Stochastic Modeling of Wind Turbines

Migran N. Gevorkyan[‡], Anastasiya V. Demidova[‡], Anna V. Korolkova[‡], Dmitry S. Kulyabov[‡§]

[‡] *Department of Applied Probability and Informatics*
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya St., Moscow, 117198, Russian Federation
[§] *Laboratory of Information Technologies*
Joint Institute for Nuclear Research
6 Joliot-Curie, Dubna, Moscow Region, 141980, Russian Federation

Email: gevorkyan_mn@rudn.university, demidova_av@rudn.university, korolkova_av@rudn.university, kulyabov_ds@rudn.university

In this paper we study statistical data collected from wind turbines located on the territory of the Republic of Poland. Sensors, mounted at three points of the wind power plant, took the readings of speed and wind direction. The readings were taken for 9 months with an interval of 10 minutes.

The general direction of research aimed at the construction of the stochastic model that predicts the change in wind speed depending on time. The aim of this work is to find the optimal distribution for the approximation of available statistical data on wind speed. We examine four distributions with heavy tails, namely, the log-normal, gamma, Weibull and beta distribution. Each distribution is parameterized by three parameters (the beta distribution has four).

For data processing, we used Python with NumPy, SciPy and Matplotlib. From the SciPy library we used statistics module that contains the function to search the parameters of the distribution provided by method of maximum likelihood.

After finding the parameters of the distributions were drawn the graphs of the density distributions, which were verified with the histogram of the frequency distribution. The obtained results allow to assert that all distributions with good accuracy can be used for the purposes of approximation. However, the study quantile-quantile graphs revealed that the Weibull distribution better approximate extreme values.

The obtained results are consistent with the data presented in the literature, where Weibull distribution is often used to approximate the distribution of the wind speed. Our future work will focus on the problem of constructing a stochastic differential equation.

The work is partially supported by RFBR grants No's 15-07-08795 and 16-07-00556. Also the publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement No 02.A03.21.0008).

Key words and phrases: wind speed, Weibull distribution, gamma distribution, beta distribution, log normal distribution.

In: K. E. Samouilov, L. A. Sevastianov, D. S. Kulyabov (eds.): Selected Papers of the VII Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems", Moscow, Russia, 24-Apr-2017, published at http://ceur-ws.org

Стохастическое моделирование ветроэнергетических установок

М. Н. Геворкян[‡], **А. В. Демидова**[‡], **А. В. Королькова**[‡], **Д. С. Кулябов**[‡§]

[‡] *Кафедра прикладной информатики и теории вероятностей*
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198

[§] *Лаборатория информационных технологий*
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, г. Дубна, Московская область, Россия, 141980

Email: gevorkyan_mn@rudn.university, demidova_av@rudn.university, korolkova_av@rudn.university, kulyabov_ds@rudn.university

В данной работе исследуются статистические данные, собранные с ветроэнергетических установок, находящихся на территории Республики Польши. Датчики, установленные в трех точках ветроэнергетической установки, снимали показания скорости и направления ветра. Показания снимались в течение 9 месяцев с интервалом 10 минут.

Общее направление исследований направленно на построение стохастической модели, предсказывающей изменение скорости ветра в зависимости от времени. Целью данной работы является поиск наиболее оптимального распределения для аппроксимации имеющихся статистических данных по скорости ветра. Мы исследуем четыре распределения с тяжелыми хвостами, а именно: логнормальное, гамма, Вейбулла и бета распределения. Каждое распределение параметризировано тремя параметрами (бета распределение — четырьмя).

Для работы с данными использовался язык Python с библиотеками NumPy, SciPy и Matplotlib. Из библиотеки SciPy мы использовали модуль статистики, который содержит функцию для поиска параметров распределения по предоставленной статистической выборке методом максимального правдоподобия.

После нахождения параметров распределений были построены графики функций плотности распределений, которые были сверены с гистограммой частоты распределения. Полученные результаты позволяют утверждать, что все распределения с хорошей точностью можно использовать для целей аппроксимации. Однако исследование квантиль–квантиль графиков выявило, что распределение Вейбулла лучше аппроксимирует экстремальные значения.

Полученные результаты согласуются с данными, изложенными в литературе, где распределение Вейбулла чаще всего используется именно для аппроксимации распределения скорости ветра. Дальнейшая наша работа будет сосредоточена на задаче построения стохастического дифференциального уравнения.

Работа частично поддержана грантами РФФИ № 15-07-08795, 16-07-00556. Также публикация подготовлена при финансовой поддержке Минобрнауки России (соглашение № 02.A03.21.0008).

Ключевые слова: скорость ветра, распределение Вейбулла, гамма распределение, бета распределение, логнормальное распределение.

1. Введение

Данная работа посвящена проблеме стохастического моделирования изменения скорости ветра, используемого для генерации электроэнергии на ветроэнергетических установках, находящихся на территории Республики Польша. В качестве первого этапа выбирается вид теоретического распределения случайной величины w, которая должна аппроксимировать скорость ветра. В данной работе мы исследовали несколько распределений на предмет аппроксимации располагаемых нами данных, а именно: логнормальное, Вейбулла, а также бета и гамма распределения [6]. Все рассматриваемые распределения имеют параметризацию коэффициентами формы, сдвига и масштаба (shape-location-scale). Для обработки статистических данных применялся язык Python 3 в связке с библиотеками NumPy, scipy.stats [1] и Matplotlib [2], а также интерактивной оболочкой Jupyter [3].

2. Описание структуры доступных для анализа данных

Статистические данные представляют собой таблицу величин в формате `csv`, составленную из следующих колонок:

1. T — время фиксации скорости и направления ветра датчиками, установленными на ветроэнергетической турбине (в формате чч:мм);
2. X_1 — мощность турбины [кВт] (отрицательные значения означают, что электроэнергия не производится, а потребляется);
3. X_2 — скорость ветра [м/с] (измеряется анемометром на вершине гондолы воздушной турбины);
4. X_3 — направление ветра [градусы] (измеряется анемометром на вершине гондолы воздушной турбины; измеряется по часовой стрелке, величина принимает значения от 0° до N°);
5. X_4 — скорость ветра на уровне 10 метров от поверхности земли [м/с];
6. X_5 — направление ветра на уровне 10 метров от поверхности земли [градусы];
7. X_6 — скорость ветра на уровне 50 метров [м/с] от поверхности земли;
8. X_7 — направление ветра на уровне 50 метров от поверхности земли [градусы].

Показатели скорости и направления ветра снимались с датчиков каждые 10 минут на протяжении 9 месяцев. В общей сложности таблица содержит 39606 записей.

Для первоначального выбора распределений, которые могут подойти для аппроксимации скорости ветра, мы построили гистограммы распределения скоростей ветров 1. Визуальная оценка этих гистограмм позволила предположить, что адекватным выбором будут распределения с так называемыми «тяжелыми хвостами».

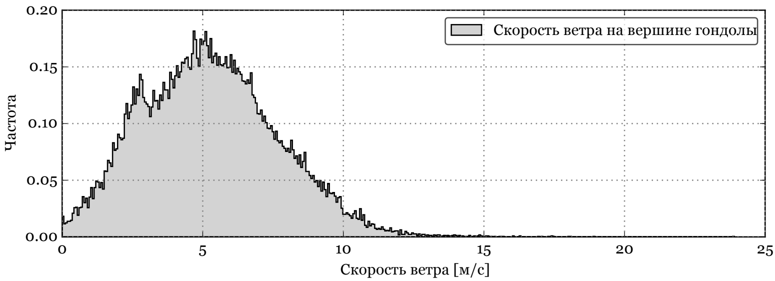

Рис. 1. Гистограмма распределения скоростей ветра, построенная по всем имеющимся данным

3. Используемые распределения

Каждое из рассматриваемых в работе распределений параметризовано тремя параметрами: α — коэффициент формы (shape), l — коэффициент сдвига (location) и s — коэффициент масштаба (scale). В случае бета-распределения добавляется еще второй коэффициент масштаба, обозначаемый буквой β. Для всех распределений параметры полагаются положительными действительными числами: $\alpha, \beta, s, l \in \mathbb{R}$, $\alpha, \beta, s > 0$, $l \geqslant 0$.

Напомним основные определения.

Случайная величина x распределена по *логнормальному закону*, если функция плотности вероятности задается следующей формулой:

$$f_{LN}(x; \alpha, l, s) = \begin{cases} \dfrac{1}{(x-l)\alpha\sqrt{2\pi}} \exp\left(-\dfrac{1}{2}\left(\dfrac{\ln(x-l) - \ln s}{\alpha}\right)^2\right), & x \geqslant l. \\ 0, \quad x < l. \end{cases}$$

Случайная величина X распределена по *Вейбуллу* [6], если функция плотности вероятности задается следующей формулой:

$$f_W(x; \alpha, l, s) = \begin{cases} \dfrac{\alpha}{s}\left(\dfrac{x-l}{s}\right)^{\alpha-1} \exp\left[-\left(\dfrac{x-l}{s}\right)^{\alpha}\right], & x \geqslant l. \\ 0, \quad x < l. \end{cases}$$

Случайная величина X подчинена *гамма-распределению*, если функция плотности вероятности задается следующей формулой:

$$f_\Gamma(x; \alpha, l, s) = \begin{cases} \dfrac{(x-l)^{\alpha-1} \exp\left(-\frac{(x-l)}{s}\right)}{s^\alpha \Gamma(\alpha)}, & x \geqslant l. \\ 0, \quad x < l. \end{cases}$$

где $\Gamma(\alpha)$ — гамма-функция.

Случайная величина X подчинена *бета-распределению*, если функция плотности вероятности задается следующей формулой:

$$f_{\mathcal{B}}(x; \alpha, \beta, l, s) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{s\Gamma(\alpha)\Gamma(\beta)}\left(\dfrac{x-l}{s}\right)^{\alpha-1}\left(1 - \dfrac{x-l}{s}\right)^{\beta-1}, & x \geqslant l, \\ 0, \quad x < l. \end{cases}$$

Если в формулах плотности вероятности логнормального и гамма-распределений положить $l = 0$, а у бета-распределения в добавок $s = 1$, то получим вид данных распределений, наиболее часто используемый в литературе [6].

4. Определение параметров распределений

В библиотеке `scipy.stats` [1] определены объекты `lognorm`, `weibull_min`, `gamma` и `beta`, которые реализуют рассматриваемые нами распределения. У каждого из данных объектов предусмотрены методы, которые реализуют функцию плотности вероятности `pdf(x, a, [b,] loc, scale)` и функцию распределения вероятности `cdf(x, a, [b,] loc, scale)`, где `x` — аргумент функции, `a`, `b` — параметры формы α (и β в случае бета-распределения), `loc` и `scale` — параметры смещения и масштаба.

Для нахождения параметров рассматриваемых распределений в `scipy.stats` предусмотрена функция `fit(data)`, которая вычисляет методом максимального правдоподобия параметры соответствующего распределения. Мы использовали данную функцию для вычисления параметров рассматриваемых нами распределений. После чего с помощью функций `pdf` и `cdf` были вычислены значения функции плотности вероятности и функции распределения.

Приведем пример кода для распределения Вейбулла. Сначала с помощью вышеупомянутой функции `fit` найдем параметры распределения Вейбулла:

```
c, loc, scale = scipy.stats.weibull_min.fit(ws1)
```

Вычислим теоретические значения распределения Вейбулла для имеющегося у нас диапазона значений случайной величины:

```
xs = np.linspace(np.min(ws1), np.max(ws1), 1000)
Weibull_PDF = scipy.stats.weibull_min.pdf(xs, c, loc, scale)
Weibull_CDF = scipy.stats.weibull_min.cdf(xs, c, loc, scale)
```

Теперь сравним доступные статистические данные с теоретическим распределением Вейбулла, построив гистограмму эмпирической плотности вероятности и теоретическую плотность вероятности, а также гистограмму эмпирической функции распределения и теоретическую функцию распределения:

```
fig21 = plt.figure(2, figsize=(10, 5))
fig22 = plt.figure(3, figsize=(10, 5))
ax21 = fig21.add_subplot(1,1,1)
ax22 = fig22.add_subplot(1,1,1)

for ax in [ax21, ax22]:
    ax.set_xlabel(r"Скорость ветра")
    ax.set_ylabel(r"Вероятность")

ax21.plot(xs, Weibull_PDF, label="Теоретическая плотность распределения",
          lw=3, color='k');
ax21.hist(ws1, bins=100, normed=1, label="Данные", color='#eaebee');
ax21.legend(loc=1, framealpha=0.8)

ax22.plot(xs, Weibull_CDF, lw=2.0, color='k',
          label="Теоретическая функция распределения");
ax22.hist(ws1, bins=100, normed=1, cumulative=True,
    histtype='stepfilled', color='#eaebee', label="Данные");
ax22.set_ybound(upper=1.05)
ax22.set_xbound(upper=20.0)
ax22.legend(loc=4, framealpha=0.8)
```

Результаты исследования представлены в графическом виде. На рис. 2 и 3 приведены результаты для распределения Вейбулла.

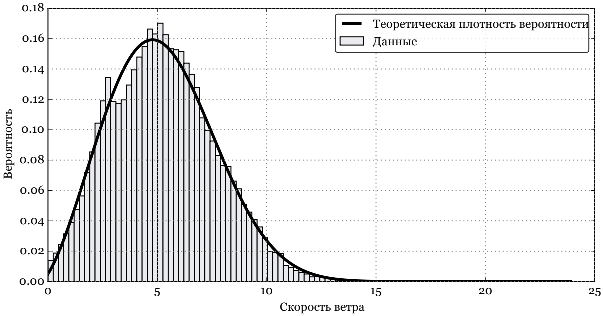

Рис. 2. Теоретическая плотность вероятности распределения **Вейбулла** в сравнении с эмпирическими данными

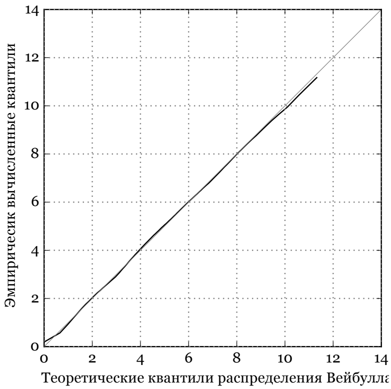

Рис. 3. Квантиль-квантиль график для статистических данных и аппроксимирующего распределения Вейбулла

Данные изображения были получены для теоретических распределений, параметры которых определялись на основе всего массива данных. Из рассмотрения квантиль-квантиль графиков (рис. 3) можно сделать вывод, что распределение Вейбулла подходит для аппроксимации имеющихся у нас данных.

5. Заключение

Результаты проделанной статистической обработки данных соответствуют результатам, изложенным в литературе [4–6], где распределение Вейбулла наиболее часто используется для аппроксимации скорости ветра. Дальнейшая наша работа будет направленна на построение стохастической модели, способной аппроксимировать скорость ветра в зависимости от времени. С другой стороны, мы рассчитываем верифицировать результаты данной работы на статистических данных с гораздо большим числом записей, чем имеющиеся в нашем распоряжении на данный момент времени.

Литература

1. Jones Eric, Oliphant Travis, Peterson Pearu et al., SciPy: Open source scientific tools for Python, 2001. URL: http://www.scipy.org/.
2. Droettboom Michael, Caswell Thomas A., Hunter John et al., Matplotlib/matplotlib: v2.0.0. 2017. Jan. URL: https://doi.org/10.5281/zenodo.248351.
3. Project Jupyter Home. 2017. URL: https://jupyter.org.
4. Fréchet Maurice René. Sur la loi de probabilité de l'écart maximum, Annales de la Société Polonaise de Mathematique (1927), pp. 93–116.
5. Weibull Waloddi, A statistical distribution function of wide applicability, Journal of Applied Mechanics, 1951, pp. 293–297.
6. Norman L. Johnson, Samuel Kotz, N. Balakrishnan. Continuous Univariate Distributions. Vol. 1 (Wiley Series in Probability and Statistics). Wiley-Interscience, 1994. ISBN: 0471584959,9780471584957.