

Modeling and predicting the membrane water content of proton exchange membrane fuel cell by using support vector regression

Jiang.Ling. Tang, Hong.Yu. He* and Hai.Ying. Liu

Department of Physics and Engineering, Guilin normal college, Guilin, Guangxi 541001, P.R. China

*Tantian62@hotmail.com

Keywords: Modeling; predicting; membrane water content; proton exchange membrane fuel cell; support vector regression.

Abstract. This study examines the use of the support vector regression (SVR) approach in modeling and predicting the membrane water content of Proton Exchange Membrane Fuel Cell (PEMFC) under two influence factors, including the impedance of single-PEM-chip and operating temperature. The leave-one-out cross validation (LOOCV) test results by SVR strongly support that the generalization ability of SVR model is high enough: mean absolute error (*MAE*) is 0.01, mean absolute percentage error (*MAPE*) is 0.15% and multiple correlation coefficients (R^2) is 1.00. This investigation suggests that the SVR approach is a promising and practical methodology to simulate the properties of fuel cell system.

Introduction

Fuel cell, as an electrochemical device, can directly convert the chemical energy to electrical energy and thermal energy without burning. Due to its merits of low operational temperature, rapid startup, little noise, high power density and light weight [1], PEMFC is believed by most experts to be the best candidate for many applications, such as low-emission vehicle, distributed power generation station and portable electrical resources, etc.

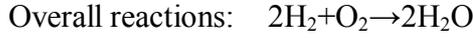
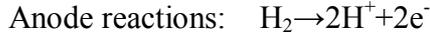
Water content in PEM of fuel cell is one of the vital factors to achieve maximum performance and durability for PEMFC, which is referred to the ratio of the number of water molecules to that of charge ($\text{SO}_3\text{-H}^+$) sites [2]. In PEMFC, the membrane must be sufficiently hydrated because its conductivity depends critically on the humidity level. On the one hand, too little water may cause the membrane drying, which would increase the ionic resistance, and exacerbates the voltage drop. On the other hand, too much water may also cause the membrane flooding by blocking of porous passages, which would lead to reduce the transport rate of reactants to the catalyst site [3]. As a result, the water content directly decides the proper operation of a PEMFC. Therefore, to keep the right amount of water in PEM, which is adequate to hydrate membrane but not disturb the transportation of the reactant gas, is very important for PEMFC to exhibit its excellent performance and stability.

Although there exists difficulty to directly monitor the water content in PEMFC in-situ and limitation of the hardware installation, various strategies to measure or estimate the membrane water content in PEM have been proposed in recent years. For example, Chen et al. [4] employed neutron radiography to observe the water distribution in anode and cathode gas diffusion layers. Mench et al. [5] utilized the gas chromatography to measure the in situ water vapor distribution in the gas channels of an operating PEMFC. Besides the hardware technologies, McKay et al. [6] proposed a lump dynamics model of PEMFC to estimate the humidity of electrode. Thawornkuno and Panjapornpon adopted the extended Luenberger observer to predict the water content in PEMFC [7]. The commercial computational fluid dynamics software (CFD) was also used to simulate the water transport and distribution in PEMFC [8, 9].

Du et al. proposed a method to indirectly measure/monitor the PEM water content based on the impedance of single-PEM-chip and its temperature via neural network [10]. Motivated by Du et al.'s above work, in this paper, we proposed the support vector regression (SVR) approach, and integrating leave-one-out cross validation (LOOCV), to model and deduce the membrane water content based on the experimentally measured impedance of single-PEM-chip and operating temperature.

Methods and materials

Description of the PEMFC system. The typical structure of a single PEMFC consists of anode, electrolyte membrane and cathode, it is shown in Fig.1. [11]. The electrochemical reactions at the anode and cathode in a single PEMFC are below:



On the anode side, hydrogen molecule splits up into two hydrogen protons and two electrons. Electrons through an external circuit and form current, while protons migrate through the electrolyte membrane. On the cathode side, the oxygen reacts with protons coming from anode and electron coming from external circuit, and finally produces with water and heat.

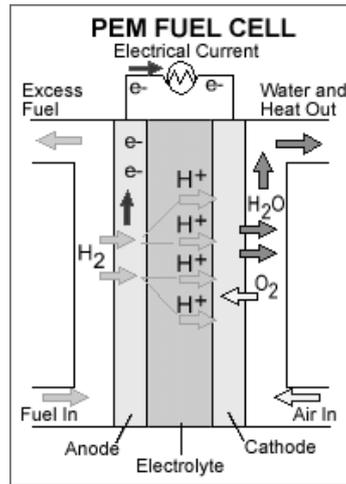


Fig.1. Structure of single PEMFC.

Theory of SVR. SVM, proposed by Vapnik and co-workers in 1995, is a statistical learning theory based on structural risk minimization principle [12]. At the present time, it has been successfully applied to solve classification and regression problems in numerous fields [13-15]. Researches show that the predictive results of SVM with many advantages, such as fast-learning, global optimization and excellent generalization ability, are generally superior to those of other pattern recognitions and regressions. In recent years, it has become a focus in machine learning research and is extensively employed in a wide range of real-world problems. When SVM is applied to regression, it is called as support vector regression (SVR).

In SVR, the basic idea is to map \mathbf{X} from the input space into a higher-dimensional feature space \mathbf{F} via a nonlinear mapping function $\Phi(\mathbf{x})$, and then to conduct linear regression in \mathbf{F} space. Therefore, SVR is to find the linear relation Eq. 1 based on a given dataset $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)$.

$$f(\mathbf{x}) = \mathbf{w} \cdot \Phi(\mathbf{x}) + b, \quad \Phi: \mathbf{R}^n \rightarrow \mathbf{F}, \quad \mathbf{w} \in \mathbf{F}. \quad (1)$$

Where \mathbf{w} is a vector for regression coefficients, b is a bias. They are estimated by minimizing the regularized risk function $R(C)$, namely:

$$\text{minimize } R(C) = 1/2 \|\mathbf{w}\|^2 + C \sum_{i=1}^n L_e(f(\mathbf{x}_i) - y_i), \quad (2)$$

$$L_e(f(\mathbf{x}_i) - y_i) = \begin{cases} 0, & \text{if } |f(\mathbf{x}_i) - y_i| < \epsilon, \\ |f(\mathbf{x}_i) - y_i| - \epsilon, & \text{if } |f(\mathbf{x}_i) - y_i| \geq \epsilon. \end{cases} \quad (3)$$

Where C is a regularized factor, n is the number of training samples, ϵ is a prescribed parameter controlling the tolerance to error, $(1/2)\|\mathbf{w}\|^2$ is used as a measurement of function flatness. The second term in Eq.2, $C \sum_{i=1}^n L_e(f(\mathbf{x}_i) - y_i)$, is the so-called empirical risk and measured by ϵ -insensitive loss

function $L_\varepsilon(f(\mathbf{x}_i)-y_i)$, which indicates that it does not penalize errors below ε . By minimizing Eq. 2, the array \mathbf{w} can be written in terms of the Lagrange multipliers and training samples as:

$$\mathbf{w} = \sum_{i=1}^l (a_i - a_i^*) F(\mathbf{x}_i) \quad (4)$$

Where l is the number of SVs. Finally, the linear relation Eq. 1 has the following explicit form:

$$f(\mathbf{x}) = \sum_{i=1}^l (a_i - a_i^*) k(\mathbf{x}, \mathbf{x}_i) + b \quad (5)$$

In Eq. 5, $k(\mathbf{x}, \mathbf{x}_i) = F(\mathbf{x}) \cdot F(\mathbf{x}_i)$ is a kernel function. Choosing different kernel function can generate different SVR models. In this paper, the radial basis kernel (Eq. 6) was utilized:

$$k(\mathbf{x}, \mathbf{x}_i) = \exp(-g \|\mathbf{x} - \mathbf{x}_i\|^2) \quad (6)$$

Dataset. The dataset used in this study was generated by Du et al. [10] and was listed in Table 1. This dataset includes the membrane water content data for 51 samples. This dataset was acquired by Du et al. via AC amplitude & frequency modulation impedance disturbance method to measure the impedance of PEM-chip at Wuhan University of Technology. The detailed experimental procedure can be referred from Ref. [10].

Modeling and predicting. In SVR model, the impedance of single-PEM-chip and operating temperature were employed as input variables, while as membrane water contents as output variables. The modeling and prediction were conducted by using the 51 samples based on SVR approach via leave-one-out cross validation (LOOCV). LOOCV is to split the n measurements into a training set of size $n-1$ and a test set of size 1, and average the squared error on the left-out measurement over the n possible way of obtaining such a partition.

Table 1 Measured and predicted PEM water contents of fuel cell under different temperature and impedance of single-PEM-chip

No.	Temperature (°C)	Impedance (mΩ)	Measured water content ^[10]	Predicted water content by LOOCV-SVR	Error	Percentage error (%)
1	23.8	28.02	1.89	1.89	0.00	0.00
2	24.1	19.59	2.69	2.70	0.01	0.37
3	24.5	19.58	2.68	2.67	-0.01	0.37
4	25.0	19.58	2.66	2.67	0.01	0.38
5	25.3	19.67	2.64	2.63	-0.01	0.38
6	25.6	14.88	3.47	3.47	0.00	0.00
7	26.1	14.92	3.44	3.44	0.00	0.00
8	26.7	14.92	3.41	3.41	0.00	0.00
9	27.1	14.92	3.39	3.39	0.00	0.00
10	27.6	12.55	4.00	4.00	0.00	0.00
11	28.2	12.51	3.98	3.98	0.00	0.00
12	28.8	12.50	3.95	3.96	0.01	0.25
13	29.4	10.39	4.71	4.71	0.00	0.00
14	30.1	10.38	4.67	4.67	0.00	0.00
15	30.8	10.42	4.61	4.61	0.00	0.00
16	31.8	9.19	5.16	5.15	-0.01	0.19
17	32.4	9.16	5.13	5.13	0.00	0.00
18	33.1	9.17	5.07	5.08	0.01	0.20
19	33.8	9.17	5.03	5.02	-0.01	0.20
20	34.8	8.04	5.65	5.65	0.00	0.00
21	35.7	8.05	5.58	5.58	0.00	0.00
22	37.0	8.04	5.49	5.49	0.00	0.00
23	37.6	8.04	5.45	5.45	0.00	0.00
24	38.5	8.03	5.39	5.39	0.00	0.00
25	39.3	8.05	5.32	5.32	0.00	0.00
26	40.8	7.16	5.87	5.88	0.01	0.17
27	41.7	7.17	5.80	5.80	0.00	0.00
28	42.6	7.17	5.73	5.74	0.01	0.17
29	43.5	7.17	5.67	5.66	-0.01	0.18
30	44.3	7.18	5.60	5.61	0.01	0.18
31	45.1	6.55	6.08	6.07	-0.01	0.16
32	46.1	6.54	6.01	6.02	0.01	0.17
33	47.1	6.52	5.95	5.96	0.01	0.17
34	48.1	6.54	5.86	5.85	-0.01	0.17
35	50.3	5.92	6.31	6.30	-0.01	0.16
36	51.3	5.92	6.23	6.24	0.01	0.16
37	52.1	5.92	6.17	6.18	0.01	0.16
38	53.0	5.91	6.12	6.12	0.00	0.00
39	53.9	5.92	6.04	6.05	0.01	0.17
40	54.8	5.38	6.58	6.57	-0.01	0.15
41	56.1	5.39	6.46	6.46	0.00	0.00
42	57.0	5.39	6.39	6.39	0.00	0.00
43	58.0	5.38	6.34	6.32	-0.02	0.32
44	58.9	5.38	6.26	6.27	0.01	0.16
45	59.8	5.35	6.23	6.24	0.01	0.16
46	60.7	5.37	6.16	6.15	-0.01	0.16
47	61.5	5.32	6.16	6.15	-0.01	0.16
48	62.1	5.32	6.11	6.12	0.01	0.16
49	62.5	5.74	5.64	5.71	0.07	1.24
50	62.4	7.67	4.23	4.20	-0.03	0.71
51	61.8	11.76	2.77	2.78	0.01	0.36

Evaluation of model's generalization performance. Mean absolute error (*MAE*), mean absolute percentage error (*MAPE*) and multiple correlation coefficients (R^2) were adopted for generalization performance evaluation.

Results and discussions

Prediction study of LOOCV test of SVR model. The indices from the 5th & 7th columns in Table 1 illustrate the membrane water contents and errors predicted by LOOCV test of SVR. The results predicted by LOOCV test of SVR versus experimental water content values are also plotted in Fig. 2. Table 2 shows the prediction performance of SVR.

From Table 1, it can be observed that, among all 51 samples, there are 21 samples with zero errors, 26 samples with absolute errors 0.01, and only 3 samples with absolute errors over 0.01, which are #43 with error -0.02, #50 with -0.03 and #49 with 0.07 respectively. Through statistic, the absolute percentage errors of 50 out of 51 samples are within a very small range of 0-0.71%. There is only one sample (#49) with the maximum absolute percentage error (1.24%) slightly exceeding 1.0%, which accounts for 1.96% of total samples.

It is visible from Fig.2 that nearly all the prediction points lie on or very close to the straight-line with slope 1. This reflects that the PEM water content values predicted by LOOCV-SVR are in quite good agreement with the measured values, and also illustrates the SVR models possess excellent forecasted ability.

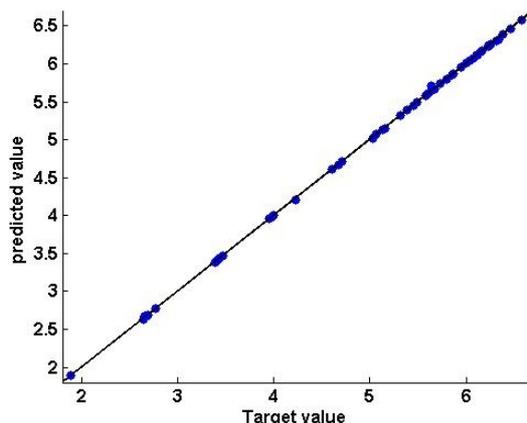


Fig.2. Comparison between target values and estimated values predicted by LOOCV test of SVR.

Table 2 lists the generalization performance by LOOCV test of SVR. From this table, it can be seen that the indices of *MAE* (0.01), *MAPE* (0.15%) and R^2 (1.00) reveal the precise of the SVR models, which is small enough to meet the demand of experiments.

	<i>MAE</i>	<i>MAPE</i> (%)	R^2
LOOCV- SVR	0.01	0.15	1.00

Statistical analysis of the SVR prediction model. Let $d=p-r$ is the difference between the predicted result achieved by the LOOCV test of SVR and the actual data by the experimental method. Where p is predicted result, r is actual data by the experimental method, and d is the predicted error. The degree of agreement between the two methods can get by calculating the bias, estimated by the mean difference \bar{d} and the standard deviation of the differences s . In Table 1, the mean difference \bar{d} is 0.000 and the standard deviation of the differences s is 0.013. One would expect most of the differences to lie between $\bar{d} - 2s$ and $\bar{d} + 2s$, which refers to as the limits of agreements. If the differences are normally distributed (Gaussian), 95% of differences will lie between these limits [16].

The upper limit of agreement is equal to:

$$\bar{d} + 2s = 0.000 + 2 \times 0.013 = 0.026.$$

The lower limit of agreement is equal to:

$$\bar{d} - 2s = 0.000 - 2 \times 0.013 = -0.026.$$

Fig. 3 shows the graph of the difference between the actual data and the predicted data via LOOCV test of SVR against the actual value. From Fig.3, it can be seen that, except one point, all other (50/51=98.04%) of the differences lie within the limit scope between the range of $\bar{d} - 2s$ and $\bar{d} + 2s$. This illustrates an excellent agreement between the estimated value predicted via LOOCV test of SVR and the actual experimental value.

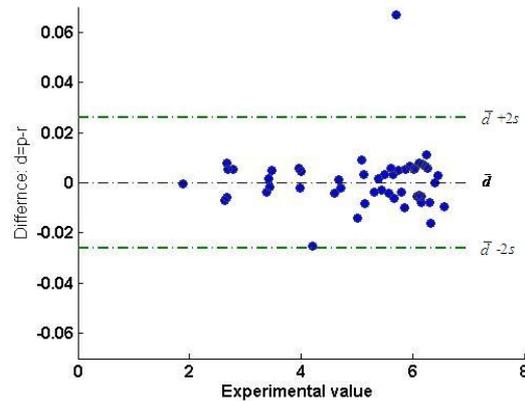


Fig.3. Difference between the experimental value and the predicted value via LOOCV test of SVR against the experimental value.

Conclusions

In this study, the support vector regression approach was used to model and predict the membrane water content of PEMFC based on two factors (the impedance of single-PEM-chip and operating temperature). The results reveal that: SVR approach may be not only a promising and practical methodology to simulate the properties of fuel cell system, but also a powerful tool to be used for optimal designing or controlling the operating process of fuel cell system.

Acknowledgements

This work was supported by the Science and Technology Research Projects for the Guangxi Universities (YB 2014471)

References

- [1] V. Mehta, J.S. Cooper: J. Power Sources Vol.114 (2003), p. 32
- [2] T.E. Springer, T.A. Zawodzinski and S. Gottesfeld: J. Electrochem. Soc Vol.138(1991), p. 2334
- [3] H. Görgün, M. Arcaç and F. Barbir: J. Power Sources Vol. 157 (2006), p. 389
- [4] Y. Chen, H. Peng, D. Hussey, D. Jacobson, D. Tran, T. Baset and M. Biernacki: J. Power Sources Vol. 179 (2007), p. 376
- [5] M. Mench, Q. Dong and C. Wang: J. Power Sources Vol. 127 (2003), p. 90
- [6] D. McKay and A. Stefanopoulou: IEEE Proceedings of 2004 American Control Conference, Boston, Massachusetts, United states, June 30–July 2, 2004; 816-821.
- [7] C. Thawornkuno and C. Panjapornpon: Chiang Mail J. Sci Vol. 35 (2008), p. 212
- [8] S. Um and C. Wang: J. Power Sources Vol. 156 (2006), p. 211
- [9] S. Shimpalee, S. Greenway, D. Spuckler and J. Zee: J. Power Sources Vol. 135 (2004), p. 79
- [10] W.C. Du, L.Y. Zhang and S.H. Quan: Telecom power technology Vol. 26 (2009), p. 14
- [11] Information on http://www.newmaker.com/art_23205.html.

- [12] V. Vapnik, *The nature of statistical learning theory*, Springer, New York, 1995.
- [13] Y.F. Wen, C.Z. Cai, X.H. Liu, J.F. Pei, X.J. Zhu and T.T. Xiao: *Corros. Sci* Vol.51(2009), p. 349
- [14] C.Z. Cai, X.J. Zhu, Y.F. Wen, J.F. Pei, G.L. Wang and W.P. Zhuang: *J. Supercond. Novel Magn* Vol. 23 (2010), p. 737
- [15] C.Z. Cai, L.Y. Han, Z.L. Ji and Y.Z. Chen: *Proteins* Vol. 55 (2004), p. 66
- [16] J.M. Bland and D.G. Altman: *Lancet* Vol. 1 (1986), p. 307