

수치 예측 알고리즘 기반의 풍속 예보 모델 학습

김세영*, 김정민*, 류광렬*

Learning Wind Speed Forecast Model based on Numeric Prediction Algorithm

Se-Young Kim*, Jeong-Min Kim*, Kwang-Ryel Ryu*

요약

대체 에너지 기술 개발을 위해 지난 20년 동안 풍력 발전에 관련한 기술들이 축적되어왔다. 풍력 발전은 자연적으로 부는 바람을 에너지원으로 사용하므로 환경 친화적이며 경제적이다. 이러한 풍력 발전의 효율적인 운영을 위해서는 시시각각 변하는 자연 바람의 세기를 정확도 높게 예측할 수 있어야 한다. 풍속을 평균적으로 얼마나 정확하게 잘 예측하는지도 중요하지만 실제 값과 예측 값의 절대 오차의 최댓값을 최소화시키는 것 또한 중요하다. 발전 운영 계획 측면에서 예측 풍속을 통한 예측 발전량과 실제 발전량의 차이는 경제적 손실을 가져오는 원인이 되므로 유연한 운영 계획을 세우기 위해 최대 오차가 중요한 역할을 한다. 본 논문에서는 풍속 예측 방법으로 과거 풍속 변화 추세뿐만 아니라 기상청 예보와 시기적인 풍속의 특성을 고려하기 위한 경향 값을 반영하여 수치 예측 알고리즘으로 학습한 풍속 예보 모델을 제안한다. 기상청 예보는 풍력 발전 단지를 포함하는 비교적 넓은 지역의 풍속을 예보하지만 풍속을 예측하고자 하는 국소지점에 대한 풍속 예측의 정확도를 높이는데 상당히 기여한다. 또한 풍속 변화 추세는 긴 시간동안 관측한 풍속을 세세하게 반영할수록 풍속 예측의 정확도를 높인다.

▶ Keywords : 풍속 예측, 수치 예측 알고리즘, 앙상블 모델 트리

Abstract

Technologies of wind power generation for development of alternative energy technology have been accumulated over the past 20 years. Wind power generation is environmentally friendly and economical because it uses the wind blowing in nature as energy resource. In order to operate wind power generation efficiently, it is necessary to accurately predict wind speed changing every moment in nature. It is important not only averagely how well to predict wind speed but also to minimize the largest absolute

•제1저자 : 김세영 •제2저자 : 김정민 •교신저자 : 류광렬

•투고일 : 2014. 10. 18, 심사일 : 2014. 12. 1, 게재확정일 : 2015. 2. 23.

* 부산대학교 전자전기컴퓨터공학과(Dept. of Computer Engineering, Pusan National University)

※ 본 논문은 미래창조과학부 산업융합원천기술개발사업으로 지원된 연구결과입니다(No.10043907).

error between real value and prediction value of wind speed. In terms of generation operating plan, minimizing the largest absolute error plays an important role for building flexible generation operating plan because the difference between predicting power and real power causes economic loss. In this paper, we propose a method of wind speed prediction using numeric prediction algorithm-based wind speed forecast model made to analyze the wind speed forecast given by the Meteorological Administration and pattern value for considering seasonal property of wind speed as well as changing trend of past wind speed. The wind speed forecast given by the Meteorological Administration is the forecast in respect to comparatively wide area including wind generation farm. But it contributes considerably to make accuracy of wind speed prediction high. Also, the experimental results demonstrate that as the rate of wind is analyzed in more detail, the greater accuracy will be obtained.

▶ Keywords : wind speed prediction, numeric prediction algorithm, ensemble of model trees

I. 서 론

마이크로그리드는 기존의 광역적 전력 시스템으로부터 독립된 소규모 전력 공급 시스템으로 공급 전력을 우선적으로 재생 에너지로부터 얻는 것을 목표로 한다. 마이크로그리드에서 사용하는 재생 에너지 중 하나인 풍력 발전은 발전 효율이 낮은 단점이 있으나 자연에서 얻는 바람의 에너지를 이용하므로 연료비가 거의 들지 않고 또한 자원이 고갈 되지 않는 이점이 있어 신재생 에너지로 각광받고 있다. 현재 마이크로 그리드에서 풍력 발전량을 예측하여 운영하지는 않지만 풍력 에너지는 기상 상태에 영향을 받아 가변적이기 때문에 안정적인 전력 공급을 위해서는 정확도 높은 미래 풍속 예측을 통해 적절한 전력 공급 계획을 세울 필요가 있다.

풍력 발전을 위한 정확도 높은 풍속 예측의 필요성은 이미 대두되었고 관련 연구가 활발하게 진행되어 왔다. 풍속을 예측하는 방법은 다음과 같이 네 가지로 구분할 수 있다. : (1) 지속적 방법(Persistence method), (2)물리적 방법(Physical method), (3)통계적 방법(Statistical method)과 인공 지능을 이용한 방법(Artificial Intelligence method), (4)앞의 방법들을 혼합한 방법(Hybrid method).[1-2]

지속적 방법은 가장 단순한 방법으로 현재 시점의 풍속이 다음 시점의 풍속과 같다고 가정하여 예측 하는 방법이다. 단기 풍속 예측에 효과적이다. 물리적 방법은 예측 지점의 지리

적 환경과 온도, 기압 등의 지역의 기상 상태를 이용하여 풍속을 예측하는 방법이다. 각종 정보들을 수집해야 하기 때문에 1시간 이내의 단기 예측에는 적절하지 않다. 또한 이 방법은 복잡한 수학적 모델을 이용하기 때문에 계산량이 많아 슈퍼컴퓨터로 처리하며 시간이 오래 걸린다. 통계적 방법과 인공 지능을 이용한 방법은 과거에 관측된 풍속을 기반으로 풍속을 예측하는 시계열 모델이다. 통계적 방법으로는 주로 ARMA(3-4)를 이용한 연구가 많으며 과거 풍속 자료 외에는 어떤 다른 자료를 필요로 하지 않는다는 장점이 있지만 예측 시점의 범위를 늘리면 정확도가 낮아진다. 풍속 예측에 사용되는 대표적인 인공 지능 방법으로는 인공 신경 회로망(Artificial Neural Network, ANN)이다. ANN 방식은 과거의 풍속 데이터에서 현재 상황과 유사한 패턴을 찾아 풍속을 예측한다. ANN의 복잡한 망의 형식을 개선하기 위해 퍼지 논리를 적용한 연구도 있다.[5] 앞서 소개한 풍속 예측 방법을 적절히 조합하여 각 각의 장점을 취하는 하이브리드 방법도 있다.[6][1-2]

본 논문에서는 풍속 변화와 상관관계가 존재하며 비교적 수월하게 얻을 수 있는 데이터와 수치 예측 알고리즘을 이용하여 풍속 예보 모델을 학습하는 방안을 제안한다. 모델을 학습하지 않고 미래 풍속을 예측하는 방법에는 기본적으로 기상청의 예보를 미래 풍속으로 보는 방법이 있다. 한국 기상청에서는 3시간 간격으로 정수 단위의 풍속 예보를 제공하고 있으며 그 정확도(proportion correct, PC)는 대략 80% 수준이다. 하지만 이 예보는 다소 넓은 지역을 대상으로 제공하므로 예보만으로는 풍력 발전 단지 내 풍속 예측의 정확도가 떨어

지는 문제가 있다. 따라서 수치 예측 알고리즘을 이용하여 풍속 예보 및 최근 관측 값 등 예측에 영향을 미치는 속성을 반영한 풍속 예보 모델을 제안하고자 한다.

풍속 예보 모델의 예측 정확도는 어떤 알고리즘과 속성을 사용하여 학습했는지에 따라 차이가 있다. 풍속은 시간의 흐름에 따라 증가하거나 혹은 감소하는 추세를 띠며 바람의 세기가 달라진다. 즉, 과거 풍속은 미래 풍속을 가늠하는데 핵심적인 정보가 될 수 있다. 이와 같은 과거 풍속의 변화 추세를 이용하기 위해서는 과거를 얼마만큼 반영할 것인지와 과거 풍속을 몇 분 단위로 반영할 것인지를 결정해야 한다. 관측 풍속과 더불어 풍속 예측에 도움이 되며 쉽게 구할 수 있는 데이터로 본 논문에서는 기상청 예보와 비슷한 시기의 풍속 경향성을 속성으로 하여 k-nearest neighbor(k-NN), locally weighted linear regression, model tree 그리고 ensemble of model trees와 같은 수치 예측 알고리즘을 이용해 풍속 예보 모델을 학습했다.

예보 모델의 성능 평가는 두 가지 측면으로 접근한다. 첫 번째로 테스트 데이터에 대해 전반적으로 얼마나 정확도 높게 예측했는지 알아보는 평균 오차와 두 번째로 전력 공급의 안정화에 기여하기 위한 최대 오차이다. 두 가지 평가에 대한 여러 풍속 예보 모델들의 성능 비교를 통해 정확도에 기여하는 속성과 수치 예측 알고리즘을 결정하고 과거 관측 풍속의 속성을 달리 반영할 실험을 통해 적절한 과거 반영 시간과 반영 단위를 알아본다.

다음 장에서는 풍속 예보 모델에 사용한 풍속 관측 자료와 기상청 예보에 대해 설명하고 3장에서 본격적으로 풍속 예측 모델의 개요와 실험에 사용한 속성 및 수치 예측 알고리즘을 소개한다. 4장에서는 속성과 수치 예측 알고리즘 별 풍속 예측 모델의 성능을 비교한 결과를 보이고 5장에서 결과에 따른 결론과 향후 연구를 제시한다.

II. 풍속 관측 자료와 기상청 예보

1. 풍속 관측 자료

풍속은 시간의 경과에 따라 변동하는데 그 변동성이 매우 크면서도 과거 풍속의 변동 추이에 따라 증가하는 추세 또는 감소하는 추세를 보인다. 따라서 가까운 과거에 관측한 풍속 정보는 풍속 예측에 중요한 정보가 된다. 실험에 사용한 풍속 관측 자료는 가파도의 2010년 8월 1일부터 2013년 7월 31일까지 1분 단위 풍속 데이터로 한국 기상청에서 제공하고 있

다. 이 중 2년 치 데이터(2010년 8월 1일 ~ 2012년 7월 31일)는 학습용 데이터로, 남은 1년 치 데이터(2012년 8월 1일 ~ 2013년 7월 31일)는 검증용 데이터로 사용하였다.

2. 기상청 예보

학습을 위해 사용한 속성 중 한 가지는 기상청 예보이다. 기상청 예보는 현재 공식적으로 제공되고 있는 유일한 풍속 예측 정보로서 다소 넓은 지역을 대상으로 하고 있으므로 발전이 이루어지는 국소 범위의 풍속을 정확히 예보하는 데 무리가 있다. 하지만 발전 지점이 속한 전체 지역의 풍력 양상을 보여주는 공신력 있는 수치로서 학습을 통한 풍속 예측의 정확도를 높이는 데 기여한다.

기상청에서 제공하는 예보는 3시간 간격이며 정수로 표현되어 있다. 또한 풍속 예보의 정확도(Proportion correct, PC)는 3시간 간격으로 관측한 풍속과 예보를 비교하여 평가하며 대략 80% 이다. 그러나 기상청에서 지난 예보 데이터는 제공하지 않으므로 앞서 설명한 예보의 특성과 정확도를 고려하여 풍속 관측 데이터로부터 추정하여 예보로 사용하였다.

한 편, 민간 기상 예보 업체에서 1시간 간격의 풍속 예보를 제공하고 있으나 예보의 정확도를 공개하지 않으므로 실험에 사용하기에 부적절하였다. 그러나 기상청 예보의 정확도로 가정하여 실험을 해본 결과 3시간 간격의 예보가 주어질 때보다 1시간 간격의 예보가 주어질 때 풍속 예보 모델의 성능이 크게 개선됨을 확인하였다. 즉, 풍속 예보 속성은 국소 범위의 풍속 예측에 기여하는 바가 크다.

III. 수치 예측 알고리즘 기반의 풍속 예보 모델

1. 풍속 예보 모델의 개요

풍속 예보 모델은 훈련용 데이터의 실제 관측 풍속과 이것을 예측하는데 관련 있는 여러 속성들 사이의 관계를 수치 예측 알고리즘으로 학습한 결과물이다. 따라서 학습에 사용한 속성들의 수치를 입력하면 관계에 의해 연산된 예측 풍속을 반환한다. 이 풍속 예보 모델에서 반환하는 예측 풍속은 마이크로그리드의 단기 운영 계획 단위를 1시간으로 가정하였을 때 미래 60분 동안의 평균 풍속으로 한다.[7]

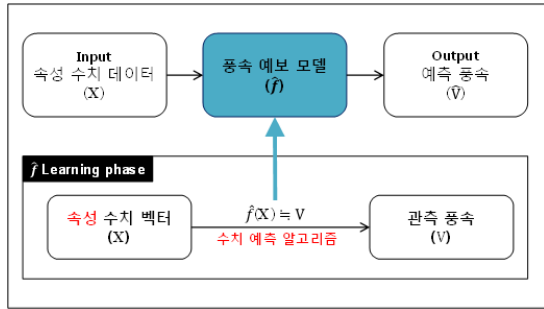


그림 1 풍속 예보 모델 학습과 적용 과정
Fig. 1. Process of Learning the wind speed forecast model and application

2. 수치 예측 알고리즘

수치 예측 알고리즘은 입력된 수치 자료와 그 자료들로 예측하고자 하는 대상의 실제 값 사이의 관계를 학습하는 방법이다. 본 논문에서 사용한 수치 예측 알고리즘은 k -nearest neighborhood(k -NN), locally weighted linear regression, model tree 그리고 ensemble of model trees 이다.

k -nearest neighborhood[8]은 non-parametric 방법으로 기억하고 있는 사례를 기반으로 예측 값을 결정한다. 사례는 속성 수치 값과 결과 수치 값으로 이루어진 벡터라고 볼 수 있다. 학습하는 단계에서 학습용 데이터의 모든 사례를 저장해 두고 예측을 위한 속성 수치 벡터가 입력되면 저장된 사례들 중 거리가 가까운 k 개의 사례들을 선택한 후 선택된 k 개의 결과 값을 평균하여 예측 값을 결정한다. locally weighted linear regression[9] 또한 k -NN과 같은 사례 기반 알고리즘 중 하나이다. k -NN은 가까운 k 개의 사례의 결과 값을 평균하여 예측 값을 결정하는 반면 locally weighted linear regression은 학습 단계에서 가까운 k 개의 사례의 결과 값에 대해 적절한 가중치를 학습하여 이 가중치 벡터를 이용하여 예측 값을 결정한다. Model tree는 수치 예측을 위한 decision tree로서 tree의 leaf node들마다 linear regression 방정식을 갖는다. Model tree를 형성하는 다양한 알고리즘 중에서 본 논문에서는 M5[10-11] 기반의 Model tree를 사용하였다. 그리고 훈련용 데이터를 여러 묶음으로 나누고 각 묶음마다 Model tree를 학습하는 것이 Ensemble of model trees 이다. 학습된 여러 개의 tree에서 얻은 결과 값을 평균한 수치를 예측 값으로 결정한다. 본 논문에서는 tree의 개수를 10개로 하여 실험하였다.

3. 속성

이 절에서는 풍속 예보 모델 학습을 위해 사용한 속성을 설명하고자 한다. 본 논문에서는 미래 풍속과 상관관계가 높은 속성 다섯 가지를 선정하였다. 자연에서 부는 바람은 매우 불규칙적으로 시시각각 변하지만 일정 시간동안의 풍속을 살펴보면 증가 혹은 감소하는 추세를 볼 수 있다. 그러므로 지난간 시간의 풍속 변화 추세는 미래의 풍속을 예측하는 중요한 속성이 될 수 있다. 또한 기상청 예보는 비록 넓은 지역을 대상으로 하지만 풍력 발전기 설치 지점을 포함하는 지역의 미래 풍속을 짐작해 볼 수 있는 믿음직한 지표가 된다.

추가적으로 예측 시점의 가까운 관측 풍속에 대한 정보뿐만 아니라 시기와 관련된 정보를 이용하여 속성으로 사용해 보았다. 실험에서 사용한 기상 데이터 관측 장소인 가파도는 전반적으로 바람이 고르게 부는 지역적 특성이 있지만 일반적으로 바람은 계절이나 시간에 따라 바람의 세기에 일정한 경향이 있다. 예를 들어 여름철 늦은 밤에는 바람이 불지 않거나 약한 반면 겨울로 갈수록 바람의 세기가 더 강해진다. 이런 경향성을 반영하기 위하여 같은 달, 같은 시간의 평균 관측 풍속을 경향 값으로 정의하고 속성으로 사용하였다. 또한 시기에 대한 직접적인 정보로 예측 시점의 월과 시간을 속성으로 추가하였다. 소개한 다섯 종류의 속성들 중 풍속 예측에 기여하는 것을 선별하기 위해 속성 선별 알고리즘으로 wrapper method의 forward selection 방식과 backward elimination 방식을 각각 적용하여 보았다.[12-13] 그 결과 모든 속성이 선별되었고 풍속 예측 정확도를 높이는 데 기여함을 확인하였다. 각 속성별로 풍속 예측에 기여하는 정도를 알아보기 위한 실험 결과를 4장에서 소개한다.

표 1. 예측 모델의 속성들
Table 1. Attributes of wind speed forecast model

| attribute | description |
|-----------------|--|
| month | 예측 시점의 해당 월 |
| time | 예측 시점의 시간을 분 단위로 환산한 시간 |
| past wind speed | 현재를 기준으로 과거 T분 동안의 관측 풍속을 h분 단위로 평균한 T/h개의 평균 풍속 |
| forecast | 예측 시점의 기상청 예보 |
| pattern | 예측 시점과 같은 월과 시간에 관측된 과거 풍속들의 평균 |

IV. 실험 및 결과

1. 풍속 변화 추세 반영

현재 시간을 T_c 라 하고 T 를 분 단위 시간이라 할 때 $T_c - T < t \leq T_c$ 인 시간 t 에 대해 해당 시간 동안의 관측 풍속을 학습을 위한 속성으로 반영하고자 한다. 시간 t 로부터 x 분 동안의 평균 관측 풍속을 $V_{t,t+x}$ 라 하고 $V_{t,t+x}^*$ 를 예측 평균 풍속이라 정의하자. T 분 동안 1분 단위로 관측된 풍속을 h 분 간격으로 평균하여 T/h 개의 평균 풍속인 $V_{T_c-T, T_c-T+h}, V_{T_c-T+h, T_c-T+2h}, \dots, V_{T_c-h, T_c}$ 을 구하고 이것을 속성으로 반영한다. 이 때 T 는 과거 반영 시간, h 는 반영 단위 시간이라 한다.

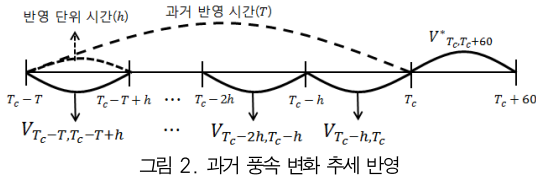


그림 2. 과거 풍속 변화 추세 반영
Fig. 2. analysis of Changing Trend of Wind Speed

본 논문에서는 T 가 60분, 120분, 180분이고 h 가 5분, 10분, 30분, 60분인 경우를 조합한 과거 풍속 변화 추세를 반영하여 풍속 예보 모델을 학습하였다.

2. 성능 평가 방법

풍속 예보 모델의 성능 평가는 두 가지 측면으로 접근한다. 첫째로 예측의 전반적인 정확도를 평가하기 위해 평균 절대 백분비 오차(Mean absolute percentage error, MAPE)와 평균 절대 편차(Mean absolute deviation, MAD)를 이용한다.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{V_i^* - V_i}{V_i} \right| \times 100(\%) \quad (1)$$

$$MAD = \frac{1}{n} \sum_{i=1}^n |V_i^* - V_i| \quad (2)$$

식 (1) 과 (2)에서 V_i 는 i 번째 검증용 데이터의 실제 풍

속 값, V_i^* 는 예측 풍속 값이고 n 은 검증용 데이터의 총 개수이다. MAPE는 실제 값에 비해 어느 정도의 오차가 발생하는지를 백분율로 나타내고 MAD는 편차로 나타낸다. 이 값들이 작을수록 풍속 예보 모델의 예측 값이 전반적으로 정확하다는 것을 의미한다.

마이크로그리드 단기 운영 계획에서 전력 공급의 안정화를 위해서는 평균적인 예측 정확도뿐만 아니라 절대 편차의 최댓값을 최소화 하는 것 또한 중요하다. 예측은 불확실한 정보이므로 예측 결과로 계산한 풍력 발전량과 실제 풍력 발전량에는 차이가 발생하게 된다. 이 차이로 인해 과잉 생산된 전기가 버려지거나 혹은 발전량 부족으로 단전이 발생하는 등 전력 공급에 차질을 빚을 수 있다. 이런 상황을 회피하기 위해 예측 풍속의 오차를 감안하여 유연한 전력 생산 계획을 세워야 한다. 따라서 최대 절대 편차(Maximum absolute deviation, Max. AD)가 작을수록 유리하다.

$$Max.AD = \max\{|V_i^* - V_i| \mid 1 \leq i \leq n\} \quad (3)$$

3. 실험 결과

3.1 비 학습 방식과 학습 방식의 정확도 비교

학습을 통해 풍속의 예측치를 구하는 것은 단순히 기상청의 풍속 예보를 확인하거나 비슷한 시기의 과거 풍속을 확인해 보는 것보다 시간이 많이 요구된다. 본 논문에서 제안하는 학습 방식을 이용한 풍속 예보 모델 학습의 필요성을 검증하기 위해 수치 예측 알고리즘으로 학습하지 않고 기상청 예보 혹은 경향성을 예측 풍속으로 하였을 때의 예측 정확도와 학습하여 얻은 예측 정확도를 비교해보았다. 학습을 통해 학습 데이터의 속성과 결과 사이의 적절한 관계를 찾아 예측 하는 것이 더 정확도가 높다.

표 2에서는 비 학습 방식과 학습 방식 각각에 대한 예측 풍속과 실제 풍속의 오차를 비교하고 있다. 3시간 단위의 기상청 예보의 정확도 보다 이 예보 값을 이용하여 학습한 풍속 예보 모델의 예측 정확도가 약간 더 좋은 것을 확인할 수 있었다. 또한 과거 풍속을 속성으로 추가하여 학습하였을 때 예측의 정확도가 상당히 개선되는 것을 알 수 있다. 비 학습방법에서 과거의 풍속 경향성으로 풍속을 예측하는 것은 상대적으로 오차가 상당히 크다. 같은 달, 같은 시간의 과거 풍속들을 평균한 값이 같은 시기의 풍속을 대변할 수 있을 정도로 정교하지는 못한 것으로 판단된다.

표 2. 비 학습 방식과 학습 방식의 비교
Table 2. Comparison of non-learning and learning style

| Measure | non-learning | | learning | | |
|-------------|--------------|---------|----------|-----------------|----------------|
| | forecast | pattern | forecast | past wind speed | all attributes |
| MAPE (%) | 38.1 | 42.12 | 36.17 | 17.74 | 16.35 |
| MAD (m/s) | 1.693 | 2.458 | 1.438 | 0.701 | 0.654 |
| MaxAD (m/s) | 13.64 | 24.58 | 10.82 | 8.96 | 8.82 |

3.2 예측 정확도에 기여하는 속성

표 3. 기상청 예보만을 반영한 풍속 예보 모델
Table 3. analysis performances of wind speed forecast model into attributes of forecast only.

| forecast | k-NN | Locally weighted linear regression | Model tree | Ensemble of model tree |
|---------------|-------|------------------------------------|------------|------------------------|
| MAPE (%) | 38.56 | 39.24 | 36.18 | 36.17 |
| MAD (m/s) | 1.574 | 1.609 | 1.437 | 1.438 |
| Max. AD (m/s) | 16.44 | 37.18 | 10.81 | 10.82 |

표 4. 과거 풍속과 시간 속성만을 반영한 풍속 예보 모델
Table 4. analysis performances of wind speed forecast model into attributes of past wind speed, month and time.

| past wind speed | T | k-NN | Locally weighted linear regression | Model tree | Ensemble of model tree |
|-----------------|-----|-------|------------------------------------|------------|------------------------|
| MAPE (%) | 60 | 20.39 | 21.85 | 18.00 | 18.00 |
| | | 0.825 | 0.885 | 0.707 | 0.707 |
| | | 14.38 | 18.97 | 8.96 | 8.96 |
| MAPE (%) | 120 | 20.62 | 23.75 | 17.89 | 17.86 |
| | | 0.842 | 0.97 | 0.703 | 0.702 |
| | | 14.71 | 18.97 | 8.94 | 8.96 |
| MAPE (%) | 180 | 20.73 | 24.36 | 17.79 | 17.74 |
| | | 0.856 | 0.997 | 0.703 | 0.701 |
| | | 14.71 | 24.09 | 8.94 | 8.96 |

표 5. 모든 속성을 반영한 풍속 예보 모델
Table 5. analysis performances of wind speed forecast model into all attributes introduced in table 1.

| all attributes | T | k-NN | Locally weighted linear regression | Model tree | Ensemble of model tree |
|----------------|-----|-------|------------------------------------|------------|------------------------|
| MAPE (%) | 60 | 18.48 | 19.42 | 16.54 | 16.51 |
| | | 0.741 | 0.788 | 0.659 | 0.658 |
| | | 8.80 | 12.39 | 8.81 | 8.82 |
| MAPE (%) | 120 | 18.02 | 19.42 | 16.40 | 16.41 |
| | | 0.725 | 0.789 | 0.654 | 0.655 |
| | | 8.81 | 12.39 | 8.79 | 8.79 |
| MAPE (%) | 180 | 18.02 | 19.42 | 16.33 | 16.35 |
| | | 0.725 | 0.789 | 0.654 | 0.654 |
| | | 8.81 | 12.39 | 8.84 | 8.82 |

표 1의 다섯 가지 속성들이 모두 풍속 예측에 기여하는지를 확인하기 위해 속성 선별 알고리즘을 적용하여 학습해본 결과 모든 속성들이 학습에 사용되었다. 각 속성들마다 예측의 정확도에 기여하는 정도를 비교해보기 위해 서로 다른 속성의 조합으로 모델을 학습하여 성능을 평가해 보았다. 이 실험에서 기상청 예보만을 반영한 풍속 예보 모델은 과거 풍속을 반영하지 않으므로 T와 h를 정할 필요가 없고, 과거 풍속을 반영한 학습은 T가 60분, 120분, 180분인 경우로 나누고 h는 60분으로 고정하여 반영하였다.

표 2와 표 3을 비교해 보면 과거 관측 풍속이 풍속 예측에 결정적인 역할을 한다는 것을 알 수 있고, 기상청 예보와 비슷한 시기의 풍속 경향성을 추가로 반영하여 학습하였을 때 모델의 성능이 약 1.5% 정도 개선되는 것을 확인하였다. 수치 예측 알고리즘으로는 k-NN과 locally weighted linear regression 보다 Model tree와 ensemble of model trees 방식의 학습이 풍속 예측의 정확도가 높다.

3.3 풍속 변화 추세 반영에 따른 정확도 비교

앞의 실험은 h를 60분으로 고정하여 과거 풍속을 반영하였다. 이번에는 h를 줄임으로써 풍속 변화 추세를 더 자세히 반영하였을 때 풍속 예보 모델의 정확도가 증가하는지를 알아보기 위한 실험을 했다. 그 결과 표 6, 표 7, 표 8에서 h를 작게 할수록 예측의 정확도가 점점 증가하며 h를 5분으로 하였을 때 예측의 정확도가 가장 높음을 확인할 수 있다.

수치 예측 알고리즘은 Model tree와 ensemble of model trees를 사용했을 때 성능이 거의 비슷하지만 Ensemble 학습 방식의 풍속 예보 모델이 근소한 차이로 예측 성능이 더 우수했다. 또한 T가 길어질수록 예측의 정확도가 증가하는 경향을 보였지만 두드러진 성능 차이를 보이지는 않았다.

그림 3은 풍속 예보 모델 중 최대 절대 편차가 가장 작았던 h=5, T=180, ensemble of model trees로 학습한 것과 기상청 예보 그리고 실제 관측 풍속(60분 평균 풍속)을 1시간 간격의 시간 흐름에 따라 나타낸 것이다. 예보는 3시간 간격으로 주어지기 때문에 편편한 구간이 생기지만 예측 풍속은 실제 풍속과 거의 유사한 추이를 보였다.

여러 실험을 통해 최대 절대 오차를 약 7m/s 까지 감소시켰지만, 8757개의 검증 데이터를 조사해본 결과 Ensemble of model trees 방식의 풍속 예보 모델로 예측했을 때 7m/s 이상의 오차가 발생한 데이터는 2건에 불과하였고 다른 알고리즘과 비교했을 때 1m/s 미만의 오차가 발생한 데이터가 가장 많았다. 절대 오차가 2m/s 이상인 데이터가

전체 데이터의 대략 1% 정도로 총 201개였다.

표 6. 풍속 변화 추세를 반영하는 방법에 따른 풍속 예보 모델의 MAPE (%) 비교

Table 6. Comparison of MAPE (%) of different kinds of wind speed forecast model according to analyzing changing trend of wind speed

| T | h | Forecast | k-NN | Model tree | Ensemble of model tree |
|-----|----|----------|-------|------------|------------------------|
| 60 | 5 | 38.1 | 15.18 | 14.02 | 13.94 |
| | 10 | | 16.74 | 14.08 | 14.02 |
| | 30 | | 18.02 | 15.08 | 15.03 |
| | 60 | | 39.15 | 16.54 | 16.5 |
| 120 | 5 | | 15.08 | 13.94 | 13.93 |
| | 10 | | 15.55 | 14.04 | 14.02 |
| | 30 | | 18.55 | 15.08 | 15.02 |
| | 60 | | 18.02 | 16.4 | 16.41 |
| 180 | 5 | | 15.08 | 13.93 | 13.92 |
| | 10 | | 15.55 | 14.04 | 14.01 |
| | 30 | | 19.14 | 15.01 | 14.92 |
| | 60 | | 18.48 | 16.34 | 16.32 |

표 7. 풍속 변화 추세를 반영하는 방법에 따른 풍속 예보 모델의 MAD (m/s) 비교

Table 7. Comparison of MAD (m/s) of different kinds of wind speed forecast model according to analyzing changing trend of wind speed

| T | h | Forecast | k-NN | Model tree | Ensemble of model tree |
|-----|----|----------|-------|------------|------------------------|
| 60 | 5 | 1.693 | 0.635 | 0.584 | 0.580 |
| | 10 | | 0.706 | 0.581 | 0.580 |
| | 30 | | 0.746 | 0.61 | 0.608 |
| | 60 | | 1.582 | 0.659 | 0.658 |
| 120 | 5 | | 0.636 | 0.582 | 0.581 |
| | 10 | | 0.651 | 0.582 | 0.581 |
| | 30 | | 0.77 | 0.612 | 0.609 |
| | 60 | | 0.725 | 0.654 | 0.654 |
| 180 | 5 | | 0.642 | 0.582 | 0.581 |
| | 10 | | 0.651 | 0.582 | 0.581 |
| | 30 | | 0.797 | 0.609 | 0.607 |
| | 60 | | 0.741 | 0.654 | 0.653 |

표 9. 예보, 풍속 예측 모델 예측 풍속의 절대 편차 분포

Table 9. Distribution of absolute deviation of forecast and wind speed forecast models

| Range of Absolute deviation (m/s) | Forecast | Model tree | Ensemble of model tree |
|-----------------------------------|----------|------------|------------------------|
| 0 ~ 1 | 3418 | 7241 | 7362 |
| 1 ~ 2 | 2444 | 1244 | 1194 |
| 2 ~ 3 | 1535 | 216 | 155 |
| 3 ~ 4 | 777 | 38 | 32 |
| 4 ~ 5 | 305 | 12 | 10 |
| 5 ~ 6 | 153 | 2 | 2 |
| 6 ~ 7 | 63 | 2 | 0 |
| 7 ~ 8 | 30 | 2 | 2 |
| 8 ~ 9 | 16 | 0 | 0 |
| 9 ~ 10 | 8 | 0 | 0 |
| 10 ~ 11 | 5 | 0 | 0 |
| 11 ~ 12 | 1 | 0 | 0 |
| 12 ~ 13 | 1 | 0 | 0 |
| 13 ~ 14 | 1 | 0 | 0 |

표 8. 풍속 변화 추세를 반영하는 방법에 따른 풍속 예보 모델의 Max.AD (m/s) 비교

Table 8. Comparison of Max.AD (m/s) of different kinds of wind speed forecast model according to analyzing changing trend of wind speed

| T | h | Forecast | k-NN | Model tree | Ensemble of model tree |
|-----|----|----------|-------|------------|------------------------|
| 60 | 5 | 13.64 | 8.61 | 7.52 | 7.52 |
| | 10 | | 14.2 | 7.74 | 7.75 |
| | 30 | | 14.49 | 7.99 | 8.11 |
| | 60 | | 14.69 | 8.81 | 8.83 |
| 120 | 5 | | 8.4 | 7.56 | 7.56 |
| | 10 | | 9.08 | 7.64 | 7.67 |
| | 30 | | 14.42 | 7.97 | 8.06 |
| | 60 | | 8.81 | 8.79 | 8.79 |
| 180 | 5 | | 8.05 | 7.45 | 7.44 |
| | 10 | | 9.08 | 7.64 | 7.66 |
| | 30 | | 14.67 | 7.8 | 8.07 |
| | 60 | | 8.8 | 8.84 | 8.83 |

V. 결론 및 향후 연구

기상청 예보, 과거 풍속의 변화 추세, 비슷한 시기의 풍속 경향 그리고 시간 관련 속성들을 모두 이용하여 학습을 하였을 때 예측의 정확도가 가장 높았다. 특히 과거 풍속의 변화

추세를 반영할 때 과거 반영 시간(T)가 길수록, 반영 단위 시간(h)가 짧을수록 더 정확한 예측을 하였다. 학습에 사용한 수치 예측 알고리즘 중에서는 모델트리와 앙상블 모델 트리를 이용하였을 때 풍속 예보 모델의 성능이 상대적으로 우수했으며, 이 중에서도 근소한 차이로 앙상블 모델 트리를 이용한 풍속 예보 모델의 예측 풍속이 정확도가 높았다.

그러나 이 실험에서 속성으로 사용한 기상청 예보는 앞서 설명했듯이 실제 기상청에서 제공한 예보가 아닌 풍속 예보 평가 방식과 정확도에 근거하여 관측 풍속으로 추정된 예보를 사용하였다. 따라서 본 논문에 기록한 예보의 성능은 실제 예보와 차이가 발생할 수 있으며 실제 예보에 따라 풍속 예보 모델의 성능 역시 실험 결과와 차이가 발생할 수 있다. 실제 기상청 예보 데이터를 이용하여 실험해 볼 필요가 있다. 기상청 예보가 정확할수록 풍속 예보 모델의 성능은 개선될 여지가 있다. 또한 풍속 예보 모델의 정확도를 개선하기 위해 풍속과 상관관계가 존재하는 온도, 습도 등의 새로운 속성 데이터를 수집하거나 학습에 반영한 비슷한 시기의 풍속 경향을 대변할 수 있는 경향 값을 더 정교하게 정의할 필요가 있다.

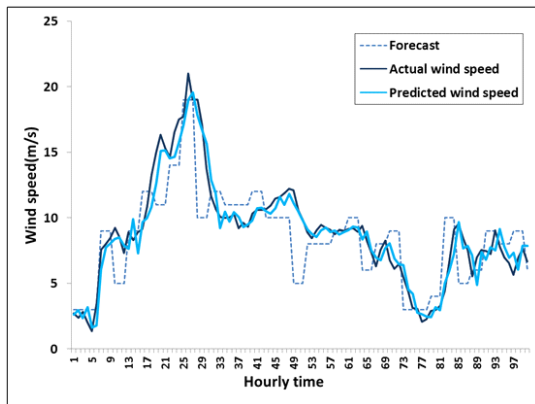


그림 4. 실제 풍속과 예보, 풍속 예측 모델의 예측 풍속 비교
 Fig. 3. Comparison between actual wind speed, forecast and predicted wind speed by wind speed forecast model

REFERENCES

[1] Y. K. Wu, and J. S. Hong, "A literature review of wind forecasting technology in the world," Proceedings of the IEEE conference on Power Tech, Lausanne, pp. 504-509, July. 2007.
 [2] W. Y. Chang, "A literature Review of Wind Forecasting Methods," Journal of Power and

Energy Engineering, vol. 2, no. 4, pp. 161-168, April. 2014.
 [3] E. Erdem, and J. Shi, "ARMA based approaches for forecasting the tuple of wind speed and direction," Applied Energy, vol. 88, no. 4, pp. 1405-1414, April. 2011.
 [4] M. Milligan, M. Schwartz, and Y. Wan, "Statistical Wind Power Forecasting models: results for U.S. Wind Farms", National Renewable Energy Laboratory, Golden, Colorado. May. 2003, NREL/CP-500-33956 (Preprint)
 [5] M. Monfared, H. Rastegar, and H. M. Kojabadi, "A new strategy for wind speed forecasting using artificial intelligent methods," Renewably Energy, vol. 34, no. 3, pp. 845-848, Mar. 2009
 [6] Z. Guo, J. Wu, H. Lu, and J. Wang, "A case study on a hybrid wind speed forecasting method using BP neural network," Knowledge-based systems, vol. 24, no. 7, pp. 1048-1056, Oct. 2011.
 [7] M. Ross, R. Hidalgo, C. Abbey, and G. Joós, "Energy storage system scheduling for an isolated microgrid," IET Renew. Power Gener., vol. 5, no. 2, pp. 117-123, Mar. 2011
 [8] S. Russell, and P. Norvig, "Artificial Intelligence : A Modern Approach Third Edition," Pearson Education, pp737-738, 2010.
 [9] C. G. Atkeson, A. W. Moore, and S. Schaal, "Locally weighted learning for control," Lazy learning. Springer Netherlands, pp. 75-113, 1997.
 [10] J. R. Quinlan, "Learning with continuous classes," 5th Australian Joint Conference on Artificial Intelligence, pp. 343-348, 1992.
 [11] Y. Wang, and I. H. Witten, "Induction of model trees for predicting continuous classes," Working paper 96/23, Hamilton, New Zealand: University of Waikato, Department of Computer Science, 1996.
 [12] R. Kohavi, and G. H. John, "Wrappers for feature subset selection," Artificial Intelligence,

vol. 97, no. 1-2, pp. 273-324, Dec. 1997.

- [13] I. Guyon, and A. Elisseeff, "An Introduction to Variable and Feature Selection," The Journal of Machine Learning Research vol. 3, pp. 1157-1182, Jan. 2003

저 자 소 개



김 세 영
 2014: 부산대학교 수학과 이학사.
 현 재: 부산대학교
 전자전기컴퓨터공학과 석사과정
 관심분야: 기계 학습, 데이터 마이닝
 Email : birdzero@pusan.ac.kr



김 정 민
 2010: 부산대학교
 전자전기컴퓨터공학부 공학사.
 현 재: 부산대학교
 전자전기컴퓨터공학과
 석박사통합과정
 관심분야: 기계 학습, 데이터 마이닝
 Email :
 jeongminkim.islab@gmail.ac.kr



류 광 렬
 1979: 서울대학교 전자공학과 공학사.
 1981: 서울대학교
 전자공학과 공학석사.
 1992: (미)미시건 대학교
 컴퓨터공학과 공학박사.
 현 재: 부산대학교
 전자전기컴퓨터공학과 교수
 관심분야: 기계 학습, 데이터 마이닝
 Email : krryu@pusan.ac.kr