International Journal
Information Technology Applications (ITA)

Volume 14, Number 2, December 2025

EMPIRICAL APPROACHES FOR NONLINEAR MODELING OF HYDROGEN FUEL CELL SYSTEMS

Rastislav Putala, Viktor Ferencey, Peter Mlynár

Abstract:

The paper presents a comprehensive methodology for the modeling and simulation of modern hydrogen fuel cells based on measured real-world data. The measured data are used to estimate optimal parameters of semi-parametric nonlinear mathematical models, which serve as a foundation for the design of control algorithms for air-cooled hydrogen fuel cells. Proton exchange membrane (PEM) hydrogen fuel cells are electrochemical devices with strongly nonlinear behavior, which poses a challenging problem for parameter identification and optimal structure design. The empirical modeling approaches presented in the paper are particularly suitable for capturing such complex dynamics without requiring detailed knowledge of the underlying physical or electrochemical processes. The paper proposes empirical mathematical models such as nonlinear Hammerstein–Wiener models (NLHW) and nonlinear autoregressive models with exogenous input (NARX), based on measured input–output data. The study demonstrates that these models enable rapid model development from measured data; however, their predictive capability may be limited when applied to datasets and operating conditions different from those used during model training. Using the MATLAB-Simulink software environment and the System Identification Toolbox, the paper demonstrates the modeling quality through numerical and graphical results and shows how these empirical models can reproduce the dynamic behavior of air-cooled PEM fuel cell systems. The proposed mathematical models are validated and compared with real measured data from an air-cooled PEM fuel cell system across various data types and datasets. The insights obtained from modeling and control of PEM fuel cells are generalizable and applicable in practice and in the development of new, modern PEM fuel cells. The paper provides an effective methodology for further practical applications in the design of optimal control algorithms and rapid prototyping.

Keywords:

Empirical approach, nonlinear behavior, electrochemical devices, predictive capability, electrochemical processes.

Introduction

A PEM fuel cell represents a progressive energy source applicable across a wide range of industries, transportation, and the energy sector. The empirical and semi empirical models are designed for particular fields and should be changed for every specific operating condition or application. The polarization curve of PEM fuel cell, which is described by 0-dimensional models are empirical and simple models [2]. The main emphasis of empirical/semi-empirical models is on the correlation between input and output. Besides, these models provide calculation for performance of PEM FC in a short time [3]; therefore, the empirical/semi-empirical models are more suitable for problems that focus on controlling. Since fuel cells exhibit strongly nonlinear behavior across their entire operating range, it is desirable to consider the implementation of nonlinear simulation models.

This is an open access article under the CC BY-NC license
https://creativecommons.org/licenses/by-nc/4.0/

Semi-empirical modeling approaches, as presented in previous studies such as *Hydrogen fuel cell system modeling approaches* [1], as well as empirical nonlinear models such as nonlinear Hammerstein–Wiener (NLHW) models and nonlinear autoregressive models with exogenous input (NARX), are commonly used for PEM fuel cell system identification.

In this paper, we focus exclusively on empirical nonlinear models, specifically NLHW and NARX models, and compare their performance using real data measured on a 500 W PEM fuel cell. The paper is organized as follows. Section 1 presents the modeling of the PEM fuel cell using NLHW models. Section 2 introduces NARX models. Section 3 describes various model quality metrics and evaluate the developed models which are compared based on those selected performance metrics.

The input variable *y1* for the nonlinear black-box models will be the current value *I* [A], and the output *u1* will be the voltage value *U* [V], i.e., a SISO (Single-Input Single-Output) model (Fig.1). The training dataset contains 553 samples, while the test dataset consists of 112 samples without short-circuiting (with decreasing current load) and 412 samples with short-circuiting (with increasing current load). The sampling period $Ts = 1$ [s] was selected based on the limitations of the measurement system. The models are valid for the current $I \in \langle 0, 40 \rangle$ [A].

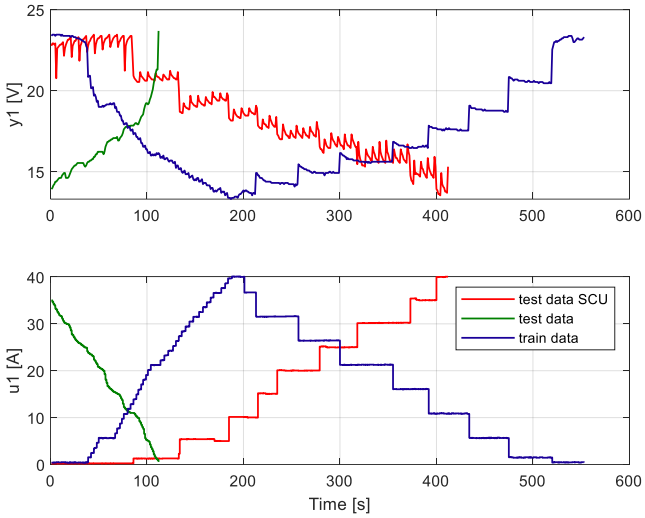


Fig.1. Input and output signals.

1 Hammerstein-Wiener Models

The nonlinear Hammerstein–Wiener model has three blocks: an input nonlinearity, a linear dynamic system, and an output nonlinearity. The linear block is a discrete transfer function and represents the dynamic component of the model as shown in (Fig.2) [4]. The input nonlinearity $f_H: \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_w}$ is called the Hammerstein function, and the output nonlinearity $f_W: \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_y}$ is called the Wiener function [5].

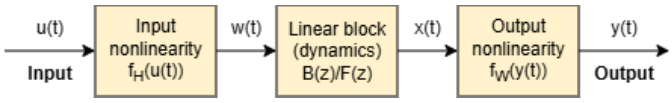


Fig.2. Block scheme of nonlinear Hammerstein-Wiener model.

The discrete Hammerstein–Wiener model can be represented mathematically as:

$$\begin{aligned} w(t) &= f_H\big(u(t)\big) \\ x(t) &= \frac{B(z)}{F(z)} w(t) \\ x(t) &= \frac{b_0 + b_1 z^{-1} + \cdots + b_n z^{-n}}{1 + a_1 z^{-1} + \cdots + a_m z^{-m}} w(t) \\ y(t) &= f_W\big(x(t)\big) \\ x(t_0) &= x_0 \\ \forall t &\in [t_0, t_f] \end{aligned} \tag{1}$$

where:

- u: $[t_0, t_f] \rightarrow \mathbb{R}^{n_u}$ are the system inputs,
- w: $[t_0, t_f] \rightarrow \mathbb{R}^{n_w}$ is the output of the input nonlinearity,
- x: $[t_0, t_f] \rightarrow \mathbb{R}^{n_x}$ is the output of the linear block,
- $B(z)$ and $F(z)$ are the numerator and denominator polynomials of the discrete linear transfer block,
- b_n and a_m are coefficients defining the dynamics of the linear block,
- y: $[t_0, t_f] \rightarrow \mathbb{R}^{n_y}$ are the system outputs after the output nonlinearity,
- $x_0 \in \mathbb{R}^{n_x}$ is the initial state of the linear block.

The input or output nonlinear block can be omitted. It is not necessary for the model structure to include both nonlinearities simultaneously. For nonlinear blocks, an estimator type is selected. Some nonlinear estimators represent simple nonlinear functions, such as a dead zone or saturation limit. Others represent more complex nonlinear functions, for example a sum of nonlinear units in a wavelet or sigmoid network [4].

The linear transfer function block is defined by model order and input delay [4]:

- *nb* – number of zeros plus one represents the length of the numerator *B*,
- *nf* – number of poles determines the order of the denominator *F*,
- *nk* – delay between input and output expressed in the number of samples.

Based on the presented theoretical background, the proposed nonlinear Hammerstein–Wiener models was identified in MATLAB Simulink using the System Identification Toolbox, and its structure is described by the following functions. We developed two models which are marked as *nlhw7* and *nlhw22*. Model identification was performed using experimental data collected from a 500 W PEM fuel cell system. The parameters of the nonlinear model obtained from identification are described in next part.

I/O nonlinearity transfer characteristics of Input nonlinearity block of developed NLHW models are shown in (Fig.3).

The input nonlinearity of *nlhw7* is piecewise linear, parameterized by the positions of 7 breakpoints:

$$\begin{bmatrix} -0.3369 & 4.1493 & 14.9848 & \cdots \\ -2.1303 & 15.8868 & 31.0225 & \\ 22.7978 & 34.9303 & 35.0576 & 45.1817 \\ 31.4648 & 31.8587 & 31.8588 & 32.1162 \end{bmatrix} \tag{2}$$

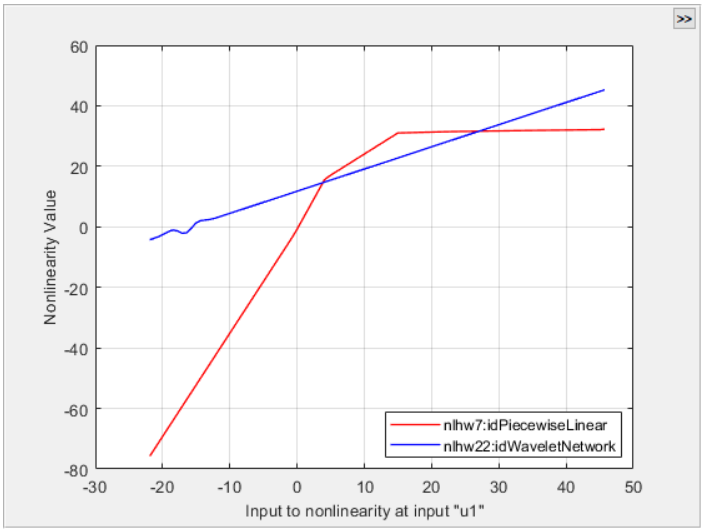


Fig.3. I/O nonlinearity transfer characteristics of Input nonlinearity block.

The input nonlinearity of *nlhw22* is a wavelet network with number of units = 2, and is defined by an equation as the sum of an output offset, a linear function, a sum of dilated and translated wavelets, and a sum of dilated and translated scaling functions [6]:

$$\begin{aligned} y(t) &= y_0 + X(t)^T \cdot P \cdot L \\ &+ \sum_{i=1}^{d_w} w_i \cdot f_w\left(b_i\left(X(t)^T \cdot Q - c_i\right)\right) \\ &+ \sum_{j=1}^{d_s} s_j \cdot f_s\left(d_j\left(X(t)^T \cdot Q - e_j\right)\right) \end{aligned} \tag{3}$$

where:

- $y(t)$ – model output,
- y_0 = 11.7796 – output offset,

linear function:

- $X(t)$ – input (regressor) vector,
- P = [0.0803] – input projection matrix,
- L = [9.1247] – vector of weights,

wavelets and scaling functions:

- Q = [0.0803] – input projection matrix,
- d_w = 1 – number of wavelet units,
- w_i = [-1.7609] – wavelet coefficient,
- b_i = [-10.0416] – wavelet dilation parameter,
- c_i = [-1.3343] – wavelet translation parameter,
- $f_w()$ – wavelet radial function,
- d_s = 1 – number of scaling units,
- s_j = [-1.9284] – scaling coefficient,
- d_j = [1203.5] – scaling dilation parameter,
- e_j = [-0.0623] – scaling translation parameter,
- $f_s()$ – scaling radial function.

Time response of the linear dynamic part (*B(z)/F(z)* parameter) of the both models are shown in figure (Fig.4), with model orders of *nb* = 2, *nf* = 4, and input delay *nk* = 1.

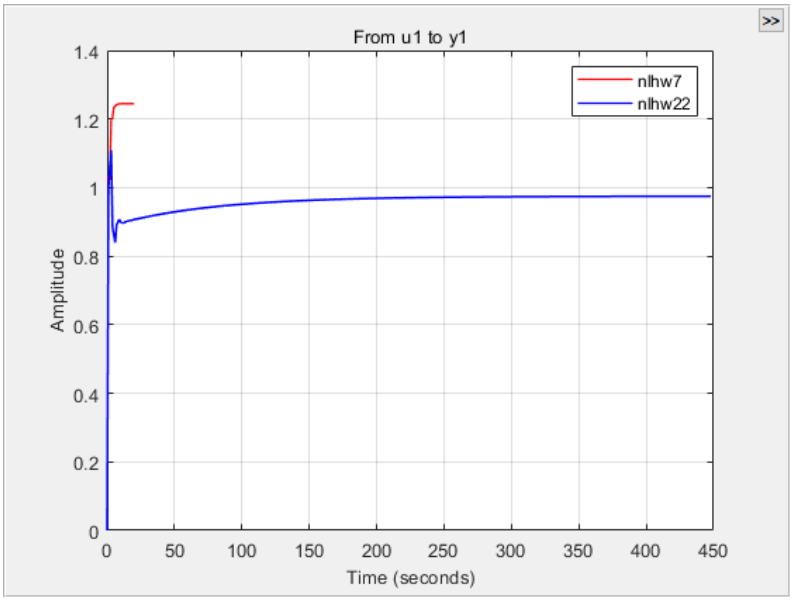


Fig.4. Time response of linear dynamic part of proposed NLHW model.

For *nlhw7* model parameters are:

$$B_{nlhw7} = [0 \quad 1.0342 \quad 1] \tag{5}$$

$$F_{nlhw7} = [1 \quad 0.9785 \quad -0.1566 \quad \cdots \quad -0.1751 \quad -0.0125] \tag{6}$$

For *nlhw22* model parameters are:

$$B_{nlhw22} = [0 \quad 1 \quad -0.9854] \tag{7}$$

$$F_{nlhw22} = [1 \quad -1.0439 \quad 0.0104 \quad \cdots \quad 0.2769 \quad -0.2284] \tag{8}$$

The output nonlinearity of *nlhw7* is piecewise linear, parameterized by the position of 7 breakpoints:

$$\begin{bmatrix} -53.6525 & -31.6267 & -9.6013 & \cdots \\ 27.1531 & 30.3081 & 27.87 & \\ 6.2962 & 38.6007 & 55.4487 & 78.5008 \\ 20.7996 & 17.2478 & -27.5852 & -63.355 \end{bmatrix} \tag{9}$$

The output nonlinearity of *nlhw22* is a wavelet network with number of units = 2 and is defined by next parameters, where offset and linear function parameters are:

- $y_0 = 5.6475$
- $P = [0.0923]$
- $L = [-3.1853]$

and wavelets and scaling functions parameters are:

- $Q = [0.0923]$
- $d_w = 2$
- $w_i = [1.1048 ; 6.791]$
- $b_i = [4.2419 ; 0.4479]$
- $c_i = [-0.0726 ; -2.842]$
- $d_s = 0$
- $s_j = []$
- $d_j = []$
- $e_j = []$

I/O nonlinearity transfer characteristics of Output nonlinearity block of developed NLHW models are shown in figure (Fig.5).

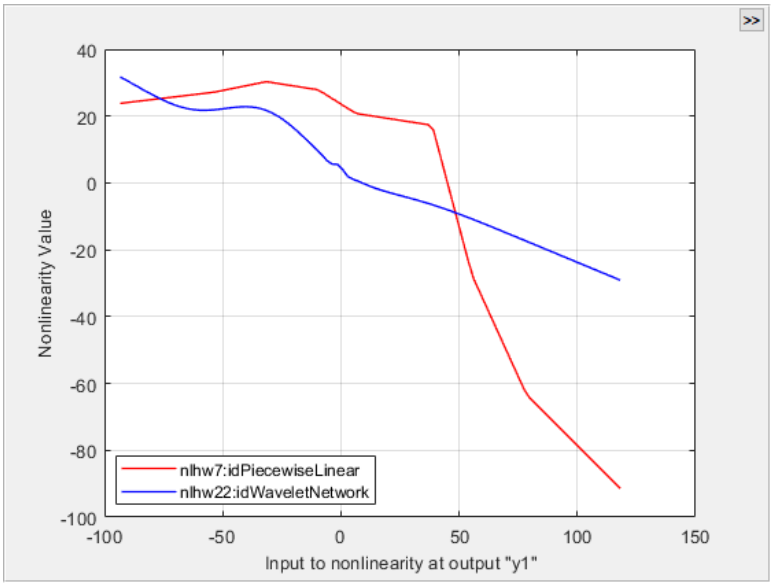


Fig.5. I/O nonlinearity transfer characteristics of Output nonlinearity block.

2 Nonlinear Autoregressive Models with Exogenous Input

Nonlinear autoregressive (ARX) models extend linear ARX models to the nonlinear case. The structure of these models (Fig.6) allows complex nonlinear behavior to be represented using flexible nonlinear functions, such as wavelet and sigmoid networks.

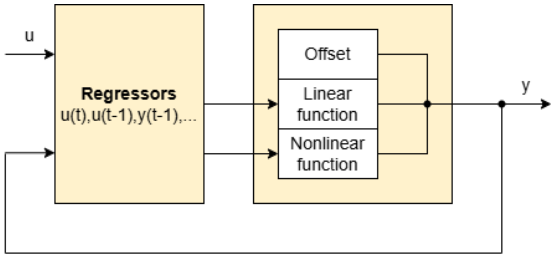


Fig.6. Block scheme of NARX model

A nonlinear ARX model predicts the current output $y_p(t)$ as a function of delayed input and output values [4]:

$$y_p(t) = F(y(t-1), y(t-2), y(t-3), \ldots , u(t), u(t-1), u(t-2), \ldots) \quad (10)$$

where $y_p(t)$ is the predicted output, $y(t-1)$, $y(t-2)$, ... are delayed outputs, and $u(t)$, $u(t-1)$, $u(t-2)$, ... are delayed inputs that serve as model regressors [4].

The function $F()$ represents a nonlinear mapping function that determines how the regressors are combined to produce the output prediction. This may include, for example, a sigmoid function, a weighted sum of wavelets, or other flexible nonlinear functions. In addition to linear lagged variables, the regressors may also include more complex nonlinear expressions such as $y(t-1)^2$, $u(t-1) \cdot y(t-2)$ or $max(u(t-1) \cdot y(t-3), -10)$ [4].

We developed two NARX models which are marked as *nlarx5* and *nlarx12* with linear regressors. For both models regressors are assigned to both linear and nonlinear output functions.

Model *nlarx5* is defined as:

- lags: $y = [1, 2]$, $u = [1, 2]$,
- wavelet network with 5 units,
- linear function: [-2.91, -0.912, -0.276, 0.0786],
- output offset y_0: 0.0795.

Model *nlarx12* is defined as:

- lags: $y = [1, 2, 3, 4, 5]$, $u = [1, 2, 3, 4, 5]$,
- sigmoid network with 10 units,
- linear function: [-2.63, 0.788, -0.165, 0.142, 0.0582, -0.0986, 0.135, -0.0017, -0.00634, 0.0225],
- output offset y_0: 0.168.

Sigmoid network is defined by an equation as the sum of an output offset, a linear function and a sum of dilated and translated sigmoid functions [7]:

$$y(t) = y_0 + X(t)^T \cdot P \cdot L + \sum_{i=1}^{n} s_i \cdot f(X(t)^T \cdot Q \cdot b_i + c_i) \quad (11)$$

where function $f(z)$ is the unit function of the sigmoid network [7], [8]:

$$f(z) = \frac{1}{1+e^{-z}} \quad (12)$$

3 Model Quality Metrics and Model Evaluation

Model quality of PEM fuel cell was evaluated using metrics such as the normalized root mean square error (NRMSE), final prediction error (FPE) and the normalized version of Akaike's Information Criterion (nAIC).

NRMSE [%] measures the agreement between the model and actual data, ranging from $-\infty$ (poor fit) to 100 (perfect fit), with 0 indicating no better fit than the data mean [4].

$$NRMSE = 100\left(1 - \frac{\| U_{real} - U_{model} \|}{\| U_{real} - \bar{U}_{real} \|}\right), \tag{13}$$

where:

- U_{real} [V] is the measured output,
- U_{model} [V] is the simulated response,
- $\bar{U}_{real}$ [V] is the arithmetic mean of the U_{real} values,
- $/\!/\cdot/\!/$ denotes the Euclidean norm of a vector.

The Final Prediction Error according to Akaike (FPE – Final Prediction Error) is defined by the following expression, where $det\left(\frac{1}{N} \cdot E^T \cdot E\right)$ describes the accuracy of the model and $\left(1 + \frac{n_p}{N}\right) / \left(1 - \frac{n_p}{N}\right)$ represents the model complexity [4].

$$FPE = det\left(\frac{1}{N} \cdot E^T \cdot E\right) \frac{1 + \frac{n_p}{N}}{1 - \frac{n_p}{N}}, \tag{14}$$

where:

- N is number of samples in the estimation dataset,
- E is the prediction error matrix,
- n_p is number of free parameters in the model.

AIC evaluates model accuracy while penalizing complexity, with lower values indicating better models where n_y is the number of outputs [4].

$$AIC = N \cdot log\left(det\left(\frac{1}{N} E^T \cdot E\right)\right) \quad \cdots \\ +2 \cdot n_p + N\left(n_y \cdot \log 2\pi + 1\right) \tag{15}$$

The normalized version of Akaike's Information Criterion is defined as follows [4]:

$$nAIC = \log\left(det\left(\frac{1}{N} \cdot E^T \cdot E\right)\right) + \frac{2 \cdot n_p}{N} \tag{16}$$

The models were trained on a dataset containing 553 samples. The fit (NRMSE) of the individual models is shown in (Fig.7) and quality of those models in (Tab.1).

Models were tested on a two test datasets with different dynamics. First test dataset contain 112 samples without active short-circuit unit (SCU). Second test dataset contain 412 samples with active SCU.

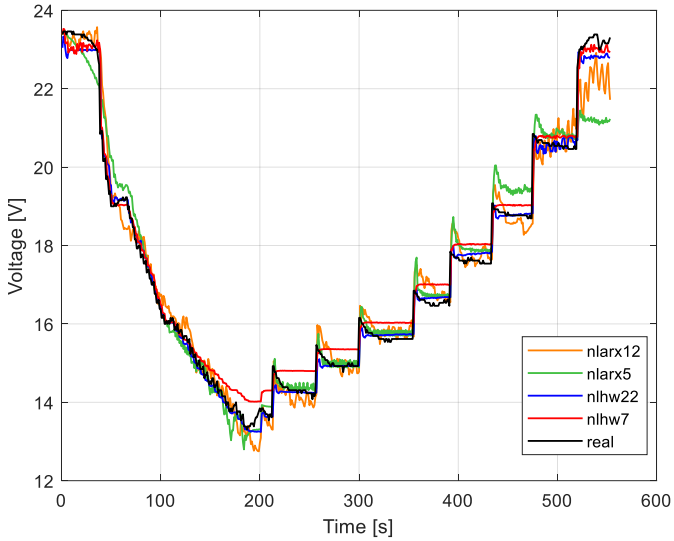


Fig.7. Time responses of developed models on training dataset.

Table 1. Quality of developed models

Model	NRMSE [%]	FPE	nAIC
nlarx12	83.6546	0.1554	-1.8710
nlarx5	78.8664	0.1024	-2.2788
nlhw22	90.3434	0.0905	-2.4028
nlhw7	87.1322	0.1702	-1.7708

In the first test without SCU (Fig.8), the models achieved the following fits to the real data (NRMSE):

- *nlarx12*: 80.3 %,
- *nlarx5*: 75.43 %,
- *nlhw22*: 87.46 %,
- *nlhw7*: 85.84 %.

Intentional short-circuiting prevents oxidation of the platinum catalyst and thus increases the performance and lifetime of a PEM fuel cell, which has a significant impact on its dynamics [9]. Short-circuiting leads to a change in the output voltage response under constant current load, as well as to a characteristic sawtooth-shaped voltage waveform. The typical duration of the short-circuit for the given fuel cell is 100 ms every 10 seconds [10]. In the second test with SCU (Fig.9), the models achieved the following fits to the real data (NRMSE):

- *nlarx12*: 59.73 %,
- *nlarx5*: 55.07 %,
- *nlhw22*: 65.79 %,
- *nlhw7*: 70.99 %.

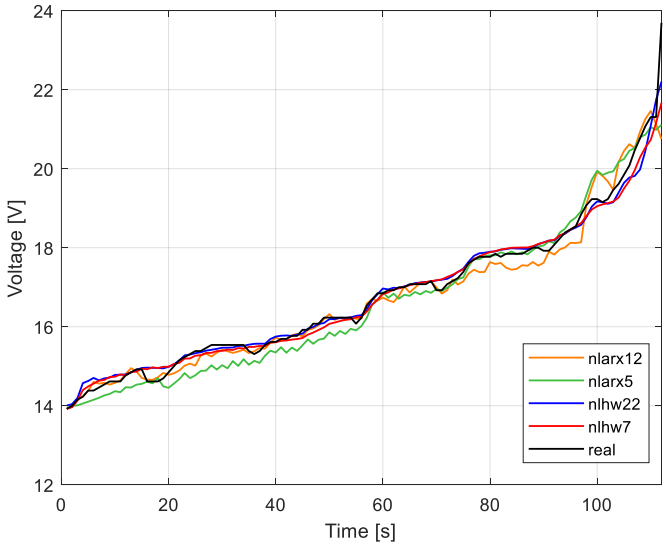


Fig.8. Time responses of measured (test data) and simulated models output.

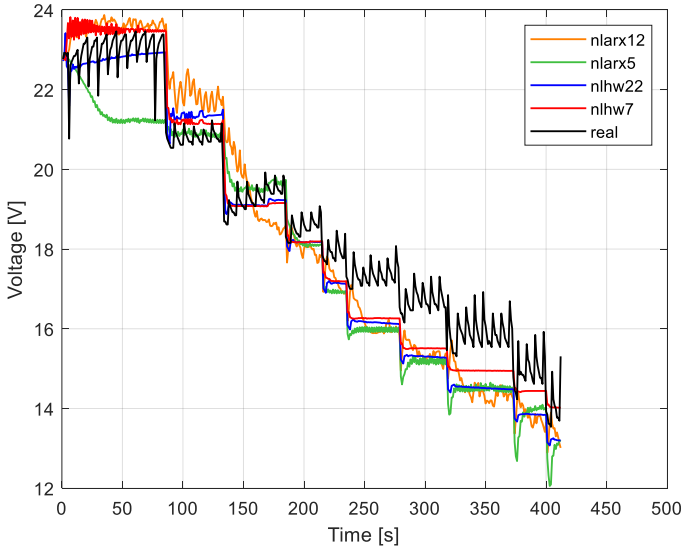


Fig.9. Time responses of measured (test data SCU) and simulated models output.

Conclusion

This paper presented an empirical approach to modeling the dynamic behavior of an air-cooled PEM fuel cell using nonlinear black-box models based on measured data. Two model structures, namely Hammerstein–Wiener (NLHW) and nonlinear autoregressive models with exogenous input (NARX), were developed and systematically compared using multiple quality metrics, including NRMSE, FPE and nAIC.

The results obtained on the training dataset indicate that NLHW models achieved superior accuracy compared to NARX models. Among them, the *nlhw22* model provided the best overall performance across the evaluated quality metrics, with a fit of 90.3434 % and low complexity, as reflected by nAIC = −2.4028.

Validation on independent test datasets revealed that all models maintained good predictive capability under operating conditions similar to the training data. However, a noticeable degradation in performance was observed when the models were applied to data with different dynamics, particularly in the presence of short-circuit operation (SCU), where the *nlhw7* model with 70.99 % fit achieved the best results among the tested models. This highlights a key limitation of empirical models, namely their reduced generalization ability outside the domain of the training dataset.

Despite this limitation, the developed models are sufficiently accurate for applications such as control design, simulation, and rapid prototyping. Their relatively simple structure and ease of identification make them attractive for practical implementation, especially in cases where detailed physical modeling is not feasible.

References

[1] R. Putala and V. Ferencey, "Hydrogen fuel cell system modeling approaches," in *2025 Cybernetics & Informatics (K&I)*, 2025, doi: 10.1109/KI64036.2025.10916424..

[2] Hamdollahi, S., & Jun, L. (2023). "A review on modeling of proton exchange membrane fuel cell," *Chemical Engineering*, 29, 61–74. https://doi.org/10.2298/CICEQ220126014H.

[3] M. Ohenoja, A. Sorsa, and K. Leiviska, "Model Structure Optimization for Fuel Cell Polarization Curves," *Computers*, vol. 7, 2018, doi: 10.3390/computers7040060.

[4] L. Ljung, *System Identification Toolbox™ User's Guide*. Natick, MA, USA: The MathWorks, Inc., 2023.

[5] Kappatou, C.D., Bongartz, D., Najman, J., et al. (2022). *Global dynamic optimization with Hammerstein–Wiener models embedded.* Journal of Global Optimization, 84, 321–347. doi:10.1007/s10898-022-01145-z.

[6] The MathWorks, Inc., "idWaveletNetwork — Wavelet network function for nonlinear ARX and Hammerstein-Wiener models," *MATLAB System Identification Toolbox Documentation*, 2026. Available: https://se.mathworks.com/help/ident/ref/idwaveletnetwork.html. (accessed 27 Apr. 2026).

[7] The MathWorks, Inc., "idSigmoidNetwork — Sigmoid network function for nonlinear ARX and Hammerstein-Wiener models," *MATLAB System Identification Toolbox Documentation*, 2026. Available: https://www.mathworks.com/help/ident/ref/idsigmoidnetwork.html. (accessed 04 May 2026).

[8] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear Black-box Modeling in System Identification: a Unified Overview," *Automatica*, vol. 31, no. 12, pp. 1691–1724, 1995. doi: 10.1016/0005-1098(95)00120-4.

[9] C. Zhang, H. Liu, T. Zeng, J. Chen, P. Lin, B. Deng, F. Liu, and Y. Zheng, "Systematic study of short circuit activation on the performance of PEM fuel cell," International Journal of Hydrogen Energy, vol. 46, no. 45, pp. 23489–23497, 2021. doi: 10.1016/j.ijhydene.2020.10.118.

[10] Horizon Fuel Cell Technologies, "H-500 Fuel Cell Stack User Manual," 2015. Available: https://www.fuelcellstore.com/manuals/horizon-pem-fuel-cell-h-500-manual.pdf. (accessed 05 May 2026).

Authors



Rastislav Putala
Faculty of Mechanical Engineering
Slovak University of Technology, Bratislava, Slovakia
rastislav.putala@stuba.sk



Viktor Ferencey
Hydrogen Fuel Cells Slovakia
Bratislava, Slovakia
v.ferencey@gmail.com
Researcher in the field of hydrogen technologies.



Peter Mlynár
Faculty of Mechanical Engineering
Slovak University of Technology, Bratislava, Slovakia
peter.mlynar@stuba.sk