Assessing risk of electricity capacity shortfalls using extreme value methods

Abstract

Although recently wind power generation plays really important role in ERCOT, its contribution to the power system is limited as the correlation between demand and wind generation is negative. This paper will conduct risk assessment by using a new statistical model which can estimate joint distribution of the conventional generation and renewable energy generations. For the estimation, three methodologies are used: the hindcast, the kernel density estimation, and the extreme value theory approaches. The model and approaches are required to compute risk adequacy metrics such as LoLE and EEU. Moreover, for uncertainty assessment, central estimation approach is utilised.

The most important result is that the extreme value theory works well in this survey from the perspective of the choice of threshold and the appropriateness of fit of the generalised Pareto distribution. Moreover, the kernel density estimation produces similar results to those in respect of the extreme value theory. Another finding is that the hindcast approach may underestimate risk adequacy values when compared with other two methodologies. However, differences of the LoLE and EEU within each year is large regardless of the chosen methodologies, which implies that year effect is huge and the uncertainty assessment would be more important than model selection.

Contents

List of Tables

List of Figures

1 Introduction

With the advancement of technology for a renewable energy, wind power generation plays an important role all over the world. This trend is true of the area covered by the Electric Reliability Council of Texas (ERCOT). The amount of installed wind generation increased dramatically from approximately 9,604 MW in 2010 to around 20,682 MW in 2017; thus the proportion of wind generation was over 20 per cent of the installed energy capacity in the region of the ERCOT in 2017. Furthermore, this upward trend is estimated to continue for a couple of years. Considering the profile of the wind units in the ERCOT, on the other hand, there is a critical issue: the demand is negatively correlated with the supply of the wind generation during the peak season, which is from June to September. To be precise, the peak of the load occurs from 2 p.m. to 8 p.m. but that of wind generation happens from 8 p.m. to 5 a.m. on the following day. This suggests that the contribution of wind power generation to the stability of the system is limited in the ERCOT. Therefore, risk assessment of the system is really important with the combination of both conventional generation and wind power generation in the area.

There have been many papers investigating the risk adequacy of the energy over decades. In early studies, the main method for conventional generation is based on the deterministic approach and later the probabilistic method is introduced for improvement of the accuracy of risk assessment in a system. The probabilistic model can consider a stochastic factor which may have a great impact on a risk metric. In response to the expanding use of renewable energy such as wind and solar power, a new risk assessment model with including renewable energy has recently been proposed. While research into this area has been developing, these models try to capture features of renewable energy whose profiles are totally different from that of conventional generation.

This paper uses the time collapse model for risk assessment in the ERCOT, mainly focusing the contribution of wind power generation. The model is currently the main capacity assessment methodology in developed countries such as the US and the UK. It is based on historical demand and wind generation data and does not consider the effect of time, since the model is used to calculate an expectation of some risk assessment metrics here. The data used in this survey, including demand, wind and conventional and weather variables, is mainly obtained through the database in the ERCOT and 12-year data series are used, from 2006 to 2017. With regard to risk adequacy, only the peak season of the data is taken from June to September and several estimations are conducted based on the data.

The 12-year series are rescaled to reflect the scenarios and future year under study, particularly 2018. With regard to rescaling, two methods are utilised for the demand and wind generation. For demand rescaling, a regression model is conducted and then the predicted demand is computed by considering the future-year scenario. On the other hand, in order to compute the rescaled wind generation, the wind factor approach is taken. The approach gives the rescaled wind power values by multiplying the installed wind unit capacity under study with the wind factor, which is the ratio of generated wind power to installed wind power units. In addition, to estimate the risk adequacy level, the distribution of conventional generation such as power units of gas, coal and nuclear requires being produced. This paper uses the two-state method, which repeats the convolution of each power unit with a probability of availability.

Among various types of the time collapse model, this paper uses the hindcast approach, which is widely accepted in academia and practice. The hindcast is based on the empirical distribution of the demand-net-of-wind based on the rescaled series of demand and wind generation. To compute the risk adequacy metrics, the distribution of the demand-net-of-wind needs to be estimated. Then, the distribution is convoluted with the distribution of conventional generation, which generates the distribution of the demand-supply balance. Based on this distribution, the loss of load expectation (LoLE) and the expectation energy unserved (EEU) can be computed.

However, like the typical issue of the statistic, the hindcast approach faces a critical problem: there are generally a small number of observations. For capacity assessment, the important point is the pair of high demand and low wind generation in the same hour. These cases are hardly observed and then would bring difficulty with accurate estimation of the distribution of the demand-net-of-wind. To address this issue, kernel density estimation and extreme value theory approaches are used. The kernel density smoothing is widely accepted as a method similar to empirical distribution and the extreme value theory leads to robust results for the interpretation of tails of the distribution. Both of them has an advantage that they do not require any statistical assumption of the relationship between demand and wind generation. In addition, the extreme value theory is thought to be a better approach considering that a parametric estimation is widely accepted in the research of energy risk assessment.

Moreover, meteorological uncertainty has to be considered, since the observed data series is not stationary. This type of uncertainty may have significant effects on the risk level of the power system, but it is extremely difficult to predict. Among the risk assessment, there are mainly two kinds of methods for addressing such uncertainty. One is the bootstrap and the other is the central estimation. So as to use the bootstrap, many series of data are required, but there are only 12 data series in this paper. Therefore there is no choice but to use central estimation which is simply equal to the mean of 12 risk metrics with a certain percentage of the confidence interval based on the $t-$distribution. This confidence interval presents the uncertainty.

By using the data of the ERCOT, the purpose of this paper is as follows:

- to investigate the appropriateness and effectiveness of the extreme value theory approach to the estimation of the demand-net-of-wind and the supply and demand balance.
- to compare the extreme value theory approach with other approaches such as the hindcast and the kernel density estimation.

An interesting result is obtained. First of all, the year effects are so large within the historical year that the differences between the chosen models can be ignored. In addition, the extreme value theory approach works well in this survey from the viewpoint of appropriateness and effectiveness. Furthermore, the LoLE and the EEU by the hindcast tend to be smaller than those by the other two approaches, which is not consistent with existing papers such as [44]. This implies that due to the limitation of the number of observations, the risk metrics are underestimated. Moreover, the results from kernel density estimation and the extreme value theory are really similar.

Following the introduction, wherein a brief summary and the purpose of this paper are described, the background and the motivation of this paper are shown in Chapter 2. In Chapter 3, the literature review of the energy risk assessment is explained. Thereafter, the mathematical model and the estimation methodologies are shown in Chapter 4. In Chapter 5, the original data and the rescaling thereof for "forward mapping" are reported. In Chapter 6, the results is shown and an analysis thereof is conducted. Then, Chapter 7 states the conclusion of this paper and proposes further research.

2 Background

2.1 Recent situation surrounding ERCOT

In the ERCOT, the contribution of wind generation to the power system has been increasing over a decade. Although the amount of installed wind generation in the area was around 9,604 MW in 2010, it dramatically increased to 20,682 MW in 2017. This trend will continue, and thus, the amount of installed generation is expected to be 24,633 MW in 2018 and 29,013 MW in 2019 according to an ERCOT estimation [18]. As a result of this, the proportion of wind generation in total installed energy capacity is approximately 20 per cent, which is the largest among renewable energy supplies in the ERCOT [42]. However, the contribution of other renewable generation is trivial compared with wind generation. To give an example, even though solar power generation in the ERCOT has the second-largest installed capacity among renewable energy, the installed capacity of solar power was only 1,007 MW in 2017, approximately only ten per cent of wind power generation. Furthermore, the contribution of solar power to the peak demand is not large, as the peak load of the ERCOT region usually occurs during the afternoon, when solar irradiances are less than around midday [18]. This implies that it is really important to make a design of the power system in the ERCOT wherein the wind power generation performs at its full potential.

2.2 Problem in ERCOT

On the other hand, the wind generation system in the ERCOT area is said to have a crucial issue. [42] describes that the contribution of wind power generation to the peak demand load is limited. The region of the ERCOT is divided into four zones: Houston Zone, North Zone, West Zone and South Zone. Because of abundant wind resources, the large majority of wind generation are located in the West zone. [42] expresses that 73 per cent of the wind units are built in the West region. The problem is that the typical profile for wind units in the area is pointed out to be negatively correlated with peak electricity demand during the peak season (from June to September). Like solar generation, furthermore, the predominant amount of wind generation is produced during off-peak hours, generally over night.

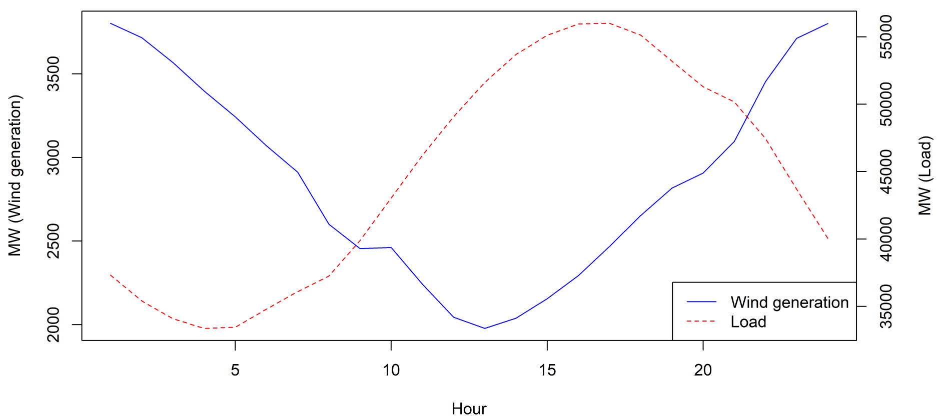

Figure 1: Aggregated wind power generation and demand by hour over the peak season (June to September) using data from 2006 to 2017.

This negative relationship is seen in Figure 1, which has two lines. The blue line is the mean of wind generations from 2006 to 2017 based on the day, and the dotted red line is the mean of all load over the same period and with the same classification of the hour. The original data derives from the webpage of the ERCOT [22]. These lines report that the amount of wind generation during daytime is relatively low. In particular, from 10 p.m. to 3 p.m. the amount is extremely low (under approximate 2,500 MW). During this time, the energy demand starts increasing and reaches its peak of around 60,000 MW at around 3 p.m. In contrast, a relatively large amount of energy is produced by the wind power units at midnight when the load decreases to its lowest level of 35,000 MW. To sum up, wind power generation is not used with their full capacity when it is most needed.

Following this situation, wind generators on the Gulf Coast of the South Zone have recently attracted many wind developers, since wind generators are much more highly correlated with the peak demand. The output of wind generation has generally a lower annual capacity factor. However, the plan to build wind units on the Gulf Coast has just started and it will take time to complete the plan.

Moreover, the ERCOT published a report [19] describing how the energy demand peak will be expected to continue increasing to 72,756 MW in 2018 and to 74,202 MW in 2019 respectively. This is based on the situation wherein the working population will grow and oil and natural gas production will be likely to continue expanding. Therefore, it is really important to conduct a capacity risk assessment for the power system in the ERCOT area while considering the contribution of wind generation [42].

3 Literature review

3.1 Capacity adequacy in power system

From the perspective of risk assessment or reliability of power systems, particularly conventional generation, there have been many papers published since 1930's. The fundamental concept of this topic is how the reserve margin — the amount of installed capacity exceeding the expected demand — is ensured in a power system. In terms of detailed methodology, see the following bibliography: [4], [5], [2], [3], [7], [8] and [37]. They present many important papers in the area of power system reliability evaluation, though I will not mention methods and approaches that the literature presents one by one in this paper. In the early period, deterministic models and techniques were widely used in practice and academia. To give an example of methodology, the reserve margin method is widely used. A certain fixed level of reserve criteria is set and a capacity adequacy test is conducted. If the installed generation capacity exceeds the demand level in respect of the fixed level, e.g. 10 per cent or 15 per cent, the system is considered adequate [34].

On the other hand, as [1] point out, this model would not be meaningful and robust, as it does not account for any probability or stochastic factors. Then, for improvement, traditional methods have been replaced with probabilistic risk approaches, which can consider stochastic factors that may have a significant impact on a power system. This shift from a deterministic to probabilistic approach would be more appropriate, as the assessment can be defined as a combination of probabilities and consequences. Many probabilistic models have been proposed and it is generally said that more accurate methodology needs to be more complicated. However, [16] proposes a method called the "Z-Method", which enables us to compute system adequacy quickly, which can be close to the system adequacy calculated through a complex model. The model has another advantage, i.e. it can consider other factors such as forced outage rates and unusual characteristics of power generation.

3.2 Risk indices

In the same way that many methodologies for power risk assessment have been developed, many indices have been created as well. When it comes to probabilistic approaches to adequacy, the loss of load probability (LoLP) and the loss of load expectation (LoLE) are often used as a risk metric. The LoLP is the probability that a system demand exceeds capacity at a given period and the LoLE concerns the expected number of hours for which a demand level is higher than that of supply per year. In addition, the expected energy unserved (EEU) is widely accepted in the research area of power risk assessment. The EEU is a metric of the amount of demand which is expected not to be served.

For the capacity value assessment, the effective load carrying capability (ELCC) and the equivalent firm capacity (EFC) are widely used. The ELCC is the additional load which the supply amount supports while maintaining the same risk level, and the EFC is the capacity which would give the same risk level. As I will discuss reliability models taking variable generation into consideration later in this section, these risk metrics can be applied to these models.

Related to these metrics, there are many indices representing the unavailability or unreliability of power units. They are used for computation of some adequacy and capacity metrics. [35] summarises briefly the background and explanations of these indices. One of the most well-known unavailability values of power plants is the forced outage rate (FOR). It is defined as follows:

$$\text{FOR} = \frac{\text{forced outage duration}}{\text{service duration} + \text{forced outage duration}}. \tag{3.1}$$

For instance, [47] estimates the distribution of conventional generation in the UK system by using a method called the capacity outage probability table. This is a method based on convolving each generator capacity and forced outage rate. The cumulative probabilities give the probability that demand will exceed the available generation [9]. For more detail, see 5.2.

3.3 Inclusion of wind generation

Due to advancement in technology and high capacity of renewable power such as wind and solar generation in various systems all over the world, probabilistic approaches taking renewable capacity into consideration have recently been developed. A risk assessment for power supply which includes renewable energy supply is not only extended to the probabilistic reliability assessment for conventional generation, as renewable energy supply highly depends on unpredictable natural resources whose profile is really different from conventional generation [30] and [33]. [17] describes the major differences between conventional and variable generation. First of all, available capacity from a variable generation unit may not be independent of that from other units or load. Although much research uses standard model in which each unit's available capacity is independent of others, this is in contrast to renewable energy units for which available capacity is dependent on weather patterns that correlate units locally and regionally. Weather patterns also influence load, resulting in possible correlations with variable generation capacity. Another difference is that capacity contributions for variable energy can vary greatly from year to year. Additionally, like the profile for generation units in the ERCOT area, depending on the month and time, the renewable resource can vary significantly.

In order to deal with these features of variable generation, various models considering renewable generation in a number of researches have been proposed. In terms of wind power, which is the main target of this paper, [13] proposes the chronological approach. It is based on the wind capacity factor, which is the ratio of average to total output of wind generation over a given period. The paper compares this chronological approach with the probabilistic approach, which is based on the combination of some distributions such as Weibull and normal distribution. Then, this paper finds that the chronological approach can approximate the capacity credit computed by the probabilistic approach as long as there is an appropriate time interval. [34] compares various models in the United States and reveals that either the ELCC or a time period based on the wind capacity factor is mostly taken. Furthermore, the ELCC for wind generation is correlated with the relation between demand and high LoLP periods. In addition, the Z-method proposed by [16] can consider wind generation. Besides, there are many specific differences between conventional generation and other variable energy than wind power. Much literature presents specific models which consider these differences according to the type of renewable variable generation. See [6], [17] and [32] for solar, [36] for tidal and [29] for wave generation.

Another issue with variable generation concerns whether the association between demand and variable generation is assumed to be independent over the studied season. While some assume this point, e.g. [10] and [40], [17] points out that this assumption is usually too strong without any evidence, since it is easily considered that when there is strong wind during a severe winter, energy demand will be high. Then, [17] proposes the new probabilistic model considering the relationship between demand and wind power generation, which is the general framework for probabilistic risk assessment in energy systems with inclusion of variable generation such as solar, wind and other renewable generation. It shows that by defining the values of conventional generation X_t, variable generation Y_t and demand D_t at point t in time ($t = 1, 2, \ldots, n$), random variable Z_t is computed through $Z_t = X_t + Y_t - D_t$ and this variable is used for risk adequacy and capacity assessments. Based on this random variable, some quantities can be calculated, such as the LoLP and LoLE by adjusting the time resolution. There are many papers taking this approach, e.g. [44] and [47]. Details of this probability model is shown in Chapter 4.

According to [17], it is necessary to clarify whether sequential models or time collapse models would be appropriate. While some literature uses time series methods such as ARMA for a capacity assessment [11], it is pointed out that when an expected value risk metric is required, a time collapse model will be enough sufficient. The time collapse model or snapshot model does not consider the relationship between time in demand and generation capacity. In other words, the marginal distribution of demand and generation of conventional and renewable generation over time are considered. In addition, it is said that the time collapse approach works well for risk assessment and is widely used in Great Britain and the United States [44]. When the time collapse model is used for risk assessment with inclusion of variable renewable generation, it is usually assumed that conventional generation and variable generation are independent. Under this assumption, joint distribution of demand and variable generation is estimated by using convolution with both of their marginal distribution. For the estimation, data of demand and conventional and renewable generation needs to be forward-mapped, which means that the data is rescaled based on the scenario and future year under study [47].

While there are various ways of the time collapse model, the hindcast approach is the most widely used in academic and practical research. This approach uses a distribution estimated by the empirical distribution of observations. The advantage of the hindcast is that the approach does not require any priori assumption about the form of both marginal and joint distribution. However, it would be difficult to quantify a risk through this approach, as there are often really little observations in the tail. As this is not only the case for risk reliability assessment in power systems, this issue prevents an accurate estimate of the tail of the distribution. It is a critical point, since it is the tail that we need to investigate for capacity evaluation in the power system [47].

3.4 Use of extreme value theory

In order to overcome this issue, [44] illustrates a method for capacity adequacy assessment by using the extreme value theory. The extreme value theory is a well-established methodology for the accurate interference of extremes of distributions which rarely happen. Therefore, it has been developed to assess the tail risk in many fields such as engineering, climatology and finance. In particular, the peaks over threshold, one of the major methods in the extreme value theory, is based on the asymptotic model and allows appropriate smoothing and extrapolation of empirical observations. [14] and [38] summarise not only theories but also many applications in a wide range of areas.

[44] applies the extreme value theory to estimate the right tail of distribution of demand-net-of-wind $(d_t - w_t)$, which is eventually convoluted with conventional generation variable X, and then convoluted distribution Z is used for reliability assessments such as the LoLE and EEU. It describes that advantages of the extreme value theory include not requiring any prior probabilistic assumptions about the relationship between demand and wind generation and to enabling us to use empirical data directly. [44] compares various capacity adequacy models by using Great Britain data and concludes that the model based on the extreme value theory is the most preferable as long as the appropriate threshold is chosen. In addition, it adds that the year effect is much greater than the impact of the choice of threshold.

3.5 Rescaling for demand and wind generation

As mentioned above, rescaling of demand and conventional and variable generation often needs to be conducted in respect of a future risk assessment in a power system. This process is really important because rescaling can combine historical observations with a future scenario of what is connected to the power system. As a result, an appropriate survey for the given year can be conducted. For example, when examining risk assessment in the UK, [47] rescales the demand by using each winter's average cold spell, which is the median outcome of winter peak demand conditional on underlying demand patterns.

To be precise, for demand there are many methods of rescaling, one of which is to build a prediction model. There are mainly two types of approaches to predicting energy demand: one is the statistical approach (including the regression model, time series model, exponential smoothing), and the other one is the artificial-intelligence-based approach (including the neural network, support vector machine, fuzzy logic). With the advent of computer power, the latter approach attract much attention because of its flexibility and capability against non-linearity. On the other hand, the artificial-intelligence-based approach is often criticised, since the techniques generally become a blackbox and some papers suggest that performance by the techniques is not entirely convincing ([15], [24] and [39]). The other method, the statistical approach, has disadvantages (compared with the artificial-intelligence-based approach) due to its limitation regarding non-linear observations of demand. However, under certain conditions such that a relationship between load and temperature is linear, this mathematical combination model works well in the prediction of demand and enables us to interpret physical phenomena easily [43]. Among various statistical models for demand prediction, a regression model is widely used all over the world. Moreover, in the context of demand rescaling, the typical method is to compute fitted values by considering future trend values.

In terms of a regression model for predicting demand series, much literature proposes various regression models such as time series approaches [26], [25], [27]. [43] provides an excellent description of how to build a regression model for predicting future load values. It reports that time factors such as the day of the week and the hour of the day and weather factors such as temperature and humidity are extremely important when considering a statistical model. In particular, many consultancy companies for power systems in the United States use temperature and humidity components for their load prediction models. Besides, it is often the case that the relationship between a load and a weather variable such as temperature is not linear, so a smoothing method is shown in [27] and [46].

In terms of wind rescaling, a similar mathematical approach can be taken. For instance, a regression model can be built by using wind generation as a dependent variable and also using some explanatory variables such as the wind speed. On the other hand, [34] describes that using the approach of the wind generation factor will bring a sufficient result.

3.6 Uncertainty assessment

Another important issue for capacity risk assessment is that of uncertainty assessment in the capacity credit, since data series such as the load and variable generation heavily rely on meteorological uncertainty and there is not usually sufficient data. [30] reports that data of multiple years is needed to obtain robust results. [45] supports this view, describing that longer historical meteorological data would be more desirable. In addition, [23] evaluates the robustness of risk value metrics by comparing various years' data in Ireland, concluding that as longer data is used, more meaningful estimates can be obtained. Therefore, one concludes that at least sufficient data can provide meaningful results. These researches also imply that the central estimate of parameters such as the LoLE would be reasonable under the assumption that the data of each year is independent and identically distributed. Here the central estimate is simple equals to the mean with a confidence interval, which represents uncertainty. In addition, the most robust way to compute the confidence interval is based on appropriate $t-$distribution [44].

An alternative way of uncertainty assessment is that of bootstrap analysis, which depends on random resampling with replacements and allows a reliability evaluation for estimates. The bootstrap method gives the mean, standard deviation and confidence interval. Many surveys in respect of power system capacity assessment use the bootstrap as their uncertainty evaluation [47]. Moreover, [27] expands the basic bootstrap method. It proposes bootstrap temperature resampling, which considers factors such as seasonal or trend patterns and the inherent serial association. However, a limitation claimed is that the bootstrap requires an adequate sample size and the method would not be meaningful without a large number of samples [44].

4 Mathematical model and estimation methodology

4.1 Probability model

This study uses the most widely accepted model for the capacity risk assessment, i.e. the time collapse or snapshot model. The model applies the excess of supply over demand at that time. The random variable can be presented as follows:

$$Z = X + W - D \tag{4.1}$$

where Z is the excess of supply over demand, X is conventional generation, W is wind generation, and D is demand. The equation is based on a stochastic process over the entire period of the study. This is needed for computation of some risk metrics [17]. Statistical analysis is based on the estimation of the joint distribution of the demand-net-of-wind $(D - W)$ and conventional generation (X), and the output of the model is the estimated distribution of the supply-demand balance (Z) over the course of a season. Random variable X and $D - W$ are assumed to be probabilistically independent. As [47] and [44] describe, this assumption is widely accepted, since there would be limited planned maintenance during a peak demand period in order to satisfy high demand. Therefore, the supply-demand balance Z can be obtained by convolving between $D - W$ and X. Moreover, this paper assumes that transmission constraints can be ignored in order to estimate the distribution of Z, as followed by [47].

4.2 Adequacy value metrics

Based on the time collapse model, the loss of load probability (LoLP) and the expected power unserved (EPU) can be defined as follows:

$$\text{LOLP} = \mathbf{P}(Z \leq 0) \tag{4.2}$$

$$\text{EPU} = \mathbf{E}(\max(-Z, 0)) = \int_{-\inf}^{0} dz(Z \leq z) \tag{4.3}$$

where $\mathbf{P}$ and $\mathbf{E}$ denote probability and expectation respectively. Furthermore, the loss of load expectation (LoLE) and the expected energy unserved (EEU) can then be defined by multiplying the length of the season. Thus:

$$\text{LOLE} = n\mathbf{P}(Z \leq 0) \tag{4.4}$$

$$\text{EEU} = n\mathbf{E}(\max(-Z, 0)) = n \int_{-\inf}^{0} dz(Z \leq z) \tag{4.5}$$

where n represents the number of hours over the peak season in this paper.

4.3 Estimation methodologies

For each year scenario, the statistical issue here is concerned with using the historical hourly two-dimensional time series of the observed demand and wind generation (d_t,w_t) of the pair (D, W) which is used for estimation of the underlying distribution of the demand-net-of-wind $D - W$. Here, t means a point in time. In order to make the distribution of supply-demand balance Z, the demand-net-of-wind distribution is required being convoluted with the distribution of conventional generation X. The distribution of supply-demand balance Z is the primary output of the time collapsed capacity assessment model.

In the survey of risk assessment, it is important to investigate the left tail of the distribution of Z, since the LoLE and EEU heavily rely on the region of the left tail wherein there is a possibility for the distribution to cover the negative value of supply-demand balance $X + W - D$. In other words, it is also important to examine the right tail of the distribution of the demand-net-of-wind $D - W$, which has a significant impact on supply-demand balance Z. In this paper, three types of methods for the energy risk assessment are utilised based on the interpretation of the left tail of the distribution of the demand-net-of-wind $D - W$: the hindcast, the kernel density and the extreme value theory. The reason as to why the latter two approaches are taken is to address the issue that there are fewer observations of pairs of high demand and low wind generation which leads to less robust estimations of the distribution of the supply-demand balance and of the computation of the risk metrics.

4.3.1 Hindcast

According to [47] and [44], the hindcast approach is the most widely accepted approach in the area of power assessment. In this estimation, the empirical historical distribution of the demand-net-of-wind $(d_t - w_t)$ is used as the predictive distribution. Then, following the convolution of conventional generation X, the distribution of supply-demand balance Z is shown by following:

$$\mathbf{P}(Z \leq z) = \frac{1}{n} \sum_{t=1}^{n} \mathbf{P}(X \leq z + d_t - w_t). \tag{4.6}$$

4.3.2 Kernel density estimation

With regard to smoothing, the kernel density estimation approach is often used as a non parametric estimation, since this approach may generate a distribution which is close to an empirical one. This is the reason as to why kernel density estimation is used in this paper. However, this approach is not common in the context of energy risk assessment. [12] applies kernel density estimation to the reliability assessment of wind generation. It describes that the estimation has advantages, i.e. the estimation is a data-driven approach and somehow flexible and requires no assumption regarding the form of the distribution. The estimator can be presented as follows:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\Big(\frac{x - x_i}{h}\Big) \tag{4.7}$$

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{4.8}$$

where h > 0 is a smoothing parameter, x_i represents the series of the demand-net-of-wind, and $K(x)$ is the kernel function. However, [12] also reports that the kernel density estimation has a limitation, i.e. the estimation can deal with small-scale and low-dimensional data series. In addition, it is said that it is better for this estimation to be used when a shape of a distribution is complex. On the other hand, distributions of the demand-net-of-wind in this paper are not so complex (Figure 8). This implies that there is no strong reason to use kernel density estimation here.

4.3.3 Extreme value theory

As introduced in literature part 3.4, the extreme value theory deals with extremes and gives an accurate interference of tails of a distribution. With regard to the extreme value theory, there are two main estimation models. The first model is called the block maxima. For an estimation by this model, the generalised extreme value distribution is fitted with data series which are created by taking the maximum value over the given time length, e.g. a half-year. However, the critical disadvantage of this method is that it requires much data. The data in this paper includes only 12 years, so it is not appropriate to use the block maxima estimation.

One of the other models in the extreme value theory is the peak over threshold. The model enables us to conduct more precise tail analyses of the distribution of the demand-net-wind.

Unlike the block maxima model, the peak over threshold approach smooths the tail of the distribution directly without any assumption about the statistical relationship between demand and wind generation, which is a huge advantage. In the model, the excesses first need to be defined:

$$Y = D - W - u \tag{4.9}$$

of $D - W$ above large enough threshold u, conditional on $D - W > u$. The excesses have a distribution known as a generalised Pareto distribution (GPD), whose function H against y is approximated by the following

$$H(y) = \mathbf{P}(Y \leq y) = 1 - \left(1 + \frac{\xi y}{\sigma_u}\right)^{-1/\xi} \tag{4.10}$$

where y is larger than zero and $1 + \frac{\xi y}{\sigma_u}$ is also greater than zero. Here there are two parameters: scale parameter σ and shape parameter ξ. They are estimated after an appropriate threshold u is chosen. It is said that scale parameter σ_u depends on the choice of threshold and increases as ξ is larger, since it is defined as follows:

$$\sigma_u = \sigma + \xi(u - \mu). \tag{4.11}$$

Shape parameter ξ is independent of threshold u for large enough u and is really important because it dominates in determining the qualitative behaviour of the distribution. If $\xi < 0$ the distribution of excesses follows a beta distribution, meaning that it has an upper bound. In contrast, if $\xi > 0$ the distribution follows a Pareto distribution and has no upper limit. If $\xi = 0$ the distribution corresponds to an exponential distribution with parameter $\frac{1}{\sigma_u}$. For more details, see [14] and [38].

According to [44], after an appropriate threshold is determined and parameters σ and ξ are obtained, the full distribution of demand-net-of-wind $D - W$ is given by its tail function while defining $v = d - w$:

$$\begin{aligned} \mathbf{P}(D - W > v) &= \mathbf{P}(D - W > u)\mathbf{P}(D - W > v | D - W > u) && (4.12)\\ &= \mathbf{P}(D - W > u)(1 - H(v - u)), \quad v \geq u. && (4.13) \end{aligned}$$

This equation means that $\mathbf{P}(D - W > u)$, which is larger than threshold u, is taken to be its empirical estimation. For values v of $D - W$ under threshold u, $\mathbf{P}(D - W > v)$ is also taken to be its empirical estimation.

As emphasised in [14] and [38], it is important to choose an appropriate threshold u. While this is a typical statistical issue, the balance between bias and variance should be considered. This means that too high a level of the threshold may produce few excesses which are estimated, leading to a good fitness of the generalised Pareto distribution but also high variance. By contrast, too low a level of the threshold may create a relatively large number of excesses, leading to high bias because the generalised Pareto distribution does not fit the excesses well. However, the choice of threshold is thought to be difficult because there has been no theoretical approach to obtain it. Rather, it has to be determined by empirical testing. Following the pragmatic approaches, this paper decides upon an appropriate level of the threshold by using the mean residual life plot and by comparing the results of the LoLE and EEU. While the details are demonstrated in Chapter 6, the mean residual life plot is based on the expectation of the generalised Pareto distribution which is obtained by following:

$$\mathrm{E}(Y) = \frac{\sigma}{1 - \xi}. \tag{4.14}$$

For a valid range of threshold u_0, the conditional expectation

$$\mathrm{E}(X-u|X>u) = \frac{\sigma_u}{1-\xi} \tag{4.15}$$

$$= \frac{\sigma_{u_0}+\xi u}{1-\xi} \tag{4.16}$$

$$\tag{4.17}$$

is plotted against u, where σ_{u_0} is equal to the scale parameter corresponding to threshold u_0. Then, for $u > u_0$, $\mathrm{E}(X-u|X>u)$ is a linear function of u. Therefore, an appropriate threshold u is decided when the line is linear.

The two parameters are obtained through maximum likelihood assuming that the observations are independent and identically distributed. This assumption may not be appropriate, since the data series of the demand-net-of-wind is not stationary on short scales. According to [44], however, the central estimates which are discussed in section 4.3.4 for uncertainty assessment can deal with this non-stationary issue; thus, the approach of maximum likelihood still provides meaningful parameters σ and ξ.

4.3.4 Considering Uncertainty factors

As mentioned in section 4.3.4, uncertainty assessment is really important because the process of the demand-net-of-wind in this paper is not a stationary process on short scales. There are possibly many factors influencing the analysis. From the perspective of time variables, not only years, days of the week and hours of the day can impact on the series of the demand-net-of-wind, but also any holiday effect, e.g. Christmas. Furthermore, weather variables such as temperature, wind speed and relative humidity can be true. All of these factors render an estimation of parameters difficult. Therefore, uncertainty has to be considered.

However, considering the fact that the time collapse model is used in this paper, the central estimates can deal with such uncertainly properly while satisfying at least one requirement. The requirement is that of sufficient long-term data series — 12-year time series of the demand-net-of-wind may be fine in respect of this requirement [44]. The central estimation simply takes the mean of a risk metric such as the LoLE and EEU and computes a confidence interval. The confidence interval represents the uncertainty mentioned above and computed based on a $t-$distribution instead of a normal distribution. This is because a $t-$distribution gives more robust results than does a normal one. Here the number of degrees of freedom is 11. In addition, when the central estimate is used in this paper, each risk metric needs to be assumed to be independent of each other. While there may be somewhat of a statistical relationship, it may not be easy to detect by using only 12-year series. The 12-year series is believed to be representative data of future scenarios.

As described in section 4.3.4, the bootstrap is a major alternative method to cope with an uncertainty. However, only a 12-year data series is used in this paper and this number is not sufficient to conduct the bootstrap.

5 Data

5.1 Demand and wind generation series

Estimation of the demand-net-of-wind is based on the data from the ERCOT. Historical hourly load data is available at [21] and data regarding the actual generated amount by the wind power units data with installed wind generation capacity is available at [22]. These series can be matched based on the hourly intervals. This paper uses the data from 2006 to 2017. According to Figure 2, the peak season can be defined as the months between June and September. Thus, this paper conduct a risk assessment by using the data over the peak season when the risk of a blackout is significantly high.

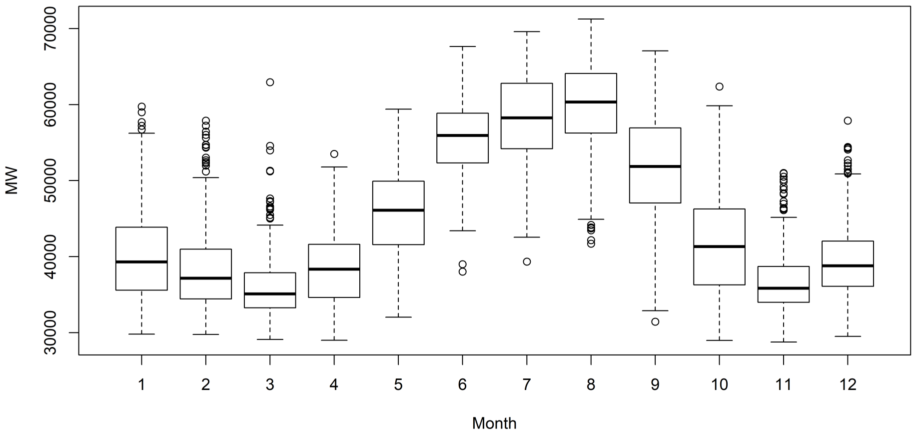

Figure 2: Box plot for demand by month from 2006 to 2017.

5.1.1 Rescaling demand series

Historical demand data needs to be rescaled to the required underlying level of the studied year by using the "forward-mapping" method. The purpose of this paper is to examine the risk assessment of 2018, so the demand and wind generation data has to be "forward-mapped" to time series appropriate to the predicted overall level of 2018. For rescaling the demand series, a regression model approach is taken in this paper. It means that an appropriate regression model is first built by considering the underlying variables, e.g. economic terms, and then predicted demand values of 2018 are computed by replacing values of the underlying variables with those of 2018. Here the hourly predicted demand series is needed; therefore, it is natural to build a regression model whose dependent variable is the hourly demand. As Figure 3 shows, however, the shape of distribution of demand over the peak period between 2006 and 2017 does not look appear to be a normal distribution; thus, it is difficult to capture the peak demands, which is essential for this paper.

Rather, the right tail of the histogram for the daily peak load over the peak season, as shown in Figure 4, satisfies the normality. When it comes to rescaling the demand, it is the right tail that is the most important for the risk assessment. Therefore, it is thought to be appropriate to build a regression model for predicting the peak demand and then make a necessary adjustment for the computation of the hourly predicted demand.

It is well known that the energy demand is highly related to weather, time and economic variables. When weather variables are used, data such as air temperature and relative humidity derives from [28]. The composite weather variable needs to be created by weighting each observatory. As it is not realistic to include all of the observatory points in the ERCOT area, only the major points are picked up in this paper. Based on the fact that the ERCOT area is divided into

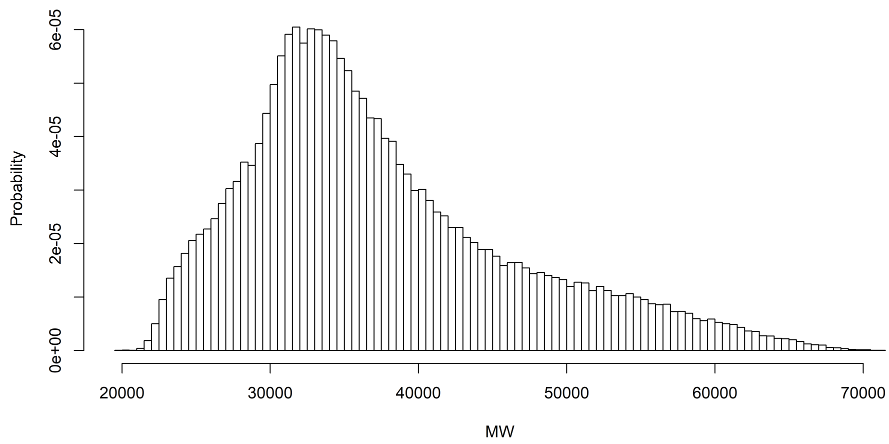

Figure 3: Histogram of hourly demand series over peak season from 2006 to 2017.

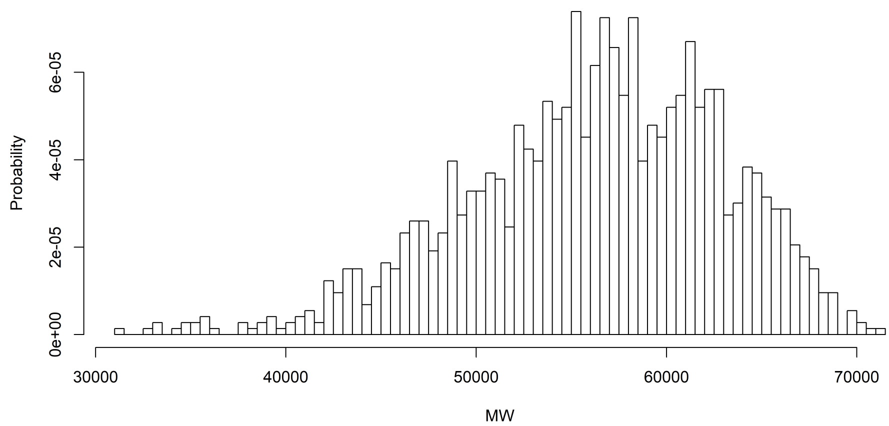

Figure 4: Histogram of daily peak demand series over peak season from 2006 to 2017.

four regions, the major cities of these four regions are picked up: Houston (Houston area), Dallas (North area), San Antonio (South area), and El Paso (West area). The weight is calculated based on the percentage of the actual energy consumption of 2017, which is 36 per cent Houston, 27 per cent Dallas, 27 per cent San Antonio and 10 per cent in El Paso [42]. In addition, one more thing should be mentioned: weather data is recorded with various time lengths, such as every hour, every half-hour and every five minutes. Here the values at the nearest time of demand and wind generation data are used for regression modelling.

Many variable combinations are tested in the calculation. By referring to [27] and [46], the following variables are used to find the appropriate regression model:

- The current temperature;
- The one-hour lagged temperature;
- The two-hour lagged temperature;
- The three-hour lagged temperature;
- The 24-hour lagged temperature;
- The current relative humidity;

- The one-hour lagged relative humidity;
- The two-hour lagged relative humidity;
- The three-hour lagged relative humidity;
- The 24-hour lagged relative humidity;
- The current wind speed;
- Smoothed hourly temperature, which is defined as follows:

$$te_h = \frac{1}{2}(te_{h-24} + to_h), \text{ where} \tag{5.1}$$
$$to_h = \frac{1}{4}(temp_h + temp_{h-1} + temp_{h-2} + temp_{h-3}) \tag{5.2}$$

  and $temp_h$ is the temperature in hour h in degrees Celsius;
- Smoothed hourly relative humidity, which is defined as follows:

$$he_h = \frac{1}{2}(he_{h-24} + ho_h), \text{ where} \tag{5.3}$$
$$ho_h = \frac{1}{4}(hmd_h + hmd_{h-1} + hmd_{h-2} + hmd_{h-3}) \tag{5.4}$$

  and hmd_h is the relative humidity in hour h in per cent;
- The day of the week;
- The holiday effect [1];
- The gross domestic product of Texas;
- The synthetic variable named t, which is simply integer, 1, 2, $\cdots$.

Then, the Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) are used to select the best model. Here the used AIC is as follows:

$$AIC = 2k - 2log(\hat{L}), \tag{5.5}$$

and the BIC is:

$$BIC = klog(n) - 2log(\hat{L}). \tag{5.6}$$

where k represents the number of parameters estimated in the model, $log(\hat{L})$ is the maximised value of the likelihood function of the model, and n concerns the number of observations.

The best model for the prediction of the daily peak demand over the peak season is below.

$$pd_t = \beta_0 + \beta_1 temp_t + \beta_2 hmd_t + \beta_3 hd_t + \beta_4 t + \sum_{j=1}^{2} \Big(\beta_5^j sin\Big(\frac{2j\pi t}{365.25}\Big) + \Big(\beta_6^j cos\Big(\frac{2j\pi t}{365.25}\Big)\Big) \tag{5.7}$$

where pd_t is the observed daily peak demand (MW), $temp$ is the temperature at the hour of the load (degrees Celsius), hmd is the relative humidity at the hour of the load (percentage), hd is a dummy variable which is equal to 1 when the day is during a holiday and at the weekend, and t represents the consecutive integer numbers from 1. The remaining variables are Fourier terms with which to capture the annual variation. The reason as to why t is chosen is that there

[1] Here a holiday includes Saturday, Sunday, Independence Day and Labour Day.

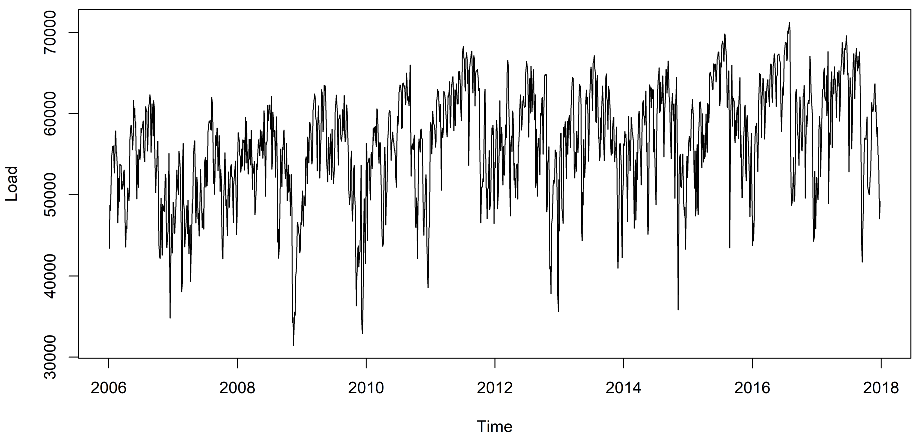

Figure 5: Daily peak demand series over peak season against period from 2006 to 2017.

is clearly an upward trend in the daily peak demand during the peak season between 2006 and 2017 (Figure 5) [2].

As expected, according to the QQ plot in Figure 6, the model does capture the right tail of the distribution of the daily peak demand over the peak season. Actually it does not capture the left tail of the distribution, but this may not be important for the risk assessment. On the other hand, the fitted hourly demands tend to be higher than the actual figures, as the regression model is based on the daily peak demand. Then, the gap between the mean of the hourly fitted demand value (55,990.92 MW) and the mean of the observed hourly demand value (44,723.55 MW) is subtracted from all of the fitted values of demand. At the same time, this manipulation leads to the extraordinary lower values in 2006 than the observed numbers in the range of relatively low demand. Therefore, the numbers of the low range in 2006 are somehow adjusted, which does not have a significant effect on the computation of the LoLE and EEU.

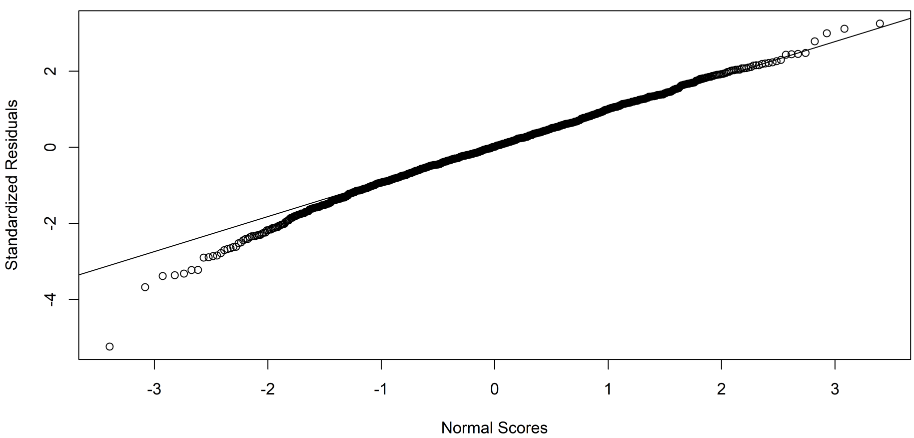

Figure 6: Normal QQ plot for the regression model for the daily peak demand over the peak season.

Then, by replacing the values of underlying variable t with the values of 2018 and substituting the other variable of each year, 12 rescaled demand series for 2018 are obtained. The mean of the highest demand of the 12 series is 72,185.9 MW with 95 per cent of the confidence interval based on t-distribution with 11 degrees of freedom (70,562.77, 73,809.04). A report published

[2] The domestic gross product of Texas is not appropriate, since it does not continue increasing from 2006 to 2017. Thus, synthetic variable t is chosen.

by the ERCOT supports the validity of this number. The ERCOT published a report called Seasonal Assessment of Resource Adequacy, which contains the prediction of the load peak of the year. For around 2018, [19] gives the peak load prediction of summer 2018, which is 72,746 MW. This is really close to the mean calculated by the model that this paper builds and also is inside of the confidence interval mentioned previously.

5.1.2 Rescaling wind generation series

In order to rescale wind power generation, the wind factor first needs to be calculated and then multiplied it with the installed wind generation capacity of 2018. The wind factor is computed through the actual wind generation divided by the installed wind generation capacity at the time. [18] reports that the estimated installed wind generation capacity would be 24,633 MW.

While there are possibilities of improvements of the rescaling for wind generation, since there has been much literature studying the estimation of wind power generation which considers more complex models e.g. [31], this paper utilises the method of rescaling, which is widely used in a pragmatic context, and its validity is supported by [13]. As [44] describes, however, there may need to be further research into how the observed pair of demand and wind, (d_t, w_t) should be treated. In this paper (d_t, w_t) is assumed to be the true pair, as they are recorded at the same hour. On the other hand, they are not actually.

5.1.3 Data visualisation of demand-net-of-wind

Figure 7 illustrates the relationship between the rescaled demand and the rescaled wind generation at the same hour over the 12-year series. The three straight lines identify these contributions to the LoLE, showing that only the most extreme values of demand D with the lowest values of wind generation W make a significant contribution to the LoLE. This figure supports the view that the hindcast approach may not be robust and then the extrapolation method may be needed to obtain appropriate risk metrics such as kernel density estimation and the extreme value theory.

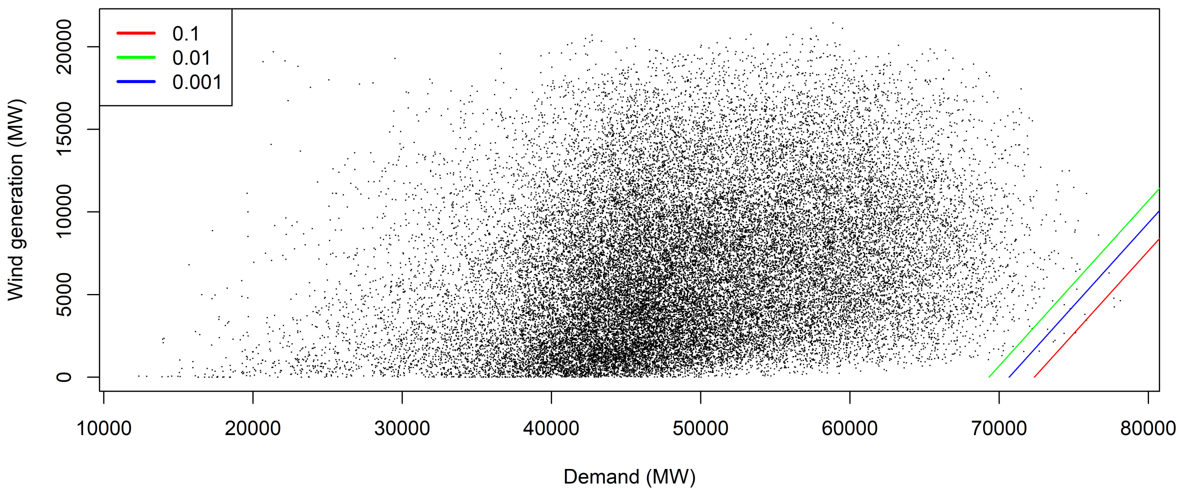

Figure 7: Scatter plot of wind generation against demand: the three lines identify the contribution to the LoLE of each point.

Type of power unit	FOR
Nuclear	0.016
Coal	0.058
Gas	0.070
Biomass	0.058
Hydro	0.000
Solar	0.000

Table 1: Forced outage rate in ERCOT.

5.2 Conventional plant data

Regarding the estimation of conventional generation, this paper uses a two-state distribution with either zero or maximum capacity [47]. This method is widely used in the context of computing risk metrics. This approach requires two types of data: the capacity of each individual unit and the availability of each unit. [20] provides a list of available power generation units and the maximum capacities during the summer of 2018 in the ERCOT area. It contains data of nuclear, coal, gas, biomass, hydro, solar and wind power units. This paper defines conventional generation as other types of generation than wind generation, and estimates conventional generation distribution. [20] reports that there are 450 power units of nuclear, coal, gas, biomass, hydro and solar in total. Capacity availability is often computed through one minus forced outage rate (1−FOR). The FOR data in Table 1 is given by Gene Preston, who is a consultant and undertakes much work with the ERCOT.

The distribution of available conventional capacity is estimated while assuming that the availability of each power plant is statistically independent and the total available conventional capacity is obtained by the following:

$$X = \sum_u X_u = X_1 + X_2 + \cdots + X_n \tag{5.8}$$

where n represents the number of units in the system. Then, each generator capacity and forced outage rate is convolved via an iterative method, which produces the distribution of available conventional generation. The result of convolution leads to the distribution of available conventional capacity X, which has a mean of approximately 74,092 MW and a standard deviation of around 1,311 MW.

6 Analysis and result

6.1 Validation of extreme value theory

In this subsection, two main issues in respect of using the extreme value theory are discussed: the choice of threshold, and the assessment of the generalised Pareto distribution fit. While the difficulty of choosing an appropriate threshold has been mentioned previously in section 4.3.3, this does not apply to this study. A wide range of thresholds are examined by using the mean residual life plot and by comparing the results of the LoLE and EEU. Then, it is found that thresholds in the range of 56 GW to 58 GW work well for all of the scenarios and all of the thresholds of this range compute really similarly to the LoLE and EEU. According to Figure 9, which shows the results of the mean residual life plots for all of scenarios, the lines between 56 GW and 58 GW appear to be straight. For the LoLE and EEU, details are shown in the three right columns of Tables 2 and 3. These columns report the LoLE and EEU based on the extreme value theory with three different thresholds u between 56 GW, 57 GW and 58 GW. These tables illustrate that the values of the LoLE and EEU are really similar among the three thresholds. Rather, the differences between each year are more significantly huge than the choice of threshold. Therefore, one can conclude that when year effects are considered, differences through the choice of threshold u can be almost ignored. This conclusion is consistent with [44].

Assessments of the validation of the generalised Pareto distribution fit are examined through the quantile plots, which illustrate the association between the ordered observations of the demand-net-of-wind above threshold u and "idealised values" predicted by the fitted generalised Pareto distribution model. One of the advantages of quantile plots is that they are an effective tool when the observed numbers have an extreme value. The quantile plots are shown as follows.

$$\left\{\left(\hat{F}^{-1}\left(\frac{i}{n+1}\right), x_{(i)}\right) : i = 1, \ldots, n\right\} \tag{6.1}$$

where $\hat{F}^{-1}$ is the inverse function of candidate model $\hat{F}$. If model $\hat{F}$ is an appropriate for the population distribution, the points of the quantile plot have to lie close to the unit diagonal. Otherwise, the candidate model is not reasonable enough to say that this is a good model. For more details see [38] and [14]. According to Figures 10, 11 and 12, the generalised Pareto distribution fits of the 12 different scenarios look fine.

Thus, one can conclude that the extreme value theory approach works well for the estimation of the right tail of the demand-net-of-wind in this survey.

6.2 Comparison between different approaches of estimation

Tables 2 and 3 show the results of the LoLE and EEU using three different approaches: the hindcast, the kernel density, and the extreme value theory. Here both of the LoLE and EEU are computed based on each historical observed condition and the central estimates of the LoLE and EEU are obtained by taking the mean with a 95 per cent confidence interval based on the $t-$distribution with 11 degrees of freedom. From these tables, it is clear that the year effect within the historical data is large. The LoLE and EEU are significantly larger in some years than in other years, regardless of what estimation approaches are used. For instance, the LoLE for 2010, 2011 and 2013 is between around 1.5 hours and approximately 3.5 hours. Meanwhile, the LoLE for the other years lies in the range of between 0 and around a half-hour. The same feature applies to the case of the EEU. The values of the EEU in 2010, 2011 and 2013 are also bigger when compared with those of the other years in all of the approaches. The main considered year effects here are weather severities such as the degree of temperature and relative humidity during summer. A meteorological record supports this view. When the ERCOT estimates a severe summer, it often uses the meteorological data series of 2011, e.g. temperature and relative humidity, because demand and supply were historically tight in this year due to extreme weather conditions [19].

While [44] describes that it is often seen that the differences of estimations between the hindcast and the extreme value theory are not huge, the table reports that the estimations of the LoLE and EEU by the hindcast tend to be smaller than those the extreme value theory. For example, even though the values themselves are not big, the differences of the LoLE are significantly huge when the LoLE is calculated by the hindcast and the extreme value theory is small, e.g. less than one hour. The differences are greater than 10 times. When the LoLE is not so small such as 1-3 hours, by contrast, the gap between the LoLE is almost double. The central estimate illustrates the good summary of these differences. The estimation by the hindcast is 0.546 and that by the extreme value theory with the threshold of 57 GW is 1.034. The gap is almost double.

This feature can apply to the case of EEU computation. There are clear differences between the hindcast and the extreme value theory approach: the EEU in respect of the hindcast is inclined to be smaller than that in relation to the extreme value theory. The central estimation by the hindcast is 0.180 and that by the extreme value theory is 0.458.

On the other hand, the differences of the LoLE and EEU between the kernel density estimation and the extreme value theory are not so big. Although almost all of the numbers of the LoLE and EEU based on the kernel density approach are slightly larger, the differences are so tiny that they can be ignored. Rather, the year effect is much larger. With regard to the LoLE, the central estimate by the kernel density is 1.048 with a confidence interval from 0.172 to 1.924 and the estimate by the extreme value theory with the threshold of 57 GW is 1.034 with a range of 0.166 and 1.903. A really similar trend can be seen in Table 3. The mean in respect of the kernel density is 0.458 and that related to the extreme value theory with the threshold of 57 GW is 0.451. Their confidence interval is $(0.022, 0.894)$ and $(0.020, 0.883)$ respectively.

The last thing which should be mentioned is that for the LoLE and EEU, all of the cases have wide confidence intervals because there is a limited amount of historical data. The most remarkable case concerns the LoLE and EEU computed by the hindcast. Their confidence intervals include negative values: $(-0.044, 1.135)$ and $(-0.025, 0.385)$ respectively. Considering real life, they could not appear. This implies that much more historical series would be needed.

To sum up these results and analysis, one can conclude that the year effect is the most dominant factor for risk metrics such as the LoLE and EEU. This is consistent with [44] and suggests that uncertainty assessment may be more important that model choice. A further finding is that while many papers such as [47] and [44] describe how the hindcast is good enough to estimate the risk metrics, the approach does not work well in this paper. My experiment illustrates that the hindcast tends to underestimate the LoLE and EEU due to the limitation of the observed numbers. Instead, the other two methods, the kernel density and the extreme value theory, can provide reasonable results by smoothing or extrapolating some unobserved values and their differences are so small that they can be ignored when the year effect is large.

Scenario	Hindcast	Kernel	EVT(56GW)	EVT(57GW)	EVT(58GW)
2006	0.025	0.518	0.504	0.509	0.509
2007	0.003	0.016	0.016	0.016	0.016
2008	0.033	0.189	0.184	0.188	0.190
2009	0.056	0.654	0.610	0.610	0.627
2010	1.615	2.586	2.555	2.575	2.575
2011	1.781	3.392	3.235	3.327	3.303
2012	0.015	0.145	0.138	0.138	0.144
2013	2.522	3.540	3.495	3.528	3.497
2014	0.000	0.000	0.000	0.000	0.000
2015	0.242	0.735	0.726	0.723	0.732
2016	0.180	0.444	0.440	0.439	0.441
2017	0.075	0.359	0.355	0.357	0.355
Mean	0.546	1.048	1.021	1.034	1.032
Lower	-0.044	0.172	0.168	0.166	0.171
Upper	1.135	1.924	1.875	1.903	1.894

Table 2: LoLE estimates: EVT represents extreme value theory. Lower is the lower bound and Upper is the upper bound of the 95 per cent confidence intervals for the mean LoLE computed for the five different approaches.

Scenario	Hindcast	Kernel	EVT(56GW)	EVT(57GW)	EVT(58GW)
2006	0.004	0.145	0.141	0.142	0.142
2007	0.000	0.004	0.004	0.004	0.004
2008	0.007	0.059	0.057	0.059	0.059
2009	0.011	0.199	0.184	0.184	0.190
2010	0.542	1.132	1.116	1.126	1.126
2011	0.575	1.600	1.512	1.563	1.550
2012	0.003	0.042	0.049	0.049	0.042
2013	0.898	1.806	1.776	1.797	1.777
2014	0.000	0.000	0.000	0.000	0.000
2015	0.056	0.249	0.245	0.244	0.247
2016	0.043	0.149	0.147	0.147	0.147
2017	0.016	0.109	0.108	0.108	0.108
Mean	0.180	0.458	0.444	0.451	0.449
Lower	-0.025	0.022	0.021	0.020	0.022
Upper	0.385	0.894	0.867	0.883	0.877

Table 3: EEU estimates: EVT represents extreme value theory. Lower is the lower bound and Upper is the upper bound of the 95 per cent confidence intervals for the mean EEU computed for the five different approaches.

7 Conclusion

By mainly following [44], which conducts an analysis of the UK, this paper investigates energy risk adequacy in the area covered by the ERCOT. For the estimation of the distribution of supply-demand balance and of the risk metrics such as the LoLE and EEU, hindcast, kernel density estimation and extreme value theory approaches are used. With regard to uncertainty, mainly in relation to meteorological factors, the method utilised is called central estimation, which simply computes the mean with the 95 per cent confidence interval based on $t-$distribution.

Being consistent with the existing literature, the extreme value theory works really well for tail analysis of the distribution of the demand-net-of-wind from the perspective of the choice of threshold. Furthermore, one can confirm the appropriateness of the generalised Pareto distribution fit. The mean residual life plots and the quantile plots of $u = 56$, $u = 57$ and $u = 58$ support this view. This leads to the robust result of the distribution of the supply-demand balance and of risk metrics such as the LoLE and EEU. Thus one can conclude that considering the feature of the extreme value theory that the method requires no statistical assumption between the demand and wind generation and can conduct an appropriate extrapolation for the tail of the distribution, the extreme value theory approach is the most preferable way. Moreover, the kernel density estimation leads to meaningful results when compared with the extreme value theory approach.

One more important conclusion of this paper is that the year effect is much larger than the differences of the choice of the approaches used. In other words, estimation of the distribution of the supply-demand balance is significantly influenced by the observation of the pair of demand and wind generation (d_t, w_t). By overwhelming the year effects, the central estimation is taken. The confidence interval explains the uncertainty, mainly with regard to meteorological aspects. Regarding this issue, more data series may be needed and then the bootstrap analysis may also be required if an appropriate number of data series can be obtained.

A different point from previous papers is that the results derived from the hindcast are not close to those obtained through the kernel density estimation and the extreme value theory approaches. Whereas previous literature reports that a result from the hindcast is meaningful enough to explain risk adequacy, the risk metrics through the hindcast in this paper tend to be underestimated compared with other approaches.

To sum up, the extreme value theory approach to the risk assessment based on renewable energy generation may be strongly recommended. The kernel density may also be true. However, the impacts of year effects on risk adequacy are more significant than the choice of models and one can say that how to deal with meteorological uncertainty is a more crucial problem. Therefore, further investigation would be needed to establish the appropriateness of the extreme value theory approach while ensuring appropriateness of dealing with uncertainty.

References

[1] Allan, R.N. and Billinton, R., 2000. Probabilistic assessment of power systems. Proceedings of the IEEE, 88(2), pp.140-162.

[2] Allan, R.N., Billinton, R., Breipohl, A.M., and Grigg, C.H., 1994. Bibliography on the application of probability methods in power system reliability evaluation: 1987-1991. IEEE Transactions on Power Systems, 9(1), pp.41-49.

[3] Allan, R.N., Billinton, R., Breipohl, A.M., and Grigg, C.H., 1999. Bibliography on the application of probability methods in power system reliability evaluation: 1992-1996. IEEE Transactions on Power Systems, (14), pp.51-57.

[4] Allan, R.N., Billinton, R. and Lee, S.H., 1984. Bibliography on the Application of Probability Methods in Power System Reliability Evaluation 1977-1982. IEEE Power Engineering Review, (2), pp.24-25.

[5] Allan, R.N., Billinton, R., Shahidehpour, S.M., and Singh, C., 1988. Bibliography on the application of probability methods in power system reliability evaluation, 1982–1987, IEEE Transactions on Power Systems, (3), pp.1555–1564.

[6] Awara, S., Zareipour, H. and Knight, A., Solar Power Capacity Value Evaluation - A Review.

[7] Billinton, R., 1972. Bibliography on the application of probability methods in power system reliability evaluation. IEEE Transactions on power Apparatus and systems, (2), pp.649-660.

[8] Billinton, R., Fotuhi-Firuzabad, M. and Bertling, L., 2001. Bibliography on the application of probability methods in power system reliability evaluation 1996-1999. IEEE Transactions on Power Systems, 16(4), pp.595-602.

[9] Billinton, R. and Allan, R.N., 1994. Reliability evaluation of power systems 2nd edition. Plenum.

[10] Billinton, R. and Huang, D., 2011. Incorporating wind power in generating capacity reliability evaluation using different models. IEEE Transactions on Power Systems, 26(4), pp.2509-2517.

[11] Billinton, R., Karki, R., Gao, Y., Huang, D., Hu, P. and Wangdee, W., 2012. Adequacy assessment considerations in wind integrated power systems. IEEE Transactions on Power Systems, 27(4), pp.2297-2305.

[12] Bo, H.U., Yudun, L.I., Hejun, Y.A.N.G. and He, W.A.N.G., 2017. Wind speed model based on kernel density estimation and its application in reliability assessment of generating systems. Journal of Modern Power Systems and Clean Energy, 5(2), pp.220-227.

[13] Castro, R.M. and Ferreira, L.A., 2001. A comparison between chronological and probabilistic methods to estimate wind power capacity credit. IEEE Transactions on Power Systems, 16(4), pp.904-909.

[14] Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. London: Springer.

[15] Darbellay, G.A. and Slama, M., 2000. Forecasting the short-term demand for electricity: Do neural networks stand a better chance? International Journal of Forecasting, 16(1), pp.71-83.

[16] Dragoon, K. and Dvortsov, V., 2006. Z-method for power system resource adequacy applications. IEEE Transactions on Power Systems, 21(2), pp.982-988.

[17] Duignan, R., Dent, C.J., Mills, A., Samaan, N., Milligan, M., Keane, A. and O'Malley, M., 2012, July. Capacity value of solar power. In Power and Energy Society General Meeting, 2012 IEEE (pp.1-6). IEEE.

[18] ERCOT, 2018. ERCOT in 2018: Challenges and Opportunities, Available at http://www.gaccsouth.com/fileadmin/ahk_atlanta/Bilder/Consulting/Federal_Projects/Energiedialog/1_MAGNESS_-_Challenges_and_Opportunities.pdf

[19] ERCOT, 2018. Final Seasonal Assessment of Resource Adequacy for the ERCOT Region (SARA) Summer 2018, Available at http://www.ercot.com/content/wcm/lists/143976/SARA-FinalSummer2018.pdf

[20] ERCOT, 2017. Report on the Capacity, Demand and Reserves (CDR) in the ERCOT Region, 2018-2027, Available at http://www.ercot.com/content/wcm/lists/143977/CapacityDemandandReserveReport-Dec2017.pdf

[21] ERCOT, Hourly Load Data Archives, Available at http://www.ercot.com/gridinfo/load/load_hist/

[22] ERCOT, Generation, Available at http://www.ercot.com/gridinfo/generation

[23] Hasche, B., Keane, A. and O'Malley, M., 2011. Capacity value of wind power, calculation, and data requirements: the Irish power system case. IEEE Transactions on Power Systems, 26(1), pp.420-430.

[24] Hippert, H.S., Pedreira, C.E. and Souza, R.C., 2001. Neural networks for short-term load forecasting: A review and evaluation. IEEE Transactions on Power Systems, 16(1), pp.44-55.

[25] Hor, C.L., Watson, S.J. and Majithia, S., 2005. Analyzing the impact of weather variables on monthly electricity demand. IEEE Transactions on Power Systems, 20(4), pp.2078-2085.

[26] Huang, S.J. and Shih, K.R., 2003. Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. IEEE Transactions on Power Systems, 18(2), pp.673-679.

[27] Hyndman, R.J. and Fan, S., 2010. Density forecasting for long-term peak electricity demand. IEEE Transactions on Power Systems, 25(2), pp.1142-1153.

[28] Iowa State University, ASOS-AWOS-METAR Data Download Available at https://mesonet.agron.iastate.edu/request/download.phtml

[29] Kavanagh, D., Keane, A. and Flynn, D., 2013. Capacity value of wave power. IEEE Transactions on Power Systems, 28(1), pp.412-420.

[30] Keane, A., Milligan, M., Dent, C.J., Hasche, B., D'Annunzio, C., Dragoon, K., Holttinen, H., Samaan, N., Soder, L. and O'Malley, M., 2011. Capacity value of wind power. IEEE Transactions on Power Systems, 26(2), pp.564-572.

[31] Li, S., Wunsch, D.C., O'Hair, E.A. and Giesselmann, M.G., 2001. Using neural networks to estimate wind turbine power generation. IEEE Transactions on Energy Conversion, 16(3), pp.276-282.

[32] Madaeni, S.H., Sioshansi, R. and Denholm, P., 2013. Estimating the capacity value of concentrating solar power plants with thermal energy storage: A case study of the southwestern United States. IEEE Transactions on Power Systems, 28(2), pp.1205-1215.

[33] Mehrtash, A., Wang, P. and Goel, L., 2012. Reliability evaluation of power systems considering restructuring and renewable generators. IEEE Transactions on Power Systems, 27(1), pp.243-250.

[34] Milligan, M. and Porter, K., 2008, June. Determining the capacity value of wind: An updated survey of methods and implementation. In Wind Power.

[35] Paska, J., 2004, September. Reliability and performance indices of power generating units in Poland. In Probabilistic Methods Applied to Power Systems, 2004 International Conference on Probabilistic Methods Applied to Power Systems (pp. 861-866). IEEE.

[36] Radtke, J., Dent, C.J. and Couch, S.J., 2011. Capacity value of large tidal barrages. IEEE Transactions on Power Systems, 26(3), pp.1697-1704.

[37] Schilling, M.T., Billinton, R. and dos Santos, M.G., 2009. Bibliography on power systems probabilistic security analysis 1968-2008. International Journal of Emerging Electric Power Systems, 10(3).

[38] Takahashi, R., and Shimura, T., 2016. Extreme Value Theory. Tokyo: Kindai Kagaku Sha.

[39] Taylor, J.W., De Menezes, L.M. and McSharry, P.E., 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. International Journal of Forecasting, 22(1), pp.1-16.

[40] Troffaes, M.C., Williams, E. and Dent, C., 2015, July. Data analysis and robust modelling of the impact of renewable generation on long term security of supply and demand. In Power and Energy Society General Meeting, 2015 IEEE (pp.1-5). IEEE.

[41] PJM Interconnection, 2003. PJM Generation Adequacy Analysis: Technical Methods, Available at https://www.pjm.com/~/media/etools/oasis/references/whitepaper-sections-12.ashx

[42] Potomac Economics, 2018. 2017 State of the Market Report For The ERCOT Electricity Markets, Available at https://www.potomaceconomics.com/wp-content/uploads/2018/05/2017-State-of-the-Market-Report.pdf

[43] Weron, R., 2007. Modeling and forecasting electricity loads and prices: A statistical approach (Vol. 403). John Wiley & Sons.

[44] Wilson, A.L. and Zachary, S., 2016. Electricity Capacity Assessment: further analysis of the demand-wind relationship.

[45] Wilson, A.L., Zachary, S. and Dent, C.J., 2015. Electricity Capacity Assessment: analysis and integration of the demand-wind relationship. National Grid 2015.

[46] Wilson, A.L., Zachary, S. and Dent, C.J., 2018. Use of meteorological data for improved estimation of risk in capacity adequacy studies. IEEE.

[47] Zachary, S. and Dent, C., 2014. Estimation of joint distribution of demand and available renewables for generation adequacy assessment. arXiv preprint arXiv:1412.1786.

Appendices

A Histograms of the demand-net-of-wind

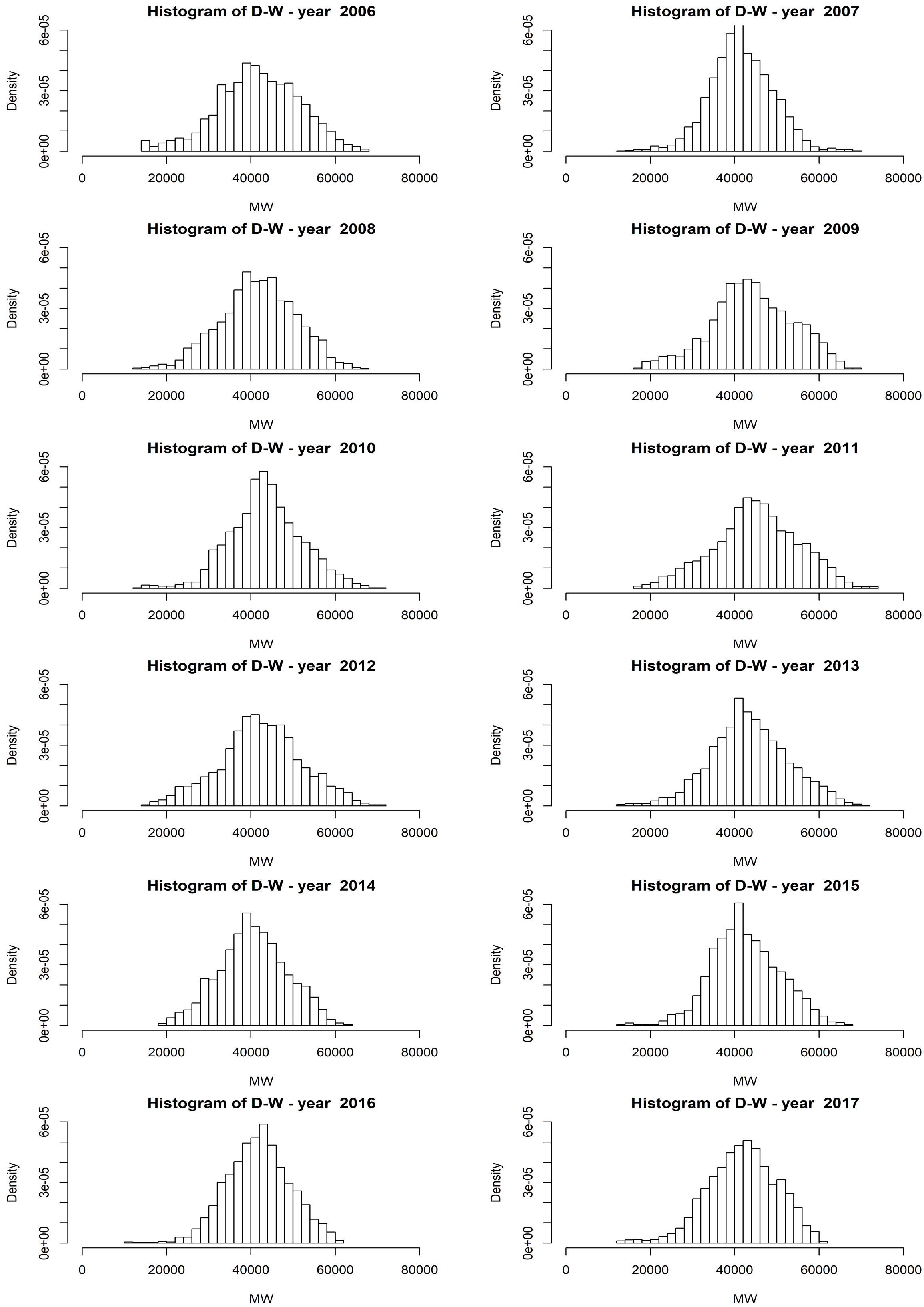

Figure 8: Histograms of the demand-net-of-wind $D - W$ in respect of the scenario from 2006 to 2017.

B Detailed results for the mean residual life plots

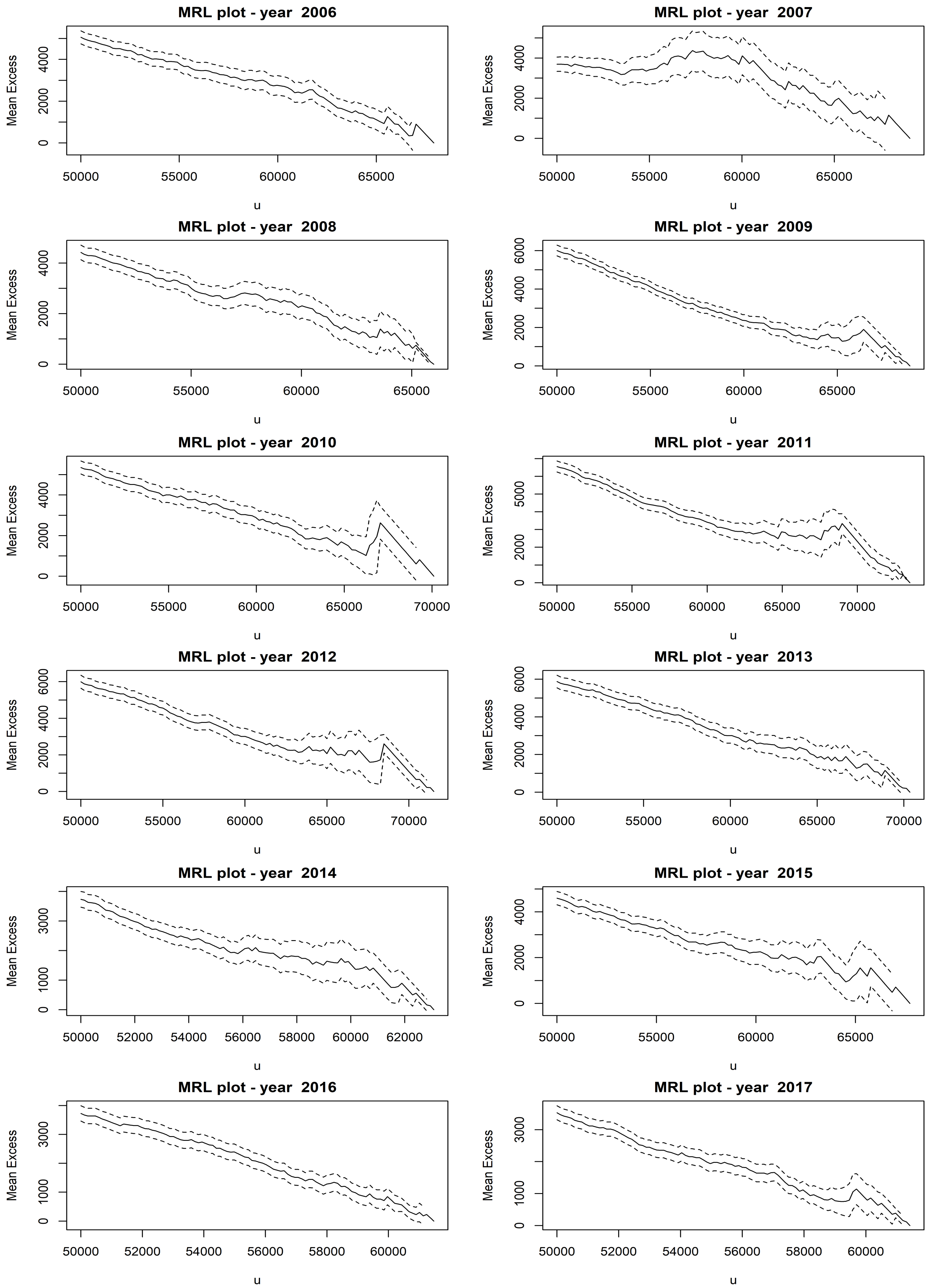

Figure 9: Mean residual life plot for $D - W$ in respect of the scenario from 2006 to 2017. The x-axis u means the threshold and the dotted lines are the 95 per cent confidence bounds.

C Detailed results for the extreme value theory analysis

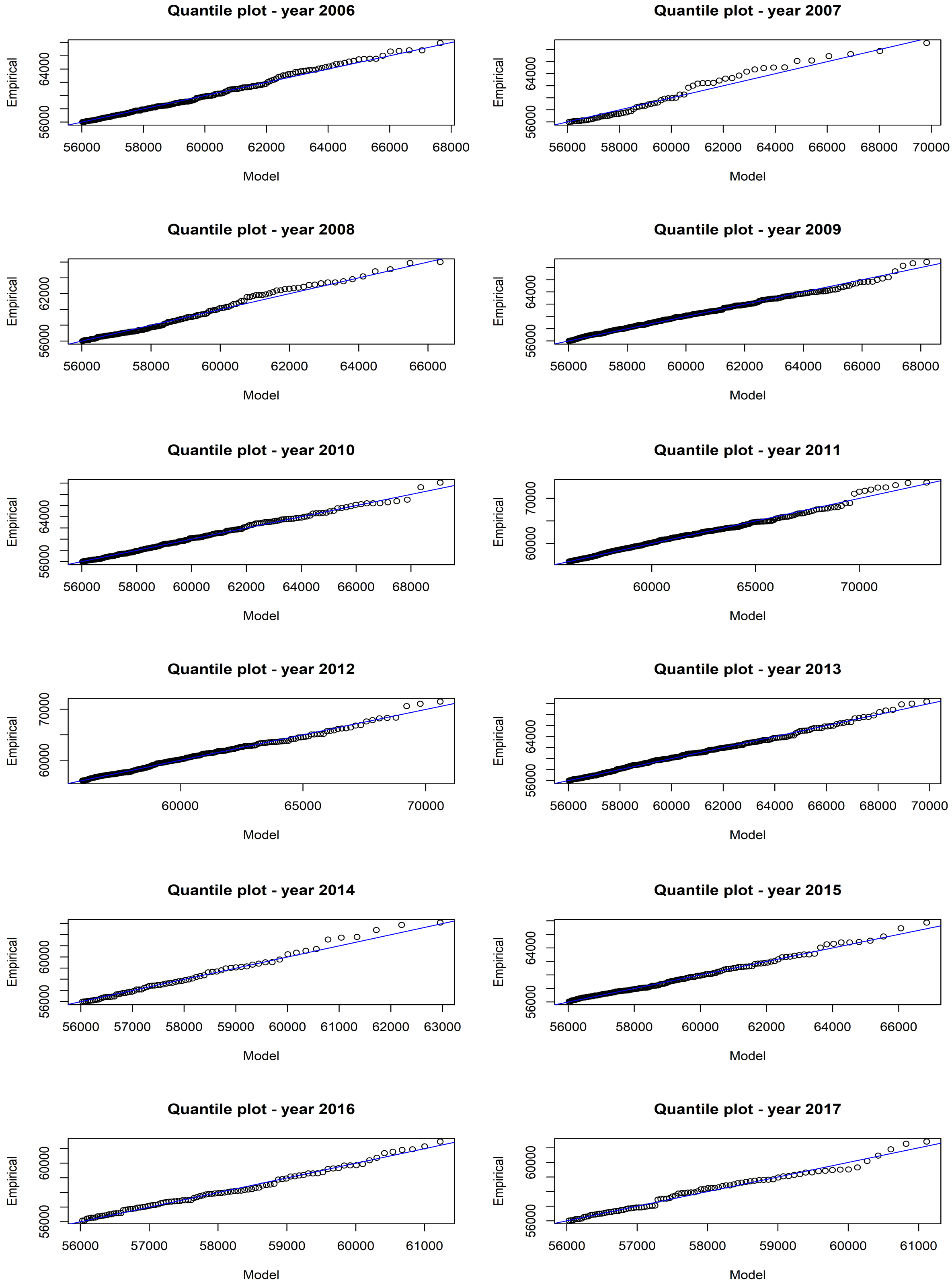

Figure 10: Quantile plots for the extreme value theory analysis in respect of the scenario from 2006 to 2017 (threshold of 56GW).

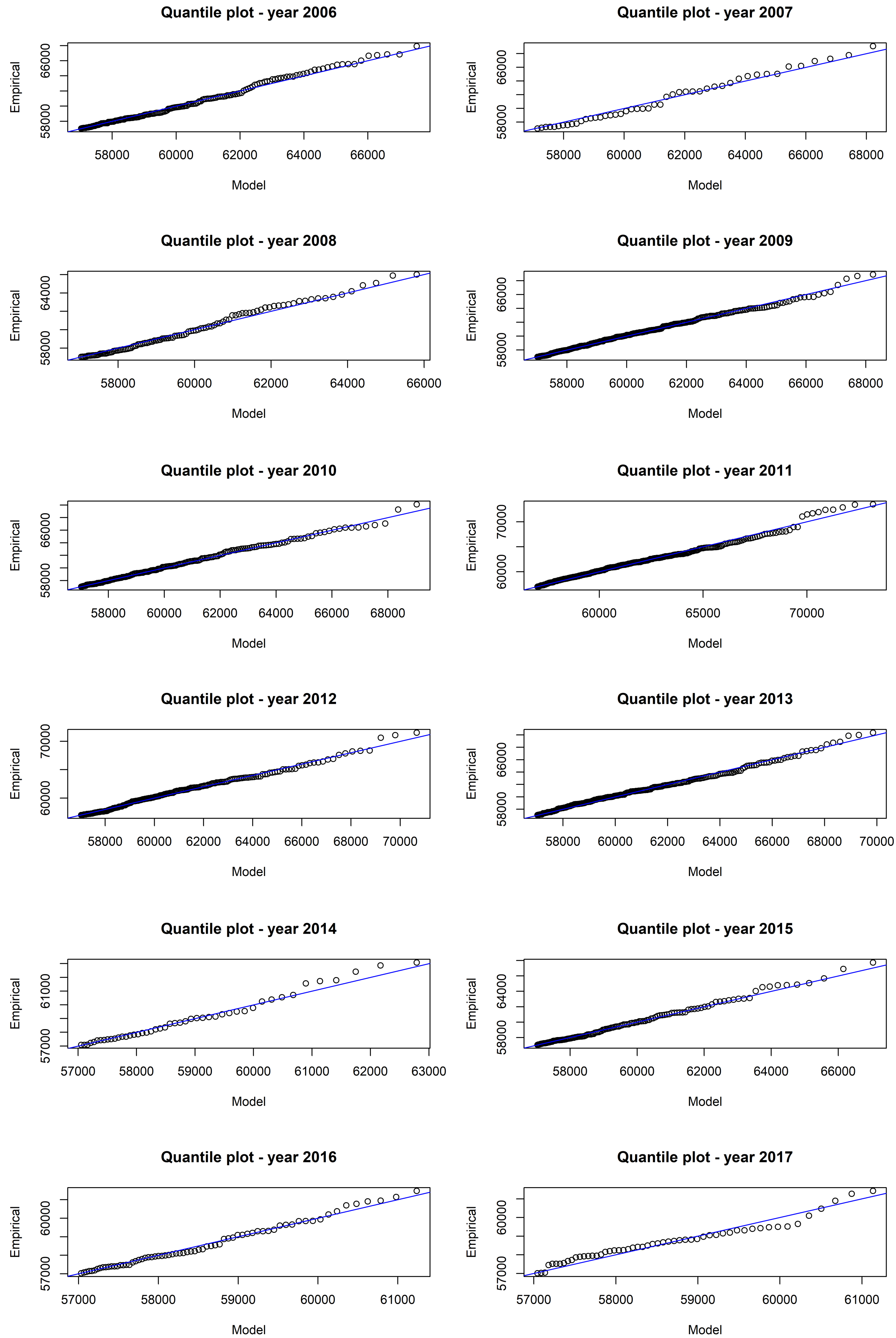

Figure 11: Quantile plots for the extreme value theory analysis in respect of the scenario from 2006 to 2017 (threshold of 57GW).

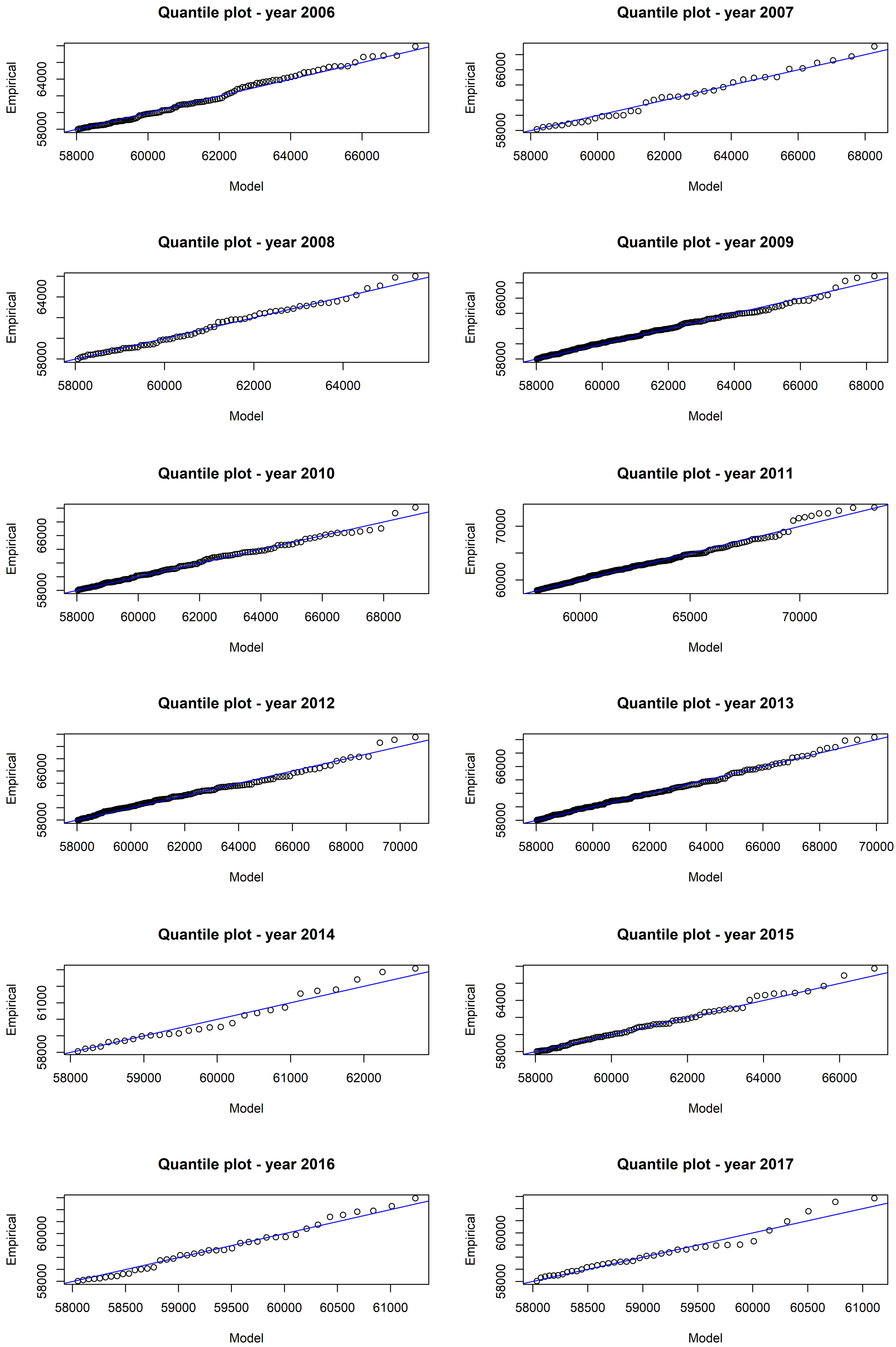

Figure 12: Quantile plots for the extreme value theory analysis in respect of the scenario from 2006 to 2017 (threshold of 58GW).