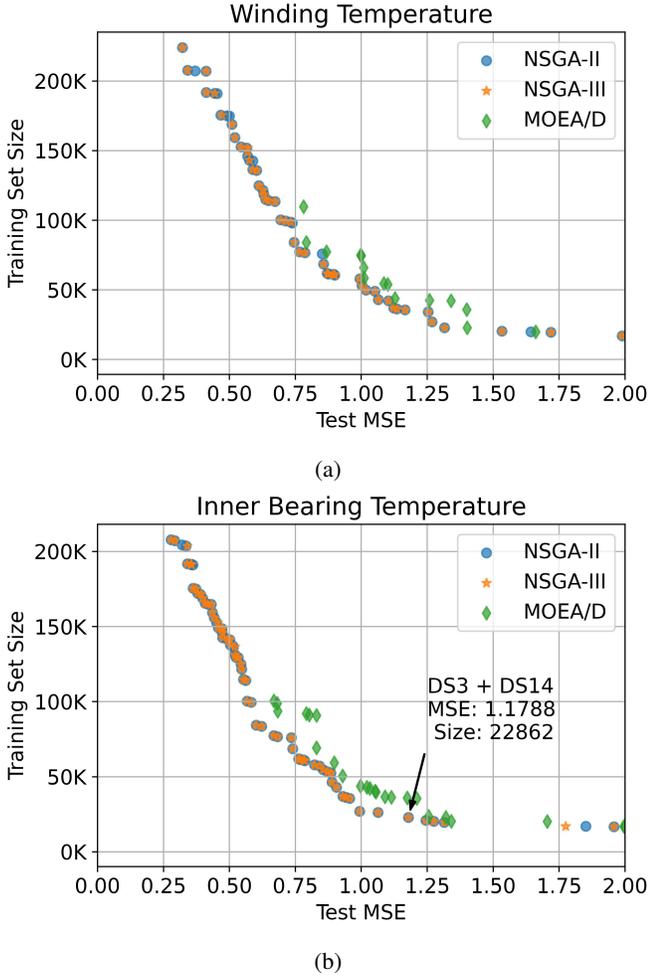


Fig. 5: Pareto-optimal solutions aggregated among all 30 independent runs for each solver for MOOP- T_w and MOOP- T_{bi} .

of the three algorithms, we proceeded to analyse the data generated by the solvers as they progressed from one generation to the other. For this, in each generation, we counted the number of instances a variable in each newly generated decision vector was marked as selected (i.e., when the real value was $x_i \geq 0.5$). The heatmaps in Fig. 8 show averages for the 30 independent runs of each solver. The intensity of the colour for a cell (j, i) represents the probability from low to high that variable x_i is selected among the individuals of generation j . We remind the reader that variables 1 to 20 represent datasets, whereas 21-30 represent EMA weights. As shown by the heatmaps, the three solvers start with a relatively equal distribution of selected variables in the first few generations but begin to lose some genetic diversity as the evolutionary process progresses. On the one hand, NSGA-II explores the decision space more widely as evidenced by the variety of variables being explored during the run. NSGA-III displays a largely similar behaviour. On the other hand, MOEA/D, does not seem to explore the solutions space much widely and only retains a few variables to the end. This loss of genetic diversity could explain why MOEA/D under-performs when compared with the non-domination-based solvers.

The variable frequency heatmaps from Fig. 8 can also provide very useful insights regarding the tackled thermal modelling use case. For example, the usage of EMA features that can capture the longest-term trend (i.e., x_{21}), medium-term trends (i.e., x_{26} and x_{27}) and short-term trends (i.e., x_{28} and x_{29}) is generally very important to LR thermal models discovered by all three solvers. It is also important to note that MOEA/D aims to compensate a limited choice of datasets with an aggressive selection of multiple EMA weights.

VII. CONCLUSIONS AND FUTURE WORK

This research has demonstrated that a novel multi-objective thermal modelling strategy recently proposed in Banda et al. [3] can employ different MOEAs to guide the discovery of

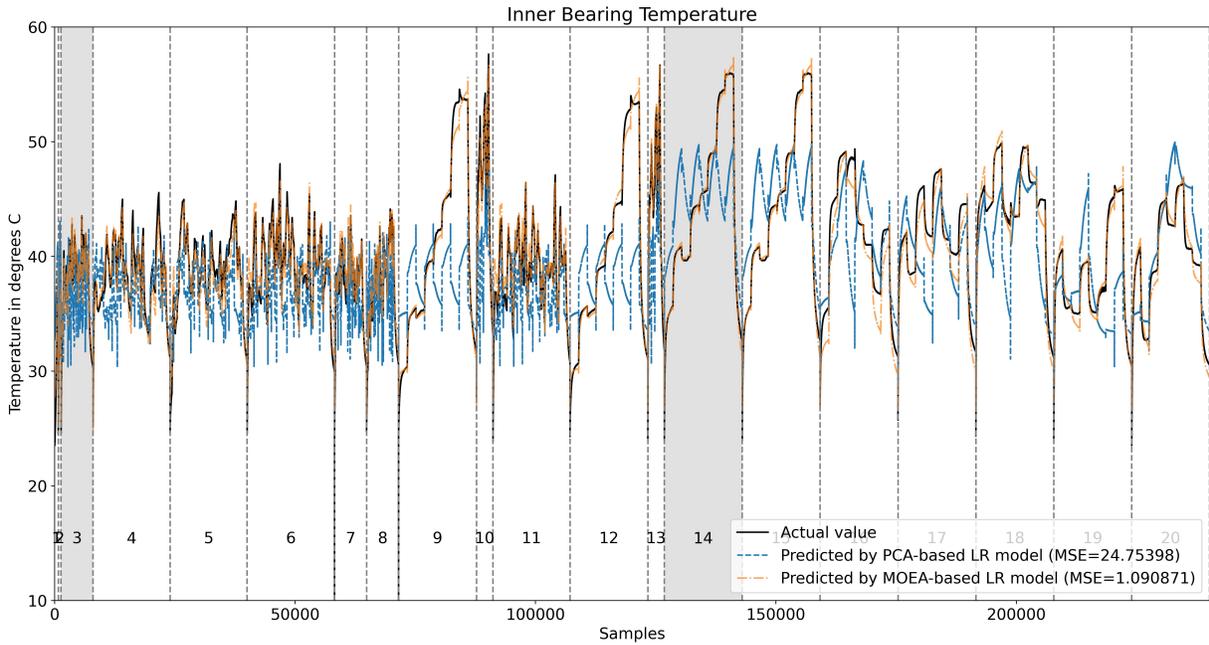


Fig. 6: Comparative predictive performance of MOEA-based and PCA-based LR models for T_{bi} across the 20 datasets. The two datasets used for training are shaded.

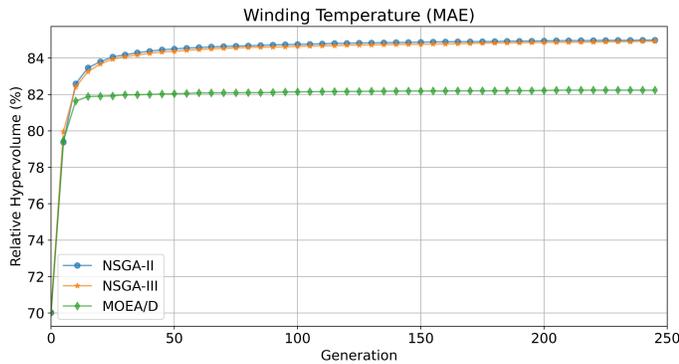


Fig. 7: Comparison of Pareto non-dominated solution sets (PNs) identified by each MOEA

high-quality Linear Regression thermal models. These models are able to estimate the temperatures of key components of an electrical machine when considering various operational scenarios. In particular our experiments have shown that non-dominance-based evolutionary strategies like NSGA-II and NSGA-III are generally able to discover the optimal data collection vs model accuracy trade-offs proposed by our industrial modelling use case. Surprisingly, MOEA/D appears unable to fully explore the decision space when using a standard parameterisation and future work will explore if insights from variable frequency heatmaps can be used to better tune the solver.

ACKNOWLEDGMENTS

This work has been supported by the COMET-K2 “Center for Symbiotic Mechatronics” of the Linz Center of Mechatronics (LCM) funded by the Austrian federal government and the federal state of Upper Austria.

REFERENCES

- [1] A. Boglietti, M. Cossale, M. Popescu, and D. A. Staton, “Electrical machines thermal model: Advanced calibration techniques,” *IEEE Transactions on Industry Applications*, vol. 55, no. 3, pp. 2620–2628, 2019.
- [2] H. Barksdale, Q. Smith, and M. Khan, “Condition monitoring of electrical machines with internet of things,” in *SoutheastCon 2018*. IEEE, apr 2018.
- [3] T. M. Banda, A.-C. Zăvoianu, A. Petrovski, D. Wöckinger, and G. Bramerdorfer, “A multi-objective evolutionary approach to discover explainability trade-offs when using linear regression to effectively model the dynamic thermal behaviour of electrical machines,” *ACM Transactions on Evolutionary Learning and Optimization*, May 2023.
- [4] D. Wöckinger, G. Bramerdorfer, S. Drexler, S. Vaschetto, A. Cavagnino, A. Tenconi, W. Amrhein, and F. Jeske, “Measurement-based optimization of thermal networks for temperature monitoring of outer rotor pm machines,” in *2020 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, Oct 2020, pp. 4261–4268.
- [5] A.-C. Zăvoianu, M. Kitzberger, G. Bramerdorfer, and S. Saminger-Platz, “On modeling the dynamic thermal behavior of electrical machines using genetic programming and artificial neural networks,” in *Computer Aided Systems Theory – EUROCAST 2019*, ser. Lecture Notes in Computer Science, R. Moreno-Díaz, F. Pichler, and A. Quesada-Arencibia, Eds., vol. 12013. Cham: Springer International Publishing, 2020, pp. 319–326.
- [6] A. Boglietti, A. Cavagnino, D. Staton, M. Shanel, M. Mueller, and C. Mejuto, “Evolution and modern approaches for thermal analysis of electrical machines,” *IEEE Transactions on Industrial Electronics*, vol. 56, pp. 871–882, 2009.
- [7] D. Wöckinger, G. Bramerdorfer, S. Vaschetto, A. Cavagnino, A. Tenconi, W. Amrhein, and F. Jeskey, “Approaches for improving lumped parameter thermal networks for outer rotor spm machines,” in *2021 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, Oct 2021, pp. 3821–3828.

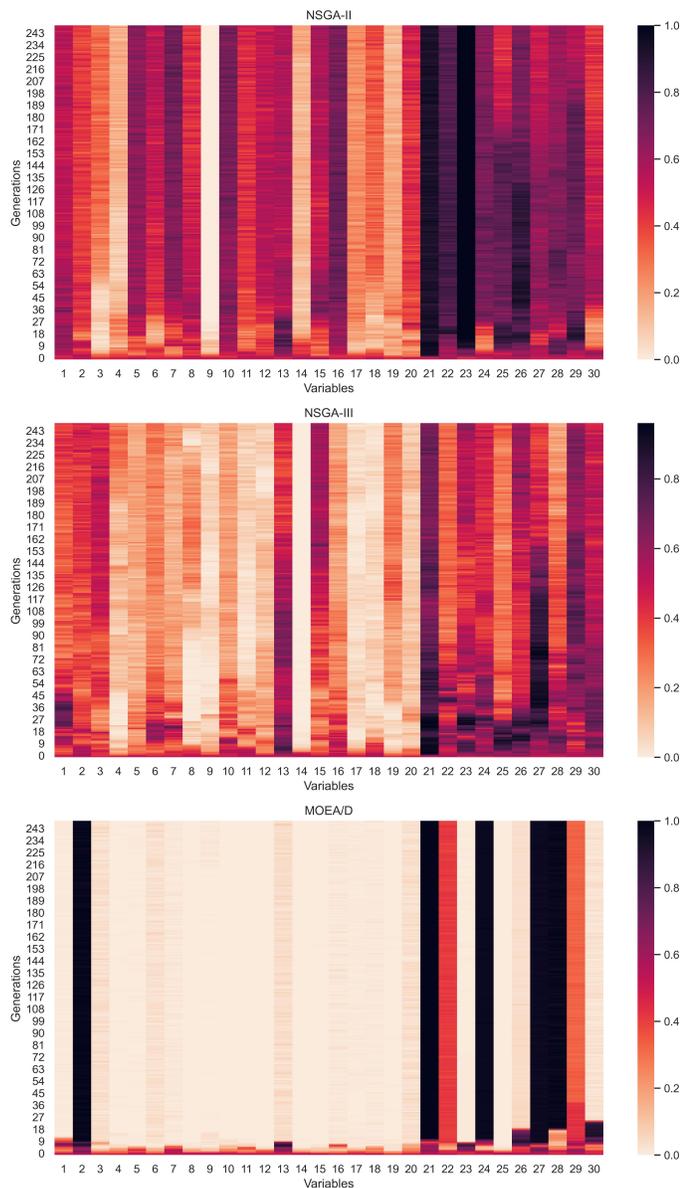


Fig. 8: Tracing the frequency of variables selected by the MOEAs in each generation.

[8] J. Nash, "Direct torque control, induction motor vector control without an encoder," *IEEE Transactions on Industry Applications*, vol. 33, no. 2, pp. 333–341, March 1997.

[9] B. Chalmers and E. Spooner, "An axial-flux permanent-magnet generator for a gearless wind energy system," *IEEE Transactions on Energy Conversion*, vol. 14, no. 2, pp. 251–257, June 1999.

[10] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *International Journal of Forecasting*, vol. 20, no. 1, pp. 5–10, 2004.

[11] W. Kirchgässner, O. Wallscheid, and J. Böcker, "Empirical evaluation of exponentially weighted moving averages for simple linear thermal modeling of permanent magnet synchronous machines," in *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*. IEEE, Jun 2019.

[12] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 182–197, 2002.

[13] K. Deb and H. Jain, "An evolutionary many-objective optimization

algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with box constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, Aug 2014.

[14] Q. Zhang and H. Li, "Moea/d: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, pp. 712–731, Dec 2007.

[15] W. M. Spears, *Evolutionary Algorithms: the role of mutation and recombination*. Springer Science & Business Media, 2013.

[16] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," *Evolutionary computation*, vol. 2, pp. 221–248, 1994.

[17] S. Verma, M. Pant, and V. Snasel, "A comprehensive review on nsga-ii for multi-objective combinatorial optimization problems," *IEEE Access*, vol. 9, pp. 57 757–57 791, 2021.

[18] L. A. Pereira, S. Haffner, G. Nicol, and T. F. Dias, "Multiobjective optimization of five-phase induction machines based on NSGA-II," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 12, pp. 9844–9853, Dec 2017.

[19] X. Sun, N. Xu, and M. Yao, "Sequential subspace optimization design of a dual three-phase permanent magnet synchronous hub motor based on NSGA III," *IEEE Transactions on Transportation Electrification*, vol. 9, no. 1, pp. 622–630, Mar 2023.

[20] C. Hernandez, J. Lara, M. A. Arjona, F. J. Martinez, J. E. Moron, R. Escarela-Perez, and J. Sykulski, "Electromagnetic optimal design of a PMSG considering three objectives and using NSGA-III," *IEEE Transactions on Magnetics*, vol. 58, no. 9, pp. 1–4, Sep 2022.

[21] R. C. Silva, M. Li, T. Rahman, and D. A. Lowther, "Surrogate-based moea/d for electric motor design with scarce function evaluations," *IEEE Transactions on Magnetics*, vol. 53, 6 2017.

[22] J. Chen and E. L. Severson, "Optimal design of the bearingless induction motor for industrial applications," in *2019 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, Sep 2019.

[23] S. Russell and P. Norvig, *Artificial Intelligence: A modern approach*, 3rd ed. Pearson Education, Inc., 2010.

[24] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," *Complex systems*, vol. 9, no. 2, pp. 115–148, 1995.

[25] K. Deb and M. Goyal, "A combined genetic adaptive search (genes) for engineering design," *Computer Science and informatics*, vol. 26, pp. 30–45, 1996.

[26] Z. Fan, W. Li, X. Cai, H. Huang, Y. Fang, Y. You, J. Mo, C. Wei, and E. Goodman, "An improved epsilon constraint-handling method in MOEA/d for CMOPs with large infeasible regions," *Soft Computing*, vol. 23, no. 23, pp. 12 491–12 510, Feb 2019.

[27] A. Benítez-Hidalgo, A. J. Nebro, J. García-Nieto, I. Oregi, and J. Del Ser, "jmetalpy: A python framework for multi-objective optimization with metaheuristics," *Swarm and Evolutionary Computation*, vol. 51, p. 100598, 2019.

[28] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[29] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. Addison-Wesley Longman Publishing Co., Inc., 1989.

[30] E. Zitzler and L. Thiele, "Multiobjective optimization using evolutionary algorithms — a comparative case study," in *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 1998, pp. 292–301.

[31] D. A. Van Veldhuizen and G. B. Lamont, "Multiobjective evolutionary algorithm research: A history and analysis," Department of Electrical and Computer Engineering, Graduate School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, Tech. Rep. TR-98-03, 1998.

[32] C. A. C. Coello, G. B. Lamont, and D. A. V. Veldhuizen, *Evolutionary algorithms for solving multi-objective problems*, ser. Genetic and Evolutionary Computation Series. Springer, 2007.

[33] E. Zitzler, L. Thiele, M. Laumanns, C. Fonseca, and V. da Fonseca, "Performance assessment of multiobjective optimizers: an analysis and review," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 2, pp. 117–132, Apr 2003.

[34] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, Mar 1947.