

Data-driven prediction of temperature variations in an open cathode proton exchange membrane fuel cell stack using Koopman operator

Da Huo, Carrie M. Hall *

Mechanical, Materials, and Aerospace Engineering Department, Illinois Institute of Technology, 10 West 35th Street, Chicago, 60616, IL, United States

HIGHLIGHTS

- Koopman theory is applied to model the proton exchange membrane fuel cell stack.
- A systematic method to train the Koopman-based model has been introduced.
- Various structures of Koopman-based models are compared and analyzed.
- Temperature variations across the fuel cell stack have been studied.
- Physics-based model and Koopman-based model are compared.

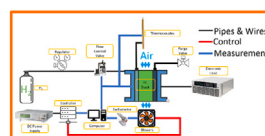
GRAPHICAL ABSTRACT

Data-driven Prediction of Temperature Variations in an Open Cathode Proton Exchange Membrane Fuel Cell Stack Using Koopman Operator

Koopman operator-based modeling offers a promising and efficient approach for accurately predicting the dynamics of open-cathode fuel cell stacks.

Physics-based Modeling

- Enhanced calibration efforts for capturing system internal dynamics.
- Increased modeling complexity for improved accuracy.



Koopman-based Modeling

- Easier to calibrate.
- Accurate prediction of overall and internal dynamics in system modeling.
- Direct application to linear optimal control algorithms.

ARTICLE INFO

Keywords:

Proton exchange membrane fuel cell (PEMFC)
Data-driven modeling
Koopman operator
Dynamic modeling
Control-oriented modeling
Physics-based modeling

ABSTRACT

In this study, a novel application of the Koopman operator for control-oriented modeling of proton exchange membrane fuel cell (PEMFC) stacks is proposed. The primary contributions of this paper are: (1) the design of Koopman-based models for a fuel cell stack, incorporating K-fold cross-validation, varying lifted dimensions, radial basis functions (RBFs), and prediction horizons; and (2) comparison of the performance of Koopman-based approach with a more traditional physics-based model. The results demonstrate the high accuracy of the Koopman-based model in predicting fuel cell stack behavior, with an error of less than 3%. The proposed approach offers several advantages, including enhanced computational efficiency, reduced computational burden, and improved interpretability. This study demonstrates the suitability of the Koopman operator for the modeling and control of PEMFCs and provides valuable insights into a novel control-oriented modeling approach that enables accurate and efficient predictions for fuel cell stacks.

* Corresponding author.

E-mail address: chall9@iit.edu (C.M. Hall).

URL: <https://sites.google.com/iit.edu/advanced-engine-control> (C.M. Hall).

<https://doi.org/10.1016/j.egyai.2023.100289>

Received 29 May 2023; Received in revised form 29 June 2023; Accepted 15 July 2023

Available online 22 July 2023

2666-5468/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The global growth of industrialization and urbanization has driven a surge in energy demand, which has raised significant concerns regarding the environmental effects of fossil fuels. These traditional energy sources are associated with several environmental problems, including air pollution, greenhouse gas (GHG) emissions, and climate change. According to the US Environmental Protection Agency, carbon dioxide (CO₂) emissions resulting from the combustion of petroleum-based products such as gasoline and diesel fuel in internal combustion engines (ICE) constitute a significant proportion of GHG emissions from the transportation sector [1]. In recent years, there has been a notable emphasis on mitigating the environmental impact of conventional fuels. To achieve this goal, research efforts have been directed towards the characterization of alternative fuels [2–6], as well as improved modeling and control. A range of studies have focused on modeling ICEs to better understand their performance and emission characteristics [7–11]. Moreover, optimizing power management strategies has been explored as a means of enhancing performance and reducing pollutants [12–16].

To address the environmental issues associate with fossil fuels, researchers have devoted significant efforts to exploring alternative energy technologies. The hydrogen fuel cell (FC) is considered a promising alternative to traditional power sources due to its high efficiency and zero emissions. The five primary categories of fuel cells are distinguished from each other based on their electrolyte composition and these include [17]: (1) Phosphoric acid fuel cell (PAFC); (2) Polymer electrolyte membrane fuel cell (PEMFC); (3) Alkaline fuel cell (AFC); (4) Molten carbonate fuel cell (MCFC); (5) Solid-oxide fuel cell (SOFC). Despite sharing the same underlying electrochemical principles, the five types of fuel cells operate at varying temperature ranges and utilize different materials. Fuel cells produce solely water as a byproduct of their electrochemical reaction, making them a viable and eco-friendly solution for energy generation. In particular, hydrogen fuel cell stacks have been widely studied as a means of producing electrical power for a variety of applications, from small portable devices to large-scale stationary power generation [18–20]. However, it is challenging to precisely simulate and predict the behavior of the stack under varying operating conditions due to the intricate electrochemical and transport processes that take place within it. Over the years, various modeling methods have been developed to enhance the comprehension and prediction of the performance of FC stacks, and a thorough review of these methods can be found in [21,22]. The study in [23] presented a physics-based dynamic model with a focus on addressing the airflow control issue. In [24], a physics-based model is developed for a 540 kW power output high-speed vehicle to predict the pressure dynamics at high current density conditions.

In recent times, there has been increased interest in the application of data-driven methods for the modeling and control of fuel cell stacks. A comprehensive survey of related literature can be found in [25,26]. In [27], the artificial neural networks (ANNs) approach was utilized as a data-driven method for modeling a 400 W FC stack. In [28], an ANN-PID controller was developed to regulate a commercial PEMFC system. However, since this kind of machine learning approach is non-linear in nature, it requires computationally intensive non-linear optimal control algorithms or the linearization of the model before utilizing linear optimal control algorithms. Additionally, as the complexity of the neural network structure grows, the training cost increases exponentially. Hence, the primary objective of this study is to identify a control-oriented modeling approach that not only provides high accuracy and requires less development effort but can also be efficiently and directly implemented with linear optimal control algorithms. To this end, this study explores the potential of the Koopman operator approach as a data-driven alternative for modeling FC stacks, which has not been comprehensively investigated previously.

Table 1

Fuel cell stack specifications.

Description	Value	Unit
Minimum operating voltage	60	V
Maximum operating current load	90	A
Maximum operating temperature	65	°C
External power supply	24/8–12	V/A
Fuel cell stack weight(wo blowers)	28	kg
Blowers weight	2	kg
Number of cells	120	–
Active area	150	cm ²

The concept of the Koopman operator, which involves representing a nonlinear dynamical system through an infinite-dimensional linear operator, was first introduced by Koopman in the early 1930s [29,30]. Since then, many researchers have studied the Koopman operator, making it a useful tool for analyzing and controlling complex systems [31, 32]. The extended dynamic mode decomposition (EDMD) method was introduced in [33] as a way to approximate the infinite-dimensional Koopman operator to a finite dimension. Since the Koopman-based model is linear, it can be integrated directly into linear model-based optimization algorithms, which are computationally efficient compared to nonlinear optimization algorithms. This attribute makes the Koopman operator a superior choice compared to other data-driven techniques [34].

The main contribution of this research work is the novel application of the Koopman operator for modeling the dynamics of the FC stack system, which has not yet been thoroughly investigated. This study investigates the potential of the Koopman-based approach by performing a comparative analysis with the physics-based model in terms of accuracy, computational time, and development complexity. A systematic method for determining the parameters for the Koopman-based model is presented. The results demonstrate that the Koopman-based model can accurately predict the temperature variation across the FC stack without any additional calibration effort. This study provides valuable insights into the development of a control-oriented modeling approach that is both accurate and efficient for FC stacks.

The paper is structured as follows: in Section 2, the experimental setup and test conditions are introduced. The physics-based model and Koopman-based models are presented in Section 3, along with an analysis of their calibration and training results. In Section 4, the performance of the models is compared and analyzed using a different dataset than the ones used in Section 3. Finally, Section 5 provides concluding remarks and outlines of our future research directions.

2. Experimental configuration

2.1. Experimental setup

In this study, a 5 kW open-cathode proton exchange membrane fuel cell stack, manufactured by Horizon, was utilized. The FC stack consists of 120 fuel cells connected in series, with each cell containing a membrane-electrode assembly sandwiched between bipolar plates. The total active area of the single-cell was 150 cm². Each individual cell is composed of 4 anode channels and 81 cathode channels. The cathode channels serve as the air supply, while the anode channels provide the hydrogen gas supply. In addition, the FC stack is equipped with a short-circuit unit (SCU) that serves as a protective mechanism against overload, low-voltage, and overheating. In the event that any of these hazardous conditions occur, the SCU will trigger a shutdown procedure to prevent potential damage or safety risks. The specifications of the FC stack utilized in this study are presented in Table 1, and the experimental setup layout is illustrated in Fig. 1.

In order to acquire and log the experimental data, various data acquisition devices are employed. The FC stack's built-in module records the FC stack's output voltage, current, and in-stack temperature. In

Table 2
List of laboratory devices used in the experiment.

Device	Model number	Description
Fuel cell stack	Horizon FCS-C5000	5 kW, 120 cells, active area 150 cm ²
Built-in controller	Horizon	Records fuel cell stack voltage, current, and in-stack temperature
Self-made circuit	Arduino Uno	Records the fuel cell stack temperature variations
Thermocouples	USP12397	Four thermocouples affixed on the opposite side of the blowers
Tachometer	8518T96	Measures rotational speed of blowers
Flow control valve	Aalborg DPC47	Records the anode intake pressure of compressed hydrogen
Self-built controller	Arduino	Controls the fuel cell stack blowers
Electronic load	Chroma 63206A-150-600	6 kW capacity, adjustable current load
DC power supply	Chroma 62006P-100-25	Maximum current of 25 A, provides constant voltage

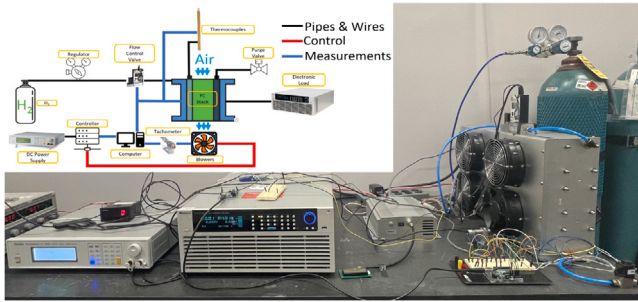


Fig. 1. Experimental setup schematic.

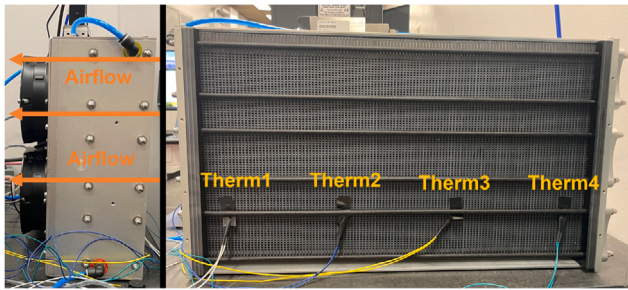


Fig. 2. Thermocouple layout used for temperature measurement of the FC stack.

addition, to investigate temperature variation across the FC stack, four thermocouples (USP12397) are attached on the opposite side of the blowers, and their readings are acquired via a self-made circuit using MATLAB. To remove unused hydrogen, back-diffused water, and nitrogen gas, a built-in purge valve is used with a fixed 10-second purging period, the duration of which varied according to the current load and temperature of the FC stack. The rotational speed of the blowers is measured using a tachometer (8518T96) and can be visually observed through a panel-meter (8518T61) as well. For the purpose of controlling the rotational speed of the blowers, a self-built controller was utilized. In addition, the anode intake pressure of compressed hydrogen is recorded by an Aalborg (DPC47) flow control valve.

The built-in module records data with a sampling time of 0.5 s, while the sampling time for the tachometer, thermocouples, and flow control valve is 1 s. Linear interpolation is employed to standardize the sampling rates and synchronize the data from various devices during the post-processing of data. To provide a constant voltage to the built-in controller and the blowers, a Chroma DC power supply (62006P-100-25), capable of delivering a maximum current of 25 A, was used. To simulate loads on the FC stack, a Chroma electronic load (63206A-150-600), with a capacity of 6 kW, was utilized. The electronic load could adjust the current load in Amperes precisely to 3 decimal places both remotely and locally. A comprehensive list of the lab devices utilized in this study is provided in Table 2. The thermocouples layout is illustrated in Fig. 2, with the left side of the figure being in proximity to the hydrogen inlet.

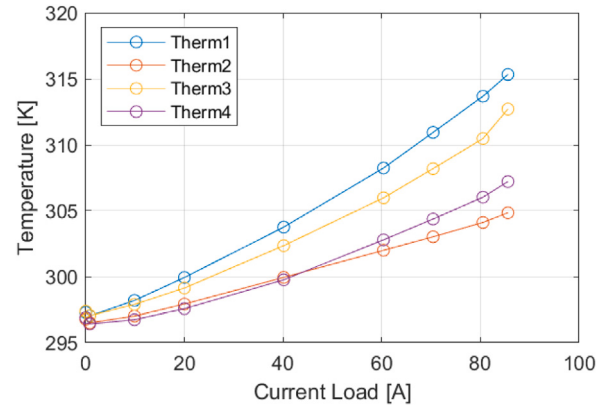


Fig. 3. Temperature variation steady-state results.

Fig. 3 shows the results of the steady-state temperature variations at the inlet side of the air flow channels, carried out under specific testing conditions of a blower maximum speed of 5800 RPM, current load spanning from 0 A to 86 A, and an ambient temperature of 294.15 K. The legend Therm1 to Therm4 corresponds to the left-to-right order of the thermocouples shown in Fig. 2. The temperature trends observed in Fig. 3 indicate a decreasing trend from left (closest to the hydrogen gas inlet) to right, with a maximum temperature variation of approximately 11 K. Notably, the thermocouple located in the middle of the fan, Therm2, exhibit the lowest temperatures due to uneven cooling effects. These observations highlight the importance of considering the cooling effects on the temperature distribution within the FC stack during the design of the FC stack coolant system. One potential solution for the uneven cooling in the FC stack system could be to implement temperature zoning control. This approach would divide the stack into different zones and apply separate cooling strategies to each zone to achieve more even cooling and temperature distribution. Despite this, it is important to note that a considerable temperature variation of up to 8 K still persists between the leftmost and rightmost thermocouples (Therm1 and Therm4). Moreover, it is crucial to highlight that a significant increase in temperature variation is observed with an increase in the current load. This observation highlights the significance of considering the coolant effect during the modeling of FC stacks and the design of the cooling controller.

2.2. Operating conditions

In this study, current load sweeps were conducted after a 30-minute warm-up period to ensure the FC stack is fully self-humidified and stable operation of the equipment. Each test condition was sustained for approximately five minutes to attain a steady-state, and the data from the last minute were averaged to obtain the steady-state values. A polarization test was conducted under the default operating conditions, with the FC stack being supervised by the built-in controller while only the current load was modified. The current load was varied from

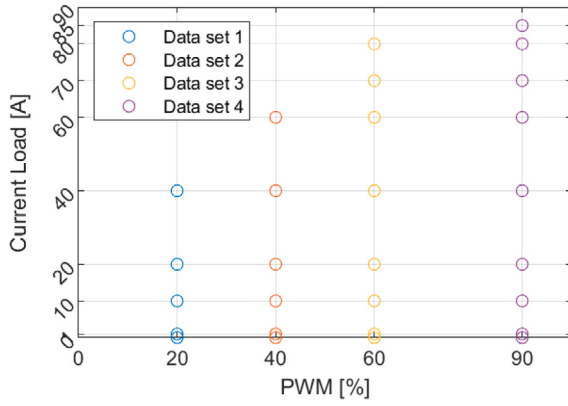


Fig. 4. Operating conditions.

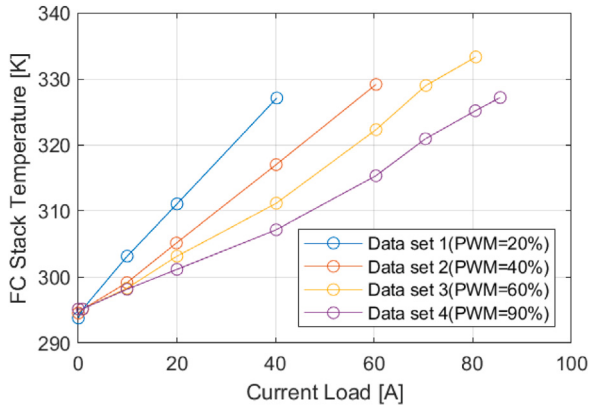


Fig. 5. Experimental data of the FC stack temperature.

0 A to 90 A, with different intervals for each load range. Specifically, smaller intervals were applied for low and high current loads, where the voltage to current gradient is relatively larger due to activation loss and concentration loss, respectively.

The thermal effects on the FC stack were investigated by varying the speed of the blowers through the pulse width modulation (PWM) signal from the self-built controller. The blower's speed was adjusted from a minimum of 1500 rpm to a maximum of 5800 rpm, and four current sweeps were conducted at a constant speed for each test set. The PWM was set to 20% for the minimum blower speed and larger than 90% for the maximum blower speed, resulting in a range of 20% to 90% PWM. The operating conditions are summarized in Fig. 4, and the corresponding stack temperature at steady state versus the current load are demonstrated in Fig. 5. As illustrated in Fig. 4, data sets 1 and 2 have fewer data points than data sets 3 and 4 because the short circuit unit (SCU) protection was activated due to the blower's reduced cooling capacity at lower speeds, causing the FC stack temperature to exceed the upper limit. The polarization test and data sets 1, 2, and 3 are utilized as calibration data sets for the physics-based model, and as training data sets for the Koopman-based model. Data set 4, which covers different PWMs and additional loads not included in the other data sets, is used to test and compare the models.

3. Modeling

In this section, an overview of the development and fundamental principles of the physics-based model and the Koopman-based model are developed, along with a discussion on their calibration results. The calibration of the physics-based model utilized data from the polarization curve test and sweep test data sets 1, 2, and 3. On the other

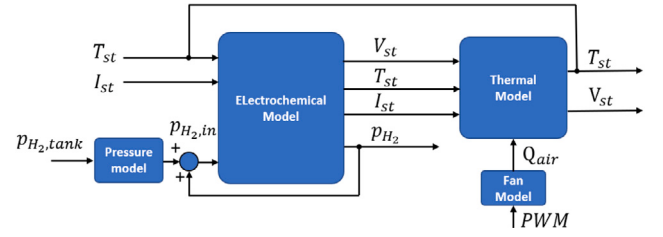


Fig. 6. Physics-based model schematic.

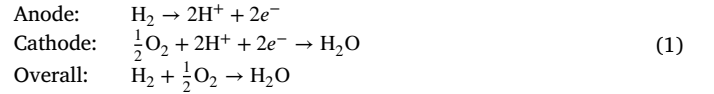
hand, the Koopman-based model was trained using sweep test data sets 1, 2, and 3. Finally, sweep test data set 4 was employed as the test data set for evaluating both the physics-based and Koopman-based models.

3.1. Physics-based model

Three submodels comprise the FC stack physics-based model, namely the electrochemical, thermal, and hydrogen inlet pressure models, as depicted in Fig. 6. The physics-based model takes four input variables, which are FC stack temperature, current load, hydrogen inlet pressure, and PWM signal, and produces two outputs, FC stack temperature and voltage. The electrochemical model was calibrated using the polarization curve test data, while the thermal and hydrogen inlet pressure models were calibrated using data sets 1, 2, and 3. As previously stated in Section 2.2, a 30-minute warm-up test is conducted prior to each experimental test. Consequently, it is assumed in this study that the membrane is fully self-humidified due to this warm-up process.

3.1.1. Electrochemical model

Fuel cells are electrochemical systems that can directly convert chemical energy into electrical energy. The electrochemical reactions occurring within a fuel cell stack can be represented by Eq. (1):



The total voltage of the FC stack, denoted by V_{st} , can be computed using Eq. (2):

$$V_{st} = N_{st} V_{cell} \quad (2)$$

where N_{st} is the total number of fuel cells, V_{cell} is the output voltage of a single cell.

Generally, the single cell output voltage can be expressed as Eq. (3):

$$V_{cell}(t) = V_{oc}(t) - V_{act}(t) - V_{ohm}(t) - V_{con}(t) \quad (3)$$

where V_{oc} is the open circuit voltage, V_{act} , V_{ohm} and V_{con} are the activation voltage loss, ohmic voltage loss, and concentration voltage loss, respectively.

The open circuit cell voltage of an FC stack is the voltage output when there is no external load connected to the system, and the full expression describing how the individual cell open circuit voltage varies with temperature, pressure, and activity under non-standard conditions can be written as Eq. (4) [17]:

$$V_{oc}(t) = E^0 + \frac{\Delta s(T_{st})}{nF} (T_{st}(t) - T_{amb}) - \frac{RT_{st}(t)}{nF} \ln \frac{1}{p_{\text{H}_2}(t) p_{\text{O}_2}^{0.5}} \quad (4)$$

where the E^0 is the non-standard reversible voltage, Δs is the sensible entropy, T_{st} is the fuel cell stack temperature, T_{amb} is the ambient temperature, R is the ideal gas constant, n is the number of moles of electrons transferred from the hydrogen gas and equal to 2, F is

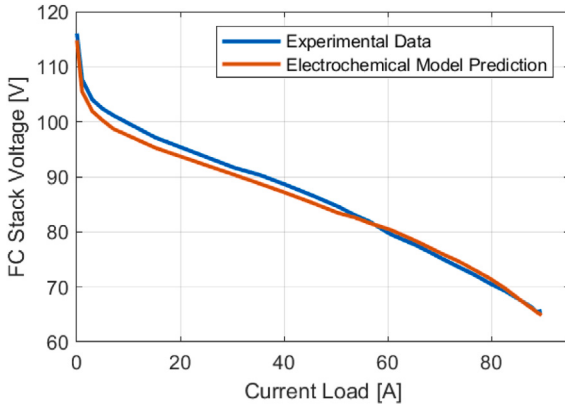


Fig. 7. Polarization curve comparison.

Faraday's constant, p_{H_2} is the hydrogen inlet pressure, and p_{O_2} is the oxygen inlet pressure and is equal to ambient pressure in this study.

The reversible voltage E^0 under non-standard conditions can be obtained by utilizing Eq. (5):

$$E^0 = -\frac{\Delta g(T_{st})}{nF} \quad (5)$$

where Δg is the non-standard-state free-energy change for the reaction [35].

The activation voltage loss can be determined from Eq. (6):

$$V_{act}(t) = \frac{RT_{st}(t)}{\alpha nF} \log\left(\frac{i(t)}{i_0}\right) \quad (6)$$

where i is the current density of the FC stack, i_0 is the exchange current density, the transfer coefficient, denoted as α , is determined by the symmetry of the activation barrier and quantifies the extent to which the electrical potential across the reaction interface affects the sizes of the forward and reverse activation barriers. The value of α falls between 0 and 1, with a value of 0.5 for symmetric reactions. In the case of most electrochemical reactions, α ranges from approximately 0.2 to 0.5 [17]. In this study α has been set to 0.5.

The ohmic voltage loss can be calculated from Eq. (7):

$$V_{ohm}(t) = i(t)R_{ohm} = i(t)(R_{ion} + R_{elec}) \quad (7)$$

where R_{ohm} is the total ohmic resistance, R_{ion} is the ionic charge transport resistance, R_{elec} is the electronic charge transport resistance and equal to 0.0007 Ω .

The value of R_{ion} can be obtained by utilizing Eq. (8) as presented in [36]:

$$R_{ion} = \frac{181.6[1 + 0.03i(t) + 0.062(\frac{T_{st}(t)}{303})^2 i^{2.5}(t)]l}{[\lambda_m - 0.634 - 3i(t)]\exp[4.18(1 - \frac{303}{T_{st}(t)})]A} \quad (8)$$

in which l is the membrane thickness, λ_m denotes the water content of the membrane, and A represents the active area.

The concentration voltage loss is given by [37] Eq. (9):

$$V_{conc} = m(\exp[ni(t)] - 1) \quad (9)$$

where the coefficients of m and n are determined through MATLAB curve fitting to the experimental polarization experimental data.

The calibration results for the polarization curve of the electrochemical model are presented in Fig. 7, where the errors are within 3%.

3.1.2. Thermal model

The FC stack thermal dynamics can be determined as Eq. (10) [38]:

$$C_t \frac{dT_{st}(t)}{dt} = P_{tot}(t) - P_{st}(t) - \dot{Q}_{coolant}(t) \quad (10)$$

in which P_{tot} is the total power released by the electrochemical reactions (W), P_{st} represents the electrical power output of the FC stack (W), and $\dot{Q}_{coolant}$ is the heat dissipated to the cooling airflow. C_t is the thermal capacitance (J/K-s) determined by data observation. Since the built-in controller and the blowers are powered by a separate DC power supply rather than consuming power generated by the FC stack in the tests, their power consumption is not considered in this equation. In both the polarization curve test and sweep tests, the initial condition of T_{st} (stack temperature) is set to match the ambient temperature. Specifically, for the polarization curve test, the ambient temperature is 297.15 K, while for the sweep tests, the ambient temperature is 294.15 K.

The total power released by the electrochemical reaction within the FC stack can be calculated from Eq. (11):

$$P_{tot}(t) = \frac{N_{st} I_{st}(t) \Delta h}{nF} \quad (11)$$

in which N_{st} denotes the total number of cells, I_{st} represents the fuel cell stack current load, Δh represents the enthalpy change of the hydrogen gas, n denotes the number of electric charges, F is the Faraday constant.

The FC stack output power is given by Eq. (12):

$$P_{st}(t) = V_{st}(t) I_{st}(t) \quad (12)$$

The variable $\dot{Q}_{coolant}$ is computed by taking into account the air flow passing through the fuel cell stack, and it can be obtained using Eq. (13):

$$\dot{Q}_{coolant}(t) = \eta_{fan} \dot{m}_{air}(t) c_p (T_{st}(t) - T_{amb}) \quad (13)$$

where η_{fan} is the fan efficiency, $\dot{m}_{air}$ is the air mass flow rate, and c_p is the specific heat coefficient of the air.

The air mass flow rate can be calculate from Eq. (14):

$$\dot{m}_{air}(t) = \rho_{air} Q_{air}(t) \quad (14)$$

where the ρ_{air} and Q_{air} are the air density and the volumetric flow rate of the air, respectively.

For arbitrary fan rotational speeds, the volumetric flow rate of air in Eq. (14) can be obtained by utilizing Eq. (15):

$$Q_{air}(t) = \omega(t) \frac{Q_{nom}}{\omega_{nom}} \quad (15)$$

in which ω is the arbitrary fan speed (RPM), Q_{nom} is the nominal volumetric flow rate (m^3/s) of the fan, and ω_{nom} is the nominal blower speed (RPM) at the operating curve, Q_{nom} and ω_{nom} are determined based on the approach suggested by [38]. In an open cathode PEMFC, the pressure drop is mainly determined by the air channel, which can be calculated by taking into account the friction factor, Reynolds number, minor loss coefficients, and hydraulic diameter. Then the intersection of the pressure drop and the fan performance curve under nominal fan speed corresponds to the nominal volumetric flow rate.

A static empirical model is utilized to establish the relationship between the input PWM command and the rotational speed of the blowers:

$$\omega(t) = \begin{cases} 1500 & \text{for PWM} \in [0,20) \\ 60.36u_{fan}(t) + 294.2 & \text{for PWM} \in [20,90) \\ 5800 & \text{for PWM} \in (90,100] \end{cases} \quad (16)$$

where, u_{fan} refers to the PWM control signal.

The model calibration was performed using three different data sets, namely, data sets 1, 2, and 3, while the validation of the thermal model was done using data set 4. The optimization of η_{fan} was carried out using the MATLAB `fminsearch` function, and the resulting value obtained was 46.59%. The steady-state data were utilized to calibrate the thermal model, and the errors were found to be within 3% for all the test conditions, as depicted in Fig. 8. To facilitate the control of the transient response, a first-order assumption was made for all the responses, and C_t was assigned a value of 5209 (J/K-s) based on experimental data.

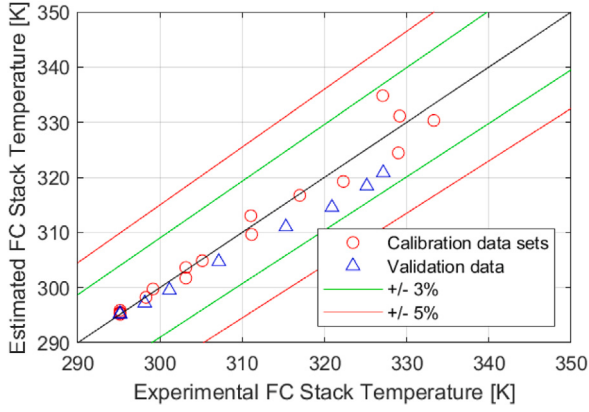


Fig. 8. Fan efficiency calibration results.

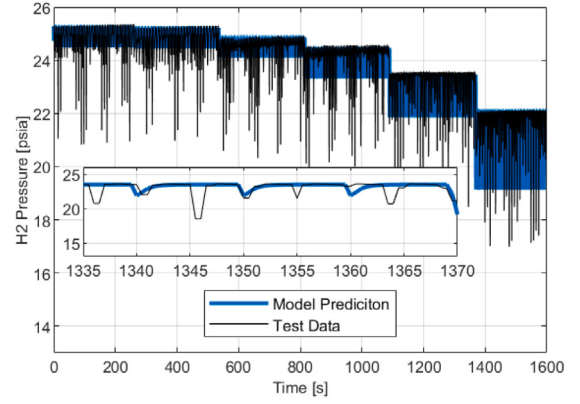


Fig. 9. Data set 2 purging result.

3.1.3. Hydrogen inlet pressure model

In fuel cell operation, the diffusion of water molecules and nitrogen molecules from the cathode air stream to the anode side, due to concentration gradient across the membrane, causes accumulation of water and nitrogen molecules in the anode channels. This accumulation can lead to voltage losses and degradation issues within the fuel cell stack [17]. To prevent this phenomenon, periodic anode purging is performed using a built-in purge valve. The purging process involves removing back-diffused nitrogen gas, water, and unused hydrogen gas from the anode channels, resulting in improved performance of the fuel cell stack. The purging cycle has a fixed period of 10 s, but the duration of purging varies based on a lookup table that considers the current load and temperature of the fuel cell stack. Eq. (17) proposed in [38] can be used to determine the hydrogen gas inlet pressure:

$$p_{H_2, \text{inlet}}(t) = p_{\text{tank}}(t)u(t) - \Delta p_{\text{purge}}u(t - t_1) + \Delta p_{\text{purge}}(1 - e^{-\frac{t-t_2}{\tau_p}}) \quad (17)$$

where p_{tank} is the hydrogen tank outlet pressure, Δp_{purge} and τ_p is the pressure drop during the purging and time constant, respectively. These variables are determined from the data. The time instances t_1 and t_2 refer to the start and the end of the purging process, respectively, and $u(t)$ is the unit step function.

While the hydrogen tank outlet pressure is set as a constant, an increase in the current load results in an increased hydrogen mass flow rate, causing a reduction in the hydrogen tank pressure. Therefore, the hydrogen tank pressure, p_{tank} , is represented as a polynomial function of the stack current load as Eq. (18):

$$p_{\text{tank}}(t) = p_{\text{tank,pre}} + aI_{st}^3 + bI_{st}^2 + cI_{st} + d \quad (18)$$

where a , b , c and d are the constant coefficients, and $p_{\text{tank,pre}}$ is the predetermined hydrogen tank pressure of 25.3 psia.

The hydrogen gas inlet pressure results of data set 2 are illustrated in Fig. 9. The model presented in this study can effectively capture the pressure drops induced by changes in the current load and the status of the purge valve. However, a slight misalignment of the pressure drop is observed due to the limited sensor sampling rate, which cannot accurately capture the transient changes. Furthermore, during each purging cycle, an additional pressure drop is observed due to the consumption of hydrogen gas by the SCU, which is responsible for protecting the FC stack from overload, overvoltage, and overheating. In the event of any hazardous conditions, the SCU initiates a shutdown procedure to prevent potential damage or safety risks.

Table 3 presents the parameters employed in the physics-based model.

Table 3

Parameters of the physics-based model.

Parameters	Value	Unit
Active area, A	150	cm ²
Thickness of the membrane, l	0.0035	cm
Electronic charge transport resistance, R_{elec}	0.00007	Ω
Charge transfer coefficient, α	0.5	–
Water content, λ_m	22	–
Reference enthalpy of H ₂ at 298.15 K, $h_{H_2}^0$	0	J/mol
Reference enthalpy of O ₂ at 298.15 K, $h_{O_2}^0$	0	J/mol
Reference enthalpy of H ₂ O at 298.15 K, $h_{H_2O}^0$	–241 830	J/mol
Reference entropy of H ₂ at 298.15 K, $s_{H_2}^0$	130.68	J/mol
Reference entropy of O ₂ at 298.15 K, $s_{O_2}^0$	205	J/mol
Reference entropy of H ₂ O at 298.15 K, $s_{H_2O}^0$	188.84	J/mol
Thermal capacitance, C_t	5209	J/K-s
Specific heat capacity of air, c_p	1006	J/kg-K
Air density, ρ_{air}	1.1839	kg/m ³
Fan efficiency, η_{fan}	0.4659	–
Fuel cell stack weight	28	kg
Blowers weight	2	kg
Number of cells, N_{st}	120	–

3.2. Koopman-based model

In this section, Koopman-based models were developed using data sets 1, 2, and 3 as the training set. To determine the optimal parameters for the modeling approach, K-fold cross-validation was performed. Five different radial basis functions, including thin plate spline, Gaussian, inverse quadratic, inverse multiquadratic, and polyharmonic spline, were trained with five different lifted dimensions (9, 13, 17, 21, and 25 dimensions) and four different prediction horizons (5, 10, 25, and 50 steps). Each training was performed 30 times, and the minimum weighted root mean square error (RMSE) was selected. The overall best Koopman-based model was chosen by comparing the weighted RMSE and computation time. Additionally, the performance of the selected Koopman-based model will be compared with the physics-based model, and the results will be presented in Section 4.

Additionally, sweep tests were conducted with a wide range of current loads and fan speeds. The current loads varied from 0 A to 85 A, while the fan speeds covered the entire RPM range. The laboratory maintained a stable ambient temperature, and a 30-minute warm-up test preceded the experimental tests. As a result, the training dataset comprehensively covered various boundary conditions. This extensive coverage ensures the model's reliability in making accurate predictions, even for interpolated test conditions, thereby guaranteeing its overall reliability.

3.2.1. Koopman-based modeling approach for PEM fuel cell stacks

Eq. (19) represents an arbitrary discrete-time nonlinear control system:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k) \quad (19)$$

where $\mathbf{x}_k$ is the states of the dynamic system at time step k , $\mathbf{x} \in \mathbb{R}^n$ with n dimensions, $\mathbf{u}_k$ is the control inputs at time step k , $\mathbf{u} \in \mathbb{R}^m$ with m dimensions, and f is the nonlinear transition mapping. The Koopman operator, denoted as $\mathcal{K} : \mathcal{F} \rightarrow \mathcal{F}$, is a linear operator that acts on nonlinear observable functions $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^N$, where $\mathcal{F}$ is a space of observables that are invariant under the action of the Koopman operator. The operator is typically infinite-dimensional and can be expressed as Eq. (20):

$$(\mathcal{K}\psi)(x) = \psi(f(x)) \quad (20)$$

where x denotes the state of the dynamic system, f is the nonlinear transition mapping, and $\psi(f(x))$ represents the observable function of the next state obtained by applying the mapping f to the current state.

The primary objective of this research is to develop a control-oriented dynamic system for the FC stack utilizing the Koopman operator. To this end, an extended dynamic mode decomposition (EDMD) approach, as described in [33], is utilized to approximate the infinite-dimensional Koopman operator into a finite-dimensional representation $\mathcal{K} \in \mathbb{R}^N$, where N is larger than n . The EDMD approach essentially performs regression on a vector of observable functions and generates the lifted state vector. The lifted state $\mathbf{z} \in \mathbb{R}^N$ can be expressed as Eq. (21):

$$\mathbf{z} = \phi(\mathbf{x}) = [\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \dots, \phi_N(\mathbf{x})]^T \quad (21)$$

For the discrete-time, the lifted state space can be expressed as Eq. (22):

$$\begin{aligned} \mathbf{z}_{k+1} &= \mathbf{A}\mathbf{z}_k + \mathbf{B}\mathbf{u}_k \\ \hat{\mathbf{x}}_k &= \mathbf{C}\mathbf{z}_k \end{aligned} \quad (22)$$

where $\hat{\mathbf{x}}$ represents the predicted states, the best-fit linear operators $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are obtained through EDMD, and $\mathbf{A} \in \mathbb{R}^{N \times N}$, $\mathbf{B} \in \mathbb{R}^{N \times m}$, $\mathbf{C} \in \mathbb{R}^{n \times N}$.

To apply the Koopman operator to our FC stack modeling problem, the data of the discrete system has the following inputs to the Koopman operator:

$$\begin{aligned} \chi &= [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k] \\ \chi^+ &= [\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{k+1}] \\ \mathbf{U} &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k] \end{aligned} \quad (23)$$

where the matrices $\chi \in \mathbb{R}^{n \times k}$ and $\chi^+ \in \mathbb{R}^{n \times k}$ are the current state and successor state, respectively, and $\mathbf{U} \in \mathbb{R}^{m \times k}$ is the control inputs.

Similarly, the state matrices in the lifted space are given as:

$$\begin{aligned} \mathbf{Z} &= [\phi(\mathbf{x}_1), \phi(\mathbf{x}_2), \dots, \phi(\mathbf{x}_k)] \\ \mathbf{Z}^+ &= [\phi(\mathbf{x}_2), \phi(\mathbf{x}_3), \dots, \phi(\mathbf{x}_{k+1})] \end{aligned} \quad (24)$$

in which $\mathbf{Z}$ and $\mathbf{Z}^+ \in \mathbb{R}^{N \times k}$.

The matrices of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ in Eq. (24) can be calculated through the best least-squares:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{B}} \|\mathbf{Z}^+ - \mathbf{A}\mathbf{Z} - \mathbf{B}\mathbf{U}\|_F \\ \min_{\mathbf{C}} \|\chi - \mathbf{C}\mathbf{Z}\|_F \end{aligned} \quad (25)$$

where $\|\cdot\|_F$ represents the Frobenius norm of a matrix, and the analytical solutions of Eq. (24) can be expressed as:

$$\begin{aligned} [\mathbf{A}, \mathbf{B}] &= \mathbf{Z}^+[\mathbf{Z}, \mathbf{U}]^\dagger \\ \mathbf{C} &= \chi\mathbf{Z}^\dagger \end{aligned} \quad (26)$$

where $\dagger$ denotes the Moore–Penrose pseudoinverse of a matrix.

In addition to investigating the effects of voltage, temperature, and hydrogen inlet pressure on the FC stack, the variation of FC stack temperature has also been studied using the Koopman operator approach.

To achieve this, the state inputs, control inputs, and outputs are defined as Eq. (27):

$$\begin{aligned} \mathbf{x} &= [\text{Temperature, Voltage, H}_2 \text{ Pressure, Therm1, Therm2,} \\ &\quad \text{Therm3, Therm4}]^T \\ \mathbf{u} &= [\text{Load, PWM, Purge Valve, Purging Time,} \\ &\quad \text{Current-step Purging Time, Current-step Purging Phase}]^T \\ \hat{\mathbf{x}} &= [\text{Temperature, Voltage, H}_2 \text{ Pressure, Therm1, Therm2,} \\ &\quad \text{Therm3, Therm4}]^T \end{aligned} \quad (27)$$

where temperature, voltage, and H_2 pressure of the state inputs and outputs denote the FC stack temperature, FC stack voltage, and FC stack hydrogen inlet pressure, respectively, and the Therm1 to Therm4 state inputs and outputs refer to the thermocouples that are connected to the FC stack in the order from right to left shown in Fig. 2. The purge valve denotes the purge valve status (On/Off), the purging time is obtained from a lookup table based on the FC stack's operating conditions, the current step purging time can vary from 0 to 0.5, and the purging phase represents the time elapsed in which purging phase within the current step since the purging time may be longer than a sample time.

As such, the lifted states of Eq. (24) can be summarized as:

$$\mathbf{z} = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, \phi_8(\mathbf{x}), \dots, \phi_N(\mathbf{x})]^T \quad (28)$$

where $\phi_N(\mathbf{x})$ represents the radial basis function with $N \geq 8$.

There are 5 radial basis functions have been investigated including thin plate spline, Gaussian, inverse quadratic, inverse multiquadratic, and polyharmonic spline. They are expressed as:

$$\begin{aligned} \phi(r)_{tps} &= r^2 \ln(r) \\ \phi(r)_{ga} &= e^{-(\gamma r)^2} \\ \phi(r)_{iq} &= \frac{1}{1+(\gamma r)^2} \\ \phi(r)_{imq} &= \frac{1}{\sqrt{1+(\gamma r)^2}} \\ \phi(r)_{phs} &= r^k \ln(r) \end{aligned} \quad (29)$$

in which the subscripts tps , ga , iq , imq , and phs correspond to the radial basis function of thin plate spline, gaussian, inverse quadratic, inverse multiquadratic, and polyharmonic spline, respectively. The parameter r denotes the distance between the data point and the randomly generated centers, and the parameters γ and k represent the inverse kernel width and the degree of polyharmonic spline, respectively.

3.2.2. Koopman-based model training results

To avoid the Koopman-based model from overfitting, the K-fold cross-validation approach was used. The training data set will be divided into K subsets or folds, training the model on K-1 folds, and validating it on the remaining fold. This process is repeated K times, with each fold serving as the validation set once. In order to mitigate the effects of the random initialization of centers, the training procedure was repeated 30 times for each RBF. The resulting weighted RMSE of the trained model over the K folds was averaged, and the model with the minimum weighted RMSE was chosen as the optimal one. The optimal parameters for each RBF model were determined in this process. Furthermore, each RBF was trained with different dimensions, namely, 9, 13, 17, 21, and 25 dimensions, and prediction horizons of 5, 10, 25, and 50 steps.

In this section, the focus is on the FC stack temperature variations. Therefore, the weighted root mean square error (RMSE) is defined by Eq. (30), which calculates the error between the predicted and experimental data:

$$RMSE_{wei} = 0.16(\hat{e}_1 + \hat{e}_4 + \hat{e}_5 + \hat{e}_6 + \hat{e}_7) + 0.1(\hat{e}_2 + \hat{e}_3) \quad (30)$$

where the notation $\hat{e}$ is used to represent the RMSE between the predicted and experimental data, while the subscript number denotes the same variables as the state inputs. Tables 4, 5, 6, and 7 present the training results of the Koopman-based models with prediction horizons of 5-step, 10-step, 25-step, and 50-step, respectively.

Table 4

5-step prediction horizon weighted RMSE train results.

Rad. basis func.	9 dim	13 dim	17 dim	21 dim	25 dim
Thin plate spl.	0.0176	0.0196	0.0178	0.0213	0.0214
Gaussian	0.0178	0.0214	0.0216	0.0221	0.0206
Inv. quad.	0.0191	0.0215	0.0212	0.0210	0.0208
Inv. multiquad.	0.0211	0.0199	0.0203	0.0208	0.0217
Polyharm. spl.	0.0202	0.0200	0.0205	0.0212	0.0218

Table 5

10-step prediction horizon weighted RMSE train results.

Rad. basis func.	9 dim	13 dim	17 dim	21 dim	25 dim
Thin plate spl.	0.0225	0.0210	0.0258	0.0261	0.0265
Gaussian	0.0231	0.0279	0.0266	0.0269	0.0260
Inv. quad.	0.0208	0.0244	0.0254	0.0265	0.0266
Inv. multiquad.	0.0227	0.0251	0.0260	0.0262	0.0260
Polyharm. spl.	0.0253	0.0268	0.0264	0.0263	0.0263

Table 6

25-step prediction horizon weighted RMSE train results.

Rad. basis func.	9 dim	13 dim	17 dim	21 dim	25 dim
Thin plate spl.	0.0278	0.0307	0.0330	0.0352	0.0360
Gaussian	0.0271	0.0365	0.0340	0.0359	0.0364
Inv. quad.	0.0297	0.0376	0.0358	0.0353	0.0356
Inv. multiquad.	0.0320	0.0339	0.0359	0.0352	0.0363
Polyharm. spl.	0.0318	0.0342	0.0354	0.0357	0.0352

Table 7

50-step prediction horizon weighted RMSE train results.

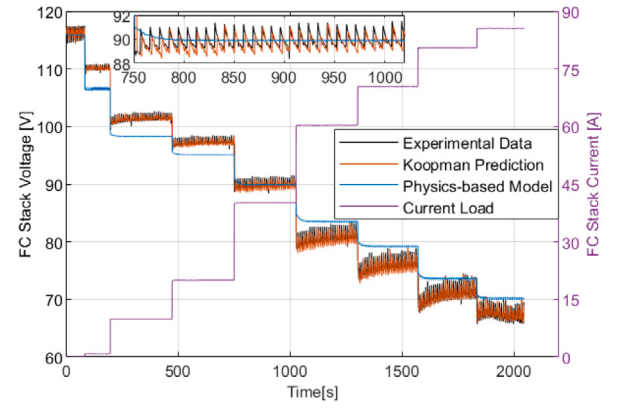
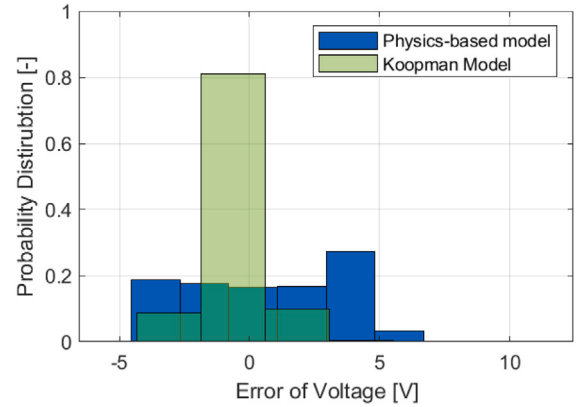
Rad. basis func.	9 dim	13 dim	17 dim	21 dim	25 dim
Thin plate spl.	0.0427	0.0504	0.0465	0.0486	0.0485
Gaussian	0.0414	0.0492	0.0489	0.0483	0.0482
Inv. quad.	0.0428	0.0498	0.0481	0.0481	0.0482
Inv. multiquad.	0.0430	0.0468	0.0480	0.0486	0.0487
Polyharm. spl.	0.0427	0.0463	0.0462	0.0489	0.0488

The training results show that for the FC stack modeling problem, with the same prediction horizon, regardless of the kernel function used, all Koopman-based models exhibit a consistent trend where smaller lifted dimensions result in better performance, which contradicts the initial intuition. This trend suggests that the system dynamics can be captured more accurately with a lower-dimensional feature space. Specifically, the Koopman-based model with the thin plate spline radial basis function and 9 dimensions exhibited the lowest weighted RMSE among all the Koopman-based models. Therefore, this model was selected for comparison with the physics-based model in the following section.

4. Results and analysis

In this section, a comprehensive analysis and comparison of the performance is provided between the physics-based model and the Koopman-based model, denoted as Koopman9 for convenience. The evaluation will be carried out on data set 4, employing a constant blower's Pulse Width Modulation (PWM) of 90%, corresponding to the maximum fan speed of 5800 RPM. The current load ranges from 0 A to 85 A. It is important to note that this specific data set 4 was not utilized during the model calibration or training processes. Our analysis will encompass a variety of aspects such as transient behavior, transient error distributions and steady-state errors for voltage, temperature, and H_2 pressure, as well as computational cost. Moreover, Koopman9's temperature variation results will also be examined by comparing them with the experimental data.

Fig. 10 presents a comparison of the voltage transient results obtained from the selected Koopman-based model (Koopman9) and the physics-based model. The Koopman model exhibits lower errors under

**Fig. 10.** Data set 4 voltage results comparison.**Fig. 11.** Data set 4 voltage transient error distribution results.

all test conditions during steady-state conditions, with both models showing steady-state errors of less than 3%. Furthermore, the Koopman model demonstrates superior performance in capturing transient dynamics compared to the physics-based model. This is due to the limitation of the physics-based model, which is limited to a first-order model and cannot fully capture the non-linear dynamic behavior of the FC stack voltage, especially in high-load regions. Furthermore, the physics-based model exhibits a larger steady-state error due to the assumption of a constant charge transfer coefficient. One possible solution is to set the charge transfer coefficient as a dynamic parameter rather than a constant parameter.

The transient error distributions shown in Fig. 11 reveal that approximately 95% of the transient errors for the Koopman model are less than ± 2 V, indicating a smaller transient error range and more concentrated transient error distribution compared to the physics-based model. The maximum transient error of the physics-based model, which occurred during the transient processes as shown in Fig. 10, is 6.68 V, further highlighting the limitations of the first-order system.

In Figs. 12 and 13, the FC stack temperature prediction results and transient error distributions of the Koopman9 model and the physics-based model are presented, respectively. Similar to the voltage results, the Koopman-based model demonstrates superior transient and steady-state performance compared to the physics-based model, exhibiting a reduced error range and a more concentrated error distribution. The physics-based model underestimates the stack temperature for all test conditions. This is due to the coolant model of the physics-based model being based on a constant efficiency for the blower, which was overestimated based on data sets 1, 2, and 3. Consequently, the model underestimates the stack temperature for data set 4 with 90% PWM, with the maximum error of 6.4 K occurring at a load of 85 A.

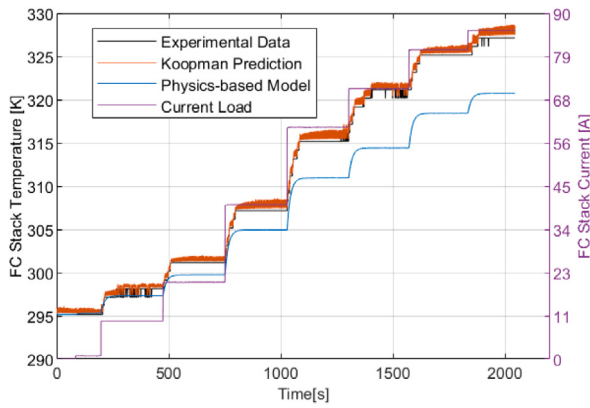


Fig. 12. Data set 4 FC stack temperature result comparison.

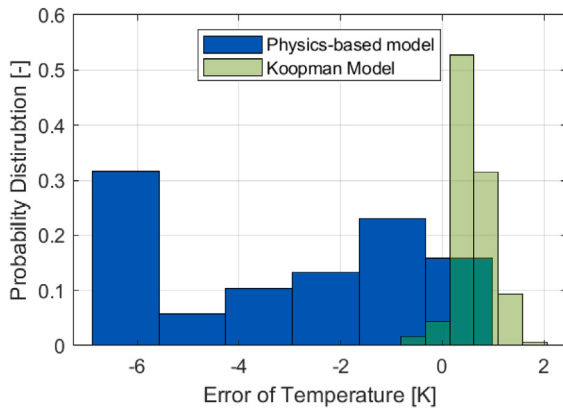
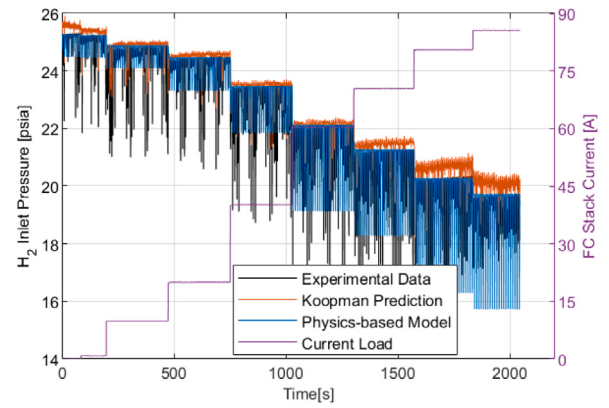
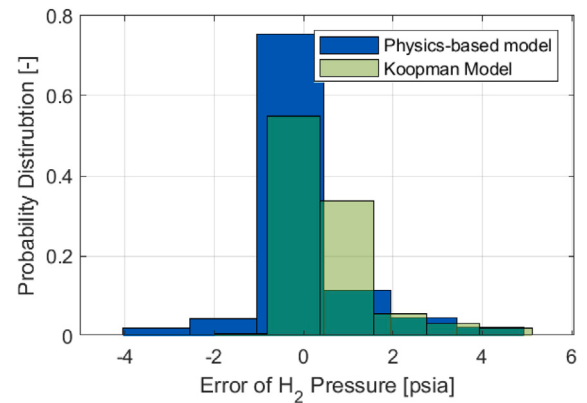


Fig. 13. Data set 4 temperature transient error distribution results.

Fig. 14. Data set 4 H₂ inlet pressure result comparison.Fig. 15. Data set 4 H₂ inlet pressure transient error distribution results.

One feasible approach to enhancing the accuracy of the temperature prediction of the physics-based model is by modeling the blower's efficiency as a function of PWM rather than as a constant value. However, this approach would result in increased model complexity and computational cost. Despite the overestimation of the temperature by the physics-based model, both the Koopman-based model and the physics-based model exhibit errors within 3%.

Figs. 14 and 15 present a comparison between the FC stack hydrogen inlet pressure prediction results and the transient error distributions obtained from the physics-based model and the Koopman9 model. Both models demonstrate a steady-state error within the 3% range, despite the transient overestimation in the Koopman. A majority of the error distribution for both models across all validation data falls within the range of 2 psia. It can be observed that the Koopman model slightly overestimates the hydrogen inlet pressure for all validation conditions. One potential approach to improving the performance of the model is to assign a higher weight to the hydrogen pressure RMSE during the training process. As in the calibration results, a misalignment of the pressure drop is also observed, indicating a need for improvement in future modeling efforts.

In contrast to the physics-based model, the Koopman-based model does not require any additional calibration effort to predict the temperature variations across the fuel cell stack. The transient temperature variation results are presented in Figs. 16, 17, 18, and 19, and are compared with the experimental data. All of the temperature prediction results for the thermocouples show a high level of accuracy, with the predicted values nearly overlapping with the corresponding experimental data. It was also observed that there was a descending trend in temperature from the left side to the right side. This phenomenon may be explained by the fact that the left side of the FC stack is closer to

the hydrogen inlet port, resulting in a relatively higher inlet pressure compared to the right side, due to the pressure drops along the stack. A higher pressure can lead to a higher open circuit voltage, as described by Eq. (4), which is resulted in a higher temperature. The temperature at the middle-left of the FC stack exhibits the lowest values, which may be attributed to uneven cooling distribution. Specifically, the thermocouple is located at the center of the fan, which is subject to a stronger cooling effect compared to other areas, resulting in lower temperatures. To achieve more precise temperature variations in the FC stack, a potential solution is to attach additional thermocouples at different locations on the stack to obtain more comprehensive temperature data. In addition, the transient error distribution results are presented in Figs. 20, 21, 22, 23 which indicating that the majority of the errors are within ± 0.5 K.

In this study, our primary goal is to find a modeling approach that could easily predict temperature variation in the FC stack. The physics-based modeling approach to capture this variation would have required additional calibration effort and resulted in a more complex structure, resulting in a longer computation time. Therefore, this option was not pursued and instead the computational cost of the Koopman-based model and the physics-based model was compared, as presented in Table 8. The Koopman9 model was found to be 48.48% faster than the physics-based model, despite the additional temperature variation predictions.

5. Conclusion and future works

In this study, a novel application of the Koopman operator is presented for control-oriented modeling of a proton exchange membrane fuel cell stack. In contrast to other data-driven approaches such as

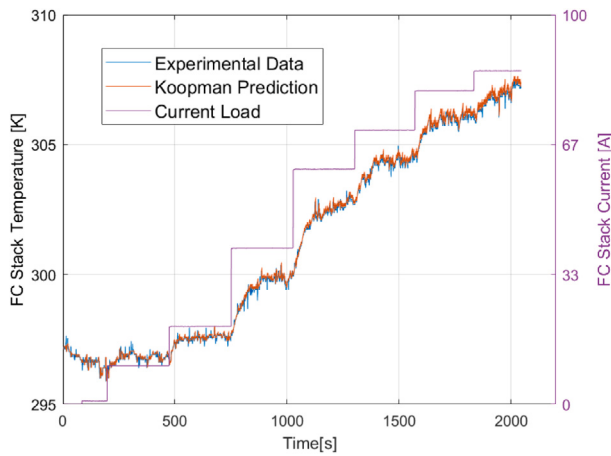


Fig. 16. Comparison of temperature variation transient results for thermocouple on the right side of the FC stack.

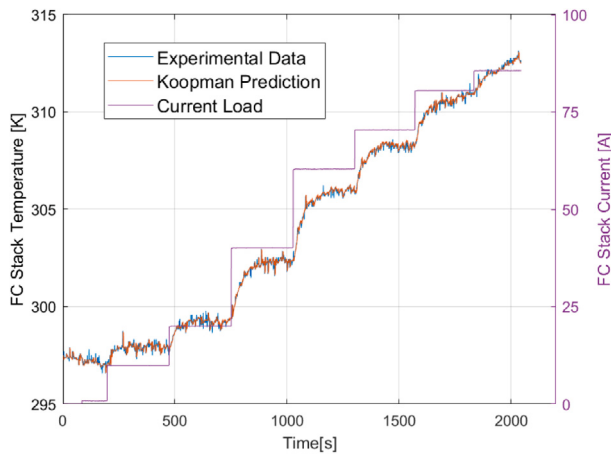


Fig. 17. Comparison of temperature variation transient results for thermocouple on the middle-right side of the FC stack.

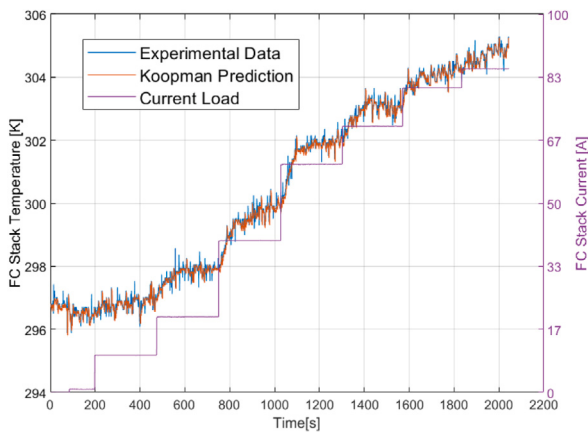


Fig. 18. Comparison of temperature variation transient results for thermocouple on the middle-left side of the FC stack.

artificial neural networks (ANNs), the Koopman operator approach stands out by eliminating the need for additional efforts in linearization and avoiding “black box” behavior that hinders interpretability. It can be directly integrated into linear control algorithms without requiring additional effort. Our study presents a systematic approach

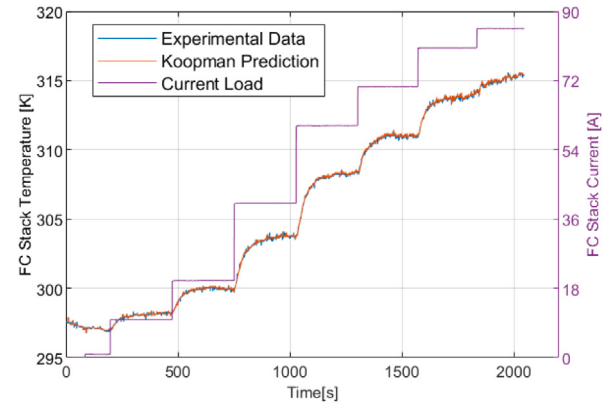


Fig. 19. Comparison of temperature variation transient results for thermocouple on the left side of the FC stack.

Table 8

Data set 4 computation cost comparison.

Model	Cost [s]
Physics-based model	0.2205
Koopman9	0.1136

to determine the Koopman-based modeling parameters, which includes the use of five radial basis functions, different prediction horizons, and K-fold cross-validation with weighted root mean square error to train the Koopman-based models. Each model structure was trained 30 times, and the minimum weighted training root mean square error model (RMSE) and its parameters were retained. The performance of the optimal Koopman-based model was compared to a baseline physics-based approach through the evaluation of the transient performance and error distribution, as well as the computation cost. Furthermore, the performance of the Koopman-based model was analyzed for temperature variation across the fuel cell stack, and results show that the Koopman-based model's accuracy is highly promising. Moreover, after conducting experiments and analyzing the results, it was found that a smaller dimension of the Koopman-based model leads to better performance in predicting the behavior of the fuel cell stack. This observation suggests that a more efficient and computationally feasible Koopman-based modeling strategy could be implemented for the control and optimization of the fuel cell stack. Based on the observations and analysis, it can be concluded that the Koopman operator can be successfully applied to the PEMFC stack modeling problem with high accuracy and computation efficiency. In addition, the Koopman-based modeling approach does not require any additional development effort compared to the physics-based approach, and it has the ability to accurately predict temperature variations. The results demonstrate that the Koopman-based control-oriented modeling approach has promising potential and may not be limited to the application of modeling the fuel cell stack.

Furthermore, based on the temperature variation results, it is suggested that the design of the FC stack coolant system and cooling controller should consider the cooling effects on temperature distribution for an open cathode proton exchange membrane fuel cell stack. Implementing temperature zoning control could be a potential solution for achieving more even cooling and temperature distribution within the FC stack for the uneven cooling situation.

Future research will explore the use of different kernel functions in the Koopman-based control-oriented modeling approach. Additionally, the impact of temperature and humidity variations within the FC stack will be investigated and the Koopman-based model will be utilized for FC stack control and efficiency optimization.

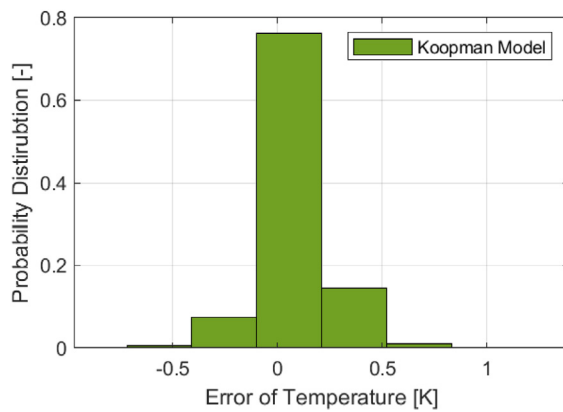


Fig. 20. Error distribution of temperature variation transient results for thermocouple on the right side of the FC stack.

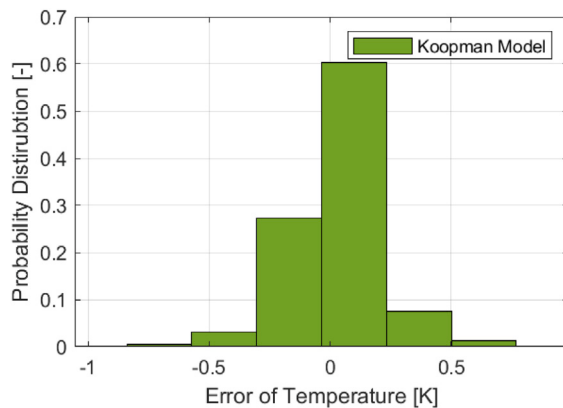


Fig. 21. Error distribution of temperature variation transient results for thermocouple on the middle-right side of the FC stack.

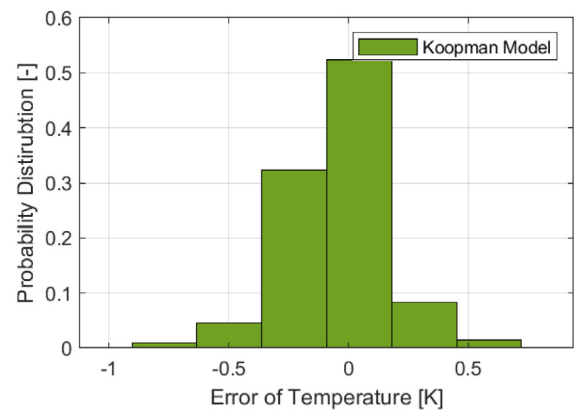


Fig. 22. Error distribution of temperature variation transient results for thermocouple on the middle-left side of the FC stack.

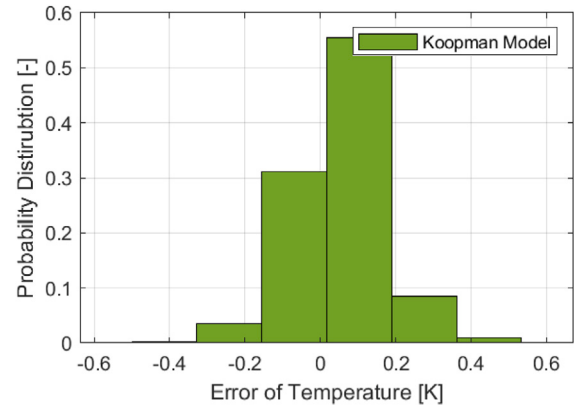


Fig. 23. Error distribution of temperature variation transient results for thermocouple on the left side of the FC stack.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Carrie Hall reports financial support was provided by National Science Foundation.

Data availability

Data will be made available on request.

Acknowledgments

This material is based upon work supported by the National Science Foundation, United States under Grant No. 2135735.

Appendix. Temperature variation transient error distribution results

See Figs. 20–23.

References

- [1] EPA. Overview of greenhouse gases. 2023, <https://www.epa.gov/ghgemissions/overview-greenhouse-gases>, last accessed 01 March 2023.
- [2] Peng Q, Rockstroh T, Hall C. The impact of fuel and injection strategy on combustion characteristics, emissions and efficiency in gasoline compression ignition operation. *Fuel* 2022;318:123548.
- [3] Cheng S, Goldsborough SS, Wagnon SW, Whitesides R, McNenly M, Pitz WJ, Lopez-Pintor D, Dec JE. Replicating HCCI-like autoignition behavior: what gasoline surrogate fidelity is needed? *Appl Energy Combust Sci* 2022;12:100091.
- [4] Sforzo B, Moon CY, Peng Q, Ilavsky J, Powell CF. Atomization performance of a simplex spray through X-ray scattering tomography. In: *AIAA SCITECH 2023 forum*. 2023, p. 2047.
- [5] Moon CY, Peng Q, Sforzo B, Kastengren A, Powell CF. X-Ray phase contrast imaging and radiography of pressure-swirl atomizing sprays in a crossflow. In: *AIAA SCITECH 2023 forum*. 2023, p. 0088.
- [6] Zhang S, Lee M, Goldsborough SS, Cheng S. Multi-stage heat release of multi-component fuels: Insights and implications for advanced engine operation. *Fuel* 2023;332:126218.
- [7] Peng Q, Rockstroh T, Hall C, Pamminger M. Zero dimension heat release modeling for gasoline, ethanol, isobutanol and diisobutylene operating in compression ignition with varying injection strategies. *Tech. rep., SAE Technical Paper*; 2023.
- [8] Mohammad A, Rezaei R, Hayduk C, Delebinski T, Shahpouri S, Shahbakhti M. Physical-oriented and machine learning-based emission modeling in a diesel compression ignition engine: Dimensionality reduction and regression. *Int J Engine Res* 2022;14680874211070736.
- [9] Pulpeiro González J, Ankobea-Ansah K, Peng Q, Hall CM. On the integration of physics-based and data-driven models for the prediction of gas exchange processes on a modern diesel engine. *Proc Inst Mech Eng D* 2022;236(5):857–71.
- [10] Schaberg P, Atkinson C. Calibration optimization of a heavy-duty diesel engine with GTL diesel fuel. *Tech. rep., SAE Technical Paper*; 2016.
- [11] Peng Q. Deep learning and model predictive methods for the control of fuel-flexible compression ignition engines (Ph.D. thesis), Illinois Institute of Technology; 2022.
- [12] Norouzi A, Heidarifard H, Borhan H, Shahbakhti M, Koch CR. Integrating machine learning and model predictive control for automotive applications: A review and future directions. *Eng Appl Artif Intell* 2023;120:105878.
- [13] Norouzi A, Shahpouri S, Gordon D, Winkler A, Nuss E, Abel D, Andert J, Shahbakhti M, Koch CR. Deep learning based model predictive control for compression ignition engines. *Control Eng Pract* 2022;127:105299.

- [14] Huo D, Meckl P. Power management of a plug-in hybrid electric vehicle using neural networks with comparison to other approaches. *Energies* 2022;15(15):5735.
- [15] Peng Q, Huo D, Hall CM. Neural network-based air handling control for modern diesel engines. *Proc Inst Mech Eng D* 2022;09544070221083367.
- [16] Peng Q, Huo D, Hall CM. A comparison of neural network-based strategies for diesel engine air handling control. In: 2022 American control conference. ACC, IEEE; 2022, p. 3031–7.
- [17] O'hayre R, Cha S-W, Colella W, Prinz FB. Fuel cell fundamentals. John Wiley & Sons; 2016.
- [18] Garche J, Jürissen L. Applications of fuel cell technology: Status and perspectives. *Electrochem Soc Interface* 2015;24(2):39.
- [19] Wilberforce T, Alaswad A, Palumbo A, Dassisti M, Olabi A-G. Advances in stationary and portable fuel cell applications. *Int J Hydrogen Energy* 2016;41(37):16509–22.
- [20] Felseghi R-A, Carcadea E, Raboaca MS, Trufin CN, Filote C. Hydrogen fuel cell technology for the sustainable future of stationary applications. *Energies* 2019;12(23):4593.
- [21] Wu H-W. A review of recent development: Transport and performance modeling of PEM fuel cells. *Appl Energy* 2016;165:81–106.
- [22] Yang B, Wang J, Zhang M, Shu H, Yu T, Zhang X, Yao W, Sun L. A state-of-the-art survey of solid oxide fuel cell parameter identification: Modelling, methodology, and perspectives. *Energy Convers Manage* 2020;213:112856.
- [23] Pukrushpan J, Stefanopoulou A, Peng H. Modeling and control for PEM fuel cell stack system. In: Proceedings of the 2002 American control conference (IEEE Cat. No.CH37301). 4, 2002, p. 3117–22 vol.4. <http://dx.doi.org/10.1109/ACC.2002.1025268>.
- [24] Hillstrom ET, Canova M, Guezennec Y, Rizzoni G. Modeling the cathode pressure dynamics in the buckeye bullet II 540 kw hydrogen PEM fuel cell system. *J Power Sour* 2013;241:33–45.
- [25] Ding R, Zhang S, Chen Y, Rui Z, Hua K, Wu Y, Li X, Duan X, Wang X, Li J, et al. Application of machine learning in optimizing proton exchange membrane fuel cells: A review. *Energy AI* 2022;100170.
- [26] Zhao J, Li X, Shum C, McPhee J. A review of physics-based and data-driven models for real-time control of polymer electrolyte membrane fuel cells. *Energy AI* 2021;6:100114.
- [27] Barzegari MM, Rahgoshay SM, Mohammadpour L, Toghraie D. Performance prediction and analysis of a dead-end PEMFC stack using data-driven dynamic model. *Energy* 2019;188:116049.
- [28] Derbeli M, Napole C, Barambones O. Machine learning approach for modeling and control of a commercial heliocentris FC50 PEM fuel cell system. *Mathematics* 2021;9(17):2068.
- [29] Koopman BO. Hamiltonian systems and transformation in Hilbert space. *Proc Natl Acad Sci* 1931;17(5):315–8.
- [30] Koopman BO, Neumann Jv. Dynamical systems of continuous spectra. *Proc Natl Acad Sci* 1932;18(3):255–63.
- [31] Mauroy A, Susuki Y, Mezić I. Koopman operator in systems and control. Springer; 2020.
- [32] Budišić M, Mohr R, Mezić I. Applied koopmanism. *Chaos* 2012;22(4):047510.
- [33] Williams MO, Kevrekidis IG, Rowley CW. A data-driven approximation of the koopman operator: Extending dynamic mode decomposition. *J Nonlinear Sci* 2015;25:1307–46.
- [34] Korda M, Mezić I. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *Automatica* 2018;93:149–60.
- [35] Kyle BG. Chemical and process thermodynamics. 1984.
- [36] Wang Y-X, Yu D-H, Chen S-A, Kim Y-B. Robust DC/DC converter control for polymer electrolyte membrane fuel cell application. *J Power Sources* 2014;261:292–305.
- [37] Hao D, Shen J, Hou Y, Zhou Y, Wang H. An improved empirical fuel cell polarization curve model based on review analysis. *Int J Chem Eng* 2016;2016.
- [38] Ishaku J, Lotfi N, Zomorodi H, Landers RG. Control-oriented modeling for open-cathode fuel cell systems. In: 2014 American control conference. IEEE; 2014, p. 268–73.