

부분최소자승법과 인공신경망을 이용한 고분자 전해질 연료전지 스택의 모델링

한인수[†] · 신현길

GS칼텍스 기술연구소
305-380 대전광역시 유성구 엑스포로 359
(2014년 8월 22일 접수, 2014년 10월 1일 수정본 접수, 2014년 10월 2일 채택)

Modeling of a PEM Fuel Cell Stack using Partial Least Squares and Artificial Neural Networks

In-Su Han[†] and Hyun Khil Shin

R&D Center, GS Caltex Corp., 359 Expo-ro, Yuseong-gu, Daejeon 305-380, Korea
(Received 22 August 2014; Received in revised form 1 October 2014; accepted 2 October 2014)

요 약

고분자 전해질 연료전지 스택의 성능 및 주요 운전 변수를 예측하기 위해 부분최소자승법과 인공신경망의 두 가지 데이터 기반 모델링 기법을 제시한다. 30 kW급 고분자 전해질 연료전지 스택 실험으로부터 확보한 데이터를 사용하여 부분최소자승 및 인공신경망 모델들을 구성한 후 각 모델의 예측 성능 및 계산 시간을 비교하였다. 모델의 복잡성을 줄이기 위해 부분최소자승법에 기초한 VIP(Variable Importance on PLS Projections) 선정기준을 모델링 절차에 포함하여, 초기 입력변수의 집합으로부터 모델링에 필요한 입력변수들을 선정하였다. 모델링 결과, 인공신경망이 스택의 평균 셀전압과 캐소드(cathode) 출구 온도를 예측하는데 있어서, 부분최소자승법 보다 우수한 성능을 보였다. 그러나 부분최소자승법 또한 입력변수와 출력변수 간에 선형적 상관관계만을 모델링 할 수 있음에도 불구하고 비교적 만족할 만한 예측 성능을 나타냈다. 모델의 정확도와 계산속도의 요구조건에 따라 두 모델링 기법은 고분자 전해질 연료전지의 설계 및 운전 분야의 성능 예측, 온라인 및 오프라인 최적화, 제어 및 이상 진단을 위해 적용될 수 있을 것으로 판단된다.

Abstract – We present two data-driven modeling methods, partial least square (PLS) and artificial neural network (ANN), to predict the major operating and performance variables of a polymer electrolyte membrane (PEM) fuel cell stack. PLS and ANN models were constructed using the experimental data obtained from the testing of a 30 kW-class PEM fuel cell stack, and then were compared with each other in terms of their prediction and computational performances. To reduce the complexity of the models, we combined a variables importance on PLS projection (VIP) as a variable selection method into the modeling procedure in which the predictor variables are selected from a set of input operation variables. The modeling results showed that the ANN models outperformed the PLS models in predicting the average cell voltage and cathode outlet temperature of the fuel cell stack. However, the PLS models also offered satisfactory prediction performances although they can only capture linear correlations between the predictor and output variables. Depending on the degree of modeling accuracy and speed, both ANN and PLS models can be employed for performance predictions, offline and online optimizations, controls, and fault diagnoses in the field of PEM fuel cell designs and operations.

Key words: Fuel Cell, Polymer Electrolyte Membrane, Back-box Modeling, Partial Least Squares, Artificial Neural Networks

1. 서 론

지구온난화가 가속화됨에 따라 친환경 운송 수단 및 고효율의 동력원의 보급 필요성이 점차 증가되고 있으며, 기존의 내연기관을 대체할 만한 수단으로 고분자 전해질(polymer electrolyte membrane, PEM) 연료전지가 각광을 받고 있다. 특히, GM과 도요타를 포함한

세계 유수의 자동차 메이커들은 연료전지 자동차 개발에 많은 연구 개발비를 투자해오고 있다. 이는 고분자 전해질 연료전지가 효율이 높고 출력밀도를 높일 수 있으며 공해물질의 배출이 없는 특징으로 인해 승용차, 버스, 지게차 및 선박을 포함한 다양한 운송 수단의 주 동력원을 대체하기에 매우 적합하기 때문이다[1].

고분자 전해질 연료전지 내부에서 일어나는 물리화학적 현상의 규명, 설계 및 운전 최적화, 제어 및 이상 진단 등을 위한 주요 수단으로 고분자 전해질 연료전지 스택(stack) 및 시스템에 대한 모델링 및 시뮬레이션 연구가 활발히 진행되어 왔다[2-9]. 고분자 전해질 연료전지는 복잡한 전기화학 반응을 동반하고 스택 내부에서 생성되는 물이 이상(two-phase) 거동을 보이는 특징으로 인해 물리화학적 메

[†]To whom correspondence should be addressed.

E-mail: ishan1969@gscaltex.com

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

커니즘을 근간으로 하는 정교한 모델 개발이 쉽지 않다. 그 동안 개발되어 온 모델은 스택의 채널 또는 매니폴드(manifold) 설계 최적화를 위한 2~3차원 전산유체역학(computational fluid dynamics, CFD) 모델[2-6]이 주를 이루고 있으며, 제어 및 최적화 응용을 위해 보다 단순화된 형태의 0~1차원 모델[7-9] 개발도 활발히 진행되어 왔다. 물리화학적 메커니즘에 기반한 모델링은 시스템의 거시적 거동에 대한 직관적인 해답을 충분히 제공하지만, 보다 정확한 예측 값을 얻기 위해서는 모델의 엄밀성을 높여야 하며, 이를 위해 많은 시간과 노력을 투자해야 한다. 특히, 예측 정확도를 높이기 위해 모델의 엄밀성을 증가시킬수록 결정해야 하는 모델 파라미터가 함께 증가하게 되어 많은 실험 및 파라미터 튜닝 작업을 수반한다.

한편, 실험으로부터 확보한 입출력 데이터를 사용하여 입력변수와 출력변수 간에 주요 상관관계를 추출하는 블랙박스 모델링과 통계적 모델링 기법이 있다. 이러한 모델링 기법들은 비교적 빠른 시간 내에 모델 구성을 가능하게 하며, 연료전지가 작동되는 전체 영역에 걸쳐 충분한 데이터가 확보될 경우, 예측 정확도가 높은 모델을 제공하는 장점이 있다. 반면에, 경험적 데이터에 근거해 미리 의도한 변수만을 예측하도록 모델이 형성되기 때문에 대상 시스템의 하드웨어가 변경되거나 예측해야 하는 변수가 바뀌는 경우 모델을 재구성해야만 한다.

인공신경망(artificial neural network, ANN)과 지지벡터머신(support vector machine, SVM)은 비선형 시스템을 모델링 하기 위해 사용되는 대표적인 블랙박스 모델링 기법으로 고분자전해질 연료전지의 성능 예측, 분석, 제어 및 최적화를 위한 정적 및 동적 모델링을 목적으로 활발히 연구되어 오고 있다[10-13]. 이외 연료전지의 운전 분석 및 진단을 위한 목적으로 통계적 모델링 기법인 주성분분석법(principle component analysis, PCA)과 부분최소자승법(partial least square, PLS)을 적용하는 방법이 몇 차례 보고된 바 있다[14-16].

본 논문에서는 선형 통계적 모델링 도구인 부분최소자승법과 대표적인 비선형 데이터 기반 모델링 기법으로 인공신경망을 적용하여 고분자전해질 연료전지 스택의 주요 작동변수의 변화에 따른 성능 및 주요 운전변수의 거동을 예측하는 방법을 제시하고자 한다. 또한, 실제 연료전지 스택에 대한 실험을 통해 확보한 데이터를 사용하여 입력변수를 선정하고 모델을 구성한 후 각 모델링 기법에 따른 예측 성능 및 계산 속도를 비교 분석한다.

2. 모델링 방법

2-1. 부분최소자승법

부분최소자승법은 실험 또는 실제 공정 운전으로부터 확보한 데이터를 활용하여 다변량 입력변수(독립변수)와 출력변수(종속변수) 간에 상관관계를 구하는 통계적 모델링 도구로 활용되고 있다[17]. 일반적으로 화학공정을 포함한 다양한 산업 분야에 걸쳐 선형 다변수 모델을 개발하기 위한 목적으로 활발히 적용되고 있으며[18-20], 입/출력 변수를 전달함수를 통해 변환하거나 인공신경망 등과 결합할 경우 비선형 다변수 모델링에도 사용할 수 있다[21]. 특히, 입력 변수들 사이에 상관관계가 매우 높아 다중 공선성(multicollinearity)이 존재하거나 데이터에 노이즈가 많이 포함되어 있더라도 일반적인 다중 회귀분석법과 비교하여 예측 신뢰성이 높은 모델을 얻을 수 있는 장점이 있다.

부분최소자승법의 기본 개념은 입출력 변수의 변화를 측정하여

얻은 데이터로부터 형성된 고차원의 데이터 공간에 대하여 각축이 서로 직교하도록 새로운 축을 차례대로 정의하여 저차원의 특성 공간(feature space)으로 압축 투영한 후 각 특성 벡터 간의 최적의 상관관계를 구하는 것이다. 따라서 입력 및 출력 변수 간에 상관관계가 매우 크거나 심지어 특이 값을 갖더라도 신뢰성이 높은 회귀 모델을 얻을 수 있다. 또한, 모델링 결과로부터 얻은 각종 계수들(score 값, loading 값 및 회귀계수)은 모델 해석에 필요한 유용한 정보를 제공한다.

모델링을 수행하기 위한 첫 번째 단계는 v 개의 입력 변수에 대하여 n 번의 샘플링을 거쳐 측정된 입력 데이터를 $(v \times n)$ 의 크기를 갖는 $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n]^T$ 의 행렬로 정리하고, 단일 출력 변수에 대하여 n 번의 샘플링을 거쳐 수집한 출력 데이터는 $\mathbf{y}=[y_1, y_2, y_3, \dots, y_n]^T$ 로 나타내는 것이다. 모델링을 수행하기 위해서는 먼저 입출력 데이터에 대하여 평균을 빼주고 이를 다시 표준편차로 나누는 데이터 mean-centering과 scaling을 수행하여 데이터 전처리를 완료한다. 다음으로 데이터 전처리가 수행된 각 \mathbf{X} 와 \mathbf{y} 에 대하여 다음과 같이 rank-one 행렬 또는 벡터의 조합으로 분해하여 outer relation을 구한다.

$$\mathbf{X} = \sum_{i=1}^a (\mathbf{t}_i \mathbf{p}_i) + \mathbf{E} \quad (1)$$

$$\mathbf{y} = \sum_{i=1}^a (\mathbf{u}_i \mathbf{q}_i) + \mathbf{f} \quad (2)$$

여기서, a 는 압축된 잠재변수(latent variable)의 개수 또는 PLS 차원을 의미한다. \mathbf{t}_i 와 \mathbf{u}_i 는 각각 입력과 출력 score 벡터를 나타내며, 각 잠재변수 i 에 대하여 차례대로 계산할 수 있다. 또한 \mathbf{p}_i 와 \mathbf{q}_i 는 loading 벡터와 값으로 각각 변수 \mathbf{X} 와 \mathbf{y} 에 대한 영향을 나타낸다. 모든 score와 loading 벡터는 잔차(residual) 행렬 \mathbf{E} 와 벡터 \mathbf{f} 를 최소화하는 방향으로 결정된다. 식 (1)과 (2)로 표현되는 outer relation 으로부터 계산한 \mathbf{t}_i 와 \mathbf{u}_i 의 선형 상관관계를 나타내는 inner relation을 다음과 같이 나타낼 수 있다.

$$\mathbf{u}_i = c_i \mathbf{t}_i + \mathbf{h}_i, \quad i = 1, 2, 3, \dots, a \quad (3)$$

회귀계수 c_i 는 각 잠재변수 i 에 대하여 잔차벡터 \mathbf{h}_i 를 최소화하는 방향으로 결정되며, 대각행렬 \mathbf{C} 로 나타낼 수 있다. 식 (1)-(3)을 조합하면 다음과 같이 PLS 예측 값 $\hat{\mathbf{y}}$ 를 얻기 위한 모델식을 구할 수 있다.

$$\begin{aligned} \hat{\mathbf{y}} &= \sum_{i=1}^a (\mathbf{t}_i c_i \mathbf{q}_i^T) = \mathbf{T} \mathbf{C} \mathbf{q}^T = \mathbf{X} (\mathbf{P}^T)^{-1} \mathbf{C} \mathbf{q}^T \\ &= \mathbf{X} \mathbf{W} (\mathbf{P}^T \mathbf{W})^{-1} (\mathbf{T}^T \mathbf{T})^{-1} \mathbf{T}^T \mathbf{y} = \mathbf{X} \mathbf{b} \end{aligned} \quad (4)$$

여기서, $\mathbf{T}(n \times a)$ 는 $[\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \dots, \mathbf{t}_a]$, $\mathbf{P}(v \times a)$ 는 $[\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \dots, \mathbf{p}_a]$, \mathbf{q} 는 $[q_1, q_2, q_3, \dots, q_a]$ 를 나타낸다. $\mathbf{W}(v \times a)$ 는 가중치 행렬 $[\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_a]$ 이며, \mathbf{b} 는 모델의 회귀계수 $[b_1, b_2, b_3, \dots, b_a]$ 를 나타낸다. 모델의 잠재변수의 개수(a)는 모델링 과정에서 결정해야 하는 유일한 튜닝 파라미터로 일반적으로 교차타당화(cross-validation)[22] 과정을 거쳐 결정되는데, 최대값은 입력변수의 수와 같으며 대체로 이보다 상당히 작은 값을 갖는다. PLS 모델의 주요 계수들(score, loading, weight 및 회귀계수)을 결정하는 방법으로 각 잠재변수에 대하여 순차적인 계산을 수행하는 NIPALS(nonlinear iterative partial least squares) 알고리즘을 적용하였다[22].

2-2. 인공신경망

인공신경망은 모델 구조를 형성하는데 유연성이 높아 다양한 산

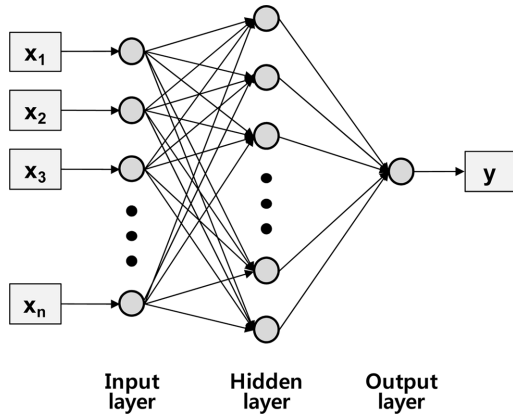


Fig. 1. Feed-forward neural network structure with one hidden layer for the PEM fuel cell stack modeling.

업 분야에 걸쳐 비선형 모델링을 위한 기법으로 활발히 사용되고 있으며, 일반적으로 충분한 양의 입력력 데이터가 제공될 경우 우수한 예측 성능을 발휘한다[19,20,23]. 인공신경망은 Fig. 1에 보인 바와 같이 뉴런(neuron) 또는 노드(node)라 불리는 수십에서 수백 개의 계산 소자가 상호 연결되어 있는 구조를 취하고 있으며, 다층 구조의 인공신경망은 입력층(input layer), 은닉층(hidden layer) 및 출력층(output layer)의 3개의 층으로 구성된다. 각 노드 간의 연결 강도를 가중치(weight)라고 하며, 적절한 학습 알고리즘을 통해 측정값과 예측값이 서로 근접하도록 가중치들을 조절함으로써 모델을 생성한다.

모델의 구조와 학습 알고리즘에 따라 다양한 형태의 인공신경망이 존재하나 본 연구에서는 Fig. 1에 보인 바와 같이 한 개의 은닉층만을 갖는 앞먹임망(feed-forward network)을 적용하였으며, 모델 예측 값은 다음과 같은 함수로 표현할 수 있다.

$$\hat{y}_k = f^{(2)}\left(\sum_{i=1}^h [w_i^{(2)} f_i^{(1)}(\mathbf{x}_k^T \mathbf{w}_i^{(1)} + b_i^{(1)})] + b_0^{(2)}\right) \quad (5)$$

여기서, h 는 은닉층에 포함된 노드의 개수, \mathbf{x}_k 는 크기 v 를 갖는 입력변수 벡터로 k 번째 측정값을 의미한다. \hat{y}_k 는 k 번째 측정 샘플에 대한 예측값이다. $b_i^{(1)}$ 와 $b_0^{(2)}$ 는 각각 은닉층과 출력층의 바이어스 값을 나타내며, $\mathbf{w}_i^{(1)}$ 와 $w_i^{(2)}$ 는 각각 입력층과 은닉층의 각 노드를 연결하는 가중치 벡터와 은닉층과 출력층을 연결하는 가중치이다. 은닉층에 포함된 노드의 전달함수는 다음과 같이 표현되는 sigmoid 함수를 사용하였으며, 출력층은 pure linear 함수($f^{(2)}(z) = z$)를 적용하였다.

$$f_i^{(1)}(z) = 1/[1 + \exp(-z)] \quad (6)$$

모델링 과정에서 결정해야 하는 튜닝 파라미터는 은닉층에 포함된 노드의 수(h)가 유일하며, 모델 계수는 가중치 행렬 $\mathbf{W} = [\mathbf{w}_1^{(1)}, \mathbf{w}_2^{(1)}, \mathbf{w}_3^{(1)}, \dots, \mathbf{w}_h^{(1)}, w_1^{(2)}, w_2^{(2)}, w_3^{(2)}, \dots, w_h^{(2)}$]과 바이어스 벡터 $\mathbf{b} = [b_1^{(1)}, b_2^{(1)}, b_3^{(1)}, \dots, b_h^{(1)}, b_0^{(2)}]$ 가 있다. 주어진 입력 및 출력 데이터에 대하여 인공신경망의 모델 계수들은 다음과 같이 측정값과 계산값의 차이를 최소화 하는 비선형 최적화를 수행하여 결정할 수 있다.

$$\min J(\mathbf{W}, \mathbf{b}) = \sum_{i=1}^h (y_i - \hat{y}_i)^2 \quad (7)$$

여기서, 모델계수인 가중치 및 바이어스 값은 역전파(back-propagation) 알고리즘을 사용하여 계산한다[24]. 식 (5)로 표현되는 모델은 은닉

층에 포함된 노드를 증가시킴으로써 일반적으로 모든 형태의 함수를 근사할 수 있는 것으로 알려져 있다[25].

3. 실험 및 데이터 수집

모델링에 필요한 데이터 수집을 위해 Greenlight사의 테스트스테이션(모델 G500)을 사용하여 Fig. 2에 보인 것과 같은 연료전지 스택 테스트 장치를 구성하였다. 실험에 사용된 스택은 최대 30 kW의 출력이 가능하며, 수소와 공기를 각각 연료와 산화제로 사용한다[26]. 스택의 수소 측은 4단의 캐스케이드(cascade) 구조를 취하고 있어 수소 재순환 장치가 불필요하며, dead-end 조건 하에서 작동하도록 설계하였다[27]. 스택에 공급되는 공기는 가습기 내에서 캐소드(cathode)에서 방출되는 잔류가스와 물질전달 및 열교환에 의하여 온도와 가습도가 상승한 후 스택에 유입된다. 수소는 무가습 상태(상대습도 0%)로 스택에 공급되며, 유량에 관계없이 레귤레이터에 의하여 일정한 공급 압력이 유지된다. 스택이 dead-end 조건으로 작동하기 때문에 스택의 수소 측 출구에 위치한 purge 밸브는 대부분 닫혀있으며 간헐적으로 열린다. 따라서 purge 밸브가 닫혀 있는 동안 스택 내부에 불순물, 질소 및 생성수 등이 축적되며, 시간이 경과할수록 수소 측 마지막 단에 위치한 셀들의 전압이 점차 낮아진다. 마지막 단의 셀전압이 미리 설정한 값에 도달할 경우 purge 밸브가 자동으로 열려 셀전압을 회복시키는 과정을 주기적으로 반복하게 된다. 스택 입구와 출구에서 측정된 수소, 공기 및 냉각수의 온도와 압력들을 포함하여 공기 및 냉각수의 유량, 전류 및 셀전압은 1초 간격으로 자동적으로 데이터베이스에 저장된다. 스택에 공급되는 냉각수의 온도는 55~65 °C의 범위에서 조절되며, 스택을 통과한 냉각수의 온도가 75 °C 이하로 유지되도록 냉각수의 유량을 제어한다. 수소는 상온에서 40~55 kPa 범위의 압력을 유지하면서 스택에 공급되며, 공기는 상압 조건에서 가습기를 통과한 후 부하에 따라 1.8~3.2의 양론비가 유지되도록 스택에 공급된다.

Table 1은 모델링을 위해 초기에 선정한 입력 및 출력 변수들을 나열한 것이다. 스택 작동 중 측정되는 여러 변수들 중 조절변수와 현재의 값을 알 수 있는 상태변수만을 초기 입력변수로 선정하였으며, 모델을 통해 예측되는 출력변수들은 스택의 성능을 나타내는 대표적인 변수로 스택의 평균 셀전압(#9)과 스택으로부터 방출되어 가습기에 열과 수분을 공급하는 습윤 공기의 상태를 나타내는 스택 캐

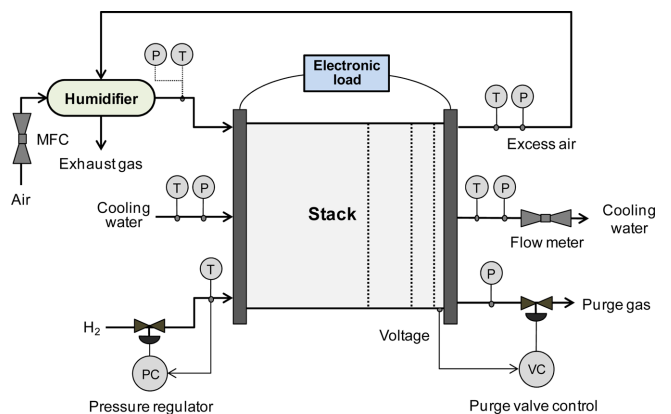


Fig. 2. Schematic of the experimental apparatus for testing the PEM fuel cell stack.

Table 1. List of the input and output variables for modeling

Variable number	Description	Type
#1	Stack current	Input
#2	Air flow rate	Input
#3	Coolant flow rate	Input
#4	Hydrogen inlet pressure	Input
#5	Air inlet pressure	Input
#6	Coolant inlet temperature	Input
#7	Hydrogen inlet temperature	Input
#8	Air inlet temperature	Input
#9	Stack average cell voltage	Output
#10	Cathode outlet temperature	Output

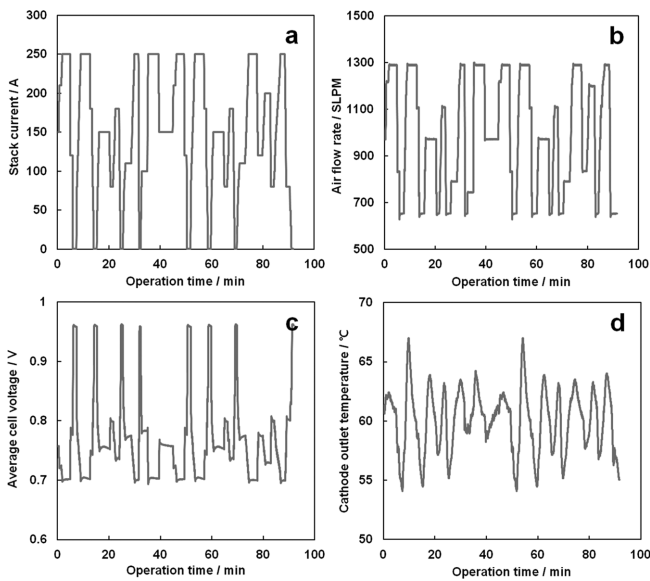


Fig. 3. Variations in the major operating variables: (a) stack current, (b) air flow rate, (c) average cell voltage, and (d) cathode outlet temperature.

스드 출구 온도(#10)이다. 스택의 작동 범위 내에서 약 90분 동안 스택의 주요 조절변수들(#1~#6)을 변화시키면서 입력변수와 출력변수의 변동을 1초 간격으로 측정하여 원시 데이터를 수집하였다. Fig. 3은 실험을 통해 측정된 주요 입출력 변수(#1, #2, #9 및 #10)의 거동을 도시한 것이다. 노이즈를 제거하고 데이터 smoothing을 위해 수집한 원시 데이터에 30초 간격으로 이동평균을 취하여 최종 5,475개의 데이터를 확보하였다.

4. 결과 및 토론

4.1. 변수 선정

입출력 데이터에 기반을 둔 통계적 모델링 또는 블랙 박스 모델링 과정에서 입력변수는 출력변수의 변화에 영향을 크게 주는 핵심변수들 위주로 선정하는 것이 모델의 예측 성능을 높이고 복잡성을 줄여 모델링에 필요한 계산 시간을 단축시키는데 유리하다. 모델의 출력값을 설명하는 입력변수들 중에 출력변수와 상관관계가 매우 미약하거나 다른 입력변수와 상관관계가 높아 다중 공선성을 유발하는 변수들을 미리 제거하는 변수 선정 과정이 최종 모델을 구성하기 전에 수행되어야 한다. 변수 선정을 위한 다양한 알고리즘들이 보고

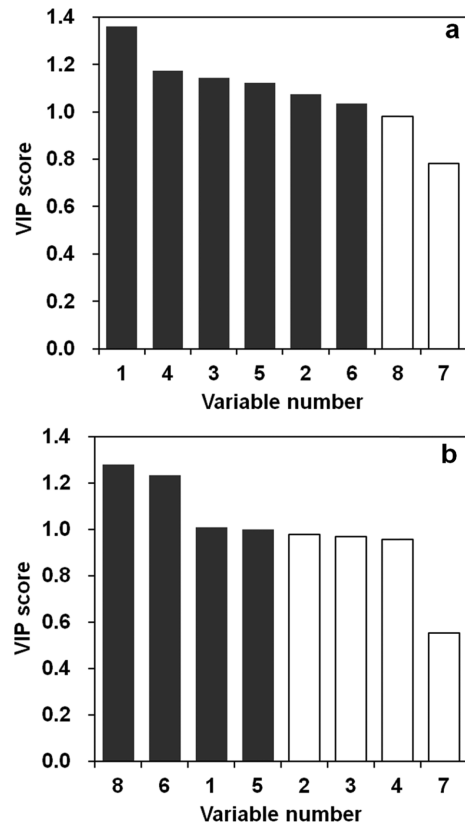


Fig. 4. VIP scores of the PLS models for predicting (a) the average cell voltage and (b) the cathode outlet temperature.

되어 있으나[28,29], 본 연구에서는 Table 1에 보인 초기 선정 변수들을 대상으로 부분최소자승법을 수행하여 계산한 VIP(Variable Importance on PLS Projections)[29] 값을 기준으로 평균 셀전압 및 캐소드 출구 온도를 예측하기 위한 입력변수들을 선정하였다. 각 입력변수에 대한 VIP score(V_j)는 다음 식에 의하여 계산할 수 있으며, 일반적으로 1.0 이상의 값을 갖는 경우 해당 입력변수를 모델링 변수에 포함시킨다[29].

$$V_j = \sqrt{v \sum_{i=1}^a (s_i w_{j,i}^2) / \sum_{i=1}^a s_i} \quad (8)$$

여기서 $s_i = c_i^T t_i$ 를 나타내며, v 는 입력변수의 수, c_i 는 PLS inner relation에 대한 회귀 계수, w 는 가중치 벡터이다. Fig. 4a와 b는 스택의 평균 셀전압과 캐소드 출구 온도 예측을 위한 PLS 모델로부터 계산한 VIP score를 각각 도시한 것이다. 그림에서 검은색 바는 VIP score가 1.0 이상인 변수들을 나타내며, 흰색 바는 그 이하인 변수를 나타낸다. 따라서 평균 셀전압 예측 모델은 Table 1에 나열한 변수 중 #1~#6을 모델의 입력변수들로 선정하였으며, 캐소드 출구 온도 예측 모델은 #1, #5, #6 및 #8을 입력변수들로 선정하였다.

4.2. 모델링 결과 및 비교

스택의 평균 셀전압 및 캐소드 출구 온도에 대한 예측 모델을 구성하기 위해 각 출력변수와 선정된 입력변수들의 조합을 갖는 데이터에 부분최소자승법과 인공신경망을 각각 적용하였다. 수집한 5,475개의 데이터 중 70%(학습데이터)는 교차타당화를 통해 모델의

구조를 확정하고 학습시키는데 사용하였으며, 30%(검증 데이터)는 모델의 예측 성능을 검증하기 위해 모델 예측값과 측정값을 비교하는데 사용하였다. 교차타당화를 통해 결정해야 하는 모델 튜닝 파라미터는 부분최소자승 모델은 잠재변수의 개수이며 인공신경망 모델은 은닉층의 노드의 개수이다. 교차타당화는 학습 데이터(전체 데이터 중 70%)만을 사용하여 튜닝 파라미터를 바꾸어가며 모델링을 수행한 후 예측 오차를 최소화시키는 잠재변수의 개수와 은닉층의 노드의 수를 찾는 과정이다. 먼저 학습 데이터를 5개의 그룹으로 분할한 후 이 중 4개의 그룹은 주어진 튜닝 파라미터를 적용하여 모델링에 사용하고, 나머지 1개의 그룹은 모델 예측 오차를 계산하는데 사용하는 방식으로 그룹을 바꾸어 가며 총 5회에 걸쳐 계산된 예측 오차를 합산하였다. 이 과정은 정수 값을 갖는 튜닝 파라미터를 바꾸어가며 반복되며, 이중 가장 작은 누적 예측 오차(prediction residual sum of squares, PRESS)를 산출하는 튜닝 파라미터를 최종적으로 선정하였다. 결정된 부분최소자승 모델의 잠재변수의 개수는 스택 평균 셀전압과 캐소드 출구 온도에 대하여 각각 5개와 4개이며, 인공신경망 모델의 은닉층 노드의 개수는 두 개의 출력변수 모두에 대하여 8개이다.

Fig. 5와 6은 각각 부분최소자승 및 인공신경망 모델에 의하여 예측된 평균 셀전압을 실험을 통해 측정된 값과 비교한 결과를 나타낸 것이다. 스택의 전체 운전 영역(전류 0~250 A 범위)에 걸쳐서 두 모델 모두 대체로 양호한 예측 성능을 보이나, 인공신경망 모델(평균 제공근 오차 0.0035 V)이 부분최소자승 모델(평균 제공근 오차 0.0086 V) 보다 상당히 우수한 예측 성능을 보이고 있음을 알 수 있다. 이는 연료전지 스택의 셀전압이 입력변수들의 변화에 대하여 비선형적 거동을 갖는 특징으로부터 발생한 것으로 판단된다. 즉, 부분최소자승 모델은 시스템의 선형적 변화만을 예측할 수 있는 한계를 갖는 반면에, 인공신경망 모델은 입력력 변수 간의 비선형 관계를 모델링 할 수 있기 때문이다. 특히, Fig. 5에서 볼 수 있는 바와 같이 부분최소자승 모델로 예측한 스택의 평균 셀전압이 개방회로전압을 포함한 저부하 영역에서 예측 오차가 다소 증가하는 경향이 있다. 이는 고분자 전해질 연료전지 스택이 고부하 영역보다 저부하 영역에서 부하 변동에 따른 셀전압의 변화 폭이 상당히 커지는 비선형적 거동의 영향으로 판단된다.

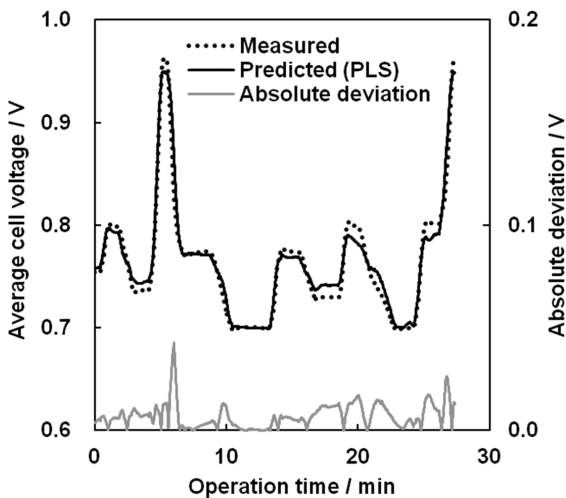


Fig. 5. Comparison between PLS predictions and experimental measurements for the average cell voltage.

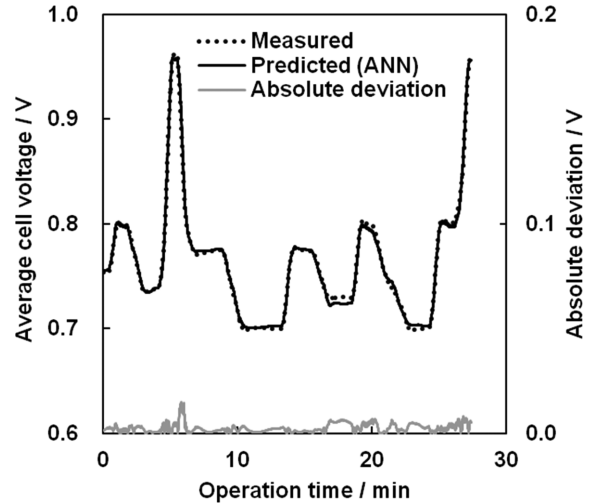


Fig. 6. Comparison between ANN predictions and experimental measurements for the average cell voltage.

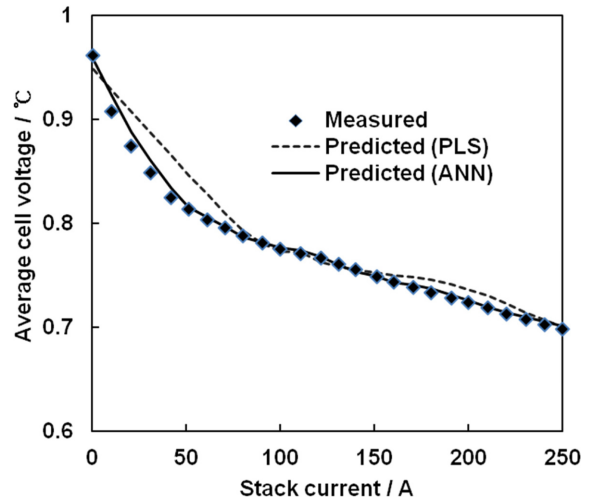


Fig. 7. Polarization curves predicted using the PLS and ANN models.

Fig. 7은 스택 실험으로부터 측정된 I(전류)-V(전압) 곡선을 각각 부분최소자승 및 인공신경망 모델로 예측한 값과 비교한 것이다. 인공신경망 모델은 측정값에 매우 근사한 예측 성능을 보이고 있으나, 부분최소자승 모델은 특히 저부하 영역에서 예측 오차가 다소 증가하는 경향을 보임을 알 수 있다.

Fig. 8과 9는 스택 캐소드 출구 온도를 부분최소자승 및 인공신경망 모델을 각각 적용하여 예측한 결과를 측정값과 비교한 그래프이다. 부분최소자승 및 인공신경망 모델 모두 평균 제공근 오차가 0.2 °C 이하(부분최소자승 모델 0.195 °C, 인공신경망 모델 0.122 °C)로 상당히 우수한 예측 성능을 나타냄을 알 수 있다. 특히, 부분최소자승법은 평균 셀전압 예측에 사용된 경우 보다 스택 캐소드 출구 온도 예측에 있어서 인공신경망 모델의 예측 성능에 근접한 결과를 보여, 선형 모델만으로도 만족스러운 예측 결과를 얻을 수 있음을 알 수 있다.

Table 2는 부분최소자승법과 인공신경망을 각각 적용하여 스택 평균 셀전압 및 캐소드 출구 온도 예측 모델을 구성하는데 소요된 컴퓨터의 계산 시간과 각 모델의 예측 정확도를 비교한 것이다. 모델링에 소요된 시간은 교차타당화를 통해 최적의 모델 튜닝 파라미

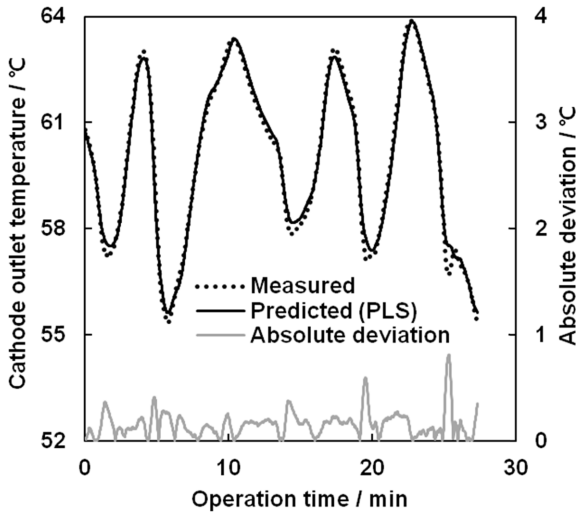


Fig. 8. Comparison between PLS predictions and experimental measurements for the cathode outlet temperature.

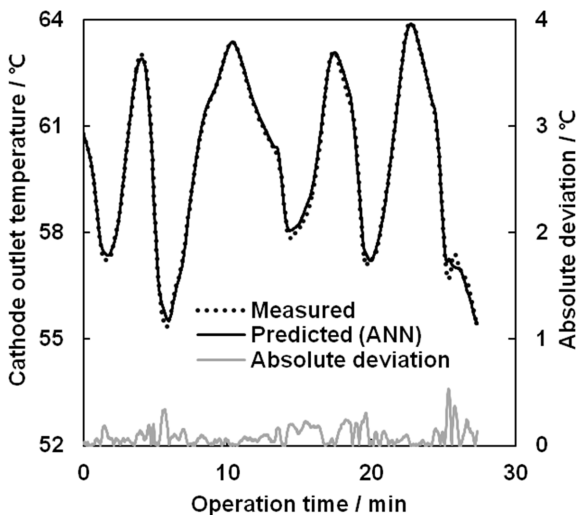


Fig. 9. Comparison between ANN predictions and experimental measurements for the cathode outlet temperature.

Table 2. Comparison of the modeling performances between partial least square and artificial neural network models

Output variables	Modeling method	Modeling time (s)	Prediction error
Average cell voltage	PLS	3.0	0.0086 V
	ANN	517.6	0.0035 V
Cathode outlet temperature	PLS	2.9	0.195 °C
	ANN	372.5	0.122 °C

터(부분최소자승 모델의 잠재변수의 수와 인공신경망 모델의 은닉층의 노드 수)를 찾기 위해, 수십 차례에 걸쳐 반복된 모델링(학습) 및 예측을 위해 수행된 시뮬레이션 시간을 합산한 것으로 Windows 8.1 PC(Intel Core i5-3337U 1.8GHz CPU)에서 계산한 결과를 나타낸다. 표에 정리된 바와 같이 모델 예측 성능은 부분최소자승 모델이 인공신경망 모델 보다 다소 떨어지나, 부분최소자승 모델 구성을 위해 필요한 계산 시간이 인공신경망 모델 구성에 필요한 시간의 0.58%(스택 평균 셀전압) 및 0.78%(스택 캐소드 출구 온도) 수준으

로 매우 적게 소요되는 장점이 있음을 알 수 있다. 따라서 부분최소자승법은 모델의 예측 성능은 다소 떨어지더라도 빠르게 모델을 갱신해야 하는 연료전지 스택의 실시간 운전 최적화 및 이상 진단과 같은 목적으로 사용하는 것이 적합할 것으로 보이며, 인공신경망은 높은 모델 예측 정확도를 요구하는 분야에 적합할 것으로 판단된다.

5. 결 론

고분자전해질 연료전지 스택의 성능 및 주요 운전 변수는 실험 데이터를 이용한 선형 다변수 통계적 모델링 기법인 부분최소자승법과 비선형 블랙박스 모델링 도구인 인공신경망을 적용하여 비교적 양호한 정확도로 예측 가능하다. 또한, 부분최소자승법에 기반한 변수 선정법을 도입하여 모델의 입력변수의 개수를 줄여 모델의 복잡성을 줄일 수 있었다.

부분최소자승법은 대상 시스템의 선형 상관관계만을 잡아내는 특징으로 인해 인공신경망 보다 예측 성능은 다소 낮으나 비교적 빠르게 모델을 생성할 수 있으며, 잠재변수의 수가 정해지면 일관된 모델링 결과를 얻을 수 있다. 반면에 인공신경망은 비선형 거동에 대한 모델링이 가능하기 때문에 비선형성이 높은 연료전지 스택의 셀전압 예측에 있어서 매우 우수한 성능을 발휘하였다. 그러나 인공신경망은 상대적으로 많은 양의 학습 데이터가 필요하며, 가중치와 바이어스를 포함한 모델 계수들을 결정하는 단계에서 지역 최소값의 존재로 최적의 모델링 결과를 얻기 위해서는 학습과정에 상당한 계산 시간이 소요되었다.

고분자전해질 연료전지 스택에 대한 부분최소자승모델 및 인공신경망 모델은 정확도 및 모델 갱신에 소요되는 시간 요구조건에 따라 운전 분석, 연료전지의 성능을 최대화하기 위한 실시간 운전 최적화, 운전 모니터링 및 이상 진단을 위한 도구로 활용될 수 있을 것으로 판단된다.

감 사

본 연구는 산업통상자원부 산하 한국산업기술진흥원과 대경지역 사업평가원의 대경광역경제권 선도산업육성사업(과제번호 R0001668)의 지원으로 수행되었습니다.

References

1. Veziroglu, A. and Macario, R., "Fuel Cell Vehicles: State of the Art with Economic and Environmental Concerns," *Int. J. Hydrog. Energy*, **36**, 25-43(2011).
2. Wang, C. Y., "Fundamental Models for Fuel Cell Engineering," *Chem. Rev.*, **104**, 4727-4766(2004).
3. Ding, Y., Bi, X. T. and Wilkinson, D. P., "Numerical Investigation of the Impact of Two-Phase Flow Maldistribution on PEM Fuel Cell Performance," *Int. J. Hydrog. Energy*, **39**, 469-480(2014).
4. Han, I.-S., Lim, J., Jeong, J. and Shin, H. K., "Effect of Serpentine Flow-Field Designs on Performance of PEMFC Stacks for Micro-CHP Systems," *Renew. Energy*, **54**, 180-188(2013).
5. Chung, H., Ha, T., Kim, H. and Han, C., "Simulation of PEM Fuel Cell with 2D Steady-State Model," *Korean Chem. Eng. Res.*, **46**, 915-921(2008).
6. Jeong, J., Han, I.-S. and Shin, H. K., "Optimal Sizing of the Manifolds

- in a PEM Fuel Cell Stack using Three-Dimensional CFD Simulations; *Trans. Korean Hydrogen & New Energy Soc.*, **24**, 386-392 (2013).
7. Guo, N., Leu, M. C. and Koylu, U. O., "Network based Optimization Model for Pin-Type Flow Field of Polymer Electrolyte Membrane Fuel Cell;" *Int. J. Hydrog. Energy*, **38**, 6750-6761(2013).
 8. Hou, Y., Yang, Z. and Wan, G., "An Improved Dynamic Voltage Model of PEM Fuel Cell Stack;" *Int. J. Hydrog. Energy*, **35**, 11154-11160 (2010).
 9. Zhao, Y. and Pistikopoulos, E., "Dynamic Modeling and Parametric Control for the Polymer Electrolyte Membrane Fuel Cell System;" *J. Power Sources*, **232**, 270-278(2013).
 10. Khadom, A. A., "Modeling of Corrosion Reaction Data in Inhibited Acid Environment using Regressions and Artificial Neural Networks;" *Korean J. Chem. Eng.*, **30**, 2197-2204(2013).
 11. Saengrungs, A., Abtahi, A. and Zilouchian, A., "Neural Network Model for a Commercial PEM Fuel Cell System;" *J. Power Sources*, **172**, 749-759(2007).
 12. Li, X., Cao, G. and Zhu, X., "Modeling and Control of PEMFC Based on Least Squares Support Vector Machines;" *Energy Conv. Manag.*, **47**, 1032-1050(2006).
 13. Zhong, Z., Zhu, X. and Cao, G., "Modeling a PEMFC by a Support Vector Machine;" *J. Power Sources*, **160**, 293-298(2006).
 14. Petrone, R., Zheng, Z., Hissel, D., Pera, M. C., Pianese, C., Sorrentino, M., Becherif, M. and Yousfi-Steiner, N., "A Review on Model-based Diagnosis Methodologies for PEMFCs;" *Int. J. Hydrog. Energy*, **38**, 7077-7091(2013).
 15. Napoli, G., Ferraro, M., Sergi, F., Brunaccini, G. and Antonucci, V., "Data Driven Models for a PEM Fuel Cell Stack Performance Prediction;" *Int. J. Hydrog. Energy*, **38**, 11628-11638(2013).
 16. Hua, J., Li, J., Ouyang, M., Lu, L. and Xu, L., "Proton Exchange Membrane Fuel Cell System Diagnosis based on the Multivariate Statistical Method;" *Int. J. Hydrog. Energy*, **36**, 9896-9905(2011).
 17. Wold, S., Sjostrom, M. and Eriksson, L., "PLS-Regression: a Basic Tool of Chemometrics;" *Chemometrics Intell. Lab. Syst.*, **58**, 109-130(2001).
 18. Han, I.-S., Kim, M., Lee, C.-H., Cha, W., Ham, B.-K., Jeong, J.-H., Lee, H., Chung, C.-B. and Han, C., "Application of Partial Least Squares Methods to a Terephthalic Acid Manufacturing Process for Product Quality Control;" *Korean J. Chem. Eng.*, **20**, 977-984(2003).
 19. Han, I.-S., Han, C. and Chung, C.-B., "Melt Index Modeling with Support Vector Machines, Partial Least Squares, and Artificial Neural Networks;" *J. Appl. Polym. Sci.*, **95**, 967-974(2004).
 20. Han, I.-S. and Han, C., "Modeling of Multistage Air-Compression Systems in Chemical Processes;" *Ind. Eng. Chem. Res.*, **42**, 2209-2218(2003).
 21. Min, K. G., Han, I.-S. and Han, C., "Iterative Error-based Non-linear PLS Method for Nonlinear Chemical Process Modeling;" *J. Chem. Eng. Japan*, **35**, 613-625(2002).
 22. Geladi, P. and Kowalski, B., "Partial Least-Squares Regression: a Tutorial;" *Anal. Chim. Acta*, **185**, 1-17(1986).
 23. Kalogirou, S. A., "Artificial Neural Networks in Renewable Energy Systems Applications: a Review;" *Renew. Sust. Energ. Rev.*, **5**, 373-401(2001).
 24. Hagan, M. T., Demuth, H. B. and Beale, M., *Neural Network Design*, PWS Publishing, Boston, MA(1996).
 25. Hornik, K., Stinchcombe, M. and White, H., "Multilayer Feed-forward Networks are Universal Approximators;" *Neural Networks*, **2**, 359-366(1989).
 26. Han, I.-S., Jeong, J., Kho, B. K., Choi, C. H., Yu, S. and Shin, H. K., "Development of a 25 kW-Class PEM Fuel Cell System for the Propulsion of a Leisure Boat;" *Trans. Korean Hydrogen & New Energy Soc.*, **25**, 271-279(2014).
 27. Han, I.-S., Jeong, J. and Shin, H. K., "PEM Fuel-Cell Stack Design for Improved Fuel Utilization;" *Int. J. Hydrog. Energy*, **38**, 11996-12006(2013).
 28. Andersen, C. M. and Bro, R., "Variable Selection in Regression - Tutorial;" *J. Chemometr.*, **24**, 728-737(2010).
 29. Chong, I.-G. and Jun, C.-H., "Performance of Some Variable Selection Methods When Multicollinearity is Present;" *Chemometrics Intell. Lab. Syst.*, **78**, 103-112(2005).