

와이블 풍속 분포 파라미터 추정을 위한 Ln-least 방법의 확률도시위치 적용

강동범* · 고경남**†

*제주대학교 대학원 풍력특성화협동과정, 박사과정

**제주대학교 대학원 풍력공학부, 부교수

An Application of the Probability Plotting Positions for the Ln-least Method for Estimating the Parameters of Weibull Wind Speed Distribution

Kang Dong-Bum* · Ko Kyung-Nam**†

*Multidisciplinary Graduate School Program for Wind Energy, Graduate School, Jeju National University, Ph.D Student

**Faculty of Wind Energy Engineering, Graduate School, Jeju National University, Associate Professor

†Corresponding author: gnkor2@jejunu.ac.kr

Abstract

The Ln-least method is commonly used to estimate the Weibull parameters from the observed wind speed data. In previous studies, the bin method has been used to calculate the cumulative frequency distribution for the Ln-least method. The purpose of this study is to obtain better performance in the Ln-least method by applying probability plotting position(PPP) instead of the bin method. Two types of the wind speed data were used for the analysis. One was the observed wind speed data taken from three sites with different topographical conditions. The other was the virtual wind speed data which were statistically generated by a random variable with known Weibull parameters. Also, ten types of PPP formulas were applied which were Hazen, California, Weibull, Blom, Gringorten, Chegodayev, Cunnane, Tukey, Beard and Median. In addition, in order to suggest the most suitable PPP formula for estimating Weibull parameters, two accuracy tests, the root mean square error(RMSE) and R^2 tests, were performed. As a result, all of PPPs showed better performances than the bin method and the best PPP was the Hazen formula. In the RMSE test, compared with the bin method, the Hazen formula increased estimation performance by 38.2% for the observed wind speed data and by 37.0% for the virtual wind speed data. For the R^2 test, the Hazen formula improved the performance by 1.2% and 2.7%, respectively. In addition, the performance of the PPP depended on the frequency of low wind speeds and wind speed variability.

Keywords: 풍력에너지(Wind energy), 바람 데이터(Wind data), 와이블 분포(Weibull distribution), Ln-least 방법(Ln-least method), 확률도시위치(Probability plotting position)

 OPEN ACCESS



Journal of the Korean Solar Energy Society
Vol.38, No.5, pp.11-25, October 2018
<https://doi.org/10.7836/kjes.2018.38.5.011>

pISSN: 1598-6411

eISSN: 2508-3562

Received: 30 May 2018

Revised: 3 September 2018

Accepted: 30 October 2018

Copyright © Korean Solar Energy Society

This is an Open-Access article distributed under the terms of the Creative Commons Attribution NonCommercial License which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

기호설명

$f(v)$: 와이블 확률밀도 함수 (unitless)

$F(v)$: 와이블 누적분포 함수 (unitless)

v : 풍속 (m/s)

k : 와이블 형상계수 (unitless)

c : 와이블 척도계수 (m/s)

X_j : j 번째 난수 (unitless)

$f(v_i)$: 실제 관측된 확률밀도 (unitless)

$\hat{f}(v_i)$: 와이블 파라미터를 이용하여 구한 추정 확률밀도 (unitless)

$\overline{f(v)}$: 실제 관측된 확률밀도의 평균 (unitless)

1. 서론

바람은 보통 계절에 따라 주기적으로 변하며, 특정 지역에서 일정 기간(보통 1년 이상) 동안 수집된 바람 데이터를 바탕으로 확률 분포도를 작성하면 어떤 크기(풍속) 또는 어떤 방향(풍향)의 바람이 얼마나 불어오는지 쉽게 파악이 가능하다^{1,2)}. 특히, 풍력 산업에서 풍속의 확률분포를 추정하는 일은 매우 중요하다. 풍력에너지는 풍속의 제제곱에 비례하므로 풍속은 풍력발전기의 발전량에 영향을 미치는 중요한 요소이다. 그러므로 특정 속도의 바람이 얼마나 불어올지를 계산해야만 풍력발전기의 정확한 발전량 추정이 가능하다. 그러므로 풍력발전기를 설치하기 전에 설치 예정 사이트에서 미리 풍속을 측정하게 되며, 이를 통계적으로 계산하여 풍속별 확률분포를 추정하게 된다.

자연풍이 갖는 풍속에 대한 확률분포는 보통 와이블(Weibull) 확률분포와 유사하므로 풍속의 도수분포함수 (frequency distribution function)로 일반적으로 와이블 분포 함수가 많이 사용되고 있다. 와이블 분포는 위에서 설명한 발전량 추정 이외에도 구조물의 풍압 및 풍하중 계산 등에 사용되며, 이외에도 고장률 함수로 기계 부품의 잔존 수명 예측 및 노후화 파악 등에 사용되기도 한다. 와이블 분포는 두 개의 파라미터로 식이 구성되는데, 이는 형상계수(shape parameter)와 척도계수(scale parameter)이다. 이 두 파라미터는 와이블 분포의 형태 및 크기를 결정한다. 와이블 분포 추정 시 실제 풍속 분포에 가까운 두 파라미터가 예측되어야 하며, 부정확한 파라미터의 예측은 풍력 자원량 및 발전량, 구조물의 풍압 및 풍하중 추정 등에 있어서 오차를 발생시키는 요인이다. 그러므로 와이블 파라미터의 예측 정확도를 높일 수 있는 방안이 마련되어야 하며 이와 관련된 연구가 필요하다.

와이블 파라미터를 추정할 때 많이 사용되는 방법 중 하나가 바로 Ln - least 방법(혹은 그래프 방법(graphical method))이다. Ln - least 방법을 적용하기 위해서는 실제 풍속 데이터로부터 풍속별 누적분포함수를 계산해야

한다. 이때, 풍속 누적분포함수를 계산하기 위하여 보통 빈 방법(bin method)을 많이 사용한다^{3,4,5}. 하지만 빈 방법 이외에도 누적분포함수를 유추할 수 있는 방법이 있으며 이는 바로 확률도시위치(PPP: probability plotting position)방법이다.

본 연구에서는 Ln - least 방법의 예측 정확도를 높이고자 경험적인 확률분포로 많이 사용되는 확률도시위치를 사용하여 풍속별 누적분포함수를 계산하였고, 이를 Ln - least 방법에 적용하였다. 본 연구의 목적은 확률도시위치를 Ln - least 방법에 적용할 때 기존의 빈 방법보다 와이블 파라미터 예측 정확도가 얼마나 향상되는지 상호 비교하는 것이다. 이를 위해 기존에 연구되었던 다양한 종류의 확률도시위치 공식을 고려하였고, 분석에 사용된 바람 데이터는 다양한 지형에서 실제 측정된 바람 데이터와 난수를 사용하여 임의적으로 추출한 가상의 바람 데이터 등 두 종류를 사용하였다. 그리고 예측 정확도 평가는 통계적인 검사 방법을 통해 수치적으로 분석하였다.

2. 와이블 분포(Weibull distribution)

와이블 분포는 확률밀도함수(PDF: probability density function), $f(v)$ 와 누적분포함수(CDF: cumulative distribution function), $F(v)$ 로 표현할 수 있으며, 아래 식 (1), (2)를 통해 계산이 가능하다³.

$$f(v) = \frac{k}{c} \left(\frac{v}{c}\right)^{k-1} \exp\left[-\left(\frac{v}{c}\right)^k\right] \quad (1)$$

$$F(v) = 1 - \exp\left[-\left(\frac{v}{c}\right)^k\right] \quad (2)$$

여기서 v 는 풍속[m/s], k 는 형상계수[unitless]이며, c 는 척도계수[m/s]이다. 형상계수 k 는 곡선의 형태를 결정하고 그 값이 감소할수록 곡선의 폭이 좁아져 뾰족한 모양으로 변하며 이로 인해 풍속변화가 적어진다. 척도계수 c 는 평균풍속이 높을수록 그 값이 크다.

3. 풍속 데이터

본 연구에서 분석된 풍속 데이터는 두 종류이며, 하나는 사이트에서 실제 수집된 ‘측정 풍속 데이터’이며, 다른 하나는 통계학에서 사용하는 난수(random number)를 활용하여 인위적으로 생성한 ‘가상 풍속 데이터’이다.

3.1 측정 풍속 데이터

측정 풍속 데이터의 대상 사이트는 Fig. 1과 같다. 사이트는 지형에 따라 총 세 곳을 선정하였고, 섬 지형으로

는 우도, 해안 지형으로는 한동, 내륙 지형으로는 청수를 선정하였다. 각 사이트에서 기상탑을 실제로 설치하여 해당 지역의 바람을 측정하였다. 자세한 사이트 정보 및 측정 현황은 Table 1과 같다. 측정 시기는 사이트 별로 상이하나 측정 기간은 모두 1년으로 동일하다. 그리고 데이터 간 동등한 비교를 위해 측정 높이는 60m로 동일화했으며, 데이터 평균화 시간은 10분이다. 연 평균풍속은 우도 8.68m/s, 한동 7.34m/s, 청수 6.46m/s 순으로 높은 값을 보였다.

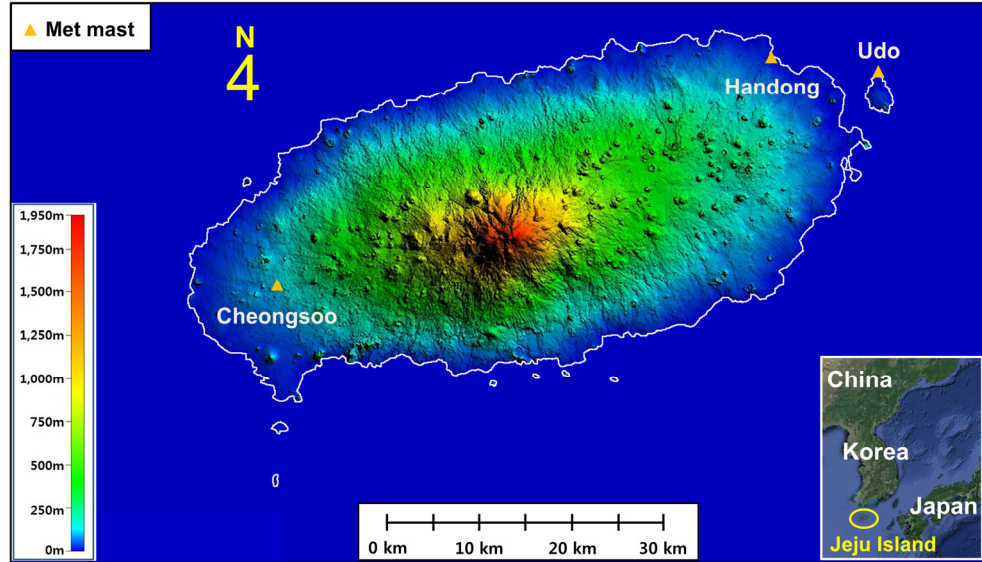


Fig. 1 The location of measurement sites.

Table 1 Sites and measurement conditions

Sites		Udo	Handong	Cheongsoo
Location	Latitude	33°31'7"N	33°32'15"N	33°18'18"N
	Longitude	126°57'14"E	126°50'5"E	126°16'1"E
Altitude [m]		20	5	130
Measurement period		Jan 1 2011 – Dec 31 2011	Jan 1 2012 – Dec 31 2012	Jan 1 2007 – Dec 31 2007
Measurement height [m]		60	60	60
Recovery rate [%]		100	100	99.8
Average wind speed [m/s]		8.68	7.34	6.46
RIX [%]		0.23	0.20	0.36
Topographical condition		Islet	Coast	Inland

해당 사이트의 지형 복잡도(terrain complexity)는 WASP(Wind Atlas Analysis and Application Program) S/W의 지형복잡지수 RIX(Ruggedness Index) 분석 결과를 제시하였다. 그 결과 내륙 지형인 청수가 0.36%로 섬 지형인 우도 0.23%, 해안지형인 한동 0.20%보다 상대적으로 높은 지형 복잡도를 보였다. 한편, 풍속 측정에

사용된 풍속계는 세 사이트 모두 컵형 NRG #40 풍속계로 자세한 제품 사양은 Table 2와 같다.

컵형 풍속계는 상대적으로 다른 센서에 비해 높은 풍속 역치(threshold)값을 갖고 있다. 그러므로 컵형 풍속계에 의해 측정된 풍속 데이터로부터 추정된 풍속 분포는 저풍속 또는 무풍속 영역에서 실제보다 낮거나 또는 높게 왜곡될 수 있다. 그러므로 이에 대한 영향을 줄이기 위해 0.5 m/s 미만의 풍속 데이터는 분석에 사용하지 않았다. 제거된 0.5m/s 미만의 풍속 데이터 비율은 우도는 0.57%, 한동은 1.74%, 청수는 1.17%이다.

Table 2 Anemometer specifications

Items	Specification
Model	NRG #40
Measuring range [m/s]	1 ~ 96
Threshold [m/s]	0.78
Accuracy	5 ~ 25 m/s: < 0.1 m/s
Operational temperature [°C]	-55 ~ 60
Type	3-cup
Classification	Class 2.4 A Class 7.7 B

3.2 가상 풍속 데이터

가상 풍속 데이터는 난수를 사용하여 통계적으로 추출하였다. 난수는 특정한 순서나 배열 또는 규칙을 갖지 않는 일련의 숫자들을 의미한다. 통계학적으로는 특정한 범위에서 균일한 분포(uniform distribution)를 갖는 모집단에서 추출된 표본들을 의미한다. 본 연구에서 난수는 특정한 범위(0과 1사이) 내에서 균일한 분포를 따르며, 생성된 난수들 사이에 독립성을 갖도록 임의적으로 생성하였으며, 난수 발생 방법으로는 합동법(congruential method)을 사용하였다. 참고로 합동법은 가장 널리 사용되는 난수 발생법 중 하나로 초기 값이 주어지면 재귀적 반복식에 의해서 새로운 난수를 발생시키게 된다. 난수 발생식은 식 (3)과 같다^{6,7)}.

$$X_i = (397204094 X_{i-1}) \bmod (2^{31} - 1) \tag{3}$$

여기서 *mod*는 *modulus*의 약자로서 앞에 제시된 숫자를 뒤에 제시된 숫자로 나눈 나눗셈의 나머지를 의미한다. 즉, $a \bmod b$ 는 a 를 b 로 나눌 때 나머지 값이다. 그리고 가상의 풍속 데이터를 얻기 위해 와이블 누적분포함수식, 식 (2)를 변형한 식 (4)를 사용하였다.

$$v = c \left[\ln \left(\frac{1}{1 - F(v)} \right) \right]^{1/k} \tag{4}$$

여기서 v 는 풍속[m/s], c 는 와이블 척도계수[m/s]이며, k 는 형상계수[unitless]이다. 식(3)을 통해 계산된 난수를 와이블 함수의 누적분포함수로 사용하였고, 와이블 파라미터(척도계수 및 형상계수)의 경우 특정 값으로 가정하여 가상의 풍속 데이터를 생성하였다. 본 연구에서는 와이블 파라미터 값이 서로 다른 총 8가지 케이스의 가상 바람 데이터를 생성하였으며, 그에 대한 설명은 Table 3과 같다. 형상계수는 1.5, 2.0을 고려하였고, 척도계수는 2.0, 5.0, 10.0, 15.0 m/s를 고려하였다. 가상 풍속 데이터의 데이터 개수는 앞에서 설명한 측정 풍속 데이터의 수집기간인 1년을 고려하여 10분 데이터의 1년 간 데이터 개수인 52,596개(= 365.25 days/year × 24 hours/day × 6 times/hour)를 고려하였다.

Table 3 Description of the virtual wind speed data

Weibull parameter values	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
Shape parameter (k) [unitless]	1.5	1.5	1.5	1.5	2.0	2.0	2.0	2.0
Scale parameter (c) [m/s]	2.0	5.0	10.0	15.0	2.0	5.0	10.0	15.0
Average wind speed [m/s]	1.80	4.49	8.98	13.47	1.77	4.42	8.83	13.25

4. 확률도시위치 공식의 Ln-least 방법 적용

4.1 Ln-least 방법(Ln-least method)

와이블 파라미터를 추정할 때 많이 사용되는 방법 중 하나는 Ln - least 방법이다. Ln - least 방법은 풍속 데이터를 로그 스케일로 보정 후 선형회귀분석을 통해 손쉽게 와이블 파라미터를 추정할 수 있다는 장점이 있으나, 다른 추정방법(모멘트 법, 최대우도법 등)에 비해 정확도가 떨어진다는 단점이 있다^{5,8,9}.

Ln - least 방법을 활용하기 위해서는 우선 풍속 데이터로부터 선형 회귀식을 유추해야 한다. 이를 위해 와이블 누적분포함수 공식인 식 (2)에 이중 자연로그함수를 취하고, 양 변을 이항하면 식 (5)와 같은 선형식을 얻을 수 있다^{10,11,12}.

$$\ln\{-\ln[1-F(v)]\} = k \ln(v) - k \ln(c) \tag{5}$$

여기서 $\ln(v)$ 를 x축에 넣고 $\ln\{-\ln[1-F(v)]\}$ 를 y축에 넣어 그래프를 그리면, 선형 그래프를 얻을 수 있는데 이 때, 직선의 기울기 값이 척도계수 k 가 되며, 직선의 y-절편 값 $k \ln(c)$ 로부터 척도계수 c 를 유추할 수 있다. 식 (5)를 적용하기 위해서는 풍속 누적분포함수인 $F(v)$ 를 계산하여야 한다. $F(v)$ 는 보통 빈 방법을 사용하여 각 풍속 빈에 대한 누적분포를 통계적으로 계산하여 얻게 되는데, 이 때 빈 방법이란 시간열 데이터를 ‘빈(bin)’이라 불리는 특정 구간으로 나누어 분류하는 것을 의미한다. Table 4는 우도에서 측정된 바람데이터를 빈 방법을 이용하여 각 풍속 빈 별 분포 및 누적분포를 계산한 예시이며, 빈 사이즈는 1m/s를 기준으로 하였다.

Table 4 The sample data set (site: Udo)

Wind speed bin [m/s]	Frequency [%]	Cumulative frequency [%]
0.5-1.5	1.76	1.76
1.5-2.5	3.58	5.34
2.5-3.5	5.04	10.38
3.5-4.5	6.23	16.61
4.5-5.5	6.96	23.57
5.5-6.5	8.33	31.90
6.5-7.5	9.29	41.18
7.5-8.5	9.29	50.47
8.5-9.5	8.87	59.34
9.5-10.5	8.93	68.27
≥10.5	31.73	100

4.2 확률도시위치(PPP: probability plotting position) 공식의 적용

확률도시위치란 확률지에 통계치를 점으로 도시하여 얻은 경험적 확률분포를 의미하며, 이미 여러 연구를 통해 다양한 확률분포 공식이 개발된 바 있다. 본 연구의 목적은 확률도시위치 공식을 Ln - least 방법에 적용하여 기존의 빈 방법을 적용한 Ln - least 방법과 비교하여 와이블 파라미터의 예측 정확도를 높이는 것이다. 본 연구에 사용된 확률도시위치 공식은 누적 확률을 대표해야 하며, 이를 위해 기존에 연구된 총 10가지의 누적 확률도

Table 5 Probability plotting position formulas

No.	Name	Cumulative probability plotting position
1	Hazen	$\frac{(i - 0.5)}{n}$
2	California	$\frac{i}{n}$
3	Weibull	$\frac{i}{(n + 1)}$
4	Blom	$\frac{(i - 3/8)}{(n + 1/4)}$
5	Gringorten	$\frac{(i - 0.44)}{(n + 0.12)}$
6	Chegodayev	$\frac{(i - 0.3)}{(n + 0.4)}$
7	Cunnane	$\frac{(i - 0.4)}{(n + 0.2)}$
8	Tukey	$\frac{(i - 1/3)}{(n + 1/3)}$
9	Beard	$\frac{(i - 0.31)}{(n + 0.38)}$
10	Median	$\frac{(i - 0.3175)}{(n + 0.365)}$

시위치 공식을 고려하였고^{6,13,14)} 자세한 정보는 Table 5에 나타났다. 여기서 n 은 풍속 데이터를 오름차순으로 정렬했을 때 그 순위이며, m 은 전체 데이터 수이다.

Table 5에서 제시된 10가지의 누적 확률도시위치 공식을 Ln - least 방법 식 (5)의 와이블 누적분포함수, $F(v)$ 에 적용하였고, 이를 통해 와이블 파라미터 k, c 값을 예측하였다. 그리고 기존의 빈 방법을 이용하여 Ln - least 방법으로 구한 k, c 값과 상호비교를 실시하여 예측 정확도를 평가하였다. 이 때, 빈 방법을 적용한 Ln - least 방법의 경우 빈 사이즈를 적게 할수록 와이블 파라미터의 예측 정확도가 높아진다는 기존 연구결과⁸⁾를 참조하여 최적의 빈 사이즈로 0.1m/s를 선정하였다.

Fig. 2는 우도의 측정 풍속 데이터를 사용하여 구한 Ln - least 방법이다. 그림에서 왼쪽 그래프는 누적분포함수를 빈 방법으로 적용한 경우이며, 오른쪽 그래프는 Hazen 확률도시위치로 적용한 경우이다. 한편, Fig. 3는 가상 풍속 데이터 Case 1 ($k=1.5, c=2.0$ m/s)을 사용하여 구한 Ln - least 방법이다. 그래프에서 보면 저 풍속 영역에서 선형 회귀선 피팅이 잘 안 되는데 이는 자연로그 함수 $\ln(x)$ 의 특성 상 x 가 1보다 적은 경우 $\ln(x)$ 가 급격히 감소하는 경향을 보이기 때문이다. 그러므로 1 m/s 보다 낮은 풍속에서 급격히 감소하는 $\ln(v)$ 의 변화를 선형 회귀선이 피팅하기에는 한계가 있다. 이러한 차이는 빈 방법보다 Hazen 확률도시위치에서 더욱 두드러진다. 그 이유는 1 m/s 미만의 풍속 데이터 개수 차이에 있다. 빈 방법의 경우 1 m/s 미만 데이터 개수가 Hazen 법에 비해 크게 적기 때문에(Fig. 2의 경우 빈 방법은 5개, Hazen 법은 385개) 이러한 피팅 불일치가 빈 방법에서는 크게 나타나지 않는다.

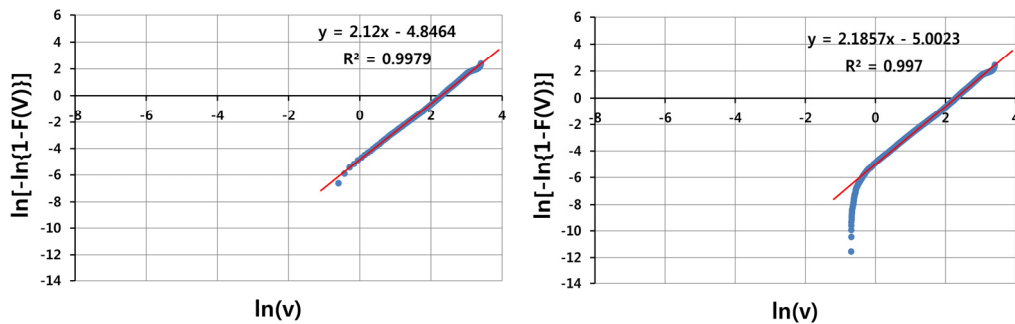


Fig. 2 The Ln-least method to estimate Weibull parameters for measured wind speed data (site: Udo) [Left: bin method (bin size: 0.1m/s), Right: PPP (Hazen method)].

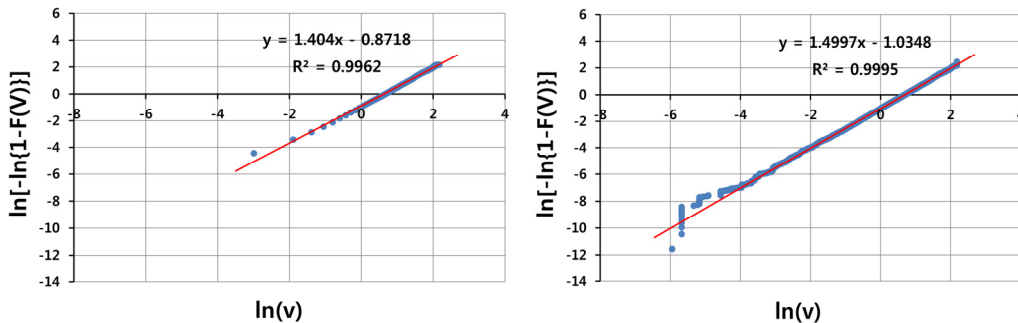


Fig. 3 The Ln-least method to estimate Weibull parameters for virtual wind speed data (known parameter values: $k=1.5, c=2.0$ m/s) [Left: bin method (bin size: 0.1 m/s), Right: PPP (Hazen method)].

5. 확률도시위치 적용 시 Ln-least 방법의 예측 정확도 비교

5.1 정확도 검사 방법

확률도시위치 및 빈 방법을 활용한 Ln - least 방법의 와이블 파라미터 예측 정확도 평가를 위해 다음 두 가지 검사 방법을 사용하였다. 첫째는 실제 값과 추정 값 사이의 오차 분석 방법으로 많이 사용되는 평균 제곱근 오차 (RMSE: root mean square error) 방법이며, 둘째는 추정 값이 얼마나 실제 값을 대표할 수 있는지 실제 값과 추정 값 사이의 적합성(goodness of fit)을 보여주는 결정계수(R^2 , coefficient of determination)이다. RMSE와 R^2 은 아래의 식 (6)과 (7)을 사용하여 계산할 수 있다^{5,8,9)}.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [f(v_i) - \hat{f}(v_i)]^2} \quad (6)$$

$$R^2 = \frac{\sum_{i=1}^n [f(v_i) - \overline{f(v)}]^2 - \sum_{i=1}^n [f(v_i) - \hat{f}(v_i)]^2}{\sum_{i=1}^n [f(v_i) - \overline{f(v)}]^2} \quad (7)$$

여기서 $f(v_i)$ 는 실제 관측된 확률밀도 값이다. 그리고 $\hat{f}(v_i)$ 는 Ln - least 방법에 의해서 예측된 와이블 파라미터를 식 (1)에 대입하여 구한 추정 확률밀도함수 값이다. $\overline{f(v)}$ 는 실제 관측된 확률밀도의 평균이다.

5.2 정확도 검사 결과

빈 방법 및 확률도시위치 공식을 적용한 Ln - least 방법의 정확도 비교를 위해 RMSE와 R^2 검사를 수행하였다. 위 검사는 빈 방법(빈 사이즈: 0.1m/s) 및 10가지 확률도시위치 공식에 모두 적용되었다.

Fig. 4는 측정 풍속 데이터에서 얻은 실제 풍속 분포와 예측된 와이블 분포를 비교하는 그림이다. 세로막대가 실제 풍속 분포이며, 검은 실선과 빨간 점선은 확률도시위치와 빈 방법이 각각 적용된 Ln - least 방법으로 추정된 와이블 분포이다. 확률도시위치 공식으로는 본 연구에서 가장 정확도가 높다고 분석된 Hazen 법이 사용되었다.

Fig. 4에서 확인할 수 있듯이 모든 사이트에서 Hazen 법이 빈 방법보다 실제 풍속 분포에 더 잘 적합 된다. 그러므로 확률도시위치 적용 시 Ln - least 방법의 예측 성능이 향상됨을 확인할 수 있다. 한편, Hazen 법으로 구한 와이블 분포와 빈 방법으로 구한 와이블 분포 사이에 불일치(discrepancy)를 발견할 수 있으며, 이들 차이가 지형에 따라서 상이하였다. 특히, 섬 지형인 우도 보다는 해안지형인 한동 또는 산악지형인 청수에서 그 불일치 정도가 증가하였다. 실제로, Hazen 법 기준으로 빈 방법과 비교할 때, 정확도 향상 정도가 RMSE 검사에서는 한동(47.9% ↑), 청수(40.7% ↑), 우도(16.1% ↑) 순으로 컸고, R^2 검사에서도 향상 정도가 한동(1.668% ↑), 청

수(1.666% ↑), 우도(0.331% ↑) 순으로 컸다. 그 이유는 풍속 데이터의 변동률 차이로 보인다. 각 사이트 별 풍속 데이터의 변동 계수(coefficient of variation) 분석 결과 우도는 47.6%, 한동은 53.7%, 청수는 50.2%였다. 즉, 풍속 변동성이 큰 한동, 청수, 우도 순으로 빈 방법과 비교한 Hazen 법의 예측 정확도 향상 정도가 커지고 있다. 그 이유는 풍속 변화가 큰 바람 조건일수록 풍속 분포에 적합한 와이블 파라미터의 예측이 어렵고, 그 정도가 확률도시위치 보다 빈 방법에서 더욱 현저해지기 때문으로 판단된다.

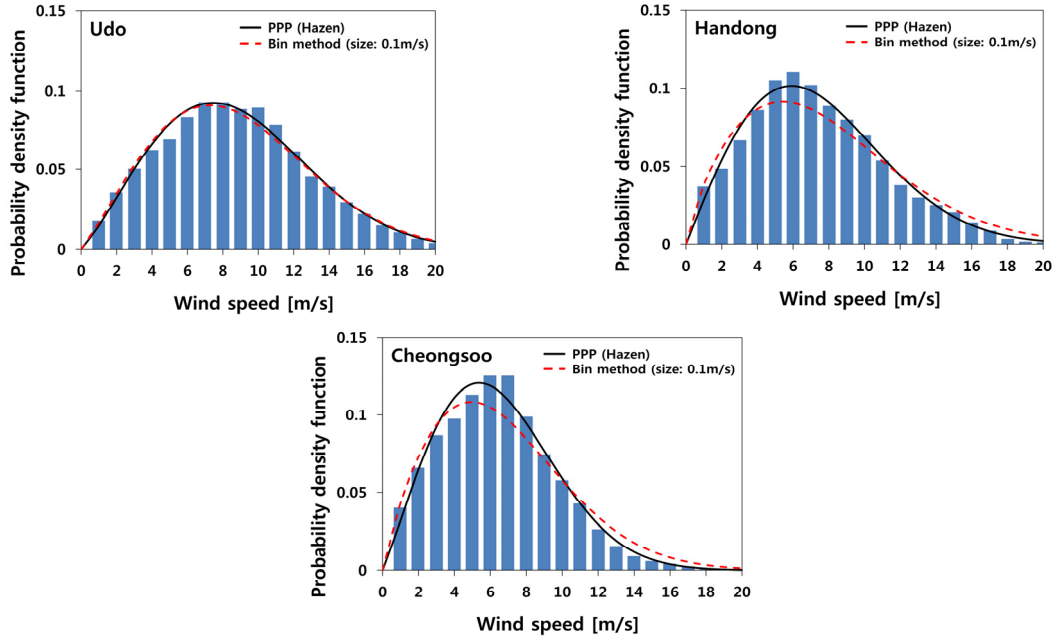


Fig. 4 A comparison between observed wind speed distributions and estimated Weibull PDFs for measured wind speed data.

Table 6은 측정 풍속 데이터에 적용된 Ln - least 방법의 정확도 검사 결과이다. 모든 사이트에서 확률도시위치가 빈 방법보다 높은 정확도를 보였다. 특히, 10가지 종류의 모든 확률도시위치 공식이 빈 방법에 비해 정확도가 높다는 결과가 나왔다. 그러므로 확률도시위치 적용 시 Ln - least 방법의 예측 성능이 향상됨을 확인할 수 있다.

가장 높은 정확도를 보이는 확률도시위치는 Hazen 법이며, 다음으로 Gringorten, Cunnane 법 순이었다. 하지만 확률도시위치에서 그 방법에 따른 예측 정확도 차이는 크지 않고 미미한 편이다. 확률도시위치 적용 시 Ln - least 방법의 예측 정확도 향상 정도는 Hazen 법을 기준으로 빈 방법과 비교할 때, 사이트 평균값 기준으로 RMSE 검사에서는 상대적으로 38.2% 높은 정확도를 보였으며, R^2 검사에서는 1.2% 높은 정확도를 보였다.

Fig. 5는 가상 풍속 데이터에서 얻은 풍속 분포와 예측된 와이블 분포를 비교하는 그림이다. Fig. 4와 비교할 때, Hazen 법과 빈 방법으로부터 예측된 와이블 분포 형태에 큰 차이는 보이지 않는다. 하지만 모든 케이스에서 Hazen 법이 빈 방법보다 실제 풍속 분포를 더 잘 적합 시킨다고 분석되었다.

Table 6 Performance results of Ln-least method for the accuracy tests (measured wind speed data)

Tests	Methods	Measured wind speed data			Avg. (Rank)
		Udo	Handong	Cheongsoo	
<i>k</i>	Hazen	2.1857	1.9872	2.0938	
	California	2.1851	1.9867	2.0933	
	Weibull	2.1851	1.9866	2.0932	
	Blom	2.1855	1.9870	2.0936	
	Gringorten	2.1856	1.9871	2.0937	
	Chegodayev	2.1854	1.9869	2.0935	
	Cunnane	2.1856	1.9871	2.0937	
	Tukey	2.1855	1.9870	2.0936	
	Beard	2.1855	1.9870	2.0935	
	Median	2.1855	1.9870	2.0936	
	Bin method (bin size:0.1m/s)	2.1200	1.7624	1.8515	
	<i>c</i>	Hazen	9.8615	8.4075	7.3214
California		9.8616	8.4074	7.3213	
Weibull		9.8620	8.4079	7.3217	
Blom		9.8617	8.4076	7.3215	
Gringorten		9.8616	8.4076	7.3214	
Chegodayev		9.8617	8.4077	7.3215	
Cunnane		9.8616	8.4076	7.3214	
Tukey		9.8617	8.4076	7.3215	
Beard		9.8617	8.4077	7.3215	
Median		9.8617	8.4077	7.3215	
Bin method (bin size:0.1m/s)		9.8355	8.7049	7.5321	
RMSE		Hazen	0.00317268	0.00338067	0.00463653
	California	0.00317597	0.00338255	0.00463910	0.00373254 (9)
	Weibull	0.00317588	0.00338354	0.00463935	0.00373292 (10)
	Blom	0.00317353	0.00338143	0.00463728	0.00373074 (4)
	Gringorten	0.00317309	0.00338104	0.00463690	0.00373034 (2)
	Chegodayev	0.00317401	0.00338187	0.00463771	0.00373120 (8)
	Cunnane	0.00317336	0.00338128	0.00463713	0.00373059 (3)
	Tukey	0.00317380	0.00338168	0.00463752	0.00373100 (5)
	Beard	0.00317395	0.00338181	0.00463765	0.00373114 (7)
	Median	0.00317390	0.00338177	0.00463761	0.00373109 (6)
	Bin method (bin size:0.1m/s)	0.00378270	0.00649098	0.00781890	0.00603086 (11)
	<i>R</i> ²	Hazen	0.9913441	0.9914118	0.9881572
California		0.9913274	0.9914036	0.9881444	0.9902918 (9)
Weibull		0.9913283	0.9913996	0.9881433	0.9902904 (10)
Blom		0.9913399	0.9914086	0.9881535	0.9903007 (4)
Gringorten		0.9913421	0.9914102	0.9881554	0.9903026 (2)
Chegodayev		0.9913375	0.9914067	0.9881514	0.9902985 (8)
Cunnane		0.9913407	0.9914092	0.9881542	0.9903014 (3)
Tukey		0.9913386	0.9914075	0.9881523	0.9902995 (5)
Beard		0.9913378	0.9914070	0.9881517	0.9902988 (7)
Median		0.9913381	0.9914071	0.9881519	0.9902990 (6)
Bin method (bin size:0.1m/s)		0.9880785	0.9751487	0.9719596	0.9783956 (11)

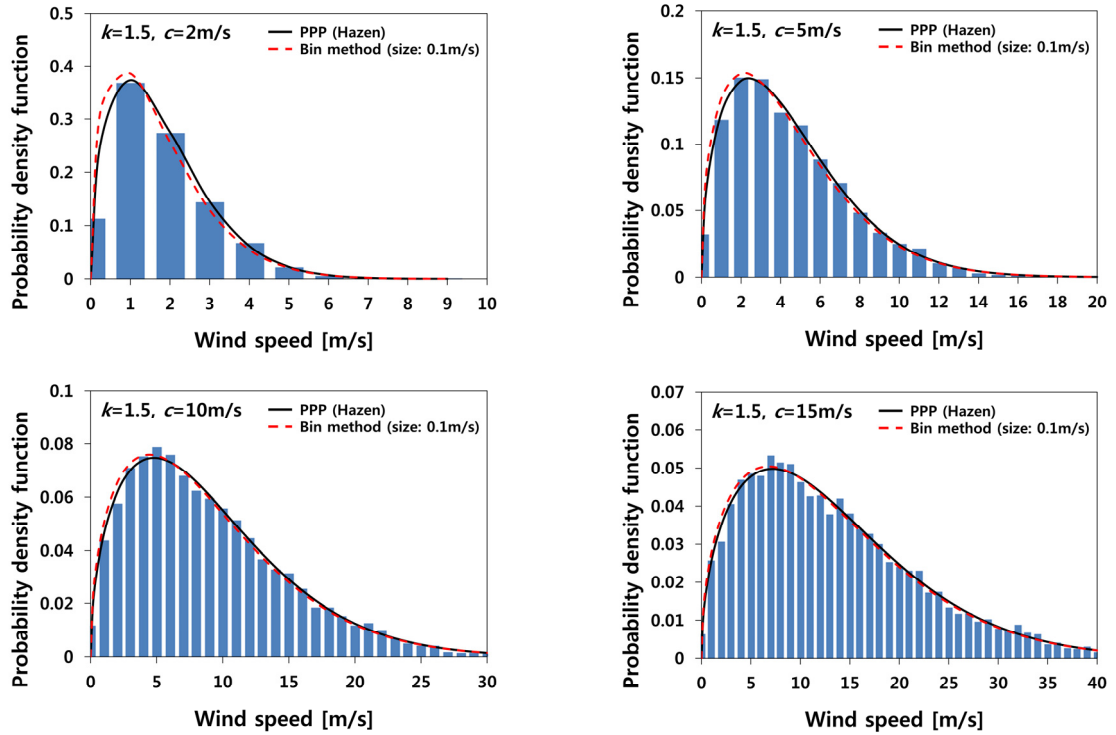


Fig. 5 A comparison between observed wind speed distributions and estimated Weibull PDFs for virtual wind speed data.

Table 7은 가상 풍속 데이터에 적용된 Ln - least 방법의 정확도 검사 결과이다. RMSE와 R^2 검사 모두 측정 풍속 데이터와 마찬가지로 모든 케이스에서 확률도시위치가 빈 방법보다 높은 정확도를 보였다. 그리고 확률도시위치 간 예측 정확도 비교 역시 측정 풍속 데이터와 마찬가지로 그 성능 차이는 크지 않았으며 정확도 순위는 Hazen, Gringorten, Cunnane 법 순이다. 가상 풍속 데이터에서 Hazen 법은 빈 방법보다 케이스 평균값 기준으로 RMSE 검사에서는 37.0%, R^2 검사에서는 2.7% 높은 정확도를 보였다.

Table 7 Performance results of Ln-least method for the accuracy tests (virtual wind speed data)

Tests	Methods	Virtual wind speed data								Avg. (Rank)
		$k=1.5, c=2.0$	$k=1.5, c=5.0$	$k=1.5, c=10.0$	$k=1.5, c=15.0$	$k=2.0, c=2.0$	$k=2.0, c=5.0$	$k=2.0, c=10.0$	$k=2.0, c=15.0$	
k	Hazen	1.4997	1.4997	1.4997	1.4997	1.9996	1.9996	1.9996	1.9996	
	California	1.4992	1.4992	1.4992	1.4992	1.9989	1.9989	1.9989	1.9989	
	Weibull	1.4991	1.4991	1.4991	1.4991	1.9988	1.9988	1.9988	1.9988	
	Blom	1.4995	1.4995	1.4995	1.4995	1.9994	1.9994	1.9994	1.9994	
	Gringorten	1.4996	1.4996	1.4996	1.4996	1.9995	1.9995	1.9995	1.9995	
	Chegodayev	1.4995	1.4995	1.4995	1.4995	1.9993	1.9993	1.9993	1.9993	
	Cunnane	1.4996	1.4996	1.4996	1.4996	1.9994	1.9994	1.9994	1.9994	
	Tukey	1.4995	1.4995	1.4995	1.4995	1.9993	1.9993	1.9993	1.9993	
	Beard	1.4995	1.4995	1.4995	1.4995	1.9993	1.9993	1.9993	1.9993	
	Median	1.4995	1.4995	1.4995	1.4995	1.9993	1.9993	1.9993	1.9993	
	Bin method (bin size:0.1m/s)	1.4040	1.4464	1.4647	1.4708	1.8396	1.8913	1.9178	1.9504	

Table 7 Performance results of Ln-least method for the accuracy tests (virtual wind speed data) (Continued)

Tests	Methods	Virtual wind speed data								Avg. (Rank)
		$k=1.5, c=2.0$	$k=1.5, c=5.0$	$k=1.5, c=10.0$	$k=1.5, c=15.0$	$k=2.0, c=2.0$	$k=2.0, c=5.0$	$k=2.0, c=10.0$	$k=2.0, c=15.0$	
<i>c</i>	Hazen	1.9937	4.9842	9.9684	14.9527	1.9953	4.9882	9.9763	14.9645	
	California	1.9938	4.9845	9.9689	14.9534	1.9953	4.9883	9.9767	14.9650	
	Weibull	1.9939	4.9847	9.9694	14.9541	1.9954	4.9885	9.9771	14.9656	
	Blom	1.9937	4.9844	9.9687	14.9531	1.9953	4.9883	9.9765	14.9648	
	Gringorten	1.9937	4.9843	9.9686	14.9529	1.9953	4.9882	9.9764	14.9646	
	Chegodayev	1.9938	4.9844	9.9689	14.9533	1.9953	4.9883	9.9766	14.9649	
	Cunnane	1.9937	4.9843	9.9687	14.9530	1.9953	4.9882	9.9765	14.9647	
	Tukey	1.9938	4.9844	9.9688	14.9532	1.9953	4.9883	9.9766	14.9649	
	Beard	1.9938	4.9844	9.9688	14.9533	1.9953	4.9883	9.9766	14.9649	
	Median	1.9938	4.9844	9.9688	14.9532	1.9953	4.9883	9.9766	14.9649	
	Bin method (bin size:0.1m/s)	1.8607	4.8146	9.7535	14.6924	1.8829	4.8352	9.7734	14.7654	
RMSE	Hazen	0.0448443	0.0076878	0.0023481	0.0015394	0.0257093	0.0038591	0.0018928	0.0016338	0.0111893 (1)
	California	0.0448939	0.0077036	0.0023537	0.0015415	0.0257443	0.0038650	0.0018944	0.0016345	0.0112039 (10)
	Weibull	0.0448936	0.0077042	0.0023541	0.0015417	0.0257428	0.0038652	0.0018946	0.0016346	0.0112038 (9)
	Blom	0.0448575	0.0076922	0.0023497	0.0015400	0.0257183	0.0038607	0.0018933	0.0016340	0.0111932 (4)
	Gringorten	0.0448507	0.0076900	0.0023489	0.0015397	0.0257137	0.0038599	0.0018930	0.0016339	0.0111912 (2)
	Chegodayev	0.0448651	0.0076947	0.0023506	0.0015403	0.0257234	0.0038616	0.0018935	0.0016341	0.0111954 (8)
	Cunnane	0.0448549	0.0076914	0.0023494	0.0015399	0.0257165	0.0038604	0.0018932	0.0016340	0.0111924 (3)
	Tukey	0.0448617	0.0076936	0.0023502	0.0015402	0.0257212	0.0038612	0.0018934	0.0016341	0.0111945 (5)
	Beard	0.0448641	0.0076944	0.0023505	0.0015403	0.0257227	0.0038615	0.0018935	0.0016341	0.0111951 (7)
	Median	0.0448633	0.0076942	0.0023504	0.0015402	0.0257222	0.0038614	0.0018935	0.0016341	0.0111949 (6)
	Bin method (bin size:0.1m/s)	0.0648455	0.0108052	0.0030711	0.0017994	0.0486960	0.0079754	0.0029872	0.0018831	0.0177579 (11)
<i>R</i> ²	Hazen	0.897637	0.979347	0.992186	0.992416	0.977275	0.996399	0.996433	0.994013	0.978213 (1)
	California	0.897365	0.979255	0.992149	0.992397	0.977191	0.996391	0.996430	0.994010	0.978148 (9)
	Weibull	0.897354	0.979250	0.992146	0.992395	0.977191	0.996391	0.996430	0.994010	0.978146 (10)
	Blom	0.897561	0.979321	0.992175	0.992411	0.977253	0.996397	0.996432	0.994012	0.978195 (4)
	Gringorten	0.897600	0.979334	0.992181	0.992414	0.977264	0.996398	0.996432	0.994013	0.978205 (2)
	Chegodayev	0.897518	0.979306	0.992169	0.992407	0.977240	0.996396	0.996432	0.994012	0.978185 (8)
	Cunnane	0.897576	0.979326	0.992177	0.992412	0.977257	0.996398	0.996432	0.994012	0.978199 (3)
	Tukey	0.897537	0.979312	0.992172	0.992409	0.977245	0.996397	0.996432	0.994012	0.978189 (5)
	Beard	0.897523	0.979308	0.992170	0.992408	0.977241	0.996396	0.996432	0.994012	0.978186 (7)
	Median	0.897528	0.979309	0.992171	0.992408	0.977243	0.996396	0.996432	0.994012	0.978187 (6)
	Bin method (bin size:0.1m/s)	0.803102	0.960180	0.986700	0.989656	0.909937	0.984123	0.991150	0.992068	0.952115 (11)

빈 방법과 비교한 Hazen 법의 예측 정확도 향상 정도는 k 가 일정할 때, c 가 증가할수록 작아지는 경향을 보인다. 일례로 $k=1.5$ 일 때, RMSE 검사에서는 $c=2.0$ m/s(30.8% ↑), $c=5.0$ m/s(28.9% ↑), $c=10.0$ m/s(23.5% ↑), $c=15.0$ m/s(14.4% ↑) 순으로 c 가 증가할수록 정확도 향상 정도가 작아진다. R^2 검사 역시 $c=2.0$ m/s(11.8% ↑), $c=5.0$ m/s(2.0% ↑), $c=10.0$ m/s(0.6% ↑), $c=15.0$ m/s(0.3% ↑) 순으로 정확도 향상 정도가 작아진다. 그 이유는 가상 풍속 데이터에서 c 값이 적어질수록 풍속의 분포 범위가 좁아지게 되며, 이는 빈 방법에서 활용 가능한 빈의 개수 역시 줄어드는 결과를 가져오게 된다. 빈의 개수가 줄어들면 그만큼 선형 회귀식인 식 (5)를 유

추하기 위한 데이터의 수가 줄어들음을 의미하며, 이는 빈 방법의 예측 성능 저하를 가져오게 된다. 그러므로 빈 방법을 활용하여 유추된 와이블 파라미터는 그 정확도가 Hazen 법에 비해서 상대적으로 떨어질 수밖에 없다.

한편, 확률도시위치 공식은 k 가 일정할 때, c 가 감소할수록 예측 정확도가 작아지는 경향을 보였다. 이는 c 값이 적어질수록 무풍($v=0$ m/s)에 가까운 풍속의 출현 빈도가 증가하기 때문이다. 와이블 분포는 공식의 특성 상 풍속이 0 m/s일 때, $f(v)$ 가 0에서 시작한다. 그러므로 실제 분포에서 0m/s에 가까운 풍속의 출현 빈도가 높다고 한다면, 예측된 와이블 분포가 여기에 적합 되기는 매우 어렵다. 본 연구에서 사용된 가상데이터의 경우 전체 데이터 개수 기준 0.5 m/s 미만 풍속 비율은 $k=1.5$ 일 때, $c=2.0$ m/s(11.3%), $c=5.0$ m/s(3.2%) $c=10.0$ m/s(1.2%), $c=15.0$ m/s(0.6%)이다. 따라서 c 값이 적을수록 실제에 가까운 와이블 분포를 추정하기 힘든 조건이며, 이에 따라서 확률도시위치 공식의 예측 정확도가 떨어지게 된다.

결론적으로 본 연구에서 가장 높은 예측 정확도를 보이는 확률도시위치 공식은 Hazen 법이었다. 하지만 분석에 사용된 바람 데이터의 종류가 제한적이므로 확률도시위치공식의 정확도 순서는 바람 데이터 조건에 따라 달라질 수 있으며 이에 대한 추가 연구가 필요하다.

6. 결론

본 연구에서는 Ln - least 방법의 예측 정확도를 높이기 위해 기존의 빈 방법을 대신하여 확률도시위치 공식을 사용하였으며, 그 성능을 상호비교 하였다. 분석에 사용된 바람 데이터는 두 종류이며, 첫째로 사이트에 실제로 설치된 기상탑에서 수집된 ‘측정 풍속 데이터’와 둘째로 난수를 활용하여 통계적으로 얻은 ‘가상 풍속 데이터’이다. 본 연구의 결과는 다음과 같다.

- (1) Ln - least 방법의 와이블 파라미터 예측 정확도 비교에서 확률도시위치가 빈 방법 보다 높은 정확도를 보였다.
- (2) 확률도시위치 공식에 따른 예측 정확도 차이는 미비하였으나, 전반적으로 가장 높은 예측 정확도를 보이는 공식은 Hazen 법이었다.
- (3) 측정 풍속 데이터에서 빈 방법과 비교한 확률도시위치의 예측 정확도 향상 정도는 지형에 따라 상이하였고 대체적으로 지형이 복잡할수록 그 향상 정도가 컸다.
- (4) 가상 풍속 데이터에서 빈 방법과 비교한 확률도시위치의 예측 정확도 향상 정도는 와이블 척도계수 값이 감소할수록 정확도 향상 정도가 커지는 경향을 보였다.
- (5) 대체적으로 저 풍속 영역의 빈도가 높고, 풍속 변화가 심한 바람 조건일수록 빈 방법에 비해 확률도시위치 공식의 예측 정확도가 높았다. 그러므로 해당 사이트의 경우 와이블 풍속 분포를 추정 시 Ln - least 방법을 사용할 때 빈 방법보다는 확률도시위치 공식을 사용할 것을 추천한다.

후기

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 「나셀 라이다를 이용한 풍력발전설비 및 풍력단지 출력성능측정 기술개발」 지원을 받아 수행한 연구 결과이며(No. 20163010024560), 산업통상자원부(MOTIE)의 재원으로 한국에너지기술평가원(KETEP)의 에너지인력양성사업 「육해상 풍력터빈 신뢰성 및 발전량 향상을 위한 O&M 기술 고급트랙」 으로 지원받아 수행한 인력양성 성과입니다(No. 20184030202200). 그리고 본 연구는 2018 한국태양에너지학회 춘계학술대회에서 발표한 ‘와이블 파라미터 추정을 위한 확률도시위치 비교 연구’를 확장 및 보충 연구하여 작성하였습니다¹⁵⁾.

REFERENCES

1. Ko, K. N., Kim, K. B., and Huh J. C., Characteristics of Wind Energy for Long-term Period (10 years) at Seoungang Site on Jeju Island, *Journal of the Korean Solar Energy Society*, Vol. 28, No. 3, pp. 45-52, 2008.
2. Song, H. S., and Kwon, S. D., Assessing Goodness-of-Fit of Weibull Distributions for Wind Resource Prediction, *Spring Conference of the Korean Solar Energy Society*, pp. 63-65, 2014.
3. Mathew, S., *Wind Energy: Fundamentals, Resource Analysis and Economics*, Springer, pp. 68-78, 2006.
4. Seguro, J. V., and Lambert, T. W., Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis, *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 85, No. 1, pp. 75-84, 2000.
5. Chang, T. P., Performance Comparison of Six Numerical Methods in Estimating Weibull Parameters for Wind Energy Application, *Applied Energy*, Vol. 88, No. 1, pp. 272-282, 2011.
6. Yahaya, A.S., Chong, S.Y., Ramli, N.A., and Ahmad, F., Determination of the Best Probability Plotting Position for Predicting Parameters of the Weibull Distribution, *International Journal of Applied Science and Technology*, Vol. 2, No. 3, pp. 106-111, 2012.
7. Huh, M., Lee, S., Cha, G., Park, J., and Yoo J., *R&statistic Computation*, Parkyeongsa, pp. 149-154, 2011.
8. Kang, D., Ko, K., and Huh, J., Comparative Study of Different Methods for Estimating Weibull Parameters: A Case Study on Jeju Island, South Korea, *Energies*, Vol. 11, No. 2, 2018.
9. Costa Rocha, P. A., de Sousa, R. C., de Andrade, C. F., and da Silva, M. E. V., Comparison of Seven Numerical Methods for Determining Weibull Parameters for Wind Energy Generation in the Northeast Region of Brazil, *Applied Energy*, Vol. 89, No. 1, pp. 395-400, 2012.
10. Ross, R., Graphical Methods for Plotting and Evaluating Weibull Distributed Data, *Proceedings of the IEEE International Conference on Properties and Applications of Dielectric Materials*, Vol. 1, pp. 250-253, 1994.
11. Deaves, D. M. and Lines, I. G., On the Fitting of Low Mean Windspeed Data to the Weibull Distribution, *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 66, No. 3, pp. 169-178, 1997.
12. Lun, I. Y. F. and Lam, J. C., A Study of Weibull Parameters Using Long-term Wind Observations, *Renewable Energy*, Vol. 20, No. 2, pp. 145-153, 2000.
13. Yahaya, A. S., Nor, N. M., Jali, N. R. M., Ramli, N. A., Ahmad, F., and Ul-Saufie, A. Z., Determination of the Probability Plotting Position for Type I Extreme Value Distribution, *Journal of Applied Sciences*, Vol. 12, No. 14, pp. 1501-1506, 2012.
14. Makkonen, L., Plotting Positions in Extreme Value Analysis, *Journal of Applied Meteorology and Climatology*, Vol. 45, No. 2, pp. 334-340, 2006.
15. Kang, D. B. and Ko, K. N., A Comparative Study on the Probability Plotting Positions to Estimate Weibull Parameters, *Spring Conference of the Korean Solar Energy Society*, pp. 102, 2018.