Performance Modelling of the Regenerative Blower for Air Supply of a SOFC System with Random Forest and Bayesian Optimization

Jiarong Zhao

North China Electric Power University <https://orcid.org/0000-0002-9946-1637>

Xin Wu (✉ wuxin@ncepu.edu.cn)

Xingyu Xiong

North China Electric Power University

Xiangchen Qian

North China Electric Power University

Research

Keywords: Solid oxide fuel cell, Air supply, Regenerative blower, Performance modelling

Posted Date: July 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-38978/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Air starvation is a serious hazard during the operation of a solid oxide fuel cell (SOFC) system. So, an accurate performance model of the air supply system is needed for the control of air flow rate and the reduction of air starvation. The performance modelling of a regenerative blower, which is used for the air supply of a 5 kW SOFC system during power generation, is proposed based on the methods of random forest (RF) and Bayesian Optimization (BO). The performance model is first built through the RF method. The BO algorithm is then used to adjust the hyper-parameters of the model. The developed model can predict the outlet air flow rate through the operation data of regenerative blower, i.e. the rotation speed of impeller, the inlet and outlet pressure, temperature and humidity, without specific geometrical parameters. A series of operation data obtained from an experimental test rig are adopted to train and validate the presented method. Finally, the prediction performance of the developed algorithm is compared with different machine learning methods, such as decision tree regression, deep neural network. Results show that the proposed method can predict the outlet air flow rate under different working conditions with minimum prediction errors within 3.5% and acceptable computational cost. Therefore, the developed algorithm is promising for performance modelling of the blower, which can be used to predict the outlet flow rate and control the air supply system at SOFC in our next research.

1 Introduction

Due to the increasingly serious greenhouse gas emission, clean coal technology has attracted great attention. Solid oxide fuel cell (SOFC) system, a high efficiency energy conversion system mainly composing of SOFC stacks, air and fuel supply system, water and heat management system, has been considered as a promising technology to combine with coal gasification to reduce the emission of greenhouse gas (Grasham et al. 2019; Choudhury et al. 2013). Coal gasification integrating SOFC will play an important role on coal industry in future (Guo et al. 2007; Krüger 2019; Chen et al. 2015), such as clean coal technology, coal-fired power generation with carbon dioxide capture and multiple cogeneration system. However, oxygen starvation may occur during power generation due to imprecise model and control of blower or compressor in the air supply system, which will lead to the degradation of cell performance, damage of electrode and electrolyte and even the whole system. Therefore, an accuracy and rapid modelling method is essential to control the air supply and SOFC system based on the structure of model predictive control and maintain the reliability, quality and efficiency of power generation of a SOFC system (Mayne 2014).

The behaviors and prevention of oxygen starvation, a serious hazard condition that will significantly influence the efficiency, durability and safety of fuel cell, has been studied and reported by many researchers (Liu et al. 2006; Gerard et al. 2010; Taniguchi et al. 2008; Vahidi et al. 2004). Fuel, such as hydrogen, methane and syngas, is electrochemically combined with oxygen to produce electrical energy and heat in fuel cell with the presence of electrolyte and electrode. The electrode half reaction and overall reaction follow the given flux stoichiometry according to electrochemical equations, which means that the fluxes of reactants should be high enough to maintain the balance of reaction (Yu et al. 2011). While

the fluxes of fuel or oxygen are insufficient to meet the requirement of electrochemical reaction, the performance of fuel cell will decrease sharply. That is vividly called starvation. There are probably many reasons contributing to this phenomenon, such as uneven flux distributions in flow field and between cells due to poor design and assembly of cell and stack, flow channel blockage due to poor water and heat management, low dynamic response of air supply due to poor control of blower or compressor (Liu et al. 2006; Gerard et al. 2010). The control of blower or compressor is a better way to reduce the oxygen starvation except for the optimization of single cell and stacks design. Voltage, current and power are the key parameters of a SOFC system, which are often used as control parameters for sub-systems, such as the flow rates and pressures of air and fuel, temperature of SOFC stacks and water management. However, Gu et al. (2012) found that flow rate oscillation may occur when the output current of fuel cell stacks is used as feedback reference to control the flow rate of the blower. Although either fuel starvation or oxygen starvation could happen in fuel cell, the latter in cathode impact the current density of all the fuel cell while the former in anode only influences the starved region (Liu et al. 2006). Furthermore, oxygen starvation even will accelerate the emphasized the degradation of electrolyte and dissolution of cathode (Gerard et al. 2010; Taniguchi et al. 2008). In summary, the performance modelling of the air supply system can be significantly necessary for the optimal control of a SOFC system.

Modelling of a SOFC system includes the modelling of fuel cell stacks and auxiliary devices, such as the blower or compressor, heat exchanger, heater, humidifier. In the past, when a SOFC system is designed and simulated, the flow rate and pressure of air are referenced from characteristic map or assumed to be ideal dynamic response or modelled simply by a nonlinear polynomial function (Zhang et al. 2015; Hong et al. 2012). However, many types of blowers or compressors can supply air to SOFC stacks, such as scroll, lobe, screw and centrifugal blowers or compressors (Tamminen et al. 2014). The performance characteristics of each kind of blowers or compressors are different, even the same kind are not similar due to manufacturing factors and environmental factors. Although many studies on blower modelling have been done, such as those methods based on theoretical analysis (Yoo et al. 2005), numerical modelling (Badami and Mura 2010) and experimental investigation (Badami and Mura 2012), more research is still needed to improve the models in order to predict the behaviors of the blower in a fuel cell system more accurately.

In recent years, there has been a great achievement in the modelling of air supply system in fuel cell system. Zhang et al. (2017) used a one-dimensional dynamic simulation model of air supply system established in the AMESim and Simulink joint simulation environment to analyse the performance of a fuel cell air supply system. Zhao et al. (2017) presented a semi-physical centrifugal compressor model using state-space equations and interior-point method, which have good agreement with the experimental data. Hering et al. (2018) designed and implemented several approaches for the characteristic map modelling of the blower in a 10 kW SOFC system based on the blower laws of similitude and different regression models. In summary, most of the researches above mentioned focus on the proton exchange membrane fuel cell (PEMFC) system for automotive applications and few literatures of air supply system in SOFC system can be found. Due to errors of theoretical model and variation of operating circumstance, the working characteristics of the centrifugal blower manufactured based on theoretical calculation or

numerical simulation will change when it is integrated into a SOFC system for air supply. On the other hand, it is difficult to obtain the specific geometry parameters of a blower for operators. Nowadays, machine learning based methods driven by a great amount of experimental data have been widely used in industry, such as wake modelling of wind turbine (Ti et al. 2020), characteristic analysis of axial compressor (Wu et al. 2019; Wu and Li 2012), fault diagnosis and modelling of SOFC system (Costamagna et al. 2016; Razbani and Assadi 2014). Thus, the performance model driven by experimental data is a better and necessary method to predict and control the air flow rate supplied to fuel cell stacks more accurately.

In this study, we propose a method to build the performance model of a regenerative blower, which can be used to predict the outlet air flow rate, based on a widely used machine learning method, random forest (RF), and Bayesian Optimization (BO). The rest of the paper is organized as follows. Section 2 describes the performance modelling method through the RF and BO (RF-BO). In Sect. 3, experimental results of the outlet air flow rate prediction are shown under the developed model using the collected data from an experimental test rig, which emulates the air supply system with the regenerative blower at a 5 kW SOFC power generation system. We also compare the prediction performance of the outlet air flow rate among the proposed method, decision tree regression, and deep neural network in Sect. 4. The conclusions are drawn in Sect. 5.

2 Methodology

Fig.1 shows the overall strategy for performance modelling of a regenerative blower using the RF method. Firstly, a series of experimental investigations are conducted to collect the operation data of a regenerative blower. Secondly, the rotation speed of impeller, the inlet and outlet pressure, temperature and humidity are chosen to be the input features of the RF model to predict the outlet air flow rate. By optimizing the hyper-parameters of the model based on the BO algorithm, the outlet flow rate of regenerative blower can be predicted by aggregating the results from a group of decision trees and the importance of the input features can be known.

2.1 RF modelling of blower performance

2.1.1 Principle of RF algorithm

RF is a bagging-based ensemble learning method, which is introduced by Breiman in 2001 (Breiman 2001). It is a combination of many decision trees that are created using the bootstrapping technique from the training data set. Each decision tree contains many decision nodes and leaf nodes and a random subset of the features is selected at each decision node in order to further increase the generalization ability. The structure of the RF algorithm is shown in Fig. 2. Let X represent the input vector containing m features with $X = \{x_1, x_2, \dots, x_m\}$, and Y represents the output scalar. The training set, S_n , containing n observations can be expressed as follows:

$$S_n = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)\}, X \in \mathbb{R}^m, Y \in \mathbb{R}. \quad (1)$$

The basic methodology for generating RF can be described as follows: for each set of the q trees, a new bootstrap sample set of the training set is created and this bootstrap sample set is used to train a decision tree. Then at each node of the decision tree, a set of features is randomly selected and the mean square error (MSE) is computed only on that set of features in order to select the optimal one. At last, the above steps are repeated until the tree and the whole forest are complete. As shown in Fig. 2, with the completion of the construction of every decision tree, the final prediction of the RF model is obtained by the majority vote for classification or the mean response for regression over the predictions of the individual trees. It is proved that the RF model has high prediction accuracy, features a high tolerance for outliers and noise data, and is not prone to overfitting (Breiman 2001). Due to the above advantages, the RF method has been widely applied in nonlinear regression problems, such as characteristics analysis of a heavy-duty gas turbine axial compressor (Wu et al. 2019), capacity estimation of lithium-ion battery (Li et al. 2018) and prediction of moisture content of biomass (Zhang et al. 2019). Performance estimation of regenerative blower exhibits nonlinear variations due to its complicated structure and is affected by various operational and environmental parameters. Therefore, the RF method is adopted to develop accurate models for the prediction of the outlet air flow rate at the regenerative blower in our case.

2.1.2 Feature importance evaluation

Since the inlet and outlet pressure, temperature and humidity are the factors that can influence the performance of a regenerative blower with specific geometry, these environmental variables and rotation speed of impeller are selected to be the input features of the RF model. However, the performance of the model will deteriorate and the computation time will increase in cases that some of the input features are not contributing to the output. In the RF method, the selection of input features can be accomplished by calculating the feature importance. It is measured using data permutation to evaluate the impact of each feature on the overall prediction performance of the model. The decrease of prediction accuracy resulting from randomly permuting the values of a feature is considered as the feature importance. The higher prediction accuracy drops, the more important a feature is, and vice versa. The importance of each feature reveals their relationship with the output. In addition, RF measures the feature importance of each feature by considering the impact of each feature as well as its multivariate interactions with other features. Feature importance can be used to select the most important features in the RF model.

2.2 Bayesian hyper-parameter optimization

Even though RF is a very user-friendly machine learning algorithm, it also has a set of parameters that must be adjusted to obtain optimal performance. BO has emerged as a practical tool for high-quality parameter selection in prediction systems. BO method provides efficient alternatives to the grid or random search of the parameter space (Snoek et al. 2012). In addition, the BO method is very useful for optimizing black-box objective functions that lack an analytical expression, which is difficult to evaluate mathematically. BO method internally maintains a Gaussian process model of the objective function and

uses objective function evaluations to train the model. One innovation in BO method is the use of an acquisition function, which is used to determine the next point to evaluate. The acquisition function can balance sampling at points that have low modelled objective functions and explore areas that have not yet been modelled well. The flowchart of the BO algorithm is shown in Fig. 3. The Gaussian model is iteratively updated until the optimal values of the parameters are found. As suggested by the developer of the algorithm (Snoek et al. 2012), the number of iterations is set to 30. Due to the above features, BO algorithm is applied to optimize the hyper-parameters of the RF model in this study.

3 Experimental Results

3.1 Regenerative blower in a SOFC system

In general, a SOFC system comprises several sub-systems, including air supply system, fuel supply system and water and heat management system, as shown in Fig. 4. Although there are many turbomachines can act as air source, regenerative blowers can seriously be considered due to its low vibrations and noises, high reliability, high compactness and relatively low manufacturing cost (Badami and Mura 2010; Badami and Mura 2012). Regenerative blower is a kind of turbomachinery with a more complicated internal flow, generally composing of an impeller with blades, a casing, an annular flow channel between them and inlet and outlet ports, as shown in Fig. 5. The inlet air is accelerated and pressurized by the centrifugal force when flow through the annular flow channel and the impeller, which is driven by an electric motor.

There are many parameters in the regenerative blower model contributing to the performance, which can be divided into three groups, i.e. rotation speed, geometry parameters and environmental parameters. The geometry parameters are very important when it is designed to meet the corresponding requirement of industry applications. However, an offset will be added to the characteristic map of a blower with specific geometry when environmental factors, such as the pressure, temperature and humidity (Wu and Li 2012), change due to the variations of altitude, season and differences between day and night. Therefore, the rotation speed of impeller, the inlet and outlet pressure, temperature and humidity are considered to be the variables that influence the performance of an off-the-shelf regenerative blower in this study.

3.2 Test rig design

In order to investigate the performance of the regenerative blower in the SOFC system and acquire the data set used for training and validating the RF model, a test rig has been designed and implemented, as shown in Figs. 6 and 7. As above mentioned, the rotation speed of impeller, the inlet and outlet pressure, temperature and humidity are the key parameters that influence the outlet flow rate of an off-the-shelf regenerative blower. Thus, several off-the-shelf sensors are installed on the test rig to obtain the key parameters. The detail of parameters and their sources are listed in Table 1. As shown in Fig. 6, pressure sensors (HH316), and temperature and humidity sensors (TH10S-B) are installed at both the inlet and outlet of blower to obtain the pressure, temperature and humidity of air. The rotation speed of blower is

controlled by a variable frequency driven (VFD) (VFD007M23A) with a range of 0–80 Hz. A vortex flowmeter (Focvor4102-01) is used for the measurement of outlet volume flow rate of the blower. Furthermore, the variation of blower load is simulated by the adjustment of the cross-sectional area at the outlet pipe with a ball valve. The analog and digital outputs of all sensors are transmitted to a data acquisition circuit for the temporary storage using RS-485 bus and can be accessed by the host computer based on the Modbus communication protocol. With the help of the data acquisition toolbox in LabView, all signals from the data acquisition device are stored on the computer and can be displayed in the graphical user interface of the PC.

Table 1
Input features

ID	abbreviation	Feature name	Source
1	TI	Inlet temperature	Temperature and humidity sensor
2	RHI	Inlet humidity	Temperature and humidity sensor
3	TO	Outlet temperature	Temperature and humidity sensor
4	RHO	Outlet humidity	Temperature and humidity sensor
5	PI	Inlet pressure	Pressure sensor
6	PO	Outlet pressure	Pressure sensor
7	RPM	Rotation speed	Variable frequency drive

In order to collect the data which can fully represent the characteristic of the blower for training and validating the RF model, a series of experimental investigations are conducted in the proposed test rig. Thirteen experimental conditions are chosen, from 1200 rpm to 4800 rpm with an interval of 300 rpm. Thirteen groups of experiments are conducted in the designed test rig and the collected data are shown in Fig. 8.

3.3 RF Model training and hyper-parameter optimization

The variable importance is calculated as the first step to build the performance model of the blower with RF model. The number of trees is set as 250, the minimum number of observations per tree leaf is set as 5 and the number of variables to select at random for each decision split is one third of the number of variables, as suggested by the reference (Breiman 2001). During the generation of RF with all collected experimental data, around one-third of the training data are left out during the bootstrap sampling, which is called “out of bag” data. The out of bag data are used to estimate the variable importance. As shown in Fig. 9, it is clear that all the input features are contributing to the performance estimation with different importance. The inlet pressure, inlet humidity, and the rotation speed of the impeller are found as the most important features to estimate the outlet air flow rate. The scores of outlet pressure, humidity and

temperature are very close and own the same importance. So, there are no input variables with zero importance and none of the variables needs to be removed.

During the training and hyper-parameter optimization of the RF model, 13 groups of experimental data are divided. Groups 1, 7 and 13, about 25% of total data (107 samples), are selected to be the validation set and the remaining 10 groups data, about 75% of total data (328 samples), are used as the training set. Five hyper-parameters of the RF model have to be optimized, which are the number of trees (numTR), the minimum number of samples per tree leaf (minLS) and required to split a decision node (minDS), the number of features to select at random for each decision split (numPTS) and the maximum depth of tree (maxDT). Although more trees are better based on the principle of bagging, the error sometimes reaches a minimum before increasing the number of trees in practice (Breiman 2001). So, the numTR is an important parameter to be optimized. The minLS minDS and maxDT represent the complexity of trees in the forest. Deep trees tend to over-fit, but shallow trees tend to underfit. In addition, numPTS is another parameter to improve the prediction performance of the RF model. In the hyper-parameter optimization, numTR is changed from 10 to 1500 and the corresponding optimized values of minLS, minDS and numPTS, maxDT under each numTR are calculated using the BO algorithm. After optimization, the BO algorithm provides the value of optimal numTR, min LS, minDS, numPTS and maxDT, i. e. 350, 1, 2, 7 and 15, respectively. In order to evaluate the prediction performance of the developed method, the statistical tools of Maximum Absolute Relative Error (MARE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are applied:

$$MARE = \max(|(y_i - \hat{y}_i) / y_i|) \quad (2)$$

$$MAE = n^{-1} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

where y_i and $\hat{y}_i$ are the predicted and sampled values of flow rate, respectively. The prediction performance and errors of the training data set after hyper-parameter optimization are shown in Figs. 10 and 11. The predicted results have a good agreement with experimental data and the MARE between the predicted and sampled values illustrated in Fig. 12.

3.4 Validation of the RF-BO method

Three groups of experimental data (groups 1, 7 and 13) are used to further test and validate the performance of the RF-BO model. As shown in Fig. 8, groups 1 and 13 are out of the range of the training data while group 7 is in the middle of the training data set. The comparisons between the predicted values and the experimental values of outlet flow rate are shown in Fig. 13. The predicted MAE, RMSE and MARE of the validation data are shown in Fig. 14. The proposed model can predict outlet air flow rate with good performance, whose MARE is within 3.5%, based on the results between Figs. 13 and 14.

4 Comparisons With Other Machine Learning Methods

In order to verify the prediction performance of the proposed RF-BO method, comparisons are conducted with other machine learning methods. The algorithms of decision tree regression (DTR), deep neural network (DNN) and RF with default values (RF-DEF) are selected. DTR builds one regression tree based on the Classification And Regression Tree (CART) algorithm (Breiman et al. 1984) and this model is selected to show the benefit of ensemble steps in the RF algorithm. DNN is widely applied in soft computing and it can approximate any nonlinear relationships between the input and output variables. In addition, RF-DEF is chosen to verify the performance of the hyper-parameter optimization. During the calculation, the minLS value of the CART model is determined by cross validation, which is set to 5 in the model. The DNN model, solved by a stochastic gradient-based optimizer, has 100 hidden layers activated by the rectified linear unit function (Glorot and Yoshua 2010). The numTR, minLS, minDS and numPTS values of the RF-DEF model are set to the default values, i. e. 100, 1, 2 and 7, respectively. The maxDT will expand until all leaves are pure or all leaves contain samples less than minLS and it is set to 11 after training. The predictions of different models are carried out on a PC (Intel Core i7 2.50 GHz with 12 GB of RAM). The MARE, MAE, RMSE and computation time, which is used for training and prediction of each model, are adopted to evaluate the performance. As shown in Table 2, DNN model has the largest prediction error even a hundred hidden layers have been constructed to fit the relationships between the flow rate and input features. The computation cost is also a significant factor in evaluating the algorithm in a practical perspective. The DNN model owns the longest time consumption due to complicated structure and numerous computations to solved the weight parameter of each neuron. CART model predicts more accurately than the DNN model, but its performance is worse than both RF-DEF and RF-BO models because of the limitation of the generalization ability of a single tree. So, the necessity of hyper-parameter optimization for the RF model is validated after comparing the results between RF-DEF and RF-BO methods. Finally, it shows that the RF-BO method has the minimum prediction errors comparing with other three methods. The computation time of the developed method can also be acceptable, since the control response time of outlet air flow rate is within 2 sec at the experimental test rig in our case.

Table 2
Predicted errors and calculation time of different models on the validation set

Model	MARE (%)	MAE (m ³ /h)	RMSE (m ³ /h)	Training time (s)	Computation time (ms)
CART	15.55	1.11	1.63	0.068	81
DNN	20.86	1.05	1.42	0.64	73
RF-DEF	8.27	0.81	1.01	0.393	199
RF-BO	3.50	0.43	0.54	0.829	427

5 Conclusions

The RF-BO method is capable of building the performance model for the regenerative blower in a 5 kW SOFC power generation system using the dynamic parameters of the blower, including the rotation speed

of impeller, the inlet and outlet pressure, temperature and humidity. The designed and implemented experimental test rig can provide reliable operation data of the blower for the training and validating the RF-BO model. The relationships between the characteristic features and the outlet air flow rate are investigated and the main factors are found through the calculated variable importance. With the help of the BO algorithm, the well-tuned RF model can accurately predict the outlet air flow rate under different rotation speeds of the blower successfully with the maximum absolute relative error (3.5%). Furthermore, the comparisons with other machine learning methods indicate that the proposed algorithm owns the best prediction accuracy at acceptable computation time without any data pre-processing procedures. Therefore, the proposed method is suitable for the prediction of air flow rate at outlet, and promising for the control of air supply system at the SOFC system and the prevention of air starvation.

Declarations

Acknowledgments

This work is supported by the National Key R&D Program of China (2017YFB0601900), Fundamental Research Funds for the Central Universities (2019MS023) from North China Electric Power University.

Declaration of interests

The authors declare that they have no known conflict of interests.

References

1. Grasham O, Dupont V, Camargo-Valero MA, García-Gutiérrez P, Cockerill T (2019) Combined ammonia recovery and solid oxide fuel cell use at wastewater treatment plants for energy and greenhouse gas emission improvements. *Appl Energ* 240:698–708
2. Choudhury A, Chandra H, Arora A (2013) Application of solid oxide fuel cell technology for power generation-A review. *Renew Sust Energ Rev* 20:430–442
3. Guo J, Bai H, Fu G (2007) The role of fuel cells in coal industry in future (in Chinese). *J China Coal Soc* 32(11):1201–1205
4. Krüger M (2019) Process development for integrated coal gasification solid oxide fuel cells hybrid power plants—Investigations on solid oxide fuel cells/gas turbine hybrid power plants run on clean coal gas. *Appl Energ* 250:19–31
5. Chen S, Lior N, Xiang W (2015) Coal gasification integration with solid oxide fuel cell and chemical looping combustion for high-efficiency power generation with inherent CO₂ capture. *Appl Energ* 146:298–312
6. Mayne DQ (2014) Model predictive control: Recent developments and future promise. *Automatica* 50(12):2967–2986

7. Liu Z, Yang L, Mao Z, Zhuge W, Zhang Y, Wang L (2006) Behavior of PEMFC in starvation. *J Power Sources* 157(1):166–176
8. Gerard M, Poirot-Crouvezier JP, Hissel D, Pera MC (2010) Oxygen starvation analysis during air feeding faults in PEMFC. *Int J Hydrogen Energ* 35(22):12295–12307
9. Taniguchi A, Akita T, Yasuda K, Miyazaki Y (2008) Analysis of degradation in PEMFC caused by cell reversal during air starvation. *Int J Hydrogen Energ* 33(9):2323–2329
10. Vahidi A, Stefanopoulou A, Huet P (2004) Model predictive control for starvation prevention in a hybrid fuel cell system. *Proceedings of the 2004 American Control Conference*
11. Yu J, Wang Y, Weng S (2011) Effect of syngas flow rate on the performance of solid oxide fuel cell (in Chinese). *J China Coal Soc* 36(10):1753–1757
12. Gu J, Lu L, Xu L, Ouyang M (2008) Modeling and control of air system for fuel cell system (in Chinese). *Chin J Mech Eng* 44(12):112–117
13. Zhang L, Li X, Jiang J, Li S, Yang J, Li J (2015) Dynamic modeling and analysis of a 5-kW solid oxide fuel cell system from the perspectives of cooperative control of thermal safety and high efficiency. *Int J Hydrogen Energ* 40(1):456–476
14. Hong SK, Dong SK, Yang JB (2012) Experimental and simulated investigation of 1 kW solid oxide fuel cell balance of power system. *J Power Sources* 214:28–32
15. Tamminen J, Viholainen J, Ahonen T, Ahola J, Hammo S, Vakkilainen E (2014) Comparison of model-based flow rate estimation methods in frequency-converter-driven pumps and fans. *Energ Effic* 7(3):493–505
16. Yoo I, Park M, Chung MK (2005) Improved momentum exchange theory for incompressible regenerative turbomachines. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy* 219(7): 567–581
17. Badami M, Mura M (2010) Theoretical model with experimental validation of a regenerative blower for hydrogen recirculation in a PEM fuel cell system. *Energ Convers Manage* 51(3):553–560
18. Badami M, Mura M (2012) Comparison between 3D and 1D simulations of a regenerative blower for fuel cell applications. *Energ Convers Manage* 55:93–100
19. Zhang Y, Bao P, Wan Y, Xu S (2017) Modeling and analysis of air supply system of polymer electrolyte membrane fuel cell system. *Energy Procedia* 142:1053–1058
20. Zhao D, Xu L, Huangfu Y, Dou M, Liu J (2017) Semi-physical modeling and control of a centrifugal compressor for the air feeding of a PEM fuel cell. *Energ Convers Manage* 154:380–386
21. Hering M, Wahl S, Meise R (2018) Characteristic blower map modeling for volume flow rate estimation in solid oxide fuel cell systems. *Flow Meas Instrum* 64:116–125
22. Ti Z, Deng XW, Yang H (2020) Wake modeling of wind turbines using machine learning. *Appl Energ* 257:114025
23. Wu Y, Wu X, Rui X, Tian R, Qiu L, Yan J (2019) Study on characteristics analysis of heavy-duty gas turbine axial compressor under off-design condition (in Chinese). *Gas Turb Technol* 32(1):12–15

24. Wu X, Li Y (2012) Self-learning based centrifugal compressor surge mapping with computationally efficient adaptive asymmetric support vector machine. *J Dyn Syst-T ASME* 134(5):1–10
25. Costamagna P, De Giorgi A, Gotelli A, Magistri L, Moser G, Sciaccaluga E, Trucco A (2016) Fault Diagnosis Strategies for SOFC-Based Power Generation Plants. *Sensors* 16(8):1–17
26. Razbani O, Assadi M (2014) Artificial neural network model of a short stack solid oxide fuel cell based on experimental data. *J Power Sources* 246:581–586
27. Breiman L (2001) Random Forests. *Mach Learn* 45(1):5–32
28. Li Y, Zou C, Berecibar M, Nanini-Maury E, Chan JCW, van den Bossche P, Van Mierlo J, Omar N (2018) Random forest regression for online capacity estimation of lithium-ion batteries. *Appl Energy* 232:197–210
29. Zhang W, Cheng X, Hu Y, Yan Y (2019) Online prediction of biomass moisture content in a fluidized bed dryer using electrostatic sensor arrays and the Random Forest method. *Fuel* 239:437–445
30. Snoek J, Larochelle H, Adams RP (2012) Practical Bayesian optimization of machine learning algorithms. 26th Annual Conference on Neural Information Processing Systems Lake Tahoe, NV, United States
31. Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and Regression Trees. *Biometrics* 40(3):358
32. Glorot X, Yoshua B (2010) Understanding the difficulty of training deep feedforward neural networks. *J Mach Learn Res* 9:249–256

Figures

Figure 1

Overall strategy for the performance modelling of regenerative blower

Figure 2

Structure of the RF algorithm

Figure 3

Flowchart of the BO algorithm

Figure 4

Diagram of the 5 kW SOFC system

Figure 5

Schematic principle of the regenerative blower

Figure 6

Diagram of the regenerative blower performance test rig

Figure 7

Implementation and installation of the regenerative blower performance test rig

Figure 8

Performance map of the blower

Figure 9

Feature importance of input features(the name of each feature isgiven in Table 1)

(a) Group 2

(b) Group 3

(c) Group 4

(d) Group 5

(e) Group 6

(f) Group 8

Figure 10

Prediction performance of the training dataset(groups 2 to 6 and group 8)using RF model after hyper-parameter optimization

Figure 11

Prediction performance of the training dataset (groups 9 to 12) using RF model after hyper-parameter optimization

Figure 12

Predicted errors of the training data set

(a) Group 1

(b) Group 7

(c) Group 13

Figure 13

Prediction performance of the validation dataset (groups 1, 7 and 13) using RF model after hyper-parameter optimization

Figure 14

Predicted MAE, RMSE and MARE of the validation data set (groups 1, 7 and 13)