GRANT journal
ISSN 1805-062X, 1805-0638 (online)
Vol. 13, issue 02
EUROPEAN GRANT PROJECTS | RESULTS | RESEARCH & DEVELOPMENT | SCIENCE

Modelling the wind speed - a comparison between 2- and 3-parameter probability distributions

Mária Michalková[1]
Ivana Pobočíková[2]
Zuzana Sedliačková[3]
Daniela Jurášová[4]

[1] Department of Applied Mathematics, Faculty of Mechanical Engineering, University of Žilina, Univerzitná 8215/1, 010 26 Žilina, Slovakia; maria.michalkova@fstroj.uniza.sk
[2] Department of Applied Mathematics, Faculty of Mechanical Engineering, University of Žilina, Univerzitná 8215/1, 010 26 Žilina, Slovakia; ivana.pobocikova@fstroj.uniza.sk
[3] Department of Applied Mathematics, Faculty of Mechanical Engineering, University of Žilina, Univerzitná 8215/1, 010 26 Žilina, Slovakia; zuzana.sedliackova@fstroj.uniza.sk
[4] Department of Building Engineering and Urban Planning, Faculty of Civil Engineering, University of Žilina, Univerzitná 8215/1, 010 26 Žilina, Slovakia; daniela.jurasova@uniza.sk

Grant: KEGA projects No. 029ŽU-4/2022 and No. 025ŽU-4/2024
Name of the Grant: KEGA No. 029ŽU-4/2022 Implementation of the principles of blended learning into the teaching of the subject Numerical Methods and Statistics, KEGA No. 025ŽU-4/2024 Implementation of new didactic tools to increase the quality of mathematics teaching in the engineering degree at technical universities
Subject: AM – Pedagogy and education

Abstract The random character of wind deserves a statistical approach to properly describe this meteorological phenomenon. This paper focuses on finding the most suitable probability distribution to characterize the wind regime in the city Nitra while considering the wind speed data covering the 2021 year. Two commonly used distributions – Weibull and gamma - in their 2- and 3- parameter forms are compared via the Kolmogorov-Smirnov test and the Anderson-Darling test, and via the criteria - Akaike's and Bayesian information criterion, respectively, coefficient of determination and the root mean square error. According to criteria's results, the 3-parameter Weibull distribution performs as the best in majority of the months throughout the year.

Keywords Distribution fitting, parameter estimation, goodness-of-fit test

1. INTRODUCTION

Modelling wind speed is crucial in various applications. In civil engineering, it's used to estimate wind loads for building design and construction. In air transport, it's considered to enhance flight safety. In renewable energy, it's used to assess wind potential at a location. As a random variable significantly influenced by time, space, local climate, and terrain [1], wind speed requires statistical methods to describe its variation.

Several probability distributions can be used to model wind speed in different areas. The 2-parameter Weibull distribution is the most common [2, 3, 4, 5]. However, it may not be suitable for all wind regimes. The 3-parameter Weibull distribution is often used, especially when there's a higher frequency of lower wind speeds, as it offers more flexibility and a better fit [6]. Beyond Weibull distributions, other options include the gamma distribution [7], lognormal distribution, Nakagami distribution [9], extreme value distribution [10], Lindley distribution [11], and more.

This paper focuses on modelling wind speed in Nitra, Slovakia, a significant administrative, industrial, and cultural center. With a small airport and wind park nearby, understanding wind conditions in Nitra is important. We compare four probability distributions: 2- and 3-parameter Weibull and gamma distributions. Our goal is to compare the fit of 2-parameter versus 3-parameter distributions as well as to identify the best overall fit. To assess the fit, we use two goodness-of-fit tests: Kolmogorov-Smirnov and Anderson-Darling. Additionally, we employ information criteria (Akaike's and Bayesian), the coefficient of determination, and the root mean square error.

The paper is organized as follows: Section 2 describes the analyzed data. Section 3 defines the probability distributions, parameter estimation method, and performance criteria. Section 4 presents the results, which are summarized in Section 5.

2. DATA DESCRIPTION

The wind speed data, analysed in the paper, were recorded at the meteorological station Nitra - Veľké Janíkovce (indicator 11968), GPS latitude 48 16" 50' [48.28056], GPS longitude 18˚ 08" 08' [18.13556], the height of 132 meters above sea level. The station is located on the outskirts of the city Nitra, within the ground of a small airport. It is surrounded by the fields; the general face of the surroundings is partially sheltered. The mast for wind measurement is within the measuring plot; it is located on the roof of the building. The standard height for measuring wind direction and speed at monitoring stations is 10 m above the ground. Vaisala automatic instruments and GILL ultrasonic instruments were used to measure wind characteristics. The data were collected from the meteorological reports within the time frame January 2021 to December 2021, included. The data were recorded at hourly intervals and split into groups referring to months.

According to descriptive statistics summarised in tab. 1 and 2, during the studied period the lowest monthly mean wind speed was observed in September with value of 3.05 m/s, while in April there was the highest mean wind speed with value of 4.93 m/s. The standard deviation varies from 2.10 m/s in September to 2.96 m/s in April. The coefficient of variation (CV) is useful for identifying months with higher variability of wind speed. According to [11], the value of CV > 40 % is classified as a very high variability and CV > 70 % indicates the extremely high variability of wind speed. The coefficient of variation ranged from 50.98 % in May to 69.01 % in March. Based on this, the results imply that the wind speed in all months can be classified as having a very high variability. Skewness and kurtosis measure the asymmetry and the peakness of the wind speed distribution, respectively. The coefficients of skewness ranged from 0.38 in April to 1.40 in September, indicating that all distributions are right skewed. Further, the wind speed data can be regarded as moderately to highly right skewed. The coefficient of kurtosis ranged from 1.99 in October to 5.28 in September. That indicates a highly leptokurtic distribution when compared to the normal distribution.

Tab. 1. Descriptive statistics of the dataset. Part 1.

	Mean	Standard deviation	Coefficient of variation	Skewness	Kurtosis
Jan.	3.988	2.490	62.455	0.618	2.4363
Feb.	4.500	2.792	62.041	0.616	2.8067
Mar.	4.047	2.793	69.008	0.619	2.2934
Apr.	4.932	2.957	59.941	0.381	2.2088
May	4.877	2.486	50.981	0.411	2.6949
June	3.204	2.117	66.084	1.098	3.7488
July	3.616	2.152	59.501	0.590	2.4271
Aug.	3.177	2.177	68.526	0.920	3.1533
Sept.	3.052	2.104	68.942	1.403	5.2820
Oct.	4.053	2.567	63.341	0.462	1.9998
Nov.	4.038	2.570	63.642	0.581	2.3892
Dec.	4.059	2.739	67.483	0.623	2.4226

Tab. 2. Descriptive statistics of the dataset. Part 2.

	Min	Max	Lower quartile	Median	Upper quartile
Jan.	0.400	11.600	1.900	3.500	5.700
Feb.	0.500	14.700	1.900	4.100	6.500
Mar.	0.300	12.300	1.550	3.350	6.200
Apr.	0.500	13.100	2.250	4.700	7.100
May	0.700	13.300	2.900	4.800	6.500
June	0.300	11.500	1.600	2.500	4.500
July	0.500	10.100	1.750	3.200	5.100
Aug.	0.400	11.900	1.400	2.500	4.700
Sept.	0.500	13.200	1.400	2.500	4.200
Oct.	0.500	10.300	1.700	3.600	6.250
Nov.	0.500	11.900	1.800	3.500	6.100
Dec.	0.400	12.700	1.600	3.400	6.300

3. METHODOLOGY

In this section, the probability distributions employed to fit the wind data are briefly characterized. Further, we provide the parameter estimates realized by the maximum likelihood method as one of the most used estimation methods. To assess the performance of each probability distribution, we apply two goodness-of-fit tests and four model selection criteria that are defined by the end of the section.

3.1 Probability distributions

The probability density function $f(x)$ of the 2-parameter Weibull distribution is given as

$$f(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right).$$

The cumulative distribution function is defined as

$$F(x) = 1 - exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right),$$

for $x > 0,\ \alpha > 0,\ \beta > 0$. Parameter α is the dimensionless shape parameter and β is the scale parameter in units of the wind speed.

The probability density function $f(x)$ and the cumulative distribution function $F(x)$ of the 3-parameter Weibull distribution are given by

$$f(x) = \frac{\alpha}{\beta^\alpha}(x-\theta)^{\alpha-1} exp\left(-\left(\frac{x-\theta}{\beta}\right)^\alpha\right),$$

$$F(x) = 1 - exp\left(-\left(\frac{x-\theta}{\beta}\right)^\alpha\right),$$

for $x \geq \theta,\ \alpha > 0,\ \beta > 0$. Same as for the 2-parameter Weibull, parameter α is the dimensionless shape parameter, β is the scale parameter in units of the wind speed. Additional parameter θ is the location parameter.

The probability density function $f(x)$ and the cumulative distribution function $F(x)$ of the 3-parameter Gamma distribution are given by

$$f(x) = \frac{1}{\Gamma(\alpha)\ \beta^\alpha}(x-\theta)^{\alpha-1} exp\left(-\frac{x-\theta}{\beta}\right),$$

$$F(x) = \frac{\gamma\left(\alpha,\ \frac{x-\theta}{\beta}\right)}{\Gamma(\alpha)},$$

for $x \geq \theta,\ \alpha > 0,\ \beta > 0$. Here $\gamma(p,x) = \int_0^x e^{-t}\ t^{p-1} dt,\ p > 0$, is the lower incomplete Gamma function. Again, α is the shape parameter, β is the scale parameter and θ is the location parameter. Setting $\theta = 0$, we obtain the 2-parameter Gamma distribution.

3.2 Maximum likelihood method

This method is based on the maximization of the likelihood function $L(x_1, x_2, \dots, x_n; \theta)$ or its logarithm $\ln L(x_1, x_2, \dots, x_n; \theta)$ where $\theta \in \Theta$ is the unknown parameter (in general, it is a vector parameter) and $x_1, x_2, \dots, x_n$ is a realization of the random sample $X_1, X_2, \dots, X_n$ of size n from the distribution with the probability density function $f(x, \theta)$. Setting the derivative of the likelihood function or the loglikelihood function with respect to the unknown parameters equal to zero, the equations for the estimates of the parameters are found. The maximum likelihood estimates of the 2-parameter Weibull are of the form

$$\frac{1}{\alpha} - \frac{\sum_{i=1}^n x_i^\alpha \ln x_i}{\sum_{i=1}^n x_i^\alpha} + \frac{1}{n}\sum_{i=1}^n \ln x_i = 0,$$

$$\beta = \left(\frac{1}{n}\sum_{i=1}^n x_i^\alpha\right)^{1/\alpha}.$$

For the 3-parameter Weibull probability distribution, the parameter estimates are found as solutions of the equations

$$\frac{1}{\alpha} = \frac{\sum_{i=1}^n (x_i-\theta)^\alpha \ln(x_i-\theta)}{\sum_{i=1}^n (x_i-\theta)^\alpha} - \frac{1}{n}\sum_{i=1}^n \ln(x_i-\theta),$$

$$\beta = \left(\frac{1}{n}\sum_{i=1}^n (x_i-\theta)^\alpha\right)^{1/\alpha},$$

$$\frac{\alpha}{1-\alpha} = \frac{1}{n}\frac{\sum_{i=1}^n (x_i-\theta)^\alpha}{\sum_{i=1}^n (x_i-\theta)^{\alpha-1}}\sum_{i=1}^n \frac{1}{x_i-\theta}.$$

The estimates for the 2-parameter gamma distribution are given by the equations

$$\beta = \frac{\bar{x}}{\alpha},$$

$$\psi(\alpha) = \ln\alpha - \ln\bar{x} + \frac{1}{n}\sum_{i=1}^{n}\ln x_i,$$

where $\psi(p) = \frac{\partial \ln\Gamma(p)}{\partial p}$, $p > 0$, is the digamma function. Here

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

The estimates for the 3-parameter gamma distribution are given by the equations [12]

$$-\frac{1}{\sigma}\sum_{i=1}^{n}\frac{\lambda + z_i}{1 + \lambda z_i} = 0,$$

$$n\left(\frac{2}{\lambda^3}\psi\left(\frac{1}{\lambda^2}\right) + 2\ln\lambda\right) +$$
$$+\sum_{i=1}^{n}\left(-\frac{2}{\lambda^3}\ln(1+\lambda z_i) + \frac{\lambda(1+z_i^2) + 2z_i}{\lambda^2(1+\lambda z_i)}\right) = 0,$$

$$\mu = \bar{x},$$

where the following reparameterization is used

$$\alpha = \frac{1}{\lambda^2},\ \beta = \sigma|\lambda|,\ \theta = \mu - \frac{\sigma}{\lambda},\ z_i = \frac{x_i - \mu}{\sigma},$$

$\lambda > 0$, $\bar{x}$ and $\psi(p)$ are defined above.

It is obvious that the parameter estimates of all distributions can be found only numerically by solving the equations in an iterative way.

3.3 Model selection criteria and goodness-of-fit tests

When the estimates of the parameters are found, one can assess the goodness-of-fit (GOF) of the model. The GOF criteria show how well the selected model fits the wind speed data. Assessing the performance of different probability distribution models is necessary to provide more accurate information about their performance and to compare these models among themselves. Here, the commonly used GOF tests - the Kolmogorov-Smirnov (*KS*) test and the Anderson-Darling (*AD*) test are employed. The GOF tests are used to decide whether the data follow the specified theoretical distribution. The *KS* test statistic represents the largest vertical difference between the theoretical and the empirical cumulative distribution function

$$D = \max_{1\le i\le n}\left[\left|\hat{F}(x_{(i)}) - \frac{i-1}{n}\right|, \left|\frac{i}{n} - \hat{F}(x_{(i)})\right|\right]$$

where $\hat{F}(x)$ is the estimated cumulative distribution function, $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ are observations in ascending order, i.e., $x_{(1)} \le x_{(2)} \le \dots \le x_{(n)}$. Function $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_{(i)} \le x)$ is the empirical distribution function, where $I(x_{(i)} \le x)$ is an indicator function assuming the value 1 if $x_{(i)} \le x$ and 0 otherwise. The null hypothesis that the data follow the distribution under test, is rejected at the chosen significance level α if the test statistic $D > D(\alpha)$, where $D(\alpha)$ is a critical value of the *KS* test. The smaller the value of the test statistic D, the better the fit.

The Anderson-Darling (*AD*) test is a modification of the *KS* test. This test is considered to be a better GOF test because it gives more weight to the tails of the distribution than does the *KS* test. The *AD* test statistic is defined as follows

$$A^2 = -n - \sum_{i=1}^{n}\frac{2i-1}{n}\left[\ln\left(\hat{F}(x_{(i)})\right) + \ln\left(1 - \hat{F}(x_{(n+1-i)})\right)\right].$$

The null hypothesis that data follow the specified distribution, is rejected at the significance level α if the test statistic A^2 is greater than the critical value of the *AD* test. Again, the smaller value of the test statistic A^2 indicates a better fit.

The application of the maximum likelihood method (MLM) for parameter estimation allows us to use the information criteria - Akaike's information criterion (*AIC*) and Bayesian information criterion (*BIC*) - to decide the GOF for the distributions. The *AIC* and the *BIC* are defined as follows [13, 14]

$$AIC = -2\ln L + 2k,$$
$$BIC = -2\ln L + k\ln n$$

where $\ln L$ is the maximum value of log-likelihood function for estimated model, *k* is number of estimated parameters and *n* is the sample size.

Further, the coefficient of determination (R^2) and the root mean square error (*RMSE*) are considered to decide on the best fitting model. The *RMSE* determines the accuracy of model by calculating average of the square difference between the observed and the predicted probabilities of the theoretical distribution. The R^2 is used to measure the linear relationship between the observed and the predicted probabilities of the theoretical distribution. The *RMSE* and R^2 are calculated by

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}\left[F_n(x_i) - \hat{F}(x_i)\right]^2\right)^{\frac{1}{2}},$$

$$R^2 = \frac{\sum_{i=1}^{n}\left[\hat{F}(x_i) - \bar{F}\right]^2}{\sum_{i=1}^{n}\left[\hat{F}(x_i) - \bar{F}\right]^2 + \sum_{i=1}^{n}\left[F_n(x_i) - \hat{F}(x_i)\right]^2}$$

where $\bar{F} = \frac{1}{n}\sum_{i=1}^{n}\hat{F}(x_i)$.

Generally, lower values of *KS, AD, AIC, BIC, RMSE* and higher value of R^2 indicate better fit of the theoretical distribution to the wind speed data as compared to the others.

4. RESULTS

The parameter estimates for all four considered probability distributions are presented in tab. 3.

Tab. 3. The parameter estimates of the applied distributions.

Month	Probability distribution	Parameter estimates	
January	2-parameter Weibull	$\hat{\alpha} = 1.658$	
		$\hat{\beta} = 4.473$	
	3-parameter Weibull	$\hat{\alpha} = 1.429$	$\hat{\theta} = 0.362$
		$\hat{\beta} = 3.982$	
	2-parameter Gamma	$\hat{\alpha} = 2.306$	
		$\hat{\beta} = 1.729$	
	3-parameter Gamma	$\hat{\alpha} = 1.773$	$\hat{\theta} = 0.319$
		$\hat{\beta} = 2.069$	
February	2-parameter Weibull	$\hat{\alpha} = 1.665$	
		$\hat{\beta} = 5.047$	
	3-parameter Weibull	$\hat{\alpha} = 1.390$	$\hat{\theta} = 0.490$
		$\hat{\beta} = 4.383$	
	2-parameter Gamma	$\hat{\alpha} = 2.299$	
		$\hat{\beta} = 1.957$	
	3-parameter Gamma	$\hat{\alpha} = 1.611$	$\hat{\theta} = 0.481$
		$\hat{\beta} = 2.495$	
March	2-parameter	$\hat{\alpha} = 1.476$	

	Weibull	$\hat{\beta} = 4.487$	
	3-parameter Weibull	$\hat{\alpha} = 1.304$	$\hat{\theta} = 0.296$
		$\hat{\beta} = 4.062$	
	2-parameter Gamma	$\hat{\alpha} = 1.896$	
		$\hat{\beta} = 2.135$	
	3-parameter Gamma	$\hat{\alpha} = 1.494$	$\hat{\theta} = 0.292$
		$\hat{\beta} = 2.514$	
April	2-parameter Weibull	$\hat{\alpha} = 1.701$	
		$\hat{\beta} = 5.529$	
	3-parameter Weibull	$\hat{\alpha} = 1.460$	$\hat{\theta} = 0.422$
		$\hat{\beta} = 4.955$	
	2-parameter Gamma	$\hat{\alpha} = 2.264$	
		$\hat{\beta} = 2.179$	
	3-parameter Gamma	$\hat{\alpha} = 1.741$	$\hat{\theta} = 0.373$
		$\hat{\beta} = 2.619$	
May	2-parameter Weibull	$\hat{\alpha} = 2.070$	
		$\hat{\beta} = 5.511$	
	3-parameter Weibull	$\hat{\alpha} = 1.820$	$\hat{\theta} = 0.443$
		$\hat{\beta} = 4.978$	
	2-parameter Gamma	$\hat{\alpha} = 3.263$	
		$\hat{\beta} = 1.495$	
	3-parameter Gamma	$\hat{\alpha} = 3.263$	$\hat{\theta} = 0$
		$\hat{\beta} = 1.495$	
June	2-parameter Weibull	$\hat{\alpha} = 1.619$	
		$\hat{\beta} = 3.602$	
	3-parameter Weibull	$\hat{\alpha} = 1.436$	$\hat{\theta} = 0.291$
		$\hat{\beta} = 3.221$	
	2-parameter Gamma	$\hat{\alpha} = 2.473$	
		$\hat{\beta} = 1.296$	
	3-parameter Gamma	$\hat{\alpha} = 1.948$	$\hat{\theta} = 0.270$
		$\hat{\beta} = 1.506$	
July	2-parameter Weibull	$\hat{\alpha} = 1.762$	
		$\hat{\beta} = 4.077$	
	3-parameter Weibull	$\hat{\alpha} = 1.431$	$\hat{\theta} = 0.476$
		$\hat{\beta} = 3.450$	
	2-parameter Gamma	$\hat{\alpha} = 2.604$	
		$\hat{\beta} = 1.389$	
	3-parameter Gamma	$\hat{\alpha} = 1.760$	$\hat{\theta} = 0.446$
		$\hat{\beta} = 1.801$	
August	2-parameter Weibull	$\hat{\alpha} = 1.536$	
		$\hat{\beta} = 3.548$	
	3-parameter Weibull	$\hat{\alpha} = 1.260$	$\hat{\theta} = 0.391$
		$\hat{\beta} = 2.996$	
	2-parameter Gamma	$\hat{\alpha} = 2.156$	
		$\hat{\beta} = 1.473$	
	3-parameter Gamma	$\hat{\alpha} = 1.456$	$\hat{\theta} = 0.382$
		$\hat{\beta} = 1.920$	
September	2-parameter Weibull	$\hat{\alpha} = 1.568$	
		$\hat{\beta} = 3.423$	
	3-parameter Weibull	$\hat{\alpha} = 1.241$	$\hat{\theta} = 0.493$
		$\hat{\beta} = 2.747$	
	2-parameter Gamma	$\hat{\alpha} = 2.404$	
		$\hat{\beta} = 1.270$	
	3-parameter Gamma	$\hat{\alpha} = 1.475$	$\hat{\theta} = 0.483$
		$\hat{\beta} = 1.742$	
October	2-parameter Weibull	$\hat{\alpha} = 1.622$	
		$\hat{\beta} = 4.537$	
	3-parameter Weibull	$\hat{\alpha} = 1.308$	$\hat{\theta} = 0.484$
		$\hat{\beta} = 3.855$	
	2-parameter Gamma	$\hat{\alpha} = 2.186$	
		$\hat{\beta} = 1.854$	
	3-parameter Gamma	$\hat{\alpha} = 1.447$	$\hat{\theta} = 0.477$
		$\hat{\beta} = 2.471$	
November	2-parameter	$\hat{\alpha} = 1.621$	
	Weibull	$\hat{\beta} = 4.521$	
	3-parameter Weibull	$\hat{\alpha} = 1.322$	$\hat{\theta} = 0.463$
		$\hat{\beta} = 3.870$	
	2-parameter Gamma	$\hat{\alpha} = 2.211$	
		$\hat{\beta} = 1.826$	
	3-parameter Gamma	$\hat{\alpha} = 1.505$	$\hat{\theta} = 0.444$
		$\hat{\beta} = 2.389$	
December	2-parameter Weibull	$\hat{\alpha} = 1.515$	
		$\hat{\beta} = 4.514$	
	3-parameter Weibull	$\hat{\alpha} = 1.278$	$\hat{\theta} = 0.393$
		$\hat{\beta} = 3.948$	
	2-parameter Gamma	$\hat{\alpha} = 1.979$	
		$\hat{\beta} = 2.051$	
	3-parameter Gamma	$\hat{\alpha} = 1.424$	$\hat{\theta} = 0.388$
		$\hat{\beta} = 2.579$	

Tab. 4 summarises the values of the goodness-of-fit criteria that allow us to choose the most accurate probability distribution among the applied ones.

Tab. 4. The GOF and the model selection criteria for the applied probability distributions.

	AIC	*BIC*	R^2	*RMSE*	*KS* test	*AD* test
			January			
W2	3308.340	3317.564	0.994	0.023	0.050	3.150
W3	3286.152	3299.988	0.995	0.021	0.043	2.524
Gam2	3312.681	3321.905	0.994	0.024	0.047	3.313
Gam3	3304.409	3318.245	0.993	0.025	0.048	3.247
			February			
W2	3149.170	3158.190	0.989	0.033	0.079	4.895
W3	3121.195	3134.726	0.985	0.037	0.071	5.542
Gam2	3156.192	3165.212	0.984	0.039	0.076	6.340
Gam3	3139.340	3152.871	0.981	0.043	0.085	6.769
			March			
W2	3422.236	3431.461	0.980	0.044	0.092	8.824
W3	3389.323	3403.159	0.983	0.041	0.077	7.477
Gam2	3422.337	3431.561	0.980	0.044	0.082	8.915
Gam3	3397.810	3411.647	0.983	0.041	0.075	7.416
			April			
W2	3495.607	3504.765	0.987	0.035	0.069	5.694
W3	3486.118	3499.856	0.984	0.040	0.083	6.979
Gam2	3520.699	3529.857	0.981	0.042	0.088	7.588
Gam3	3517.992	3531.729	0.977	0.046	0.095	8.637
			May			
W2	3399.878	3409.102	0.996	0.019	0.042	1.860
W3	3396.047	3409.883	0.993	0.025	0.052	3.128
Gam2	3429.212	3438.436	0.986	0.035	0.071	5.341
Gam3	3431.212	3445.048	0.986	0.035	0.071	5.341
			June			
W2	2891.865	2901.024	0.977	0.045	0.099	8.454
W3	2852.689	2866.427	0.984	0.037	0.085	5.559
Gam2	2857.850	2867.008	0.982	0.041	0.093	6.659
Gam3	2839.810	2853.548	0.987	0.034	0.078	4.555
			July			
W2	3107.345	3116.569	0.990	0.030	0.063	4.603
W3	3071.579	3085.415	0.992	0.028	0.053	3.552
Gam2	3106.697	3115.921	0.990	0.031	0.058	4.815
Gam3	3090.157	3103.993	0.990	0.030	0.060	4.285
			August			
W2	3025.086	3034.311	0.982	0.041	0.093	7.335
W3	2961.509	2975.345	0.991	0.028	0.064	3.599
Gam2	3005.979	3015.204	0.985	0.038	0.083	6.470
Gam3	2965.695	2979.531	0.992	0.027	0.056	3.386
			September			
W2	2843.958	2853.117	0.983	0.038	0.079	7.142
W3	2748.740	2762.478	0.995	0.021	0.054	1.995
Gam2	2801.463	2810.622	0.986	0.035	0.083	5.374
Gam3	2745.727	2759.465	0.995	0.020	0.051	1.728

October						
W2	3354.628	3363.852	0.982	0.042	0.085	8.813
W3	3315.600	3329.436	0.984	0.040	0.077	7.629
Gam2	3361.959	3371.183	0.982	0.043	0.079	8.921
Gam3	3334.001	3347.837	0.983	0.040	0.077	7.753
November						
W2	3239.357	3248.515	0.986	0.037	0.076	5.884
W3	3205.517	3219.254	0.988	0.034	0.074	5.019
Gam2	3243.262	3252.421	0.985	0.038	0.076	6.214
Gam3	3223.374	3237.112	0.986	0.036	0.077	5.500
December						
W2	3407.498	3416.722	0.980	0.044	0.091	8.546
W3	3365.978	3379.814	0.983	0.040	0.083	7.074
Gam2	3408.548	3417.772	0.980	0.044	0.085	8.729
Gam3	3377.748	3391.584	0.983	0.040	0.088	7.090

Comparing the performance of all probability distributions, the 3-parameter Weibull distribution generally provided the best fit for most months. In February, the 3-parameter Weibull obtained the best results in terms of information criteria and the *KS* test, whereas the 2-parameter Weibull achieved the best values of the R^2, *RMSE* and the *AD* test. In April, the 3-parameter Weibull obtained the best results in terms of information criteria; however, according to the goodness-of-fit tests, the R^2 and *RMSE*, the 2-parameter Weibull provided the best fit. In May, the best fit is obtained by the 2-parameter Weibull. The 3-parameter gamma distribution performed as the best one in June and September (according to all criteria). In August, 3-parameter gamma distribution achieved the best results according to the goodness-of-fit tests, the R^2 and *RMSE*. According to the information criteria, the best fit is obtained by the 3-parameter Weibull.

When we compare the performance of the 2- and 3- parameter Weibull distribution, we can see that the 3- parameter distribution provided more accurate approximation than the 2-parameter distribution in majority of months. Similarly, the 3- parameter gamma distribution fitted the data better in comparison to the 2-parameter gamma distribution.

5. CONCLUSION

In the paper, we fitted the wind speed in the city Nitra by four probability distributions (2-parameter and 3-parameter Weibull, 2-parameter and 3-parameter gamma) to identify the probability distribution most suitable for modelling. All of them provided accurate enough fit; however, the Weibull probability distribution outperformed the gamma distribution in most months. The 3-parameter Weibull distribution obtained the best results in October to December, in January and in July. The 2-parameter Weibull distribution beat the rest of the distributions in April and May. The 3- parameter gamma distribution excelled in June and September.

From the comparison between the performances of the 2- and 3-parameter probability distributions, we can conclude that the 3-parameter distributions obtained better results than the 2- parameter ones. This indicates that the presence of the location parameter improves the results significantly.

Sources

1. AKPINAR, E.K., AKPINAR, S. Determination of the Wind Energy Potential for Maden. *Energy Conversion and Management*, 45, 2004. pp. 2901-2914.
2. CELIK, J.N. A Statistical Analysis of Wind Power Density based on the Weibull and Rayleigh Models at the Southern Region of Turkey. *Renewable Energy*, 29, 2003. pp. 593-604.
3. LUN, I.Y.F, LAM, J.C. A Study of Weibull Parameters using Long-term Wind Observations. *Renewable Energy*, 20, 2000. pp. 145-153.
4. SEGURO, J.W., LAMBERT, T.W. Modern Estimation of the Parameters of the Weibull Speed Distribution for Wind Energy Analysis. *Journal of Wind Engineering and Industrial Aerodynamics*, 85, 2000. pp. 75-84.
5. STEVENS, J.M., SMULDERS, P.T. The Estimation of the Parameters of the Weibull Wind Speed Distribution for Wind Energy Utilization Purposes. *Wind Engineering*, 3(2), 1979. pp. 132-145.
6. POBOČÍKOVÁ, I., SEDLIAČKOVÁ, Z., MICHALKOVÁ, M. Application of Four Probability Distributions for Wind Speed Modeling. *Proceedia Engineering*, 192, 2017. pp. 713-718.
7. MORGAN, E.C., LACKNER, M., VOGEL, R.M., BAISE, L.G. Probability Distributions for Offshore Wind Speeds. *Energy Conversion and Management*, 52(1), 2011. pp. 15-26
8. ALAVI, O., MOHAMMADI, K., MOSTAFAEIPOUR, A. Evaluating the Suitability of Wind Speed Probability Distribution Models: A Case Study of East and Southeast Parts of Iran. *Energy Conversion and Management*, 119, 2016. pp. 101-108.
9. SARKAR, A., SINGH, S., MITRA, D. Wind Climate Modeling using Weibull and Extreme Value Distribution. *International Journal of Engineering, Science and Technology*, 3(5), 2011. pp. 100-106.
10. KANTAR, Y.M., USTA, I., ARIK, I., YENILMEZ, I. Wind Speed Analysis using the Extended Generalized Lindley Distribution. *Renewable Energy*, 118, 2018. pp. 1024-1030.
11. HARE, W. *Assessment of Knowledge on Impacts of Climate Change - Contribution to the Specification of Art. 2 of the UNFCCC: Impacts on Ecosystems, Food Production, Water and Socio-economic Systems*. External expertise report for German Advisory Council on Global Change. Berlin, 2003 104 p.
12. HIROSE, H. Maximum Likelihood Parameter Estimation in the Three-Parameter Gamma Distribution. *Computational Statistics & Data Analysis*, 20, 1995. Pp. 343-354
13. AKAIKE, H. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6), 1974. pp. 716-723.
14. SCHWARZ, G. Estimating the Dimensions of a Model. *The Annals of Statistics*, 6(2), 1978. pp. 461-464.