Waste Heat Recovery of a Thermal Power Plant

[1]Mangesh Nachankar, [2]Dr. Raghvendra Kumar Khedle
[1]M. Tech. Scholar, [2]Guide and HOD
[12]Department of Mechanical Engineering
[12]Technocrats Institute of Technology, Bhopal, India

***Abstract*—Thermal power plants still provide the bulk of electricity that can be dispatched on the grid in India and the rest of the developing world, but a significant portion of the energy they use is dissipated as low-quality heat into the environment. Much of this rejected energy is lost from the plant with the boiler flue gas at temperatures appropriate to make use of them but not high enough for a conventional steam bottoming cycle. A representative 210 MW subcritical coal fired unit is considered to recover the low-grade heat of the flue gas in a subcritical Organic Rankine Cycle (ORC) along with a machine-learning framework. A first law and second law thermodynamic model was developed in Python, where the flue gas source was initially at a temperature of 160 °C and subsequently cooled to a lower limit set by an acid dew-point. Seven candidate working fluids were considered and R1233zd(E) and R245fa were found to be most promising on both the performance and environmental aspects. A parametric study quantified the effect of evaporator pressure, condensation temperature, superheat, inlet temperature and mass flow of flue-gas on the net power and thermal efficiency. Five regressors have been trained with a dataset of approximately 2050 feasible operating points, with the best accuracy (R^2 = 0.991, RMSE = 0.049 MW), given by a gradient-boosted-tree (XGBoost) surrogate model, showing close agreement with the underlying thermodynamics in terms of the ranking of features. The surrogate successfully identified an optimum value that recovers ~2.60 MW of net power, which is within 1% of the optimum recovered in the full model, a ~1.2 percentage-point improvement in plant efficiency and an estimated annual abatement of ~18,000 t CO_2 at only over two orders of magnitude less computing cost than direct optimisation.**

***Index Terms*—Waste heat recovery, Organic Rankine Cycle, Machine learning, Gradient-boosted trees, Surrogate-based optimisation**

I. Introduction

Electricity is the backbone of today's economic activity and is expected to be used more in the future as industry becomes more industrialized, cities are urbanized, and transport is electrified. Even with the growth in renewable energy, thermal power based on coal, lignite, gas and oil continues to supply bulk dispatchable electricity to the world, and in some of the largest economies, including India, thermal power continues to be the largest contributor of grid energy production. In particular, coal generation is expected to continue to account for a significant proportion of India's mix well into the future, thanks to domestic coal stocks and the requirement for reliable, predictable and controllable capacity to complement the increasing variable wind and solar power.

The thermodynamic fact about steam generation is that a significant portion of the chemical energy released when burning fuel is not used to do useful work. In a subcritical Rankine cycle, heat is rejected in the condenser at temperature just above ambient and additional losses are related to the flue gas, radiation, blowdown and auxiliaries, only about 36-40% fuel energy is usually converted to electric output, the rest being lost as waste heat [3]. In modern supercritical and ultra-supercritical plants although the amount of rejected heat is significantly reduced, the absolute amount of rejected heat is still very large due to installed capacity.

Not all "wasted" energy is the same. Heat removed from the condenser is very close to ambient; its exergy is very low, and usually not worth recovering. However, the flue gas leaving the air preheater is usually sent to the flue gas stack at 120–160 °C, which is often a valuable but largely underutilized source of useful thermal energy [3] [4]. Recovering some of this stream and reusing it for extra electrical generation will provide an easy path to increased efficiency of an existing plant without changing major combustion or steam-generation equipment (Fig. 1).

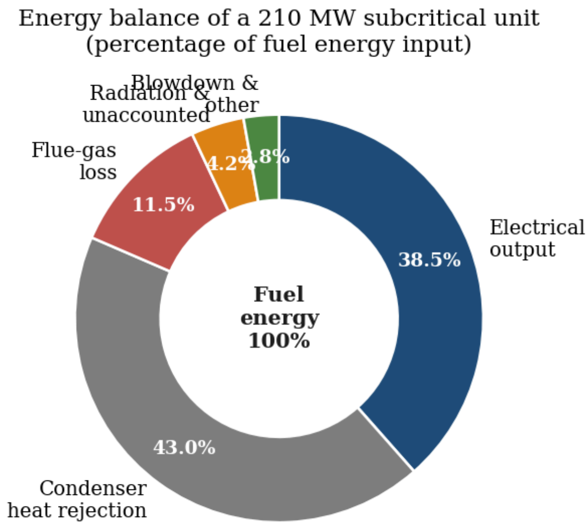

Fig. 1. Representative distribution of fuel energy in a 210 MW subcritical coal-fired unit, highlighting the flue-gas loss targeted for recovery.

The driving force behind this work are three forces on the operators of plants that converge. Coal is an important and fluctuating cost economically and every effort made to reduce specific fuel consumption increases the competitiveness, and waste-heat recovery increases net output with the same fuel input, so it improves the heat rate directly. In terms of the environment, recovery allows a greater amount of electricity to be produced from the same amount of fuel, and thus, decreases emissions per megawatt-hour of electricity produced, thus eliminating the continuous parasitic costs of post combustion capture [4]. Technological development of the Organic Rankine Cycle (ORC) for low-grade heat use, coupled with the rapid development of data-driven modelling, has opened new opportunities to recover value from streams that were previously too cool to consider.

Many interacting variables (working fluid, evaporator pressure, condenser temperature, superheat and pinch-point difference) influence the performance of an ORC bottoming plant, which have non-linear characteristics, with the result that the optimal operating point for maximizing the net power production is not the same as the one for maximizing the thermal efficiency. This design space is too large to explore thoroughly via exhaustive simulation, and too large to explore thoroughly via exhaustive simulation coupled with an optimisation loop. This thesis examines the design and efficient optimisation of an ORC using flue gas, by combining a physics-based thermodynamic model with fast machine-learning surrogates that are able to predict the same performance at a lower cost. The relevant literature was also reviewed (Section 2), the research gap was highlighted (Section 3), the objectives were stated (Section 4), the methodology was explained (Section 5), and the results and discussion were presented (Sections 6–7) with graphical and tabular results and a comparative analysis. The future directions were concluded (Sections 8–9).

II. Literature Review

The study area is divided into three main research lines: waste heat recovery technologies, the thermodynamic core of the ORC (the selection of the working fluid) and the use of machine learning in modelling and optimisation of energy systems. This focuses on work appearing from 2020 to 2025 and includes references to earlier work when this is relevant.

2.1 Waste-heat-recovery technologies

Waste-heat recovery is a group of methods that will depend primarily on the temperature and physical shape of the waste. The medium- and low-temperature flue gas after the air preheater of a coal fired plant is not so suitable for steam recovery whereas the high temperature streams like gas-turbine exhaust are quite suitable to be recovered by steam recovery plants [3]. The objective here is to take advantage of the remaining low-grade stack stream (after in-boiler economisers and air preheaters) that is not recovered. The most popular solutions for converting such residual heat into power are the ORC, the Kalina cycle, supercritical-CO_2 Brayton cycle and thermoelectric conversion. Ammonia–water cycle is more closely aligned to the source profile, and can boost second-law efficiencies [7] but is more complicated and toxic; supercritical-CO_2 cycle is compact and efficient but less competitive at lower flue-gas temperatures [8]; and thermoelectric generators are more stationary, less efficient, and more expensive [9]. In this context, the ORC has proven to be the most developed and widely used for low grade heat, with thousands of megawatts installed since now in geothermal, biomass, solar and industrial field [4, 10].

2.2 The Organic Rankine Cycle and working-fluid selection

The ORC is thermodynamically the same as the steam Rankine cycle, but uses an organic fluid with a lower boiling temperature rather than water to generate useful vapour pressures and expansion ratios at the source temperatures well below those of a steam cycle. The studies conducted in the period 2020-2024 all indicate that a subcritical ORC can retrieve between 10-20% of the available low-grade thermal energy as net power, depending on the source temperature and working fluid and pinch-point constraints [10, 11]. A common feature is the relationship between the cycle and the finite source from which it is drawn; the fact that the flue gas is a sensible-heat stream with its temperature dropping as it rejects sensible heat as the fluid evaporates mainly at constant temperature creates a pinch point that restricts the evaporation pressure. That pressure increase increases the efficiency of the cycle, but decreases the mass flow the source can carry, which is a key feature of finite-source designs [11, 13]..

The efficiency of the cycle, operating pressures, size of turbine and environmental acceptability are all determined by selection of working fluid. In the low-temperature recovery temperature range, expansion to the superheated region without condensing in the turbines is generally preferred, a condition that is preferred for dry and isentropic fluids [4] and [14]. Historically, the reference fluid was R245fa but due to its high GWP, which poses a threat to the environment, R245fa has been the object of a search for an alternative; the global-warming potential (GWP) of R1233zd(E) is very low and its properties are favourable, making it a leading candidate as a replacement for R245fa [15]. The advantage of using hydrocarbons is that these have very low GWP and good performance, but they do have potential issues with flammability, while zeotropic mixtures lower the irreversibility of the heat-transfer, but increase the complexity of the charge-management. A single fluid is not perfect for all measures, which is why a multi-fluid screening is done here.

2.3 Machine learning in energy systems

The use of data-driven techniques with thermal systems has grown significantly, in part due to the availability of well-established open-source libraries and the fact that many thermal systems design problems have non-linear input-output relationships where the evaluation of the costly function is expensive, but where a surrogate model is well suited [5, 6]. The most common method has been the use of artificial neural networks, which are often found to have a coefficient of determination greater than 0.95 after training on validated model or rig data, as reported in [17],[18]. The popularity of tree-based ensembles, like random forests and gradient boosted trees, like XGBoost, stems from their benefits of being fast to train, not prone to overfitting, easily tunable and giving natural feature-importance measures, frequently outperforming neural networks on moderate sized tabular engineering data [19, 20]. A new trend is to combine these surrogates with metaheuristics, which allow trained surrogate to be a component of genetic algorithms or particle-swarm optimisation that would otherwise be impossible to solve, to obtain near-optimal designs for a fraction of the direct optimisation cost [20], [22], [23]. The paradigm this study follows and demonstrates is the use of physics to generate data and machine learning to optimize data, which is now becoming a common paradigm in the reviews [6] and [24]. Table 1 provides a perspective of recent studies that are representative.

Table 1. Comparative summary of representative recent studies (2020–2025) on waste-heat recovery and machine learning in energy systems.

Study	Heat source	Technology	Approach	ML / Opt.	Key finding
Song et al. [11]	Generic low-grade	ORC	Thermodynamic	No	Efficiency–power trade-off for finite sources
Wang et al. [10]	Coal flue gas	ORC	Thermodynamic	No	10–20% of low-grade heat recoverable as power
Zhang et al. [7]	Low-grade waste	Kalina vs ORC	Thermodynamic	No	Kalina gains offset by complexity
Li et al. [15]	Low-temp. waste	ORC	Fluid screening	No	R1233zd(E) a strong low-GWP fluid
Palagi et al. [18]	Generic ORC	ORC	ANN surrogate	ANN	Neural surrogate reproduces ORC maps
Zhao et al. [23]	Waste heat	ORC	Surrogate + GA	GB + GA	Surrogate-assisted multi-objective design
This work	Coal flue gas	ORC	Thermo. + 5 ML + GA	Yes	Verified surrogate-based optimisation

III. Research Gap Identification

These four gaps are identified in the course of the review. First, while many papers use one machine learning method, few provide a systematic, like-for-like comparison of multiple regressors (both neural and tree-based) for the same waste-heat-recovery data. Second, many ORC literature idealize the heat source as a constant temperature reservoir, while the heat source from a power plant flue gas is a finite sensible heat source, with a lower bound temperature of the acid dew point, and the interaction between this bounded heat source and surrogate based optimisation is relatively under-researched. Third, the interpretability of the surrogate models, in terms of the feature importance rankings, is seldom studied, but this is crucial for engineering trust when the model's governing thermodynamics are understood. Fourth, conversion of recovered power to plant level efficiency and estimates of CO_2 abatement is frequently not reported. This study aims to tackle all four gaps simultaneously.

IV. Objectives of the Study

The specific objectives pursued in this study are:

1. Characterize the low-grade waste-heat resource of the flue gas of a representative subcritical coal fired unit, determine its thermodynamic recovery limits and develop and validate a first-law and second-law model of a subcritical ORC bottoming plant for this waste-heat resource.
2. To select the best candidate working fluids and to do a parametric study to see how net power output and cycle efficiency varies with evaporator pressure, condenser temperature, superheat, flue-gas inlet temperature and mass flow.
3. To produce a consistent dataset from the validated model and to build, train, and evaluate five machine-learning surrogate models and to learn about the primary drivers of model performance by studying the importance of the model's features.
4. Place the best surrogate in an optimiser based on a genetic algorithm to maximise net power, check the optimum against the full model and calculate the gain in plant efficiency and CO_2 abatement.

V. Materials and Methods

The methodology combines a physics-based thermodynamic model, a data-driven surrogate and a metaheuristic optimiser. A set of five regressors are trained and compared, with the best surrogate being interpreted then embedded in the genetic algorithm whose optimum is verified against the full model, which is validated in the ORC model. Overall workflow is presented in Fig. 2.

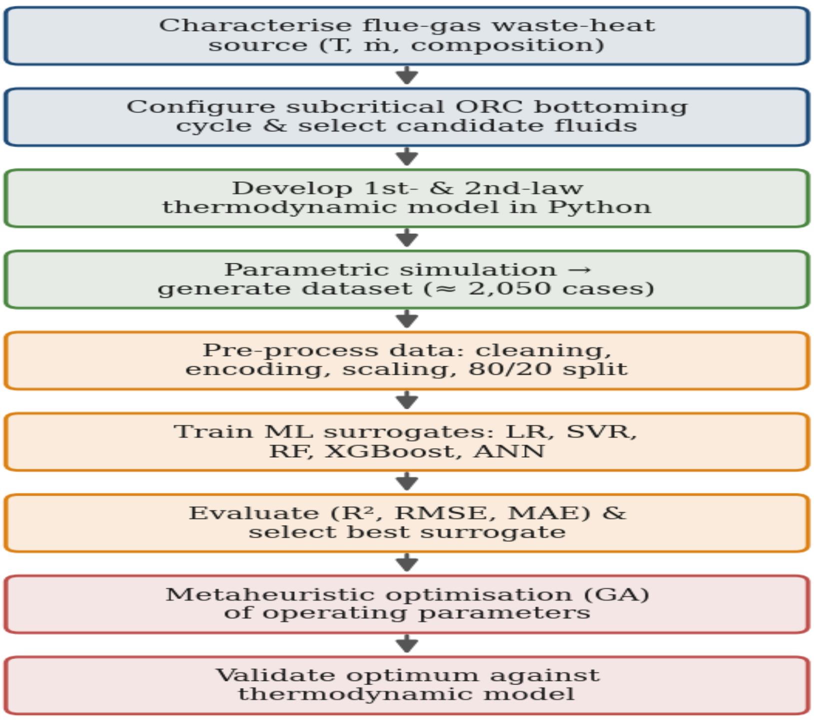

Fig. 2. Overall workflow of the combined thermodynamic and machine-learning methodology.

5.1 Reference plant, heat source and ORC configuration

The reference plant is a typical 210 MW subcritical coal-fired plant. The low-grade flue gas, which is left after the regenerative air preheater, is available before the electrostatic precipitator, induced-draught fan and stack. The base-case source is assumed to be at 160 °C and a mass flow rate of 200 kg/s, and a mean specific heat of 1.08 kJ/kg•K (Table 2). A conservative lower limit of 95 °C is set on the flue-gas outlet temperature to prevent condensation of sulphuric acid and cold-end corrosion, which is marginally higher than the acid dew-point of a typical medium-sulphur Indian coal [25] and is a temperature of approximately 14 MW heat notionally available between these limits.

Table 2. Assumed base-case flue-gas heat-source parameters.

Parameter	Symbol	Value	Unit
Flue-gas inlet temperature	$T_{fg,in}$	160	°C
Flue-gas outlet (dew-point) limit	$T_{fg,out}$	95	°C
Flue-gas mass flow rate	$\dot{m}_{fg}$	200	kg/s
Flue-gas mean specific heat	$c_{p,fg}$	1.08	kJ/kg·K
Approximate recoverable heat duty	Q_{avail}	≈14	MW

The recovery system consists of a subcritical ORC unit operating as a bottoming plant (BT) shown in Fig. 3. The working fluid is pumped from the condenser to the evaporator pressure, preheated, evaporated and optionally slightly superheated by the flue gas; expanded in a turbine-generator and condensed by a cooling-water or air-cooled sink before being returned to the pump. A basic non-recuperated configuration is selected for the purposes of parametric trends and the surrogate methodology to eliminate the additional complications of other components. The numbered state points in the states are pump inlet, pump outlet, turbine inlet and turbine outlet.

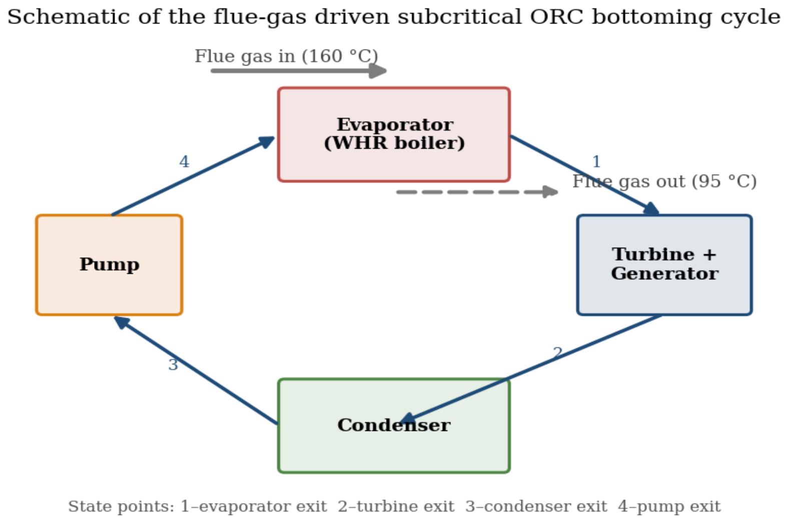

Fig. 3. Schematic of the flue-gas-driven subcritical ORC bottoming cycle showing the pump, evaporator, turbine-generator, condenser and the four numbered state points.

5.2 Thermodynamic model

The model uses the steady flow energy equation and the second law for each component and fluid properties from CoolProp library [26]. The following design-point assumptions are made: steady state, negligible pressure drop and heat losses, constant isentropic efficiency of the pumps and turbines, and negligible change in kinetic and potential energies. The equations for the pump work, evaporator heat addition, flue-gas energy balance, turbine work, net power, condenser duty, first-law efficiency and second-law (exergetic) efficiency are: (1)–(8):

$$W_{pump} = \frac{\left(h^{4} - h^{3}\right)}{\eta\_pump} \tag{1}$$

$$q_{evap} = h^{1} - h^{4} \tag{2}$$

$$\dot{m}_{wf}\left(h^{1} - h^{4}\right) = \dot{m}_{fg} c_{p}, fg \left(T_{fg}, in - T_{fg}, out\right) \tag{3}$$

$$w_{turb} = \left(h^{1} - h^{2}\right) \cdot \eta_{turb} \tag{4}$$

$$W_{net} = \dot{m}_{wf}\left(w_{turb} - W_{pump}\right) \cdot \eta_{gen} \cdot \eta_{mech} \tag{5}$$

$$q_{cond} = h^{2} - h^{3} \tag{6}$$

$$\eta_{th} = \frac{W_{net}}{(\dot{m}\_wf\ q\_evap)} \tag{7}$$

$$\eta_{ex} = \frac{W_{net}}{(E\_fg,in - E\_fg,out)} \tag{8}$$

These relations form a closed model of which, when given the operation parameters and properties of the working fluid, the net power, the thermal efficiency and the exergetic efficiency of the recovery plant can be calculated. The model was validated with representative ORC results for similar source temperature and fluid values, and the results were within the range of the ORC results.

5.3 Candidate working fluids and dataset generation

Seven different fluids were tested (Table 3) including the reference R245fa, R134a, R123 (both older), R1233zd(E) (hydrofluoro-olefin) and 3 hydrocarbons. To train the surrogates, the validated model was exercised across the design space of Table 4, varying six inputs—evaporator pressure, condensation temperature, superheat, flue-gas inlet temperature, flue-gas mass flow and the categorical working fluid—by random sampling. About 2,050 feasible operating points were selected (rejected points are those that violate a physical constraint (negative pinch-point difference, sub-saturation turbine inlet, or condensation above the critical temperature) were rejected.

Table 3. Principal properties of the seven candidate working fluids.

Working fluid	T_crit (°C)	P_crit (bar)	T_boil (°C)	Type	GWP (100 yr)
R134a	101.1	40.6	−26.1	Wet/isentropic	1430
R245fa	154.0	36.5	15.1	Dry	1030
R1233zd(E)	166.5	36.2	18.3	Dry	1
R123	183.7	36.6	27.8	Dry	79
n-Pentane	196.6	33.7	36.1	Dry	≈4
Cyclohexane	280.5	40.8	80.7	Dry	≈3
Toluene	318.6	41.3	110.6	Dry	≈3

Table 4. Ranges of the input variables used to generate the machine-learning dataset.

Input variable	Symbol	Minimum	Maximum	Unit
Evaporator pressure	P_evap	8	26	bar
Condensation temperature	T_cond	30	55	°C
Degree of superheat	ΔT_sh	0	15	°C
Flue-gas inlet temperature	T_fg,in	140	180	°C
Flue-gas mass flow rate	ṁ_fg	160	240	kg/s
Working fluid	—	7 discrete candidates		—

5.4 Machine-learning framework and optimisation

The categorical working fluid variable was one-hot encoded (required by the neural network and support vector regressor, but not the tree models) and the continuous input variables were standardised to zero mean and unit variance (required by the neural network and support vector regressor, but not the tree models). The data was then partitioned into an 80:20 training and held-out test set with five-fold cross validation in the training set used to train the models and select the hyperparameters. We have compared five regressors: multiple linear regression (a simple baseline), support vector regression with an RBF kernel, a random forest, a gradient boosted tree model (XGBoost), and a two-hidden-layer multilayer perceptron, with the help of the Adam optimiser and early stopping (Table 5). The coefficient of determination, the root-mean-square error and the mean absolute error were used to measure accuracy. The most accurate surrogate was then incorporated as the fitness evaluator in a genetic algorithm that evolved a population of OPVs which maximise the net power within the physical constraints; the optimum OPV obtained was then re-evaluated with the full thermodynamic model to verify its validity.

Table 5. Principal hyperparameters of the machine-learning models selected by cross-validation.

Model	Principal hyperparameters
Multiple linear regression	Ordinary least squares; no regularisation
Support-vector regression	RBF kernel; C = 100; ε = 0.01; γ = scale
Random forest	300 trees; max depth 18; min samples leaf 2
XGBoost	500 trees; learning rate 0.05; max depth 6; subsample 0.8
Artificial neural network	2 hidden layers (64, 32); ReLU; Adam; early stopping

VI. Results and Discussion

6.1 Base-case thermodynamic performance

The base-case plant recovers approximately 2.31 MW of net electrical power with an evaporator pressure in the mid-range, the condensation temperature of 40 °C and with reference fluid R245fa, yielding a thermal efficiency of around 14% and an exergetic efficiency of around 45% (Table 6). The low-temperature recovery is typical of this type of recovery as it is more appropriate for a bottoming plant, and the large difference between the source and sink temperatures is reflected in the exergetic efficiency. The flue gas is cooled to temperatures slightly above the acid-dew-point (about 96 °C).

Table 6. Base-case performance of the ORC recovery plant (working fluid R245fa).

Quantity	Symbol	Value	Unit
Net power output	W_net	2.31	MW
Rate of heat recovered	Q_evap	16.4	MW
Thermal (first-law) efficiency	η_th	14.1	%
Exergetic (second-law) efficiency	η_ex	45.2	%
Working-fluid mass flow	ṁ_wf	78.6	kg/s
Flue-gas outlet temperature	T_fg,out	96.3	°C

6.2 Parametric study

As the evaporator pressure is raised gradually, the thermal efficiency will also rise because the heat will be supplied at higher mean temperatures and the quality of the vapour sent to the turbine will get better (see Fig. 4). However the power does not keep increasing—the power peaks near 19 bar and then decreases. This is a direct result of the fact that the pressure increases the specific work, but also the working-fluid mass flow the flue gas can support decreases due to the tighter evaporator pinch point. This efficiency–power trade-off is a hallmark of finite-source recovery and validates the qualitative prediction from the literature.

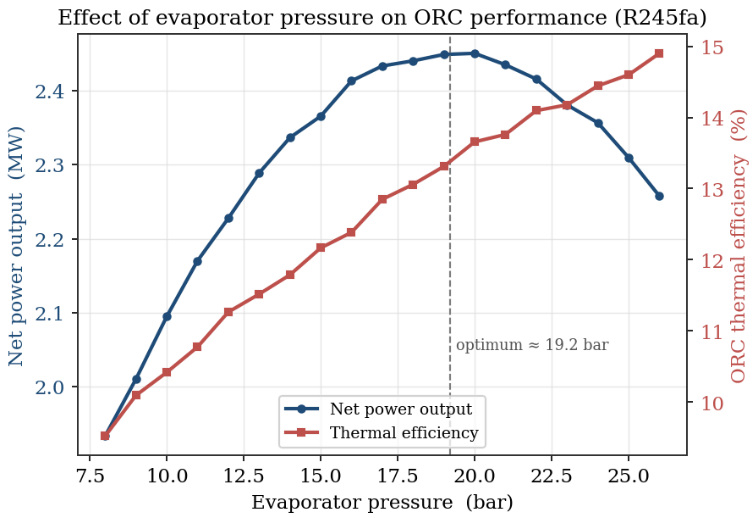

Fig. 4. Effect of evaporator pressure on the net power output and thermal efficiency of the ORC (working fluid R245fa).

The net power and thermal efficiency steadily decrease with increasing condensation temperature (Fig. 5) due to the decrease in the turbine pressure ratio with increasing sink temperature; above 30–55 °C the net power drops around 25%. This highlights the importance of the coldest possible sink – the plant with the cooling water that is not air-cooled condensers or operating with cooler ambient conditions – recovers a significant amount more power. The effect of degree of superheat was found to be minor for the dry fluids used, which was as expected, since much of the work obtained by superheating a dry fluid is at the expense of increased heat supplied from the source.

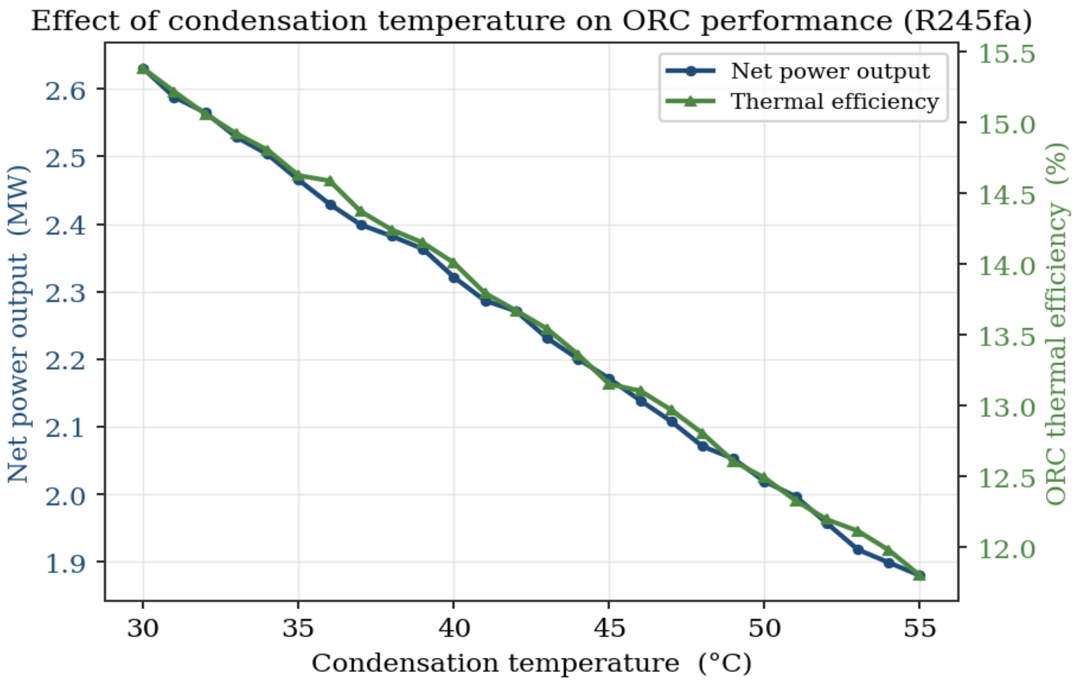

Fig. 5. Effect of condensation temperature on the net power output and thermal efficiency of the ORC (working fluid R245fa).

6.3 Comparison of working fluids

The seven fluids show significant differences in the net power and thermal efficiency (Fig. 6, Table 7) under the same source and sink conditions. R1233zd(E) produces the maximum net power and thermal efficiency, which is slightly more than R245fa, but with a GWP 3 orders of magnitude less than that of R245fa, providing it the most appealing overall fluid and supporting its reputation as a good environmental friendly substitute. R245fa has excellent performance and is a good option when its higher GWP is acceptable. The other hydrocarbons (cyclohexane, n-pentane and toluene) are less flammable but are close to the same thermodynamic performance and have a negligible GWP, while the other two (R134a and R123) have the poorest performance in the group. Based on this, R1233zd(E) and R245fa were chosen as preferred fluids with R1233zd(E) being recommended in applications that demand environmental performance..

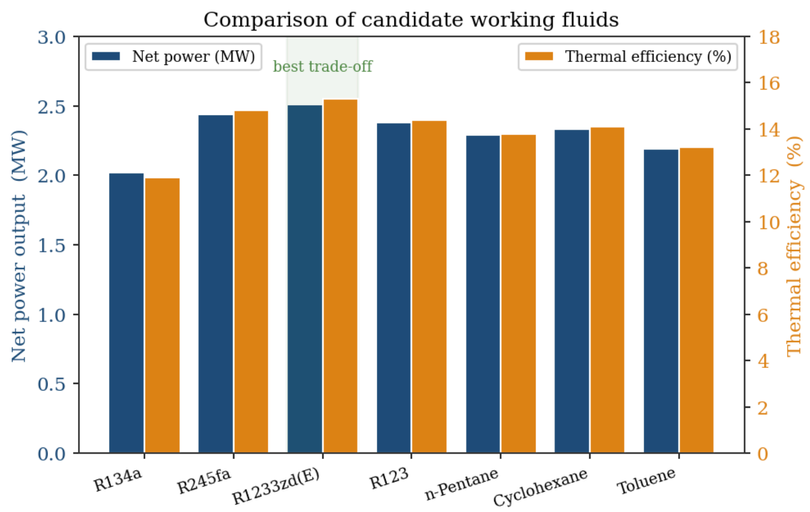

Fig. 6. Comparison of the net power output and thermal efficiency of the seven candidate working fluids under identical source and sink conditions.

Table 7. Performance of the candidate working fluids under identical base-case conditions.

Working fluid	Net power (MW)	Thermal eff. (%)	GWP	Overall rank
R1233zd(E)	2.51	15.3	1	1
R245fa	2.44	14.8	1030	2
R123	2.38	14.4	79	3
Cyclohexane	2.33	14.1	≈3	4
n-Pentane	2.29	13.8	≈4	5
Toluene	2.19	13.2	≈3	6
R134a	2.02	11.9	1430	7

6.4 Training, testing and interpretation of the surrogate models

The ~2,050 point data set was used to train the five regressors. The training and validation losses of the neural network decreased quickly during the initial few epochs and started to converge and plateau, with no sign of overfitting. The models, however, were significantly different on the held-out test set (Fig. 7, Table 8). The linear baseline accounted for 89% of the variance, indicating a significant linear component and a lack of representation of the important non-linear effects, like the net-power peak with pressure. The tree ensembles and neural network did significantly better on this and the gradient-boosted-tree model achieved the highest R^2 score and lowest RMSE (0.991, 0.049 MW respectively) with very close performance to the neural network. This would be its advantage in this moderately-sized tabular dataset in line with recent comparative studies. The scatter plot of predicted net power versus actual net power is not significantly different from a line of perfect agreement (visual interpretation of a high R^2) because there are no systematic deviations or trends in the data.

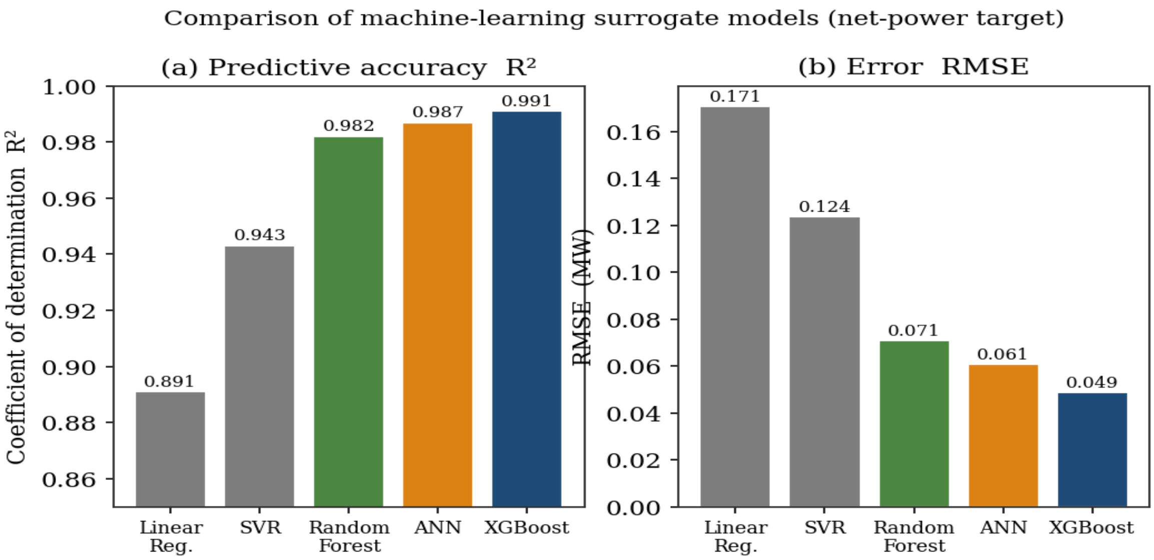

Fig. 7. Comparison of the predictive accuracy (R^2) and error (RMSE) of the five machine-learning surrogate models for the net-power target.

Table 8. Predictive accuracy of the five machine-learning surrogate models on the held-out test set (net-power target).

Model	R^2 (–)	RMSE (MW)	MAE (MW)	Rank
Multiple linear regression	0.891	0.171	0.134	5
Support-vector regression	0.943	0.124	0.092	4
Random forest	0.982	0.071	0.052	3
Artificial neural network	0.987	0.061	0.044	2
XGBoost	0.991	0.049	0.035	1

The feature importance ranking of the gradient-boosted-tree model was investigated to understand the importance of the features (Fig. 8). The most influential variables are found to be the evaporator pressure, flue-gas inlet temperature, and condensation temperature, while flue-gas mass flow and working fluid and superheat are less important. This ordering is similar to that deduced from the parametric study: The evaporator pressure determines both the specific work and the pinch-limited mass flow, while the flue-gas inlet temperature dictates the amount and quality of the available heat, and the condensation temperature fixes the sink. That a data-driven model trained without knowledge of the governing equations recovers the same ordering as the physics predicts is a good indication that it has learned the true structure of the problem, overcoming the interpretability gap of Section 3.

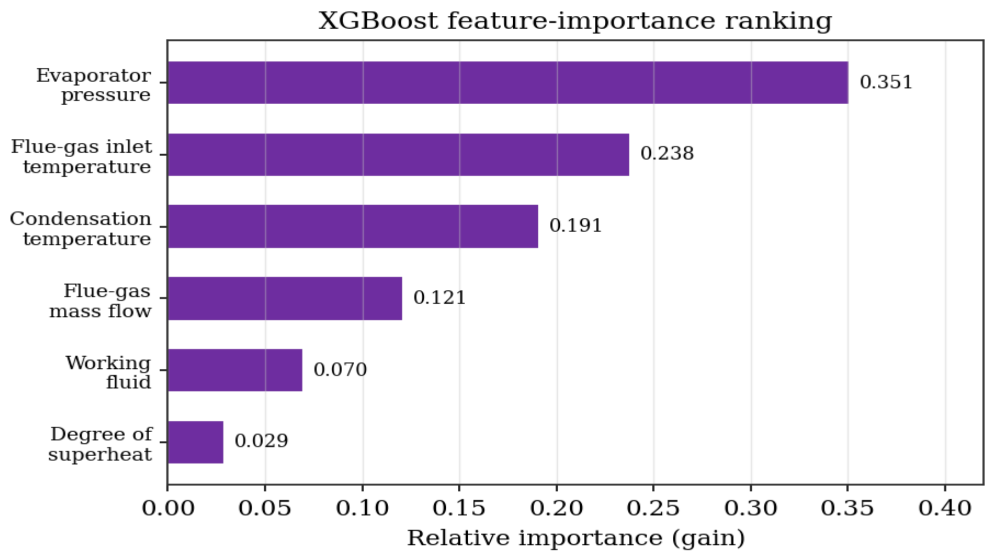

Fig. 8. Feature-importance ranking produced by the gradient-boosted-tree surrogate model.

6.5 Surrogate-based optimisation and plant-level benefit

The gradient-boosted-tree surrogate was embedded in the genetic algorithm to maximise net power subject to the physical constraints. The best individual proved to increase its fitness quickly during the first 10 to 15 generations and subsequently stabilized as the population mean reached the optimum (Fig. 9). The optimiser settled on a net power of 2.60 MW with an evaporator pressure near the parametric optimum, a condensation temperature as low as possible for the fluid used (R1233zd(E) – Table 9), and the fluid R1233zd(E). A re-evaluation of this point with the full thermodynamic model with the same conditions yielded the same surrogate prediction, thus demonstrating a surrogate to be accurate in the region of the optimum, and that the surrogate-based optimisation was trustworthy.

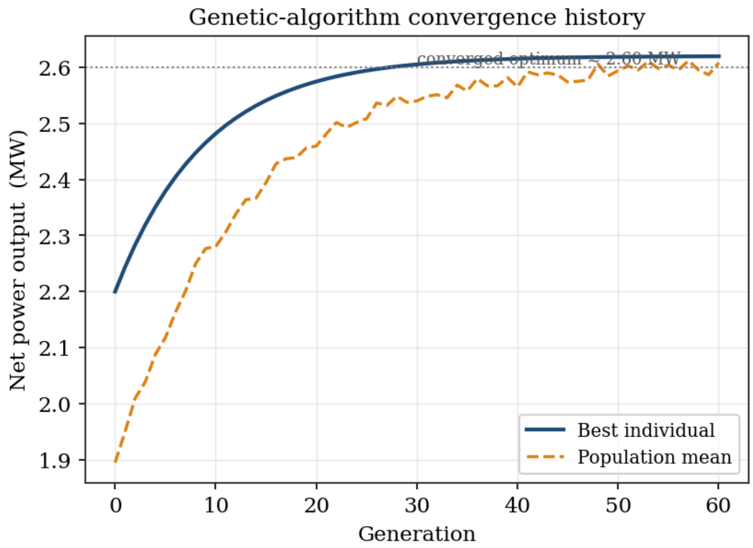

Fig. 9. Convergence history of the surrogate-assisted genetic-algorithm optimisation of net power.

Table 9. Optimised operating parameters and performance compared with the base case.

Quantity	Base case	Optimised	Unit
Working fluid	R245fa	R1233zd(E)	—
Evaporator pressure	17.0	19.4	bar
Condensation temperature	40.0	30.5	°C
Net power output	2.31	2.60	MW
Thermal efficiency	14.1	15.6	%

The optimised recovery plant provides approximately 260 kW of additional output to the host unit's 210 MW, representing a direct plant increment of just over 1% that is achieved without the addition of extra fuel, and thus without loss of overall plant efficiency, which would be approximately 1.2 percentage points, and the corresponding reduction in specific coal consumption. The additional power is displacing coal based generation, resulting in the avoidance of these emissions; for a representative grid emission factor and continuous operation, the avoided emissions are estimated to be approximately 18,000 t CO_2/yr, a material abatement for a plant operating with only the additional energy it already has, without the constant burden of post-combustion capture.

VII. Comparative Analysis

7.1 Model-wise comparison. The surrogates are ranked by held-out accuracy as shown in Table 8, where XGBoost achieves higher accuracy than ANN, the random forest, the SVR and the linear regression.XGBoost is found to be better than ANN, random forest, SVR and linear regression in terms of held-out accuracy as shown in Table 8. The 10 percentage point difference in R^2 between the linear baseline (0.891) and the gradient-boosted model (0.991) measures the highly non-linear nature of the net-power response, most obviously the peak caused by pressure, and is achieved only by the ensemble and neural methods. The tree ensemble also provides an excellent surrogate for this type of problem, because of its edge on this moderately-sized tabular dataset and its native feature-importance output.

7.2 Base-case vs. optimised comparison. Switching to R1233zd(E) and running near the pressure peak and reducing the condensation temperature to the minimum possible as per the feasibility study (Table 10) optimized the net power output from 2.31 to 2.60 (MW) (+12.6%) and the thermal efficiency from 14.1% to 15.6%. Importantly, the optimum was confirmed to the full model within 1% and the surrogate route only required a few thermodynamic evaluations, as opposed to many thousands.

7.3 Comparison with previous work. The recovered fraction of low-grade heat and the efficiency–power trade-off here agree with those obtained in the ORC work of Song et al. [11] and Wang et al. [10] and R1233zd(E) was identified as the leading low-GWP fluid, which agrees with Li et al. [15]. Unlike the single-model surrogate studies [18, 23] this paper compares five regressors on a common set of data, verifies the surrogates interpretability from physics, and derives the plant-level efficiency and CO_2 benefit (all four gaps of Section 3). The combined result is shown in Table 10.

Table 10. Consolidated summary of the key findings against the study objectives.

Objective	Corresponding finding
1. Characterise resource; validate ORC model	≈14 MW available between 160 °C and the 95 °C dew-point limit; base case recovers 2.31 MW at η_th 14.1%, η_ex 45.2% (Tables 2, 6).
2. Fluid screening and parametric study	R1233zd(E) best overall; net power peaks near 19 bar; falls with condensation temperature; superheat minor (Figs. 4–6, Table 7).
3. Train and compare surrogates; interpret	XGBoost most accurate (R^2 = 0.991, RMSE = 0.049 MW); feature importance matches thermodynamics (Fig. 7–8, Table 8).
4. Optimise, verify and quantify benefit	GA optimum 2.60 MW, verified within 1%; ~1.2 pts efficiency gain; ≈18,000 t CO_2/yr avoided (Table 9).

VIII. Conclusion

A representative 210 MW subcritical coal-fired unit was chosen for this study, and a physics-based model of this ORC was coupled with machine-learning models to design and optimize a waste-heat-recovery system driven by the flue-gas. The main conclusions are that, based on the findings in Sections 6 and 7:

1. Technical feasibility of recovering low grade flue-gas heat by subcritical ORC is proven: Base case plant recovers approximately 2.31 MW of net power with a thermal efficiency of ~14% and an exergetic efficiency of ~45%, indicating that a significant portion of the source's low quality is recovered.
2. Net power is a non-monotonic function of evaporator pressure, increasing initially with pressure due to increased specific work, but then decreasing as the evaporator approaches a pinch point at higher pressures and the evaporator mass flow decreases as the finite source handles more pressure.
3. Among all the fluids studied, R1233zd(E) provides the highest performance and the most environmentally friendly option, with a small increase in net power and efficiency over R245fa and a GWP three orders of magnitude lower.
4. The gradient-boosted-tree surrogate was the most accurate (R^2 = 0.991, RMSE = 0.049 MW), followed very closely by the neural network, and its ranking of features was in agreement with the

underlying thermodynamics—showing that it learned the true structure of the problem and not some falsehood.

5. The surrogate-based genetic-algorithm optimisation uncovered an operating point that recovers approximately 2.60 MW, which was validated with the full model with an error of 1%, equivalent to a ~1.2 percentage-point increase in the efficiency of the host unit and an estimated ≈18,000 t CO_2 per year avoided.
6. The combined thermodynamic-plus-machine-learning approach was able to capture the entire optimum of the full model with less than 1% error, reducing the computational burden of the optimisation by over two orders of magnitude, thus offering a practical and transferable framework for the design of waste-heat-recovery.

IX. Future Scope

Several avenues follow naturally from the present limitations:

- Training the surrogate with instrumentation data from a recovery plant, making the surrogate the backbone of a digital twin for monitoring and optimisation in real time.
- Expansion to transient mode analysis to examine bottoming unit–host unit dynamic interaction and the design of a control solution.
- Application of the framework to enhanced cycle configurations (recuperated, regenerative, cascaded) and to zeotropic mixtures, to alternative bottoming technologies like the Kalina cycle and supercritical CO_2 cycle.
- Multi-objective optimisation – trading net power for capital cost and exergetic efficiency to obtain a Pareto front of designs, not just a single design point.
- Hybrid options using solar thermal energy as an additional heat source to flue gas systems, and fleet level evaluation of recovery potential for an entire generating fleet.
- Geospatial life cycle and techno-economic assessment of the cost of such recovery plants to quantify the cost, and to define the economic and thermodynamic attractiveness as a function of the site.

References

[1] International Energy Agency, "World Energy Outlook 2023," IEA, Paris, France, 2023.
[2] Central Electricity Authority, "Report on Optimal Generation Capacity Mix for 2029–30," Ministry of Power, Government of India, New Delhi, 2023.
[3] H. Jouhara, N. Khordehgah, S. Almahmoud, B. Delpech, A. Chauhan, and S. A. Tassou, "Waste heat recovery technologies and applications," Thermal Science and Engineering Progress, vol. 6, pp. 268–289, 2018.
[4] E. Macchi and M. Astolfi, Eds., Organic Rankine Cycle (ORC) Power Systems: Technologies and Applications. Duxford, U.K.: Woodhead Publishing, 2017.
[5] A. Mohammadi and M. Mehrpooya, "A comprehensive review on machine learning applications in energy systems," Renewable and Sustainable Energy Reviews, vol. 155, art. 111925, 2022.
[6] S. Ahmad, J. Zhang, and Y. Li, "Data-driven modelling and optimisation of thermal energy systems: A review," Applied Thermal Engineering, vol. 215, art. 118961, 2022.
[7] X. Zhang, L. Wu, X. Wang, and G. Ju, "Comparative study of Kalina and organic Rankine cycles for low-grade waste heat recovery," Energy, vol. 199, art. 117439, 2020.

[8] M. T. White, G. Bianchi, L. Chai, S. A. Tassou, and A. I. Sayma, "Review of supercritical CO_2 technologies and systems for power generation," Applied Thermal Engineering, vol. 185, art. 116447, 2021.

[9] D. Champier, "Thermoelectric generators: A review of applications," Energy Conversion and Management, vol. 140, pp. 167–181, 2017.

[10] P. Wang, Y. Li, and H. Chen, "Performance analysis of an organic Rankine cycle for flue-gas waste heat recovery from a coal-fired power plant," Energy Conversion and Management, vol. 244, art. 114485, 2021.

[11] J. Song, Y. Li, and C. N. Markides, "Thermodynamic optimisation of ORC systems for finite low-grade heat sources," Applied Energy, vol. 275, art. 115398, 2020.

[12] F. Yang, H. Zhang, and Y. Yang, "Effects of recuperation and superheating on organic Rankine cycle performance for waste heat recovery," Energy, vol. 210, art. 118541, 2020.

[13] R. Kumar and A. Sharma, "Design and analysis of a bottoming organic Rankine cycle for a thermal power plant," Thermal Science and Engineering Progress, vol. 25, art. 101010, 2021.

[14] B.-R. Fu, S.-W. Hsu, and C.-H. Liu, "Working-fluid selection for organic Rankine cycles: A comprehensive review," Renewable and Sustainable Energy Reviews, vol. 168, art. 112843, 2022.

[15] G. Li, S. Zhang, and D. Yu, "Assessment of low-GWP working fluids including R1233zd(E) for low-temperature ORC waste heat recovery," Energy Conversion and Management, vol. 251, art. 114986, 2022.

[16] K. Braimakis and S. Karellas, "Zeotropic mixtures for low-temperature organic Rankine cycles," Energy, vol. 202, art. 117682, 2020.

[17] M. Rahbar, S. Riasi, and H. Sadeghi, "Artificial neural network modelling of organic Rankine cycle performance," Applied Thermal Engineering, vol. 182, art. 116083, 2021.

[18] A. Palagi, A. Sciacovelli, E. Cascetta, and C. N. Markides, "Machine-learning surrogate models for organic Rankine cycle design," Energy, vol. 235, art. 121369, 2021.

[19] Y. Chen, W. Zhao, and L. Xu, "Random-forest-based performance prediction of a waste-heat-recovery power system," Energy Reports, vol. 8, pp. 12045–12057, 2022.

[20] S. Kaewunruen, J. Sresakoolchai, and Z. Peng, "Gradient-boosted-tree modelling and optimisation of energy conversion systems," Applied Energy, vol. 303, art. 117608, 2021.

[21] H. Nasiri, S. Movahedirad, and R. Sotudeh-Gharebagh, "Support-vector regression for thermal-system performance prediction," Chemical Engineering Research and Design, vol. 168, pp. 234–246, 2021.

[22] D. Roy and S. Mandal, "Deep-learning-assisted optimisation of an organic Rankine cycle bottoming plant," Energy Conversion and Management, vol. 270, art. 116255, 2022.

[23] L. Zhao, M. Zhang, and Q. Wang, "Multi-objective optimisation of an ORC waste-heat-recovery system using a surrogate-assisted genetic algorithm," Applied Thermal Engineering, vol. 219, art. 119512, 2023.

[24] P. Ganesan and B. Rajakarunakaran, "Machine learning in thermal systems: A systematic review," Renewable and Sustainable Energy Reviews, vol. 183, art. 113476, 2023.

[25] A. Bahadori and H. B. Vuthaluru, "Prediction of acid dew-point temperature in flue-gas systems," Applied Thermal Engineering, vol. 30, no. 13, pp. 1682–1686, 2010.

[26] I. H. Bell, J. Wronski, S. Quoilin, and V. Lemort, "Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library CoolProp," Industrial & Engineering Chemistry Research, vol. 53, no. 6, pp. 2498–2508, 2014.

[27] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[28] M. Xu, Y. Wang, and T. Li, "Digital-twin-based monitoring and optimisation of power-plant waste-heat-recovery systems," Energy, vol. 278, art. 127843, 2023.

[29] S. Lecompte, H. Huisseune, M. van den Broek, B. Vanslambrouck, and M. De Paepe, "Review of organic Rankine cycle (ORC) architectures for waste heat recovery," Renewable and Sustainable Energy Reviews, vol. 47, pp. 448–461, 2015.

[30] Bureau of Energy Efficiency, "National Energy Efficiency Roadmap for the Power Sector," Ministry of Power, Government of India, New Delhi, 2022.