

기상자료를 이용한 기계학습모델 기반 태양광 발전량 예측

서영민 · 이병준 · 최윤영[†]

경북대학교 건설환경공학과

Machine Learning Model-based Photovoltaic Power Generation Forecasting Using Meteorological Data

Young-Min Seo · Byeong-Jun Lee · Yun-Young Choi[†]

Dept. of Constructional and Environmental Engineering, Kyungpook National University

(2017년 5월 1일 접수, 6월 21일 심사 및 수정 완료, 6월 23일 채택)

Abstract

This study investigates the performance of machine learning-based photovoltaic power generation forecasting models using meteorological data. Three machine learning models, artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS) and generalized regression neural network (GRNN) models, are developed utilizing photovoltaic power generation and meteorological data, and then the model performances are compared based on statistical model efficiency indices. The machine learning models with photovoltaic power generation and meteorological variables, especially solar radiation and sunshine duration, as inputs perform better than those with only photovoltaic power generation variable. ANN models yield better forecasting results than ANFIS and GRNN models. Among all models, ANN3 model with photovoltaic power generation, solar radiation and sunshine duration as input variables produces the best performance based on model efficiency indices. Therefore, ANN model using solar radiation and sunshine duration among meteorological variables can be an effective tool for reliable photovoltaic power generation forecasting.

Key words : Photovoltaic power generation forecasting, Artificial neural network, Adaptive neuro-fuzzy inference system, Generalized regression neural network

요지

본 연구에서는 기상자료를 이용한 기계학습 기반 태양광 발전량 예측모델의 성능을 조사하였다. 태양광 발전량 및 기상자료를 활용하여 세 가지 기계학습모델, 즉 artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), generalized regression neural network (GRNN) 모델을 구축하였으며, 통계학적 모델성능 평가지표에 근거하여 모델성능을 비교하였다. 태양광 발전량 및 기상변수, 특히 일사량 및 일조시간을 입력으로 하는 기계학습모델들은 태양광 발전량만을 이용한 기계학습모델보다 더 우수한 성능을 나타냈다. ANN 모델은 ANFIS 및 GRNN 모델보다 더 우수한 예측결과를 산출하였다. 모든 비교모델 중에서 태양광 발전량, 일사량 및 일조시간을 입력변수로 하는 ANN3 모델은 모델성능 평가지표에 근거하여 가장 우수한 성능을 나타냈다.

[†]Corresponding author

따라서 기상변수 중 일사량과 일조시간을 활용하는 ANN 모델은 신뢰성 있는 태양광 발전량 예측을 위한 효과적인 도구가 될 수 있다.

핵심용어 : 태양광 발전량 예측, 인공지능망, 적응 뉴로퍼지 추론 시스템, 일반화귀신경망

I. 서론

세계 인구 및 경제 규모가 확대됨에 따라 이용 가능한 에너지 자원은 점차 고갈되고 있는 실정이다. 따라서 재생가능한 에너지 개발과 지속가능한 사회건설은 현재 인류에게 가장 중요한 문제 중 하나로서 대두되고 있다. 재생 가능한 에너지 자원은 사회적 에너지 수요에 부합하는 잠재적 지속가능한 해법을 제공한다.¹⁾ 특히, 재생 가능한 에너지 중 태양광 발전은 현재 수요와 공급 측면에서 가장 급속한 증가를 보여주고 있다.^{2~4)} 태양광 발전량은 태양 에너지에 의존하는 그 특성상 기상조건 (기온, 전운량, 일사량 등) 및 대기조건 (먼지, 기타 이물질 등)에 따라 발전량의 변동이 심하며,^{4,5)} 강한 비선형 특성을 가지기 때문에 단일 변수만을 이용하여 그 값을 신뢰성 있게 예측하기는 매우 어렵다.

최근 artificial neural network (ANN), radial basis function neural network (RBFNN), support vector machine (SVM) 등을 포함하는 기계학습모델은 태양광 발전 시스템에 대한 비선형 모델링에 성공적으로 적용되어 왔다.^{1,4~9)} 기계학습모델을 이용한 태양광 발전량 예측과 관련된 주요 연구들을 살펴보면, Fernandez-Jimenez 등⁶⁾은 기상예측모델과 ANN 모델을 이용하여 단기 태양광 발전량을 예측하였다. Shi 등⁷⁾은 기상분류 및 SVM 모델을 활용하여 태양광 발전량을 예측하였다. Cha 등⁸⁾은 태양광 발전요소 데이터 및 퍼지 알고리즘을 이용하여 발전효율 예측모델을 설계하였다. Giorgi 등⁹⁾은 웨이블릿 (wavelet), ANN, SVM을 이용하여 태양광 발전량을 예측하였다. Lee⁴⁾는 기상 정보의 이동평균을 고려한 태양광 발전량 예측을 위하여 ANN, SVM, RBFNN 모델을 적용하였다. Lee와 Lee⁵⁾는 기계학습을 이용한 태양광 발전량 예측 및 결합 검출 시스템 개발을 위하여 ANN, SVM, 딥러닝 (deep learning)

모델을 적용하였다. Cervone 등¹⁾은 ANN과 analog ensemble (AnEn)을 이용한 단기 태양광 예측모델을 제시하였다. 이와 같이 기존의 연구들은 주로 외국의 기상자료와 태양광 발전량 자료를 이용하여 예측모델을 구성하고 있다는 점에서 국내 기상조건 및 태양광 발전 상황에 적합한 예측모델 개발이 필요함을 시사하고 있다.

따라서 본 연구에서는 국내 태양광 발전소의 태양광 발전량과 인근 기상관측소에서 관측된 기상자료를 이용한 기계학습 기반 태양광 발전량 예측모델을 비교하였다. 영암태양광발전소의 태양광 발전량과 인근 목포 기상대의 기상자료를 활용하여 세 가지 기계학습모델, 즉 ANN, adaptive neuro-fuzzy inference system (ANFIS), generalized regression neural network (GRNN) 모델을 구축하였으며, 통계학적 모델평가지표를 이용하여 태양광 발전량 예측에 대한 모델성능을 비교·평가하였다.

II. 자료 및 방법

2.1. 분석자료

Fig. 1은 본 연구에서 선정한 태양광 발전소 및 기상 관측소의 위치를 나타내고 있다. 본 연구에서는 기상자료를 이용한 기계학습 기반 태양광 발전량 예측모델을 구축하기 위하여 연구대상으로서 영암태양광발전소 (2차)를 선정하였다. 전남 영암군 삼호읍에 위치한 영암태양광발전소 (2차)는 2009년에 준공되어 한국농어촌공사에 의해 현재 운영되고 있으며, 1491.6 kW의 발전용량을 가지고 있다 (<http://www.knrec.or.kr>, accessed on April 23, 2017). 태양광 발전량 자료는 행정자치부에 의해 운영되고 있는 공공데이터포털 (<http://www.data.go.kr>, accessed on April 23, 2017)에 공개된 2013 ~ 2015년의 시간단위 자

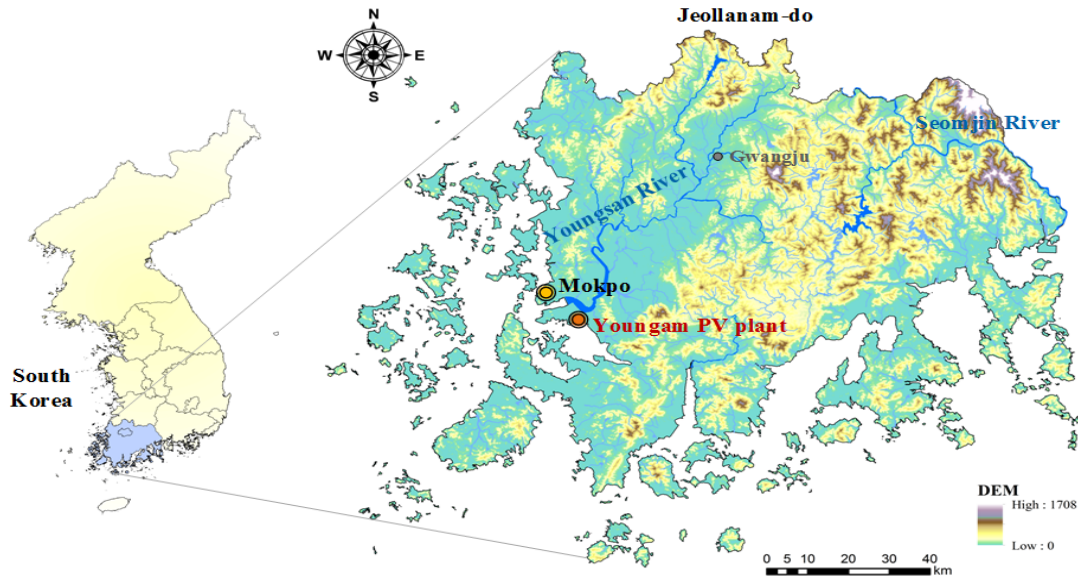


Fig. 1. Location map for photovoltaic power plant and weather station.

료를 활용하였다. 기상자료는 영암태양광발전소 (2차)에 가장 인접해 있는 목포 기상대 (기상청)로부터 동일한 기간에 대하여 수집되었다. 기상자료는 기온, 강우량, 풍속, 습도, 증기압, 이슬점 온도, 일조시간, 일사량, 전운량, 시정, 지면온도가 수집되었으며, 태양광 발전량 자료와 마찬가지로 시간단위 자료가 수집되었다. 이 중 태양광 발전량과 상관성이 높은 지면온도, 일사량, 일조시간, 습도를 입력집합 구성에 활용하였다. 선택된 자료 중 2013년 1월 ~ 2015년 3월 사이의 자료 (75%)는 모델 학습자료로 사용되었으며, 2015년 4월 ~ 2015년 12월 사이의 자료 (25%)는 모델 검증자료로 사용되었다.

2.2. 분석방법

2.2.1. ANN 모델

ANN 모델은 복잡한 비선형 및 다차원 관계를 모델링하기 위한 다층계산시스템이다. Fig. 2는 ANN 모델 구조의 예를 나타내고 있다. ANN 구조 중에서 가장 널리 적용되고 있는 다층퍼셉트론 (multilayer perceptron, MLP)은 일반적으로 3개의 층, 즉 입력층 (input layer), 은닉층 (hidden layer), 출력층 (output layer)으로 구성된다.

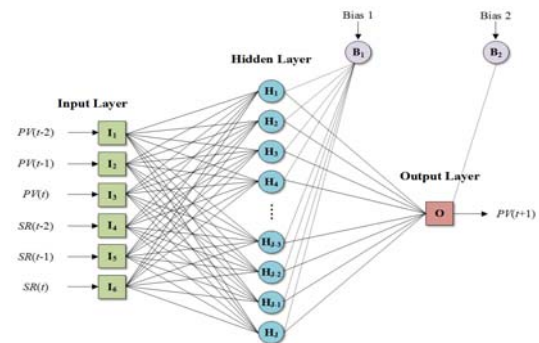


Fig. 2. ANN model architecture.

Günther와 Fritsch¹⁰⁾에 따르면, J 개의 은닉 뉴런(hidden neuron)을 가지는 3층 구조의 MLP는 식 (1)과 같이 출력값을 산정한다.

$$\begin{aligned} o(\mathbf{x}) &= f\left(w_0 + \sum_{j=1}^J w_j \cdot \left(w_{0j} + \sum_{i=1}^n w_{ij} x_i\right)\right) \\ &= f\left(w_0 + \sum_{j=1}^J w_j \cdot f\left(w_{0j} + \mathbf{w}_j^T \mathbf{x}\right)\right) \end{aligned} \quad (1)$$

여기서, w_0 는 출력뉴런(output neuron)의 절편, w_{0j} 는 j 번째 은닉뉴런의 절편, w_j 는 j 번째 연결가중치(connection weight), $\mathbf{w}_j = (w_{1j}, \dots, w_{nj})$ 는 j 번째 연결가중치 벡터, $\mathbf{x} = (x_1, \dots, x_n)$ 은 입

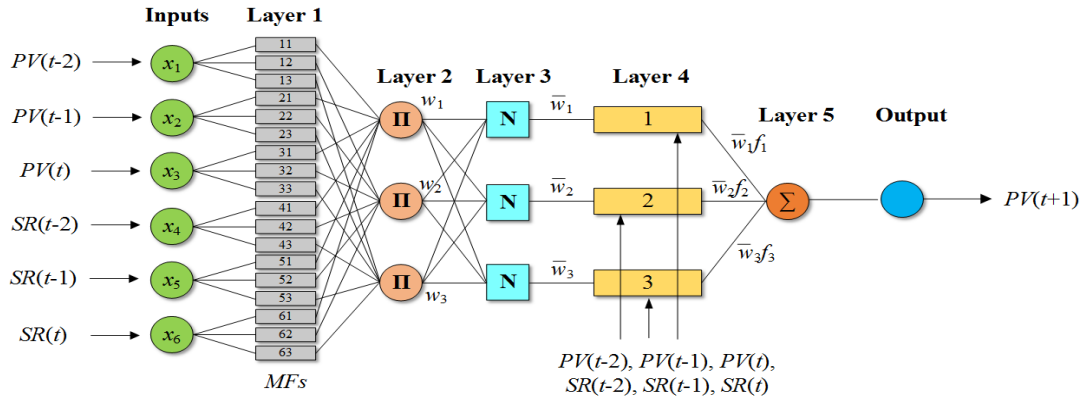


Fig. 3. ANFIS model architecture.

력벡터이고 f 는 활성화 함수(activation function)이다. ANN에 대한 자세한 내용은 Günther와 Fritsch¹⁰⁾를 참조할 수 있다.

2.2.2. ANFIS 모델

퍼지규칙기반 시스템 (fuzzy rule-based system)의 일종인 ANFIS는 퍼지추론시스템 (fuzzy inference system, FIS)과 적응신경망 (adaptive neural network)을 결합한 수학적 모델이다. ANFIS 모델에 대한 자세한 내용은 Jang 등¹¹⁾을 참조할 수 있다.

Fig. 3은 ANFIS 모델 구조의 예를 나타내고 있다. ANFIS 모델은 일반적으로 5개의 층으로 구성된다. 첫 번째 층은 퍼지화 단계 (fuzzification stage)로서 소속함수 (membership function)를 이용하여 크리스프 값 (crisp values)을 언어항 (linguistic terms)으로 변환한다. 두 번째 층은 t-norm 연산자를 이용한 추론단계이고, 세 번째 층은 퍼지 IF-THEN 규칙에 대한 강도비를 계산한다. 네 번째 층에서는 결과부 매개변수들이 산정된다. 마지막으로 다섯 번째 층은 네 번째 층에서 산정된 모든 신호들을 합산함으로써 최종 결과를 산출한다. k 개의 입력과 m 개의 소속함수를 가지는 ANFIS 모델의 각 층은 각각 식 (2)~(6)과 같은 계산을 수행한다.¹¹⁾

$$O_{1,j} = \varphi_{A_{ij}}(x_i), \quad (2)$$

$$i = 1, 2, \dots, k, \quad j = 1, 2, \dots, m$$

$$O_{2,j} = w_j = \prod_{i=1}^k \varphi_{A_{ij}}(x_i) \quad (3)$$

$$O_{3,j} = \bar{w}_j = \frac{w_j}{\sum_{q=1}^m w_q} \quad (4)$$

$$O_{4,j} = \bar{w}_j f_j = \bar{w}_j \left(\sum_{i=1}^k p_{ji} x_i + r_j \right) \quad (5)$$

$$O_5 = \sum_{j=1}^m \bar{w}_j f_j = \frac{\sum_{j=1}^m w_j f_j}{\sum_{j=1}^m w_j} \quad (6)$$

여기서, O_{1-5} 는 층 1 ~ 5의 출력, $\varphi_{A_{ij}}$ 는 소속 함수, A_{ij} 는 언어라벨 (linguistic label), w_j 는 규칙강도 (rule strength), \bar{w}_j 는 정규화 규칙강도 (normalized rule strength), p_{ji} 와 r_j 는 결과부 매개변수이다.

2.2.3. GRNN 모델

RBFNN을 개선한 모델인 GRNN 모델은 복잡한 비선형 관계를 모델링하기 위한 비선형 회귀 기반 범용근사기 (universal approximator)이다. GRNN 모델은 원패스 학습 알고리즘 (one-pass learning algorithm)에 기초하기 때문에 ANN 모델과는 달리 반복적인 학습과정을 필요로 하지 않는다.¹²⁾

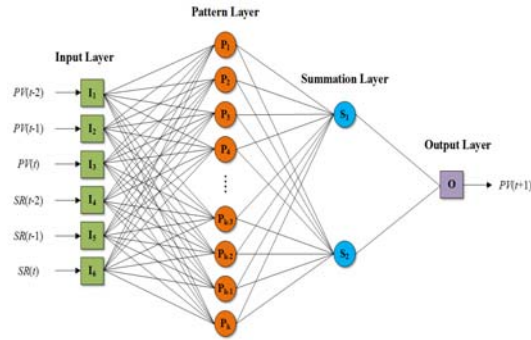


Fig. 4. GRNN model architecture.

Fig. 4는 GRNN 모델 구조의 예를 나타내고 있다. GRNN 모델은 4개의 층, 즉 입력층(input layer), 패턴층(pattern layer), 합산층(summation layer), 출력층(output layer)으로 구성된다. GRNN 모델의 첫 번째 층인 입력층은 입력벡터를 입력 받는 입력노드들로 구성된다. 두 번째 층인 패턴 층은 식 (7)과 같이 패턴 가중치와 입력벡터 사이의 유클리드 거리(Euclidean distance)를 산정하고, 다음으로 식 (8)과 같은 radial basis function (RBF) 커널(kernel)을 이용하여 노드 출력을 산출한다.¹²⁾

$$D_j = \sqrt{\sum_{i=1}^n (w_{ij} - x_i)^2}, \quad (7)$$

$$i = 1, 2, \dots, n, \quad j = 1, 2, \dots, k$$

$$f(D_j) = \exp\left(\frac{-D_j}{2\sigma^2}\right) \quad (8)$$

여기서, D_j 는 패턴 가중치(pattern weight)와 입력벡터 사이의 유클리드 거리, x_i 는 입력벡터의 원소, n 은 학습패턴의 개수, w_{ij} 는 패턴 가중치, k 는 입력벡터의 크기, f 는 RBF 커널, σ 는 평활 인자(smoothing factor)이다.

세 번째 층인 합산층은 두 개의 노드, 즉 S_1 과 S_2 로 구성되며, 식 (9)~(10)과 같이 패턴층의 출력에 대한 가중합 및 비가중합을 각각 산정한다.¹²⁾

$$S_1 = \sum_{j=1}^k y_j f(D_j) \quad (9)$$

$$S_2 = \sum_{j=1}^k f(D_j) \quad (10)$$

여기서, y_j 는 S_1 노드와 패턴층 사이의 연결 가중치이다. 마지막으로, 출력층은 S_1 노드의 출력을 S_2 노드의 출력으로 나누어 최종 출력값을 산출한다. GRNN 모델에 대한 자세한 모델은 Specht¹²⁾를 참조할 수 있다.

2.2.4. 통계학적 모델성능 평가지표

본 연구에서 구축한 태양광 발전량 예측모델의 성능은 통계학적 모델성능 평가지표를 이용하여 비교·평가되었다. 통계학적 모델성능 평가지표는 효율성계수(coefficient of efficiency, CE), 일치지수(index of agreement, d), 결정계수(co-efficient of determination, r^2), 제곱근 평균제곱오차(root-mean-square error, RMSE), 평균 절대오차(mean absolute error, MAE), 평균 고차오차(mean higher order error, MS4E)를 적용하였다. 각 지표들에 대한 정의는 식 (11)~(16)과 같다.¹³⁾

$$CE = 1 - \frac{\sum_{i=1}^N (PV_i - PV_i^*)^2}{\sum_{i=1}^N (PV_i - \overline{PV})^2} \quad (11)$$

$$d = 1 - \frac{\sum_{i=1}^N (PV_i - PV_i^*)^2}{\sum_{i=1}^N (|PV_i^* - \overline{PV}| + |PV_i - \overline{PV}|)^2} \quad (12)$$

$$r^2 = \left| \frac{\sum_{i=1}^N (PV_i - \overline{PV})(PV_i^* - \overline{PV})}{\sqrt{\sum_{i=1}^N (PV_i - \overline{PV})^2 \sum_{i=1}^N (PV_i^* - \overline{PV})^2}} \right|^2 \quad (13)$$

$$RMSE = \left\{ \frac{1}{N} \sum_{i=1}^N [PV_i^* - PV_i]^2 \right\}^{0.5} \quad (14)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |PV_i^* - PV_i| \quad (15)$$

Table 1. Combination of input variables for model configuration

Input sets	Input variables	Output variables
Set 1	$PV(t-2), PV(t-1), PV(t)$	$PV(t+1)$
Set 2	$PV(t-2), PV(t-1), PV(t), SR(t-2), SR(t-1), SR(t)$	$PV(t+1)$
Set 3	$PV(t-2), PV(t-1), PV(t), SR(t-2), SR(t-1), SR(t), SD(t-2), SD(t-1), SD(t)$	$PV(t+1)$
Set 4	$PV(t-2), PV(t-1), PV(t), SR(t-2), SR(t-1), SR(t), SD(t-2), SD(t-1), SD(t), H(t-2), H(t-1), H(t)$	$PV(t+1)$
Set 5	$PV(t-2), PV(t-1), PV(t), SR(t-2), SR(t-1), SR(t), SD(t-2), SD(t-1), SD(t), H(t-2), H(t-1), H(t), ST(t-1), ST(t)$	$PV(t+1)$

Note. SR : solar radiation, SD : sunshine duration, H : humidity, ST : surface temperature

$$MSAE = \frac{1}{N} \sum_{i=1}^N (PV_i - PV_i^*)^4 \quad (16)$$

여기서, PV_i^* 는 태양광 발전량 예측치, PV_i 는 태양광 발전량 관측치, \overline{PV} 는 태양광 발전량 관측치의 평균, \widehat{PV} 는 태양광 발전량 예측치의 평균이며, N 은 태양광 발전량 관측자료의 개수를 나타낸다.

III. 결과 및 고찰

3.1. 모델 구축

태양광 발전량 예측을 위한 기계학습모델의 입력변수는 Seo 등¹⁴⁾이 설명한 통계학적 방법을 활용하여 결정하였다. 먼저 기상자료와 태양광 발전량 자료 사이의 피어슨 상관계수 (Pearson's correlation coefficient, r)를 산정하여 그 절대값이 0.5 이상인 기상변수들을 선별하였다. 다음으로 자기상관함수 (autocorrelation function), 부분자기상관함수 (partial autocorrelation function), 교차상관함수 (cross correlation function), 평균상호정보 (average mutual information)를 적용하여 각 기상변수별 최적 시차 (lag)를 결정하였다. 그 결과, 상관분석으로부터 지면온도 ($r = 0.501$), 일사량 ($r = 0.899$), 일조시간 ($r = 0.801$), 습도 ($r = -0.529$)가 선택되었으며, 최적 시차 결정 및 입력변수 조합을 통해 Table 1과 같이 5개의 입력변수 집합이 결정되었다. 선택

된 입력변수들은 기계학습모델의 효과적인 학습을 위하여 $[0, 1]$ 의 범위로 정규화 (normalization)되었다.¹³⁾

본 연구에서는 ANN 모델을 이용한 태양광 발전량 예측을 위하여 3개의 층으로 구성된 MLP 구조의 ANN 모델을 적용하였다. ANN 모델링에서 은닉뉴런의 개수는 Seo 등¹⁴⁾에서 제시된 시행착오법을 적용하여 결정되었다. 즉, 은닉뉴런의 개수를 점진적으로 변화시키면서 모델의 평균 제곱오차 (mean square error (MSE)) 값을 산정하고, MSE 값이 최소가 되는 은닉뉴런의 개수를 최적 은닉뉴런 개수로 결정하였다. 뉴런 출력을 산정하기 위한 활성화 함수는 ANN 모델링에서 일반적으로 적용되고 있는 로지스틱 시그모이드 함수 (logistic sigmoid function)를 적용하였으며, ANN 모델은 역전파 알고리즘 (back-propagation algorithm)을 이용하여 학습되었다.

태양광 발전량 예측을 위한 ANFIS 모델링에서, FIS는 퍼지 c-평균 군집 (fuzzy c-means clustering, FCM) 알고리즘¹⁵⁾을 이용하여 생성되었다. 시행착오적 분석을 바탕으로 소속함수의 유형 및 개수로서 입력노드별 3개의 가우시안 소속함수 (Gaussian membership function)를 적용하였으며, 기타 모델 설정은 MathWorks¹⁶⁾에서 제시하고 있는 기본설정을 활용하였다. ANFIS 모델은 경사하강최적화 (gradient descent optimization)와 최소제곱법 (least-squares method)을 결합한 하이브리드 학습 알고리즘 (hybrid learning algorithm)을 이용하여 학습되었다.

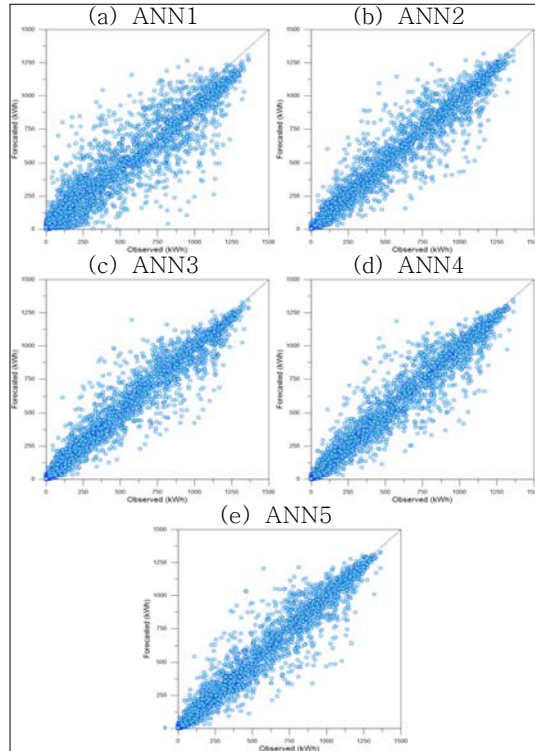


Fig. 5. Scatter plots for ANN models.

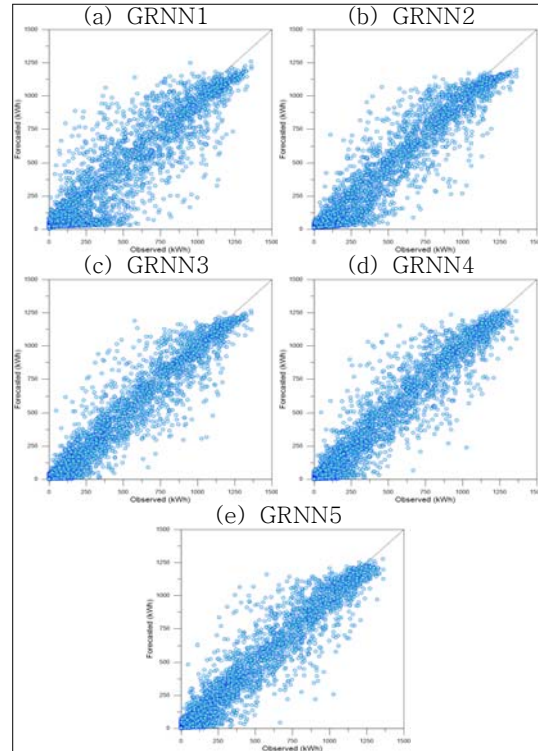


Fig. 7. Scatter plots for GRNN models.

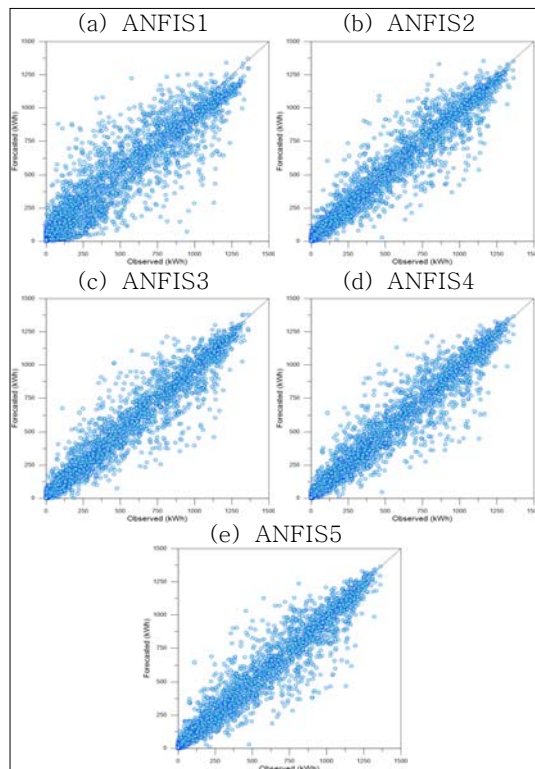


Fig. 6. Scatter plots for ANFIS models.

한편, GRNN 모델링에서 가장 중요한 단계는 RBF 커널의 평활인자를 결정하는 것이다. 평활 인자는 시행착오법과 최적화 알고리즘을 이용하여 결정될 수 있다. 본 연구에서는 시행착오법을 이용하여 평활인자를 결정하였다. 즉, 평활인자 값을 단계적으로 변화시키면서 모델의 MSE를 구하여 그 값이 최소가 되는 평활인자를 최적 평활 인자로 결정하였다.

3.2. 모델성능평가

본 연구에서 태양광 발전량 예측 모델의 성능은 산점도 (scatter plot) 및 통계학적 모델성능 평가지표를 이용하여 비교·평가되었다. Figs. 5~7은 ANN, ANFIS, GRNN 모델에 대한 태양광 발전량의 예측치와 관측치를 도시한 산점도를 나타내고 있다. 산점도에서 1:1 경사 (수평:수직) 직선 (즉, $y=x$ 직선) 주위에 자료점들이 가까이 위치할수록 모델 성능이 우수함을 나타낸다. Fig. 5로부터 ANN1 모델 (입력집합 Set 1을 입력으로 사용한 ANN 모델)은 다른 ANN 모델에

Table 2. Performance evaluation for ANN, ANFIS and GRNN models

Models		CE	d	r^2	RMSE (kWh)	MAE (kWh)	MS4E ($\times 10^8$ kWh ⁴)
ANN	ANN1	0.914	0.976	0.917	107.202	58.381	17.973
	ANN2	0.961	0.990	0.961	71.996	33.461	5.228
	ANN3	0.964	0.991	0.964	69.485	33.428	4.326
	ANN4	0.963	0.990	0.963	70.322	33.444	4.619
	ANN5	0.964	0.991	0.964	69.552	33.217	4.630
ANFIS	ANFIS1	0.909	0.976	0.909	110.036	62.257	18.882
	ANFIS2	0.958	0.989	0.958	74.579	34.654	5.429
	ANFIS3	0.957	0.989	0.958	75.228	34.808	5.746
	ANFIS4	0.958	0.989	0.958	75.205	35.376	5.621
	ANFIS5	0.958	0.980	0.958	74.862	34.723	5.513
GRNN	GRNN1	0.892	0.971	0.893	119.845	70.419	21.457
	GRNN2	0.935	0.983	0.936	93.150	52.489	8.986
	GRNN3	0.945	0.986	0.946	85.251	41.626	8.259
	GRNN4	0.942	0.985	0.943	87.542	43.952	8.447
	GRNN5	0.943	0.985	0.944	86.828	44.752	8.262

Note. For example, ANN1 stands for ANN model using input set, Set 1.

비해 1:1 경사 직선 주위에 자료점들이 더 넓게 퍼져 있으며, ANN2 ~ ANN5 모델은 유사한 분산도를 나타냄을 확인할 수 있다. 이로부터 ANN2 ~ ANN5 모델은 ANN1 모델에 비해 우수한 모델 성능을 나타냄을 알 수 있으며, Figs. 6 ~ 7로부터 ANFIS 및 GRNN 모델 또한 이러한 경향을 나타냄을 알 수 있다. 즉, ANFIS2 ~ ANFIS5 모델 및 GRNN2 ~ GRNN5 모델은 ANFIS1 및 GRNN1 모델과 비교하여 더 우수한 모델 성능을 각각 나타내는 것으로 분석되었다.

한편, 본 연구에서는 모델성능의 상세한 비교를 위하여 통계학적 모델성능 평가지표에 근거하여 비교 모델들의 모델성능을 정량적으로 비교하였다. Tables 2는 통계학적 모델성능 평가지표를 이용한 태양광 발전량 예측 모델의 모델성능 평가결과를 나타내고 있다. Table 2로부터 ANN1 모델은 다른 ANN 모델에 비하여 낮은 모델성능을 가지는 것으로 분석되었다. 즉, ANN1 모델의 CE, d , r^2 값은 다른 ANN 모델보다 낮게 나타난 반면, RMSE, MAE, MS4E 값은 높게 나타났으며, 이러한 경향은 ANFIS 및 GRNN 모델에서도 동일하게 나타났다. ANN1, ANFIS1, GRNN1 모델은 입력집합 Set 1 (Table 1)을 이용하여 구축된 모델로서 태양광 발전량만을 입력변수로 사

용한다. 이러한 결과로부터, 기계학습 모델의 입력변수로서 태양광 발전량뿐만 아니라 기상자료를 포함하는 것은 태양광 발전량만을 입력변수로 가지는 모델보다 모델 성능을 향상시킬 수 있음을 알 수 있다.

Table 2로부터 ANN3 및 ANN5 모델은 다른 ANN 모델보다 CE, d , r^2 값은 높은 반면, RMSE, MAE, MS4E 값은 낮게 나타나는 것으로 분석되었다. ANN3과 ANN5 모델의 모델성능 평가지표는 거의 비슷한 값을 가지는 것으로 나타났지만 ANN3 모델이 ANN5 모델보다 약간 더 나은 모델 성능을 보여주었다. 즉, ANN3 모델은 ANN 모델 중에서 가장 우수한 모델성능을 나타내는 것으로 분석되었다. Table 2로부터 ANFIS2 모델은 다른 ANFIS 모델에 비해 높은 CE, d , r^2 값을 가지는 반면, 낮은 RMSE, MAE, MS4E 값을 나타내는 것으로 분석되었다. ANFIS2 모델은 ANFIS 모델 중에서 가장 우수한 모델 성능을 나타내었다. 그리고 GRNN3 모델은 다른 GRNN 모델에 비해 CE, d , r^2 값은 높게 나타난 반면, RMSE, MAE, MS4E 값은 낮은 결과를 보여주었다. 즉, GRNN3 모델은 GRNN 모델 중에서 가장 우수한 모델성능을 나타내는 것으로 분석되었다. 이러한 결과로부터, ANN3, ANFIS2, GRNN3

모델은 각각 ANN, ANFIS, GRNN 모델 중에서 가장 우수한 모델성능을 나타냄을 알 수 있다. ANFIS2는 입력집합 Set 2, ANN3 및 GRNN3은 Set 3을 이용하여 구축된 모델이다. Set 2는 태양광 발전량과 일사량으로 구성되고, Set 3는 태양광 발전량, 일사량, 일조시간으로 구성된다. 따라서 태양광 발전량 예측모델에서 태양광 발전량, 일사량, 일조시간을 입력변수로 가지는 기계학습모델은 다른 모델에 비해 우수한 모델 성능을 나타냄을 알 수 있으며, 지면온도와 습도는 모델 성능에 미치는 영향이 미미함을 알 수 있다.

Tables 2로부터 ANFIS1 모델을 제외한 모든 ANFIS 모델은 GRNN 모델보다 우수한 모델 성능을 나타내었으며, ANN1 모델을 제외한 모든 ANN 모델은 ANFIS 모델보다 우수한 모델 성능을 나타내었다. 즉, 태양광 발전량 예측모델은 대체로 ANN, ANFIS, GRNN 모델 순서로 우수한 모델 성능을 나타냄을 알 수 있다. 또한 ANN3 모델은 모든 비교모델 중에서 가장 우수한 모델 성능을 나타내는 것으로 분석되었다. 따라서 비교모델 중에서 태양광 발전량 예측을 위한 최적 모델은 입력변수로서 태양광 발전량, 일사량, 일조시간을 가지는 ANN3 모델인 것으로 나타났다. 비록 비교모델 간의 성능이 현격하게 차이 나는 것은 아니지만 시간단위 예측에서 발생할 수 있는 오차의 누적 또는 전파를 고려했을 때 가장 우수한 성능을 가지는 모델을 실무측면에서 적용하는 것이 합리적이라 판단된다. 또한 태양광 발전량 예측에서 성능의 차이가 결과적으로 경제적인 효과로 귀결되므로 시간단위 예측에서의 오차의 누적과 예측에 따른 경제적 효과를 고려하여 다른 모델보다 우수한 태양광 발전량, 일사량, 일조시간을 입력변수로 가지는 ANN3 모델을 적용하는 것이 합리적인 것으로 판단된다.

IV. 결 론

본 연구에서는 기상자료를 이용한 기계학습 기반 태양광 발전량 예측모델을 비교하였다. 이를 위하여 영암태양광발전소의 태양광 발전량과 목

포 기상대의 기상자료를 활용하여 세 가지 기계학습모델, 즉 ANN, ANFIS, GRNN 모델을 구축하였으며, 통계학적 모델평가지표를 이용하여 태양광 발전량 예측에 대한 모델성능을 비교·평가하였다. 그 결과, 기계학습 모델의 입력변수로서 태양광 발전량뿐만 아니라 기상자료를 포함하는 것은 태양광 발전량만을 입력변수로 가지는 모델보다 모델 성능을 향상시킴을 확인할 수 있었다. 이 중 태양광 발전량, 일사량, 일조시간을 입력변수로 가지는 기계학습모델이 다른 모델들에 비해 우수한 모델 성능을 나타내었다. 태양광 발전량 예측모델의 성능은 대체로 ANN, ANFIS, GRNN 모델의 순서로 우수한 모델 성능을 나타내었다. 비교모델 중에서 태양광 발전량 예측을 위한 최적 모델은 입력변수로서 태양광 발전량, 일사량, 일조시간을 가지는 ANN3 모델인 것으로 분석되었다. ANN3 모델은 시간단위 예측에서의 오차의 누적 및 예측에 따른 경제적 효과를 고려했을 때 다른 모델보다 적용성이 우수하다고 판단된다. 따라서 기상자료로서 일사량 및 일조시간을 이용한 ANN 모델은 신뢰성 있는 태양광 발전량 예측을 위한 효과적인 도구로 활용될 수 있을 것으로 판단된다.

References

1. Cervone, G., Clemente-Harding, L., Alessandrini, S., Monache, L. D., Short-term photovoltaic power forecasting using artificial neural networks and an analog ensemble, *Renewable Energy*, 108, 274~286, 2017.
2. Kim, J. Y., June, S. J., Lee, J. E., Environmental assessment in the siting of solar and wind power plants, *Journal of Korean Society of Environmental Technology*, 12(2), 141~147, 2011.
3. Kim, S. M., Kim, G. H., A study on the new renewable energy industry cluster network: focused on photovoltaic industry

- in Jeonnam region, Journal of Korean Society of Environmental Technology, 13(2), 145~154, 2012.
4. Lee, H. J., Use of the moving average of the current weather data for the solar power generation amount prediction, Journal of Korea Multimedia Society, 19(8), 1530~1537, 2016.
 5. Lee, S. M., Lee, W. J., Development of a system for predicting photovoltaic power generation and detecting defects using machine learning, KIPS Transactions on Computer and Communication Systems, 5(10), 353~360, 2016.
 6. Fernandez Jimenez, L. A., Muñoz Jimenez, A., Falces A., Mendoza Villena, M., Garcia Garrido, E., Lara Santillan, P. M., Zorzano Alba, E., Short term power forecasting system for photovoltaic plants, Renewable Energy, 44, 311~317, 2012.
 7. Shi, J., Lee, W. J., Liu, Y., Yang, Y., Wang, P., Forecasting power output of photovoltaic systems based on weather classification and support vector machines, IEEE Transactions on Industry Applications, 48(3), 1064~1069, 2012.
 8. Cha, W. C., Park, J., Cho, U., Kim, J. C., Design of generation efficiency fuzzy prediction model using solar power element data, The Transactions of the Korean Institute of Electrical Engineers, 63(10), 1423~1427, 2014.
 9. Giorgi, M. D., Congedo, P., Malvoni, M., Photovoltaic power forecasting using statistical method: impact of weather data, IET Science, Measurement and Technology, 8(3), 90~97, 2015.
 10. Günther, F., Fritsch, S., Neuralnet: training of neural networks, R Journal, 2(1), 30~38, 2010.
 11. Jang, J. S. R., Sun, C. T., Mizutani, E., Neuro fuzzy and soft computing: a computational approach to learning and machine intelligence, Prentice-Hall, New Jersey, 335~345, 1997.
 12. Specht, D. F., A general regression neural network, IEEE Transactions on Neural Networks, 2(6), 568~576, 1991.
 13. Dawson, C. W., Wilby, R. L., Hydrological modelling using artificial neural networks, Progress in Physical Geography, 25(1), 80~108, 2001.
 14. Seo, Y., Kim, S., Kisi, O., Singh, V. P., Daily water level forecasting using wavelet decomposition and artificial intelligence techniques, Journal of Hydrology, 520, 224~243, 2015.
 15. Bezdec, J. C., Pattern recognition with fuzzy objective function algorithms, Plenum Press, New York, 65~79, 1981.
 16. MathWorks, Fuzzy logic toolbox user's guide, The MathWorks, Inc. [http://www.mathworks.com/help/pdf_doc/fuzzy/fuzzypdf\(2016\).2016](http://www.mathworks.com/help/pdf_doc/fuzzy/fuzzypdf(2016).2016).