Identifying the key system parameters of the organic Rankine cycle using the principal component analysis based on an experimental database

Yan, Dong; Yang, Fubin; Yang, Fufang; Zhang, Hongguang; Guo, Zhiyu; Li, Jian; Wu, Yuting

Published in:
Energy Conversion and Management

Link to article, DOI:
[10.1016/j.enconman.2021.114252](https://doi.org/10.1016/j.enconman.2021.114252)

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Yan, D., Yang, F., Yang, F., Zhang, H., Guo, Z., Li, J., & Wu, Y. (2021). Identifying the key system parameters of the organic Rankine cycle using the principal component analysis based on an experimental database. *Energy Conversion and Management*, 240, [114252]. <https://doi.org/10.1016/j.enconman.2021.114252>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Identifying the key system parameters of the organic Rankine cycle using the principal component analysis based on an experimental database

Dong Yan^{a,*}, Fubin Yang^{a,*}, Fufang Yang^{b,*}, Hongguang Zhang^a, Zhiyu Guo^a, Jian Li^c, Yuting Wu^a

^a Key Laboratory of Enhanced Heat Transfer and Energy Conservation of MOE, Beijing Key Laboratory of Heat Transfer and Energy Conversion, Faculty of Environment and Life, Beijing University of Technology, Beijing 100124, China

^b Center for Energy Resources Engineering (CERE), Department of Chemical and Biochemical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

^c Key Laboratory for Thermal Science and Power Engineering of MOE, Beijing Key Laboratory for CO₂ Utilization and Reduction Technology, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Keywords:

Organic Rankine cycle
Experimental analysis
Principal component analysis
Machine learning
Key parameter subset

ABSTRACT

The organic Rankine cycle (ORC) is a promising technology for medium-and-low temperature heat utilization. However, the mechanism of how system parameters affect output have been investigated very little in the experimental aspect. Experimental investigation on the impact of each system parameter on system performance requires decoupling these system parameters. In this work, a series of experiments are conducted on a 10 kW scale ORC experiment setup. Statistical analysis is performed to identify a key parameter subset based on an experimental database. 6 system parameters, including temperature (T_e) and pressure (p_e) at the evaporator outlet, temperature (T_c) and pressure (p_c) at the condenser inlet, expander shaft efficiency (η_{SSE}), and working fluid pump efficiency (η_p) are obtained. Combined with the ORC net power output and thermal efficiency, an experimental database of system operation conditions is constructed. Subsequently, the principal component analysis (PCA) of ORC is conducted based on the experimental database. Prediction models are developed using multi-linear regression (MLR), back propagation artificial neural network (BP-ANN), and support vector regression (SVR). Finally, accounting for the prediction performance of models and system parameter inter-correlation behavior, the key parameter subset is determined with the exhaustive feature selection method. The results imply that the key parameter subset is (p_e , η_p , p_c , η_{SSE}). Further removing or including more system parameters would reduce the accuracy of prediction models. In addition, the MLR models are slightly less accurate than the more sophisticated BP-ANN and SVR models.

1. Introduction

Organic Rankine cycle (ORC) is an efficient and affordable technology for the utilization of low-and-medium temperature heat source from the waste heat and renewable energy [1,2]. Because of its advantages of versatility, high efficiency and simple configuration, ORC is considered a promising technology in many applications, such as geothermal [3], solar [4], biomass [5], internal combustion engines waste heat [6–8], and industrial waste heat [9,10].

Efforts have been made to improve ORC system performance by optimizing system parameters in both theoretical and experimental approaches. Yang et al. [11] investigated the effect of working fluid characteristic parameters on the ORC thermodynamic performance

based on a corresponding state approach. Unamba et al. [12] investigated the influence of heat source temperature, working fluid flow rates and pressure ratios on the ORC overall efficiency based on experiments. Dickes et al. [13] analyzed the influence of fluid charge on the ORC system performance and pointed out that the fluid charge and receiver size are the key parameters. Sarkar [14] evaluated the effect of pinch point design and working fluid mass flow on the subcritical-supercritical ORC system performance. Imran et al. [15] studied the effect of working fluid superheating on ORC system thermal efficiency based on experiments. Branchini et al. [16] performed a parametric investigation of ORC system and found the evaporation pressure was the most important key parameter which could improve system efficiency. Meng et al. [17] investigated the influence of centrifugal pump efficiency on ORC system performance and found that increasing the pump efficiency could

* Corresponding authors.

E-mail addresses: yangfubin@bjut.edu.cn (F. Yang), fufya@kt.dtu.dk (F. Yang).

Nomenclature

T	Temperature ($^{\circ}\text{C}$)
p	Pressure (bar)
$\dot{m}$	Mass flow rate (kg s^{-1})
h	Specific enthalpy (J kg^{-1})
s	Specific entropy ($\text{J kg}^{-1} \text{K}^{-1}$)

Abbreviations

SSE	Single screw expander
ORC	Organic Rankine cycle
RMSE	Root mean square error
R^2	Determination coefficient
PCA	Principal component analysis
PC	Principal component
MLR	Multivariable linear regression

SVR	Support vector regression
BP-ANN	Back propagation artificial neural network

Subscripts

e	Evaporator outlet
c	Condenser inlet
1-6	State point

Greek

η_{SSE}	Expander shaft efficiency
η_{P}	Working fluid pump efficiency
η_{ORC}	ORC system thermal efficiency
$\dot{W}_{\text{net}}$	ORC system net power output (kW)
$\dot{W}_{\text{SSE}}$	Single screw expander power output (kW)
$\dot{W}_{\text{P}}$	Power consumption of working fluid pump (kW)

Fig. 1. Schematic diagram of the ORC test system.

improve system performance. Xi et al. [18] conducted an orthogonal experiment to analyze the sensitivity of parameters on ORC system thermal efficiency and proposed that the expander inlet temperature was the dominant factor. On the one hand, there are many system parameters that are related to the ORC system performance. It is inevitable that there is redundant information as many parameters are obtained.

On the other hand, the ORC system performance cannot be determined by only 1 or 2 key parameters. Therefore, to obtain accurate ORC system performance based on as few system parameters as possible, it is necessary to identify the key system parameter subset based on a statistical approach.

Statistical methods can be used to interpret the relationship between the system parameters and outputs. Efforts have been made to identify the key ORC system parameters using statistical method. Bademlioglu et al. [19,20] adopted two statistical methods: Taguchi and ANOVA, to obtain the contribution ratios and importance order of system parameters in ORC system. Kalina et al. [21] presented the grey box and black box regression models to identify the optimal operation condition of ORC. However, for any thermal system, the system parameters are highly coupled, and the system parameters cannot vary independently. Principal component analysis (PCA) is a widely used statistical method that transforms raw multi-dimensional observation data into a set of orthogonal vectors. It can discover the hidden patterns in the variables and decouple these variables by constructing a few principal components that contain as much variables information as possible. Furthermore, The PCA can reduce the dimension of the database. Recently, PCA has received much attention in thermodynamic research [22–24].

This paper aims to identify the key parameter subset in the ORC system based on an experimental database. 6 system parameters, including temperature (T_e) and pressure (p_e) at the evaporator outlet, temperature (T_c) and pressure (p_c) at the condenser inlet, expander shaft

Fig. 2. Photograph of the ORC test system.

Table 1

Types and parameters of the components of the working fluid circulation.

Component	Type	Parameter
Expander	Single screw expander	Star-wheel diameter: 117 mm
Pump	Centrifugal pump (CR 3–25)	Flow rate: 3 m ³ /h, stage: 25, delivery head: 122 m
Condenser	Shell-and-tube heat exchanger	Heat exchange area: 21.6 m ²
Evaporator	Shell-and-tube heat exchanger	Heat exchange area: 20 m ²

Table 2

Parameters of measurement sensors in the ORC test system.

Measurement sensors	Measurement range	Accuracy
Temperature sensor	−50–150 °C	±0.75 °C
Pressure sensor	0–2 MPa	±0.25%
Mass flow rate	0–3000 kg/h	±0.2%
Rotation speed sensor	0–8000 rpm	±0.5%
Torque sensor	0–100 N.m	±0.5%

efficiency (η_{SSE}), and working fluid pump efficiency (η_p) are selected as variables. Correlations among these ORC system parameters are investigated, and the PCA is performed. The net power output and thermal efficiency are selected as ORC system performance indicators. Regression models are developed using a few machine learning methods. Finally, according to the data correlation and prediction performance of the regression models, the key parameter subset is identified using the exhaustive feature selection method. The key parameter subset can benefit the design and control of the ORC system.

2. Experiment

A 10 kW scale ORC test system is employed in the present work. Fig. 1 shows the system schematic diagram. The system consists of 3 circulations: conductive oil circulation, working fluid circulation, and cooling water circulation. Fig. 2 shows a photograph of the ORC test system.

2.1. Conductive oil circulation

In the experimental setup, the conductive circulation serves as the heat source. It consists of 3 major components: boiler, pump, and heat exchanger. In the boiler, the conductive oil is heated by an electric heater. Then, it flows into the heat exchanger and rejects heat to the working fluid circulation. Finally, it is pumped back into the boiler.

2.2. Working fluid circulation

The working fluid circulation consists of 4 components: pump, evaporator, expander, and condenser. Table 1 shows the types and parameters of the components. The working fluid is R123 [25]. It is pumped into the evaporator and heated to superheated vapor. Then, it flows into the expander and drives the machinery. Finally, it cools down in the condenser and enters the working fluid pump. Table 2 shows the main parameters of the sensors used in this circulation. The locations of temperature and pressure sensors are depicted in Fig. 1.

Table 3

Variation ranges of ORC system test data.

	p_c (bar)	T_c (°C)	p_e (bar)	T_e (°C)	η_p (%)	η_{SSE} (%)	$\dot{W}_{net}$ (kW)	η_{ORC} (%)
Maximum	10.61	121.40	2.33	49.20	29.26	69.84	4.96	5.72
Minimum	4.18	80.40	1.07	24.08	7.94	46.04	1.59	3.56

Table 4

Uncertainties of the net power output and thermal efficiency.

	$\dot{W}_{net}$ [kW]	η_{ORC}
Maximum	0.1850	0.00334
Minimum	0.1605	0.00170
Average	0.1719	0.00244

Fig. 3. T-s diagram of the ORC system.

2.3. Cooling water circulation

The cooling water circulation serves as the heat sink. It consists of 3 major components: pump, cooling water tower, and condenser. This circulation absorbs heat from the working fluid circulation, and rejects heat to the environment.

2.4. Experimental procedure

A series of experiments are conducted on the ORC test system. The aim is to investigate how the system performance is affected by the system parameters including p_c , T_c , p_e , T_e , η_p and η_{SSE} . The rotation speed of the expander of the ORC system is controlled within the range of 3000 ± 20 rpm, in order to match the 50 Hz electricity grid frequency. The operation conditions are changed by adjusting the conductive oil circulation and the working fluid pump. During the experiment, the pressure and the temperature at the evaporator outlet changed from the initial state ($T = 80.40$ °C, $p = 4.18$ bar) to the end state ($T = 121.40$ °C, $p = 10.61$ bar). At the same time, all the measurement data are recorded. Table 3 shows the variation ranges of the ORC system experimental data.

A total of 2636 steady experimental data points are collected. It takes about 220 h to obtain these experimental data points. The mass flow rate in the system components is used for determining whether the experiment is at steady-state. When the difference of mass flow rate between the inlet of expander and the inlet of evaporator is less than 50 g/s, the ORC system is considered to be at steady state. The combined uncertainty is,

$$\Delta Y = \sqrt{\sum_i^n \left(\frac{\partial Y}{\partial X_i} \Delta X_i \right)^2} \quad (1)$$

where ΔY represents the uncertainty of target variable Y , and ΔX_i

Fig. 4. Framework of calculation and training processes.

represents the uncertainty of measured variable X . Table 4 shows the uncertainties of the net power output and thermal efficiency.

3. Modeling

3.1. Thermodynamic model

This section describes the thermodynamic framework of the ORC system. Fig. 3 shows the T - s diagram. The ORC system consists of 4 processes: the heat absorption process (5-1) in the evaporator, the expansion process (1-2) in the single screw expander, the condensation process (2-4) in the condenser, and the compression process (4-5) in the working fluid pump.

The expander shaft efficiency is,

$$\eta_{SSE} = \frac{\dot{W}_{SSE}}{\dot{m}(h_1 - h_2)} \quad (2)$$

The working fluid pump efficiency is,

$$\eta_P = \frac{q_v(p_5 - p_4)}{\dot{W}_P} \quad (3)$$

The net power output is,

$$\dot{W}_{net} = \dot{W}_{SSE} - \dot{W}_P \quad (4)$$

The thermal efficiency is,

Table 5
Correlation coefficients between the system parameters.

	p_e	T_e	p_c	T_c	η_P	η_{SSE}
p_e	1	0.992	0.975	0.978	0.478	0.147
T_e		1	0.979	0.974	0.452	0.198
p_c			1	0.982	0.332	0.299
T_c				1	0.402	0.148
η_P					1	-0.548
η_{SSE}						1

$$\eta_{ORC} = \frac{\dot{W}_{SSE} - \dot{W}_P}{\dot{m}(h_1 - h_5)} \quad (5)$$

3.2. Principal component analysis

PCA is a classic statistical method to analysis datasets that are correlated between the dimensions. It extracts the important information from the variables, and transforms the original variables into a set of new orthogonal variables, i.e., principal components (PCs) [26]. The PCs represent a rank of magnitudes of variable variance on the orthogonal axis. They are ordered so that the first PC accounts for the most variance in the dataset, and that the following PCs account for the most variances on the directions that are orthogonal to the obtained PCs subsequently.

The procedure of PCA is described as follows [27]. First, a new matrix $\mathbf{X}^*$ is standardized from the raw data matrix $\mathbf{X}$. Then, the correlation coefficient matrix $\mathbf{R}$ is calculated based on $\mathbf{X}^*$. Then, the eigenvalues ($\lambda_1, \lambda_2, \dots, \lambda_p$) in the order of magnitudes are obtained by solving the characteristic equation $|\lambda \mathbf{I} - \mathbf{R}| = 0$ with Jacobi algorithm and the eigenvector is found corresponding to the eigenvalue. Finally, the PCs contribution percentage, the cumulative contribution percentage, and the PC factor loadings are calculated.

In the ORC system, the two pairs of system parameters (p_e and T_e , p_c and T_c) are always strongly correlated. The PCA reduces the dimension of the experimental database. Thus, the “basic structure” of the system parameters can be captured, i.e., the key parameters can be identified.

3.3. Regression models

This section introduces the regression models. In a previous work [28], it has been found that the traditional thermodynamic ORC regression model may reduce the accuracy as the irreversible losses, such as pressure drops, heat losses and mechanical losses, are neglected. Multivariable linear regression (MLR) is commonly used regression model with simple structure. Back propagation artificial neural network (BP-ANN) and support vector regression (SVR) are popular methods to develop machine learning regression. In this work, MLR, BP-ANN and SVR models are developed to correlate the experimental data. The flow chart of specific calculation and training processes is shown in Fig. 4. These data are randomly divided into training (75%), validation (15%) and test (15%) sets. The training set is used for training the correlations between input and output variables. The validation set is used for tuning the intermediate parameters of BP-ANN and SVR models during the training process. In the MLR model, no validation set is needed; the training and validation sets are combined (85%) and used for training. The test set is used as new data for testing the prediction performance. The invariables are the 6 system parameters, i.e., p_e , T_e , p_c , T_c , η_P , and η_{SSE} . The responses are the net power output and thermal efficiency. Details of the 3 models are shown below.

- (1) MLR: MLR is the straightforward approach that uses a linear combination of the 6 invariables to predict the responses.
- (2) BP-ANN: BP-ANN is a multi-layer feedforward neural network trained according to the error back propagation algorithm [29]. It has a lot of advantages, such as self-adaption, self-learning, nonlinear mapping, high accuracy, and has been widely used in ORC research [30–32]. In this work, the MATLAB neural network toolbox is used. The number of hidden node layer is 1. The number of hidden nodes is 9. The learning rate is 0.01. The training function is Levenberg-Marquardt.
- (3) SVR: SVR is an important branch of support vector machine, and it minimizes the total deviation of all sample points from the hyperplane [33]. The SVR is a powerful tool that allows one to decide how tolerate one is about regression errors. In this work, the libsvm-3.24 toolbox [34] is used. The grid method is selected.

Table 6
PCA factor loadings and eigenvalues.

	PC1	PC2	PC3	PC4	PC5	PC6
p_c	0.4870	-0.0030	-0.0609	-0.4561	-0.6646	-0.3306
T_c	0.4864	0.0336	-0.0040	-0.5482	0.6661	0.1342
p_c	0.4804	0.1349	-0.1002	0.3275	-0.2415	0.7585
T_c	0.4813	0.0303	-0.2912	0.5898	0.2370	-0.5279
η_p	0.2399	-0.6414	0.7083	0.1704	0.0050	-0.0163
η_{SSE}	0.0791	0.7539	0.6322	0.0853	0.0056	-0.1357
Eigenvalue	4.1788	1.5336	0.2551	0.0220	0.0066	0.0039

Fig. 5. Contribution percentage and cumulative contribution percentage of the PCs.

3.4. Exhaustive feature selection

Exhaustive feature selection is an effective technique to find the optimal subset of relevant features. For each number of parameters (1 to 6), the key parameter subset is identified by comparing all the different combinations.

4. Results and discussion

4.1. Correlation analysis and PCA

Before the PCA is implemented, the correlations between the system

Fig. 6. R^2 of MLR, BP-ANN and SVR models.

Fig. 7. RMSE of the MLR, BP-ANN and SVR models for the net power output and thermal efficiency.

parameters (raw data) are investigated. In this work, the 4 system parameters, p_c , T_c , p_c and T_c , are denoted as the working fluid state parameters, and the other 2 system parameters, η_p and η_{SSE} , are denoted as the equipment operating efficiency parameters. Table 5 shows the correlation coefficients between the 6 system parameters. The closer the absolute value of the correlation coefficient is to 1, the stronger the correlation between the system parameters is. In contrast, the closer the correlation coefficient is to 0, the weaker the correlation between the system parameters is. The results show that there are strong correlations between the working fluid state parameters. As expected, the correlations between the equipment operating efficiency parameters, and either of them with the working fluid state parameters are not as pronounced. The correlation coefficient between the equipment operating efficiency parameters is -0.548, and the correlations between the equipment operating efficiency parameters and the working fluid state parameters are even lower, implying that the equipment operating efficiency parameters are quite independent.

In PCA, the 1st principal component (PC) contains the most variance of the data. The 2nd PC contains the most variance in the remaining dimensions, etc. Table 6 shows the PCA factor loadings and eigenvalues. PC1, PC2, PC3, PC4, PC5 and PC6 represents the 1st, 2nd, 3rd, 4th, 5th, and 6th PCs, respectively. Fig. 5 shows the contribution percentage and cumulative contribution percentage of the PCs. The cumulative information contribution rate of PC1 and PC2 is up to 95.21%. According to the PCA, the PC1 factor loadings of the ORC working fluid state parameters are much larger than those of the equipment operating efficiency parameters, while the PC2 factor loadings of the ORC working fluid state parameters are much smaller than those of the equipment operating efficiency parameters. Thus, the PC1 can be viewed as the working fluid state parameters indicator; the PC2 can be viewed as the equipment operating efficiency parameters indicator.

Table 7

Correlation coefficients between each PC and ORC system performance.

PCA factors	$\dot{W}_{net}$	η_{ORC}
PC1	0.9149	0.6162
PC2	-0.3448	-0.4857
PC3	0.1047	0.1603
PC4	-0.1394	-0.1040
PC5	-0.0657	-0.0737
PC6	-0.0703	-0.1821

Fig. 8. Used system parameters.

4.2. Model evaluation

The determination coefficient (R^2) and root mean square error (RMSE) are 2 important indicators to evaluate the performance of regression models. Figs. 6 and 7 show R^2 and RMSE of the MLR, BP-ANN and SVR models. All the net power output models achieve good prediction (the R^2 of each regression model is more than 0.995, and the RMSE of each regression model is less than 0.055). R^2 of the MLR thermal efficiency model ($R^2 = 0.9772$) is slightly smaller than that of the BP-ANN and SVR models. RMSE of the MLR net power output and thermal efficiency models are larger than that of BP-ANN and SVR models. SVR is the most accurate among the models. Although slightly less accurate, the MLR models are comparable to the more sophisticated models.

The correlation coefficients between each PC and system performance is shown in Table 7. The net power output is highly related to

Fig. 9. R^2 of the PCA-MLR models using different numbers of system parameters. **Fig. 10.** RMSE of the PCA-MLR models for (a) the net power output and (b) thermal efficiency using different numbers of system parameters.

PC1. The correlation coefficients with the other PCs are much smaller. The much smaller correlation coefficient implies that the impact of the system parameters is largely captured by PC1. However, one has to include PC2 in any model because the thermal efficiency requires the extra variance information, especially that of the equipment operating efficiency parameters.

4.3. Identification of the key parameter subset

For a real ORC system, many parameters can be obtained from the installed sensors. It is straightforward that information from more than 1 or 2 of these parameters is needed to account for variances on the many dimensions. However, many of these parameters are strongly correlated with each other, as discussed in the previous sections. Furthermore, some of the parameters are not as important as others. Most importantly, a model with many parameters is very likely to be over-trained. Consequently, to obtain an accurate predictive model, it is desired to identify a key parameter subset.

As discussed in Section 4.1, most of the system parameters information is extracted in PC1 and PC2. In this section, MLR models are developed using PC1 and PC2 as invariables. The models are denoted as

Fig. 11. (a) Relative deviation of the net power output and (b) absolute deviation of the efficiency of the PCA-MLR models.

PCA-MLR. In this approach, the key parameter subset is identified. In the following context, all the results are of the test set, i.e., predictions on new data. Fig. 8 shows the used system parameters in the PCA-MLR models using 1–6 systems parameters. Fig. 9 shows R^2 of the PCA-MLR models for the net power output and thermal efficiency using different numbers of system parameters. PC1 and PC2 are used as invariables of the PCA-MLR models as mentioned above. In the particular case in which only 1 system parameter is used, only PC1 is used, obviously. The feature selection is exhaustive, i.e., all the combinations among the 6 system parameters are compared upon removal of each system parameter. One should note that early removal does not indicate low importance of the removed parameter. The removal sequence is because of both the importance of the system parameters on system performance and their correlation with other system parameters. As discussed above, T_c and T_e are removed as the number of variables is reduced from 6 to 4 because they are strongly correlated with p_c and p_e . With 4 system parameters, both R^2 reaches the maximum. Compared to R^2 of MLR models using all 6 system parameters as invariables, R^2 of the PCA-MLR models is slightly higher. Both R^2 of the PCA-MLR models decrease as the number of the system parameters is reduced from 4 to 1. Fig. 10 shows RMSE of the PCA-MLR models for the net power output and thermal efficiency using different numbers of system parameters. The behavior is similar to that of R^2 . Thus, the key parameter subset can be identified as $(p_e, \eta_p, p_c, \eta_{SSE})$. Fig. 11 shows that the relative deviation of the PCA-MLR net power output model and the absolute deviation of the PCA-MLR thermal efficiency model. It presents that the relative deviation of MLR model for the net power output is between -5% and 5% , while the absolute deviation of MLR model for the thermal efficiency is between -0.15% and 0.15% .

5. Conclusions

This paper has identified the key parameter subset in the ORC system based on an experimental database obtained from a 10 kW scale ORC system. Correlations between the system parameters are investigated, and decoupled using PCA. Prediction models are developed using MLR, BP-ANN, and SVR. Key parameters are identified based on an exhaustive search. The major conclusions are:

- (1) The determined key parameter subset consists of 4 system parameters: $(p_e, \eta_p, p_c, \eta_{SSE})$. The selected key parameter subset includes information from all 4 major components. Further removal of parameters would result in loss of information (and thus reduced accuracy). Including more system parameters would result in over training (and thus reduced accuracy).
- (2) The MLR models are slightly less accurate than the more sophisticated BP-ANN and SVR models. Upon replacing the system parameters with the first 2 PCs of the key parameter subset as invariables, the PCA-MLR model is of simple structure and acceptable accuracy. The relative deviation of the net power output is between -5% and 5% . The absolute deviation of the thermal efficiency is between -0.15% and 0.15% .
- (3) The working fluid state parameters present stronger correlation between each other than the equipment operating efficiency parameters. The correlation coefficients of working fluid state parameters are all larger than 0.974, while the correlation coefficients of the equipment operating efficiency parameters with the other system parameters are smaller than 0.6.
- (4) Increasing the evaporating pressure of the ORC system, reducing the condensing pressure, and improving the working fluid pump efficiency and the expander shaft efficiency can improve ORC system performance.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was sponsored by the National Natural Science Foundation of China (Grant Nos. 51906119 and 51776005), the Beijing Natural Science Foundation (Grant No. 3192014), the National Postdoctoral Program for Innovative Talents (Grant No. BX20200178), China Postdoctoral Science Foundation (Grant No. 2020M680548), and supported by State Key Laboratory of Engines, Tianjin University (Grant No. K2020-08). Jian Li thanks the supports of Shuimu Tsinghua Scholar Program (Grant No. 2020SM013).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.enconman.2021.114252>.

References

- [1] Imran M, Pili R, Usman M, Haglund F. Dynamic modeling and control strategies of organic Rankine cycle systems: methods and challenges. *Appl Energy* 2020;276: 115537.
- [2] Zhai HX, Shi L, An QS, Wang SL, An BL. Key parameter influence mechanism and optimal working fluid screening correlation for trans-critical organic Rankine cycle with open type heat sources. *Energy* 2020;214:119084.
- [3] Ahmadi A, El Haj Assad M, Jamali DH, Kumar R, Li ZX, Salameh T, et al. Applications of geothermal organic Rankine Cycle for electricity production. *J Cleaner Prod* 2020;274:122950.

- [4] Gogoi TK, Hazarika P. Comparative assessment of four novel solar based triple effect absorption refrigeration systems integrated with organic Rankine and Kalina cycles. *Energy Convers Manage* 2020;226:113561.
- [5] Oyekale J, Petrollese M, Heberle F, Brüggemann D, Cau G. Exergetic and integrated exergoeconomic assessments of a hybrid solar-biomass organic Rankine cycle cogeneration plant. *Energy Convers Manage* 2020;215:112905.
- [6] Chatzopoulou MA, Markides CN. Thermodynamic optimisation of a high-electrical efficiency integrated internal combustion engine – Organic Rankine cycle combined heat and power system. *Appl Energy* 2018;226:1229–51.
- [7] Baldasso E, Mondejar ME, Andreasen JG, Rønnenfelt KAT, Nielsen BØ, Haglind F. Design of organic Rankine cycle power systems for maritime applications accounting for engine backpressure effects. *Appl Therm Eng* 2020;178:115527.
- [8] Shi L, Shu G, Tian H, Deng S. A review of modified Organic Rankine cycles (ORCs) for internal combustion engine waste heat recovery (ICE-WHR). *Renew Sustain Energy Rev* 2018;92:95–110.
- [9] Loni R, Najafi G, Bellos E, Rajaei F, Said Z, Mazlan M. A review of industrial waste heat recovery system for power generation with Organic Rankine Cycle: recent challenges and future outlook. *J Cleaner Prod* 2021;287:125070.
- [10] Anastasovski A, Rasković P, Guzović Z. A review of heat integration approaches for organic rankine cycle with waste heat in production processes. *Energy Convers Manage* 2020;221:113175.
- [11] Yang F, Yang F, Chu Q, Liu Q, Yang Z, Duan Y. Thermodynamic performance limits of the organic Rankine cycle: Working fluid parameterization based on corresponding states modeling. *Energy Convers Manage* 2020;217:113011.
- [12] Unamba CK, White M, Sapin P, Freeman J, Lecompte S, Oyewunmi OA, et al. Experimental investigation the operating point of a 1-kW ORC system. *Energy Procedia* 2017;129:875–82.
- [13] Dicks R, Dumont O, Lemort V. Experimental assessment of the fluid charge distribution in an organic Rankine cycle (ORC) power system. *Appl Therm Eng* 2020;179:115689.
- [14] Sarkar J. Generalized pinch point design method of subcritical-supercritical organic Rankine cycle for maximum heat recovery. *Energy* 2018;143:141–50.
- [15] Imran M, Haglind F, Lemort V, Meroni A. Optimization of organic rankine cycle power systems for waste heat recovery on heavy-duty vehicles considering the performance, cost, mass and volume of the system. *Energy* 2019;180:229–41.
- [16] Branchini L, De PA, Peretto A. Systematic comparison of ORC configurations by means of comprehensive performance indexes. *Appl Therm Eng* 2013;61:129–40.
- [17] Meng F, Zhang H, Yang F, Hou X, Lei B, Zhang L, et al. Study of efficiency of a multistage centrifugal pump used in engine waste heat recovery application. *Appl Therm Eng* 2017;110:779–86.
- [18] Xi H, Zhang H, He Y-L, Huang Z. Sensitivity analysis of operation parameters on the system performance of organic rankine cycle system using orthogonal experiment. *Energy* 2019;172:435–42.
- [19] Bademlioglu AH, Canbolat AS, Kaynakli O. Multi-objective optimization of parameters affecting Organic Rankine Cycle performance characteristics with Taguchi-Grey Relational Analysis. *Renew Sustain Energy Rev* 2020;117:109483.
- [20] Bademlioglu AH, Canbolat AS, Yamankaradeniz N, Kaynakli O. Investigation of parameters affecting Organic Rankine Cycle efficiency by using Taguchi and ANOVA methods. *Appl Therm Eng* 2018;145:221–8.
- [21] Kalina J, Świerzewski M. Identification of ORC unit operation in biomass- fired cogeneration system. *Renewable Energy* 2019;142:400–14.
- [22] Parhizkar T, Rafieipour E, Parhizkar A. Evaluation and improvement of energy consumption prediction models using principal component analysis based feature reduction. *J Cleaner Prod* 2021;279:123866.
- [23] Kouadri A, Hajji M, Harkat M-F, Abodayeh K, Mansouri M, Nounou H, et al. Hidden Markov model based principal component analysis for intelligent fault diagnosis of wind energy converter systems. *Renewable Energy* 2020;150:598–606.
- [24] Yu W, Zhao F, Yang W. Qualitative analysis of particulate matter emission from diesel engine fueled with Jet A-1 under multivariate combustion boundaries by principal component analysis. *Appl Energy* 2020;269:115068.
- [25] Lemmon EW, Bell IH, Huber ML, McLinden MO. NIST Reference Fluid Thermodynamic and Transport Properties Database (REFPROP); 2018.
- [26] Abdi Hervé, Williams Lynne J. Principal component analysis. *Wiley Interdiscip Rev Comput Stat* 2010;2(4):433–59.
- [27] James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. New York: Springer Science+Business Media New York; 2013.
- [28] Yang Fubin, Cho Heejin, Zhang Hongguang, Zhang Jian, Wu Yuting. Artificial neural network (ANN) based prediction and optimization of an organic Rankine cycle (ORC) for diesel engine waste heat recovery. *Energy Convers Manage* 2018; 164:15–26.
- [29] Rumelhart David E, Hinton Geoffrey E, Williams Ronald J. Learning representations by back-propagating errors. *Nature* 1986;323(6088):533–6.
- [30] Yang Fubin, Zhang Hongguang, Hou Xiaochen, Tian Yaming, Xu Yonghong. Experimental study and artificial neural network based prediction of a free piston expander-linear generator for small scale organic Rankine cycle. *Energy* 2019;175: 630–44.
- [31] Feng Yong-qiang, Liu Yu-Zhuang, Wang Xin, He Zhi-Xia, Hung Tzu-Chen, Wang Qian, et al. Performance prediction and optimization of an organic Rankine cycle (ORC) for waste heat recovery using back propagation neural network. *Energy Convers Manage* 2020;226:113552.
- [32] Luo Xianglong, Wang Yupeng, Liang Junwei, Qi Ji, Su Wen, Yang Zhi, et al. Improved correlations for working fluid properties prediction and their application in performance evaluation of sub-critical Organic Rankine Cycle. *Energy* 2019;174: 122–37.
- [33] Vapnik VN. The nature of statistical learning theory. New York, NY: Springer New York; 2000.
- [34] Chang CC, Lin CJ. LIBSVM: a Library for support vector machines. *ACM Trans Intell Syst Technol* 2011;2(3):27.