

군산풍력발전단지의 풍력발전량 단기에측모형 비교에 관한 연구[†]

이영섭¹ · 김진² · 장문석³ · 김현구⁴

¹²동국대학교 통계학과 · ³⁴한국에너지기술연구원

접수 2013년 4월 21일, 수정 2013년 5월 15일, 게재확정 2013년 5월 20일

요약

최근 신재생에너지와 대체에너지의 필요성이 증가함에 따라 환경오염과 온실효과를 초래하지 않는 풍력에너지 개발에 많은 연구와 투자가 이루어지고 있다. 풍력에너지는 무공해 에너지이며 자원량이 무한대이고 바람이 부는 곳이라면 어디에서든지 전력생산이 가능하다. 그러나 풍력에너지는 바람에 크게 의존하며 불규칙적인 특성이 있어 효율적인 풍력발전이 어렵다는 단점이 있다. 이러한 이유로 풍력발전에 있어서 정확한 풍력발전량 예측은 매우 중요한 요소이다. 본 연구에서는 이러한 풍력발전량의 효율적인 예측을 위해 군산 풍력단지의 자료를 이용해 시계열모형인 ARMA모형과 데이터마이닝 기법 중 신경망모형을 사용하여 풍력발전량을 예측하고 비교분석 하였다. 그 결과 신경망모형 적합결과가 ARMA모형 적합결과 보다 더 좋은 예측력을 나타내었다.

주요용어: 군산풍력단지, 시계열모형, 신경망모형, 풍력발전량 예측.

1. 서론

최근 화석에너지의 고갈에 대비한 친환경적인 대체 에너지원으로 풍력에너지가 주목을 받고 있다 (Kim과 Kim, 2011). 세계적으로 재생에너지 중에서 풍력발전이 가장 빠르게 성장하고 있다. 풍력에너지는 무공해 청정에너지이며 자원량이 무한대이고 바람이 부는 곳이라면 어디에서든지 전력생산이 가능하다. 또한 풍력발전은 온실가스와 같은 환경오염물질이 배출되지 않기 때문에 공해물질 저감효과도 매우 크다. 1Gwh 에너지의 생산에 필요한 발전 단지의 면적을 보면 풍력은 1,355m²/Gwh, 석탄은 3,642m²/Gwh, 태양열은 3,561m²/Gwh, 그리고 태양광은 3,237m²/Gwh로 풍력발전이 가장 작은 면적을 사용하기 때문에 국토의 사용측면에서 가장 효율적이라 할 수 있다 (Hwang 등, 2012).

풍력에너지는 바람에 강하게 의존하며, 불규칙적인 특성으로 인해 효율적인 풍력발전이 어렵다. 따라서 정확한 풍력발전량 예측이 풍력발전에 있어서 중요한 과제이다. 풍력발전량의 예측은 에너지생산과 전력시장에서의 안정성을 보장하기 위한 중요한 요소이기 때문에 세계적으로 여러 가지 통계적 기법을 기반으로 풍력발전량 예측모형이 개발되고 있다. 우리나라도 향후 서해상에 대규모 풍력발전단지를

[†] 본 연구는 2012년 지식경제부 재원으로 한국에너지기술연구원 (KETEP)의 연구비 지원을 받아 수행한 연구과제 결과입니다 (과제번호 2011T100100307).

¹ 교신저자: (100-715) 서울시 중구 필동로 1길 30, 동국대학교 통계학과, 교수. E-mail: yung@dongguk.edu

² (100-715) 서울시 중구 필동로 1길 30, 동국대학교 통계학과, 석사과정.

³ (695-971) 제주도 제주시 구좌읍 해맞이해안로 200, 한국에너지기술연구원 제주글로벌연구소 풍력연구실, 책임연구원.

⁴ (305-343) 대전광역시 유성구 가정로 152, 한국에너지기술연구원 신재생에너지자원센터, 책임연구원.

개발할 예정으로 개발 시에는 풍력발전예보시스템 구축이 필요하다. 풍력발전예보시스템이란 풍력발전 단지에 의해 생산되는 전력량을 통계적 기법을 이용한 예측모형을 통해 사전에 예측하는 것을 말한다 (Abraham과 Ledolter, 2005).

본 연구에서는 이러한 풍력발전예보시스템 구축을 위한 사전단계로 인근 서해안인 군산풍력발전단지의 풍력발전량에 대한 단기예측모형을 구축하고자 한다. 풍력발전량은 자료의 특성상 시계열자료에 해당되므로 시계열예측모형인 ARMA모형을 통해 예측가능하다. 따라서 본 연구에서는 ARMA모형을 통해 풍력발전량을 예측해 보고, 더불어 데이터마닝기법 중의 하나인 신경망 (neural network) 기법을 적용하여 풍력발전량 단기예측모형을 구축하고자 한다. 특히 풍력에너지의 변동특성을 은닉계층 내에서 자동으로 구성하는 장점이 있는 다층퍼셉트론 (multi-layer perceptron) 신경망을 사용하여 풍력발전량을 예측한다. 주어진 군산 풍력발전단지의 전력생산량 자료를 각각 모형구축용 자료 (training set)와 모형평가용 자료 (test set)로 분할하여, 모형구축용 자료를 통해 ARMA모형과 신경망모형을 구축하고, 모형평가용 자료를 통해 구축된 모형들을 비교 평가한다.

2. 자료 및 연구 방법

2.1. 풍력 관측 자료

풍력발전량 단기예측모형 구축을 위해 군산풍력발전단지의 2008년 3월 1일부터 2010년 2월 28일까지 시간당 풍력발전량 자료를 이용하였다. Figure 2.1은 1시간 단위 풍력발전량을 나타낸 것으로 그림에서와 같이 모형의 구축과 평가를 위해 2008년 3월 1일부터 2009년 2월 28일까지의 1년간 자료를 이용하여 모형을 구축하였으며, 나머지 2009년 3월 1일부터 2010년 2월 28일까지의 1년간 자료를 이용하여 모형을 평가하였다.

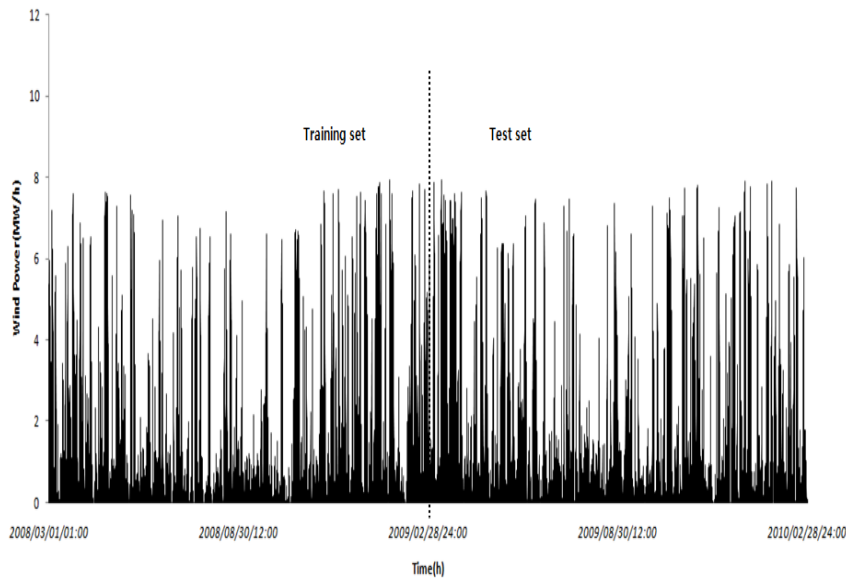


Figure 2.1 Time series of Gunsan wind power generation

2.2. 연구방법

본 연구에서는 시계열예측모형으로 ARMA모형을 구축하고, 데이터마이닝 기법인 신경망모형을 적용해 풍력발전량 단기예측모형을 구축하였다. ARMA모형은 시간이 흘러도 변하지 않는 정상적인 (stationary) 현재와 과거간의 선형관계를 찾아내어 예측을 시도하는 것이다 (Lee, 2012). Box 등(1994)이 제안한 ARMA(p,q)모형은 식 (2.1)과 같이 과거 p 차까지의 관측값이 현재의 관측값에 y_t 에 영향을 미치는 자기회귀 (autoregressive) 과정을 가지며 차수가 q 인 이동평균 (moving average) 과정을 갖는 모형이다.

$$y_t = \delta + \sum_{i=1}^p \Phi_i y_{t-i} + \sum_{j=1}^q \Theta_j \varepsilon_{t-j} + \varepsilon_t \quad (2.1)$$

여기서 δ 은 상수항이고 Φ_i 는 i 번째 자기회귀계수, Θ_j 는 j 번째 이동평균계수, ε_t 는 t 시점에서의 오차항, ε_{t-j} 는 이전 시점 $t-j$ 의 무작위 오차, p 와 q 는 각각 자기회귀와 이동평균의 순서를 나타낸다.

신경망모형에는 여러 가지 다양한 모형이 있는데 그 중에서도 가장 널리 사용되는 모형은 다층퍼셉트론 신경망이다. 은닉층이 1개인 다층퍼셉트론 신경망의 구조는 Figure 2.2와 같이 입력층 (input layer)과 1개의 은닉층 (hidden layer), 그리고 출력층 (output layer)으로 이루어져 있다. 각 은닉층의 노드와 출력 노드는 시그모이드 (sigmoid) 함수와 같은 비선형 활성화함수 (activation function)를 사용한다. 다층퍼셉트론은 은닉층의 은닉마디가 입력층으로부터 전달되는 입력값들의 선형결합을 비선형 함수로 처리하여 출력층으로 전달함으로써 상당한 유연성을 가질 수 있게 된다(Potts, 2000; Powers와 Ali, 2000). 또한 분류 값과 실제 값의 차이를 줄여나가는 반복과정을 통해 최종 분류 값에 도달하게 된다 (Oh 등, 2011).

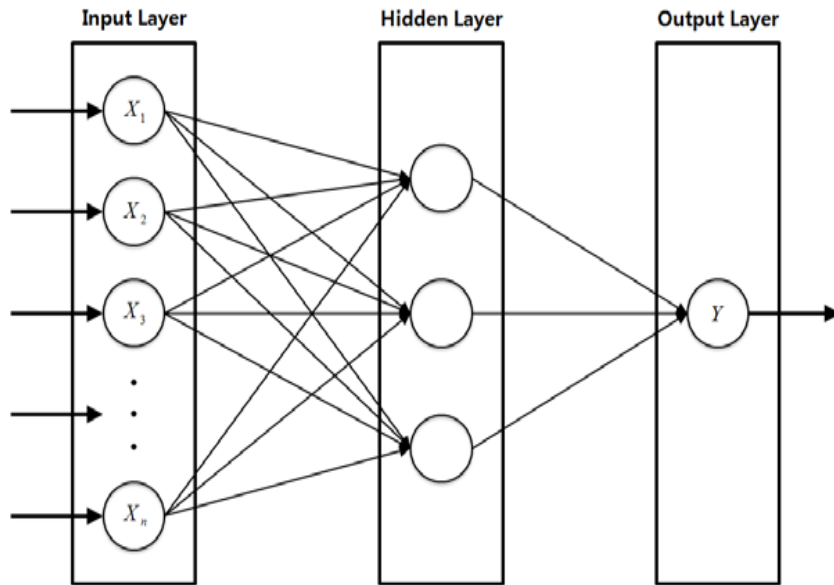


Figure 2.2 Structure of multi-layer perceptron

본 연구에서는 1개의 은닉층을 가진 다층퍼셉트론 신경망모형을 사용하였다. 활성화함수로 쌍곡 탄젠

트 (hyperbolic tangent) 함수를 사용 하였고 합성함수와 최종활성함수로는 선형결합함수를 사용하였다. 쌍곡 탄젠트함수는 식 (2.2)와 같다 (Guo 등, 2012).

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (2.2)$$

신경망모형을 적용하기 위해 More와 Deo (2003)의 연구에서처럼 단일시계열자료를 재구성하였다. 첫 번째 관측 자료에서 임의의 시점 t 의 값을 종속변수 (output)로 했을 때 독립변수 (input)는 $t - 1$, $t - 2$ 시점의 값이 되고 두 번째 자료에서 $t + 1$ 시점의 값이 종속변수일 때 독립변수는 첫 번째 자료에서의 종속변수였던 t 시점의 값과 $t-1$ 시점의 값이 된다. Table 2.1은 이러한 방식으로 재구성된 풍력발전량 자료의 일부로 첫 번째 관측 자료에서 종속변수값 7.78은 두 번째 관측 자료에서의 독립변수값이 되고, 두 번째 종속변수값 5.52는 다시 세 번째 관측 자료에서의 독립변수값이 되는 구성을 갖는다.

Table 2.1 Examples of data pairs

pair numbers	input		output
1	4.47	6.86	7.78
2	6.86	7.78	5.52
3	7.78	5.52	3.39
4	5.52	3.39	5.48
5	3.39	5.48	2.67

Table 2.2 Information of forecasting models

model	input	output	the number of nodes
NN(1,1)	P_t	P_{t+1}	1
NN(1,2)	P_t	P_{t+1}	2
NN(1,3)	P_t	P_{t+1}	3
NN(1,4)	P_t	P_{t+1}	4
NN(1,5)	P_t	P_{t+1}	5
NN(2,1)	P_t, P_{t-1}	P_{t+1}	1
NN(2,2)	P_t, P_{t-1}	P_{t+1}	2
NN(2,3)	P_t, P_{t-1}	P_{t+1}	3
NN(2,4)	P_t, P_{t-1}	P_{t+1}	4
NN(2,5)	P_t, P_{t-1}	P_{t+1}	5
NN(3,1)	P_t, P_{t-1}, P_{t-2}	P_{t+1}	1
NN(3,2)	P_t, P_{t-1}, P_{t-2}	P_{t+1}	2
NN(3,3)	P_t, P_{t-1}, P_{t-2}	P_{t+1}	3
NN(3,4)	P_t, P_{t-1}, P_{t-2}	P_{t+1}	4
NN(3,5)	P_t, P_{t-1}, P_{t-2}	P_{t+1}	5

신경망모형은 현시점으로부터 1시간 후의 풍력발전량을 예측하기 위해 현재시점의 풍력과 과거시점인 한 시간 전의 풍력발전량과 두 시간 전의 풍력발전량을 각각 차례대로 독립변수에 추가시켜가며 모형을 구축해 보았다. 그리고 은닉층의 노드 수는 최소 1개부터 최대 5개까지 늘려가며 각각의 경우에 대해 모형을 구축하였다. 편의를 위해 종속변수와 독립변수 및 노드의 수에 따른 각각의 신경망모형을 Table 2.2에 정리하였다.

3. 모형 평가 및 비교

군산풍력발전단지의 1시간 단위 풍력발전량자료로 ARMA모형과 신경망모형을 구축하였다. ARMA 모형은 잔차의 ACF (autocorrelation function), PACF (partial autocorrelation function) 등을 확인하

여 가장 적합한 모수 p , q 를 찾았다. 여러 가지 모형을 적합해본 결과 모형평가용 자료의 RMSE (root mean square error) 값이 가장 작으며, 잔차분석을 통해 가장 적합한 것으로 나타난 ARMA(1,2)모형이 선택되었다. 또한 Table 2.2에서 정리했던 독립변수와 노드의 개수를 변화시킨 15개의 신경망모형을 각각 적합시켜 보았다. 모형구축용 자료와 모형평가용 자료에 대한 모형적합 결과는 Table 3.1과 같다. ARMA(1,2)모형의 모형구축용 자료, 모형평가용 자료 각각에 대한 RMSE값과 15개의 신경망모형 중 모형구축용 자료의 RMSE값 기준으로 상위 다섯 개의 가장 적합한 모형을 선택하여 각각의 자료에 대해 RMSE값을 비교한 것이다.

Table 3.1 RMSE for training set and test set

model	the number of estimated parameters	RMSE	
		training set	test set
ARMA(1,2)	3	0.5888	0.6385
NN(2,5)	21	0.5827	0.6323
NN(3,5)	25	0.5798	0.6323
NN(3,4)	21	0.5809	0.6333
NN(3,3)	16	0.5839	0.6338
NN(2,4)	17	0.5827	0.6342

Table 3.1을 보면 ARMA(1,2)모형의 RMSE값은 모형구축용 자료에서 0.5888, 모형평가용 자료에서 0.6385인 것으로 나타났다. 예측결과가 좋은 상위 5개의 신경망 모형들의 RMSE값을 보면 모형구축용과 모형평가용 자료 모두에서 ARMA(1,2)모형보다 작은 RMSE값을 나타내는 것을 알 수 있다. 신경망모형 중 가장 예측력이 좋은 모형은 모형평가용 자료를 기준으로 했을 때 RMSE값이 0.6323인 NN(2,5)모형과 NN(3,5)모형이다. 모형의 효율성을 고려한다면 추정된 모수 개수가 보다 적은 2개의 독립변수를 사용한 NN(2,5)모형이 군산풍력발전단지의 풍력발전량 단기예측에 있어 가장 적합한 모형이라 할 수 있다.

예측모형들의 좀 더 구체적인 비교를 위해 정확도의 향상률을 측정해보았다. Table 3.2를 보면 persistent모형과 다른 두 모형의 RMSE값을 나타냈었다. 일반적으로 모형 평가 기준을 정할 때 서로 다른 분석기법에 의해 추정된 값을 persistent모형을 기준으로 증가한 비율을, 모형을 비교하는 척도로 고려할 수 있다 (Sfetsos, 2002). 따라서 정확도 향상률은 persistent모형을 기준으로 측정가능하다. persistent모형은 1시간 후 풍력발전량을 예측하기 위해 특정 모형을 가정하지 않고 단순히 현시점의 실제값을 1시간 후의 예측값으로 사용하는 것을 말한다. 각 예측모형들의 RMSE값을 이용하여 persistent모형 대비 예측정확도가 얼마나 향상되었는지를 비교해 본 결과 Table 3.2와 같이 나타났다. 시계열 예측 모형 ARMA(1,2)모형은 persistent모형 대비 2.61%의 향상률을 보였으며, 신경망모형 중 가장 예측력이 좋은 NN(2,5)모형의 경우 persistent모형에 비해 3.55%의 향상률을 보였다.

Table 3.2 Performance of forecasting models (test set)

model	RMSE	% improvement from persistent
persistent	0.6556	-
ARMA(1,2)	0.6385	2.61
NN(2,5)	0.6323	3.55
NN(3,5)	0.6323	3.55
NN(3,4)	0.6333	3.40
NN(3,3)	0.6338	3.33
NN(2,4)	0.6342	3.26

군산풍력발전단지의 풍력발전량 예측에 가장 적합한 것으로 판단된 NN(2,5)모형을 이용한 풍력발전

량 예측값을 실제값과 비교해보면 Figure 3.1과 같다. 모형평가용 자료의 일부로서 2009년 8월 4일 13시부터 2009년 8월 25일 5시까지의 실제값과 예측값을 비교해 보면 모형적합이 타당하게 이루어졌음을 알 수 있다.

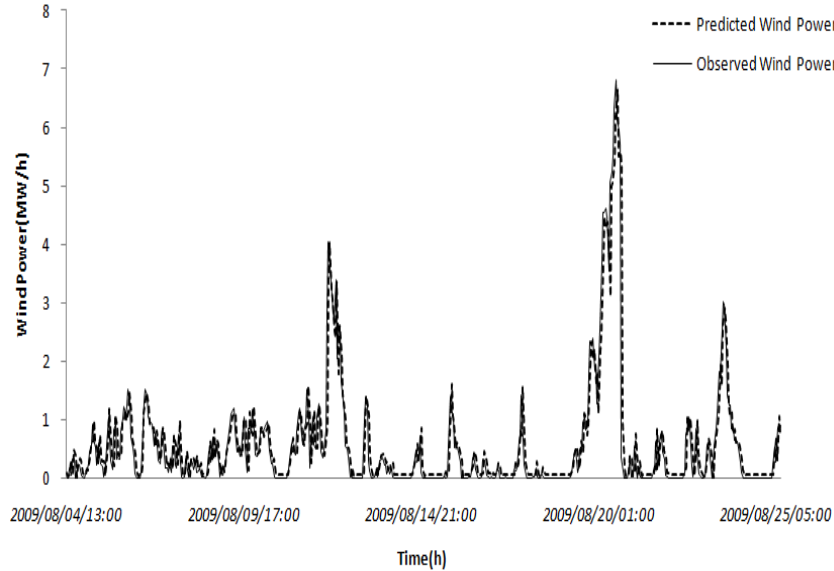


Figure 3.1 Time series of observed wind power and predicted wind power

4. 결론 및 향후과제

바람의 불규칙적인 특성으로 인해 어려운 전력시장에서의 안정적인 거래와 풍력발전단지의 효율적인 운영에 있어서 정확한 풍력발전량 예측은 매우 중요하고 필요한 요소이다. 이를 위하여 본 연구에서는 군산풍력발전단지의 1시간 단위 전력생산량자료를 기반으로 시계열예측모형인 ARMA모형과 데이터마이닝 기법 중 신경망모형을 사용하여 풍력발전량 단기예측모형을 구축하였다. 모형구축용 자료를 이용하여 모형을 구축하고, 구축된 모형을 모형평가용 자료에 적합 시켜 ARMA모형과 신경망모형을 비교하였다. 그 결과 독립변수로 현재시점과 과거시점인 한시간전과 두시간전의 풍력발전량을 독립변수로 하고 은닉층 노드의 개수가 5개인 신경망모형이 정확도 측면에서 가장 적합한 것으로 나타났다.

향후에는 현시점으로부터 1시간 후의 풍력에 대한 예측 뿐 만아니라 3~6시간 이후의 풍력에 대한 예측으로 확장시켜 나가고자 한다. 또한, 바람의 불규칙적인 특성을 고려하기 위하여 풍속 및 풍향 등의 기상학적 요소들을 추가로 이용해 더 좋은 예측력과 효율성을 가진 풍력발전량 예측모형을 구축해 볼 필요가 있다. 또한, 현재 비교한 시계열모형과 신경망모형의 비교를 위하여 RMSE가 작은 모형을 모형선택의 기준으로 하였다. 따라서 시계열자료의 정상성 여부, 모형 평가 측도의 다양성 등 모형의 고도화 작업을 위한 좀 더 심도 있는 분석은 향후과제로 남겨두고자 한다.

본 연구에서는 군산풍력발전단지를 대상으로 분석하였는데 향후 풍력이 생산되는 다른 지역에 대해서도 풍력발전량 예측모형을 확장 적용하여 우리나라의 실정에 맞는 모형을 구축하고 활용함으로써 우리나라 풍력발전 산업 성장에 기여할 수 있을 것이다.

References

- Abraham, B. and Ledolter, J. (2005). *Statistical methods for forecasting*, Wiley, New York.
- Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994). *Time series analysis, forecasting and control*, 3rd Ed., Prentice Hall, Englewood Cliffs, New Jersey.
- Guo, Z., Zhao, W., Lu, H. and Wang, J. (2012). Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. *Renewable Energy*, **37**, 241-249.
- Hwang, M. Y., Jin, C. H., Yun, U., Kim, K. D. and Ryu, K. H. (2012). Building of prediction model of wind power generation using power ramp rate. *Journal of the Korea Society of Computer and Information*, **17**, 211-218.
- Kim, S. Y. and Kim, S. H. (2011). Study on the prediction of wind power generation based on artificial neural network. *Journal of Institute of Control, Robotics and System*, **17**, 1173-1178.
- Lee, H. (2012). Analysis of time series models for consumer price index. *Journal of the Korean Data & Information Science Society*, **23**, 535-542.
- More, A. and Deo, M. C. (2003). Forecasting wind with neural networks. *Marine Structure*, **16**, 35-49
- Oh, K. J., Kim, T. Y., Jung, K. and Kim, C. (2011). Stock market stability index via linear and neural network autoregressive model. *Journal of the Korean Data & Information Science Society*, **22**, 335-351.
- Potts, W. J. E. (2000). *Neural network modeling course notes*, SAS Institute Inc., Cary, NC.
- Powers, J. and Ali, M. M. (2000). Application of neural networks in aluminum corrosion. *Journal of the Korean Data & Information Science Society*, **1**, 157-172.
- Sfetsos, A. (2002). A novel approach for the forecasting of mean hourly wind speed time series. *Renewable Energy*, **27**, 163-174.

A study on comparing short-term wind power prediction models in Gunsan wind farm[†]

Yung-Seop Lee¹ · Jin Kim² · Moon-Seok Jang³ · Hyun-Goo Kim⁴

¹²Department of Statistics, Dongguk University

³⁴Korea Institute of Energy Research

Received 21 April 2013, revised 15 May 2013, accepted 20 May 2013

Abstract

As the needs for alternative energy and renewable energy increase, there has been a lot of investment in developing wind energy, which does not cause air pollution nor the greenhouse gas effect. Wind energy is an environment friendly energy that is unlimited in its resources and is possible to be produced wherever the wind blows. However, since wind energy heavily relies on wind that has unreliable characteristics, it may be difficult to have efficient energy transmissions. For this reason, an important factor in wind energy forecasting is the estimation of available wind power. In this study, Gunsan wind farm data was used to compare ARMA model to neural network model to analyze for more accurate prediction of wind power generation. As a result, the neural network model was better than the ARMA model in the accuracy of the wind power predictions.

Keywords: ARMA models, neural network models, wind farm, wind power forecasting.

[†] This work was supported by the New & Renewable Energy of the Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant funded by the Korea government Ministry of Knowledge Economy (No. 2011T100100307).

¹ Corresponding author: Professor, Department of Statistics, Dongguk University-Seoul, Seoul 100-715, Korea. Email: yung@dongguk.edu

² Graduate student, Department of Statistics, Dongguk University-Seoul, Seoul 100-715, Korea

³ Principal researcher, Wind Energy Laboratory, Korea Institute of Energy Research, Jeju 695-971, Korea.

⁴ Principal researcher, New & Renewable Energy Resource Center, Korea Institute of Energy Research, Daejeon 305-348, Korea.