CHAPTER-VI

Estimation of load demand distribution

6.1. Introduction

In the contemporary world, technology involvement with human being has been significantly increased. This increased technology demand needs a reliable power source for functioning. High technology penetration is increasing power consumption level and showing a growing trend. Conventionally, the major contributor is fossil fuels for generation of electricity. With technological ease and change in human living, multiple new alternative energy sources have been developed for the sustainable growth. Adverse environmental effect and limited reservoir are restricting consumption of fossil fuels in recent times. Renewable energy sources are developing as more prominent and sustainable alternative sources of energy. Dispersed power generation flexibility and environment-friendly sources of renewable energy are making this more demandable. In the digital world, demand-side management with decentralized power generation system is becoming more controllable and advantageous. Evolution of smart grid is making this more feasible and controllable to increase penetration of distributed generation through renewable energy sources. Smart grid creates transparency and better communication system for demand-side management. Onsite power generation flexibility with renewable distributed generation technology gives freedom to reach out in rural and remote areas. This is a major factor for the development of developing nations where the majority lives in far-flung areas without access to electricity. India is also taking advantage of this technological shift to reach out their rural areas and rising as a major economy. Kumar and Ravikumar (2016) studied the feasibility of renewable energy sources in Indian urban buildings with the design of microgrids regarding challenges, opportunities, and techno-economic feasibility analysis. This also helps in nature balance to reduce carbon emission and maintain ecology system.

Modern smart grid technology gives a platform to develop more transparent and flexible system to increase integration of renewable energy sources, direct customer involvement secures information sharing and reduces power losses in modern power distribution system (Fan et al., 2013). The smart grid is the transparent, secure, user-friendly and computer-operated system. It helps in the decision-making process to maintain optimal flow and achieve demand-supply harmony with minimum financial and environmental losses. Rahimi and Ipakchi (2010) have beautifully shown the importance of smart grid in modern power distribution system in Figure-6.1. From the Figure-6.1, we can see factors affected by the smart grid to enhance power distribution system. This shows smart grid is prompted by several economical, political, environmental, social, and technical factors.

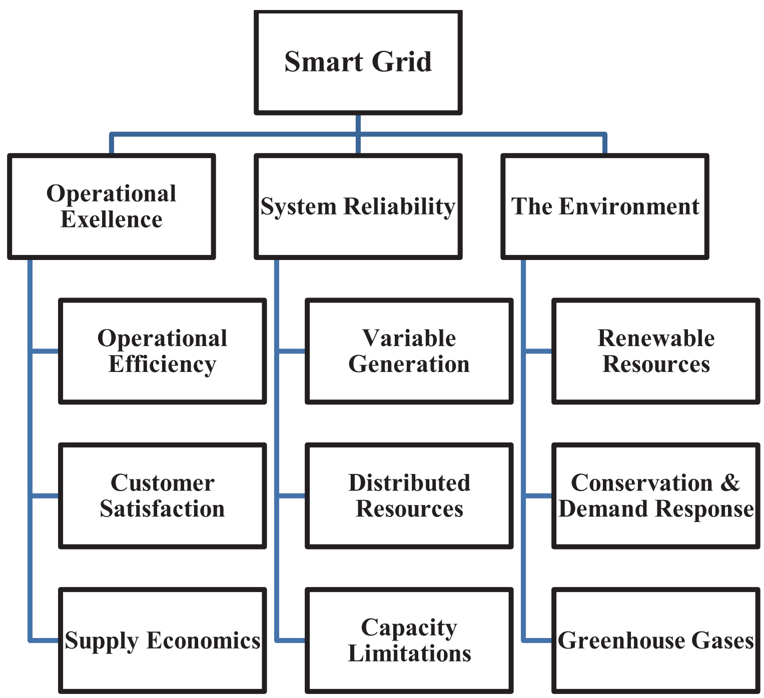

Figure 6.1: Three-way benefits of smart grid (Rahimi and Ipakchi, 2010)

Demand side management with the smart grid is an important function to develop better electricity distribution control and management, good infrastructure and management of integrated decentralized renewable energy resources (Gelazanskas and Gamage, 2014). Demand side management will lead to developing better managerial decision in optimal utilization of energy resources. With the effective utilization of demand-side management, we can reduce chances of developing under-utilized electrical infrastructure regarding generation capacity, transmission lines, and distribution networks. It helps to analyse demand uncertainty with statistical analysis to maintain continuity of supply as per demand.

We have collected the hourly load demand data of the year 2014 from the EWSS center at BHU. After the collection of data for the future prediction of load demand, we applied Lognormal, Gamma and Weibull probability distributions via R programming language software to analyze these three distributions graphically and computationally. Because of the continuous distribution of load demand data, we have taken all these three distributions from the continuous probability distribution family. Also, we considered only those probability distributions which will vary in between zero and infinity because load demand can never be negative. With these assumptions, we proposed the Lognormal, Gamma and Weibull distributions amongst the other probability distributions for estimating the load demand uncertainty at BHU campus. For finding the best-fitted distribution out of these three distributions, we have conducted the Goodness-of-fit test to find the best-fitted probability distribution of the given annual hourly load demand data. For this, we used three Goodness-of-fit test methods, which are the Kolmogorov-Smirnov test, Anderson-Darling test, and Cramer-von Mises tests. This will help in the managerial decision-making process for optimal power flow and will reduce losses and environmental degradation.

A Realistic modeling could be possible if we solve these problems with such methods. With mathematical programming techniques, we can perform the economic dispatch, while meeting all renewable distributed generation units and system constraints. In applications of the electric power system, specialized power systems that do not always rely upon three-phase AC power are found in aircraft, electric rail systems, ocean liners, and automobiles. In modern technology, most refrigerators, air conditioners, pumps and industrial machinery use AC power whereas most computers and digital equipment use DC power. Apart from these applications, one of the largest appliances connected to an electric power system is the HVAC unit, and ensuring this unit is adequately supplied is an important consideration in electric power systems.

6.2. Demand side management in smart grid system

The concept of smart grid started with the notion of advanced metering infrastructure to improve demand-side management, energy efficiency, and a self-healing electrical grid to improve supply reliability and respond to natural disasters or malicious sabotage (Fang et al., 2012). Soares et al. (2017) developed the stochastic model for energy resources management considering demand response in smart grids. However, several developments have led to the expansion of the initially perceived scope of the smart grid, and are helping shape the new face of the electricity industry. These include:

a) Emphasis on environmental protection, including renewable generation (wind, solar, etc.) and demand response (DR);
b) The drive for better asset utilization, including operating closer to the "knee of the curve" while maintaining reliable system operation; and
c) The need for enhanced customer choice.

Smart grid evolved as a source of sustainable future for the modern power sector. It provided a new space for modern technological advancement and helping to develop more user-friendly environment. Figure-6.2 schematically depicts factors about the new emerging smart grid paradigm is based on an interconnection of three subsystems (Samantaray, 2014):

1) An electric power system which accomplishes the generation, transmission, distribution, and consumption of electricity;
2) A communication system which establishes the connectivity for information exchange among different systems and devices; and
3) An information system which stores and processes data information for decision making on power system operation and management.

In Figure-6.2, the smart grid shows the integration of electric power systems, communication technologies, and information technologies. Integration of communication and information technologies is to achieve security, fault diagnosis, and intelligent monitoring. Integration of information and electric power system is to get plug-in hydro electric vehicle, intelligent agent and integration of distributed energy resources. Integration of electric power system and communication system is to have demand response, electric storage and standard interoperation.

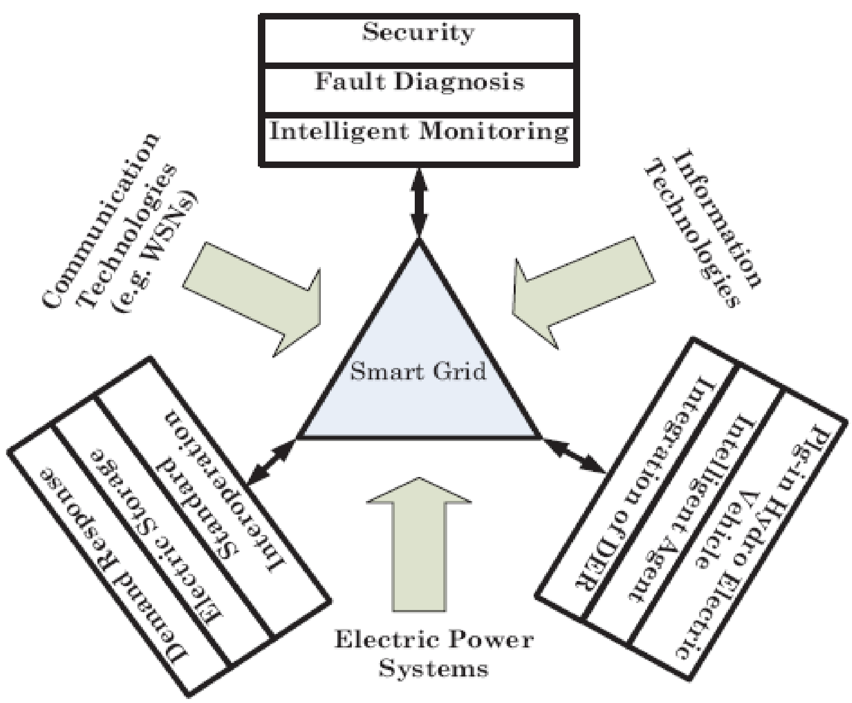

Figure 6.2: Technological factors affected by smart grid (Samantaray, 2014)

Smart grid gives freedom to power-sharing between multiple power sources and utilities to create the competitive environment. It helps to reduce monopoly of the centralized system and increase the involvement of multiple players. Multiple service providers can participate in the electricity market to provide electricity services to customers and utilities. It allows a customer to utilize renewable energy sources in their residential or commercial buildings integrated with grid and share surplus power with a grid to reduce net consumption load. Two-way communications are playing a pivotal role in protected information sharing to boost transparency in billing. Compared to conventional power grid systems, more renewable-energy-based distributed generation units and energy storage devices are integrated into the smart grid. As a result, the traditional electricity consumers are gradually transformed into electricity "prosumers" who not only consume energy but can also produce energy and feed it to the power grid. Therefore, the basic assumption of unidirectional electricity delivery in the traditional electric power system is no longer practical. Bidirectional energy flows need to be established between electricity customers and power distribution systems. Fang et al. (2012) and Farhangi (2010) shown in their work advantages of smart grid over the conventional power grid. Moreover, some renewable distributed generation units, energy storage devices, and utilities nearby can be interconnected as a micro-grid, which can operate in either a grid-connected mode or an islanded mode for reliability enhancement while reducing transmission and distribution losses.

The conventional electric grid system is a network that acts as a link for transmission, distribution, and control of electric power from power producers to consumers. Increasing power demand and complexity in managing power grid, generation and capacity limitation needs for the development of highly reliable, self-regulating and efficient grid system which will allow the integration of renewable distributed power generation with grid supply to meet the prosumers demand.

Smart grid comes with smart metering techniques, digital sensors, and intelligent control systems with analytical tools to automate, monitor and control the two-way flow of energy during the operation from power to plug. The smart grid helps the power utilities and grid to have a digital intelligence to the power system network.

Many issues contribute to the incapability of conventional grid to competently meet the demand for consistent power supply. Advantages of smart grid over conventional grid system have been shown in tabular form in Table-6.1.

Table 6.1: Comparison between conventional and smart grid (Hossain et al., 2013; Farhangi, 2010)

Conventional Grid	**Smart Grid**
Electromechanical, solid state	Digital/Microprocessor
One-way and local two-way communication	Global/integrated two-way communication
Centralized generation	Accommodates distributed generation
Few sensors	Sensors throughout
Limited protection, monitoring and control systems	WAMPAC, Adaptive protection
'Blind'	Self-monitoring
Manual restoration	Automated, 'self-healing.'
Check equipment manually	Monitor equipment remotely
Limited control system contingencies	Pervasive control system
Estimated reliability	Predictive reliability
Failure and blackouts	Adaptive and islanding
Failure and blackouts	Adaptive and islanding

Demand side management depends on secured information sharing between multiple channels (Liang et al., 2014). For information sharing, three kinds of communication networks can be established in the smart grid. A wide area network (WAN) facilitates the communications among bulk generators and transmission facilities for wide-area situational awareness. A neighbourhood area network (NAN) or field area network (FAN) supports the communications among distribution substations and field electrical devices for power distribution and microgrid operation. Home area networks (HAN), business area networks (BAN), and industrial area networks (IAN) can be deployed within residential, commercial, and industrial buildings, respectively, for communication among electrical appliances for the DSM purpose. Mostly nature of data becomes stochastic with the uncertainty of renewable energy sources (solar and wind) and load demand which needs statistical analysis and trend analysis. The research and development on smart grid communication networks have been extensively carried out.

The smart grid communication network architectures, performance requirements, research challenges, state-of-the-art technologies, development aspects, and experimental studies have been discussed in the previous literature (Fan, 2013; Gao et al., 2012; Hossain et al., 2012; Yan et al., 2013). As more and more electric devices in the critical power infrastructure are interconnected via communication networks, cyber security has an immediate impact on the reliability of the smart grid. Furthermore, increased connectivity of electrical appliances at the customer side can enable personal information collection, which may invade customer privacy. The cyber security requirements, network vulnerabilities, attack countermeasures, secure communication protocols and architectures, and privacy issues in the future smart grid have been surveyed in previous works (Liu et al., 2012; Wang and Lu, 2013).

Based on information acquired via the communication system, the information system can make optimal decisions on electric power system operation and transmit the control signals via the corresponding communication networks. Although basic information management functionalities are already in place in the traditional bulk generation and transmission systems based on the supervisory control and data acquisition (SCADA) systems, developing an advanced information management system in the context of the smart grid is technically complex due to the following challenges (Liang et al., 2014):

- The output of renewable energy resources is intermittent, which results in large variations in power supply. Although a large body of studies have been carried out to forecast such an uncertain output, the stochastic nature of renewable power generation should be addressed in smart grid planning and operation;
- The buffering effect of energy storage devices not only introduces more state variables in power system operation but also requires accounting for the inter-period buffer state transitions over the entire time frame (within a week) under consideration. Efficient management schemes should be designed for energy storage devices at a low computational complexity;
- Customer behaviour patterns in the presence of DSM are more dynamic than in the traditional electricity grid, which leads to large variations in load demand. The main reason is that the usage of electrical appliances can be shifted over time to electricity customers in response to electricity prices. Moreover, different customers can collaborate with each other to reduce their overall energy bills, based on the information obtained via FAN/NAN communications;
- EV drivers can select different charging locations in response to electricity prices, which can lead to large variations in charging demand and poor accuracy of charging demand estimation. Further, high EV mobility can result in a highly dynamic energy storage capacity of the electric power system, taking account of the random nature of route and commute schedules of EV drivers.

To address these technical challenges, first, we establish proper stochastic models to characterize randomness in demand side. Then, we incorporate the stochastic models in the system-level information management to facilitate smart grid planning and operation.

6.3. Stochastic models

6.3.1. Statistical Analysis

Demand side management modeling requires analysis of load demand data for the given period. To reduce the expenses and time required to process long-term load demand data, it is desirable to use statistical probability distribution for describing the load demand variations. The primary tool to describe load demand characteristics is a probability density function which gives the mathematical formulation of given probability distribution. Probability density functions of a theoretical probability distribution predicted by estimating its parameters. The parameters of probability distribution which describes load demand uncertainty with theoretical distributions are estimated using statistical data of the given period. Many probability distributions have been proposed in recent past. In this work, statistical analysis of Lognormal, Gamma and Weibull distributions are used to identify the pattern of load demand distribution. For the simulation and computational work, we used R programming language software. With the computational result of R programming language software, we analyzed our distribution models graphically and analytically along with the goodness-of-fit test results. Also, parameters of each distribution have been calculated from the same software. Here are the mathematical formulations of Lognormal, Gamma and Weibull distributions:

6.3.1.1. Lognormal distribution

The lognormal distribution is a probability distribution of a random variable, whose logarithm is normally distributed. Lognormal probability distribution function with μ as location and σ as scale parameters is given by

$$\ln(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left[-\frac{(\ln x-\mu)^2}{2\sigma^2}\right] \quad (6.1)$$

Lognormal CDF is written as

$$LN(x;\mu,\sigma) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\ln x-\mu}{\sigma\sqrt{2}}\right)\right] \quad (6.2)$$

Where mean and variance of the distribution are

$$E[X] = e^{\mu+\frac{1}{2}\sigma^2} \quad (6.3)$$

$$Var[X] = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2} \quad (6.4)$$

6.3.1.2. Gamma distribution

Probability density function of Gamma distribution with α (shape) and β (rate) parameters is

$$f(x;\alpha,\beta) = \frac{\beta^{\alpha}x^{\alpha-1}e^{-x\beta}}{\Gamma(\alpha)} \quad (6.5)$$

Gamma CDF is written as

$$F(x;\ \alpha,\beta) = \frac{\gamma(\alpha,\beta x)}{\Gamma(\alpha)} \quad (6.6)$$

Where mean and variance of the distribution is

$$E[X] = \frac{\alpha}{\beta} \quad (6.7)$$

$$Var[X] = \frac{\alpha}{\beta^2} \quad (6.8)$$

6.3.1.3. Weibull distribution

Probability density function of two parameters Weibull distribution with c being the scale parameter and k the shape parameter is

$$f(x;c,k) = \left(\frac{k}{c}\right)\left(\frac{x}{c}\right)^{k-1} exp\left[-\left(\frac{x}{c}\right)n^k\right] \quad (6.9)$$

Weibull CDF is written as

$$F(x; c, k) = 1 - e^{-\left(\frac{x}{c}\right)^k} \tag{6.10}$$

Where mean and variance of the distribution with Γ (Gamma function) are

$$E[X] = c\Gamma(1 + \frac{1}{k}) \tag{6.11}$$

$$Var[X] = c^2 \left[\Gamma\left(1 + \frac{2}{k}\right) - \left(\Gamma\left(1 + \frac{1}{k}\right)\right)^2\right] \tag{6.12}$$

6.3.2. Goodness-of-fit tests

The fitness of the theoretical probability density functions with actual data analyzed by the goodness-of-fit test. This gave the differences of the theoretical distribution function with actual data variation to choose more fitted one. With this, we have reduced the chances for failure in the future predictability of demand variation and pattern recognition. In reality, there is no one having correct distribution. Probability density function is the most used tool to estimate and model the load demand distribution. Many researchers show the use of Normal distribution as load demand estimation to use them in demand-side management (Bo and Li, 2009; Liu et al., 2011). However, to assume a Normal distribution of a particular case without any statistical analysis or investigation may result in errors in their application or determining theoretical distribution analysis of load demand data. Normal distribution function varies in between minus infinity to plus infinity and in practicality load demand cannot be negative. For overcoming such issues and identification of fittest probability density function, we considered continuous probability distribution that will lie between zero and infinity. In a case of BHU campus Lognormal, Gamma, and Weibull probability distributions have been investigated with the aid of three goodness-of-fit tests namely; Kolmogorov-Smirnov, Anderson-Darling and Cramer-von Mises tests. Brief information on selected goodness-of-fit tests is given below:

6.3.2.1. Kolmogorov - Smirnov test

The Kolmogorov-Smirnov test is useful when parameters of the probability distribution do not have to be estimated. We can estimate the proposed distribution with the help of modified Kolmogorov-Smirnov test, where the parameters are estimated with the given data. It has been proven that the Kolmogorov-Smirnov test uses empirical cumulative distribution function. For the given *n* ordered data points $X_1, X_2, \ldots, X_n$, the empirical cumulative distribution function is defined as (Daniel, 1990)

$$E_n = n(i)/N \tag{6.13}$$

Where n(i) is the number of points less than X_i and X_i are ordered between smallest and largest value. This increases by 1/N for each value of ordered data between smallest to largest.

The advantage of this test is that the distribution of Kolmogorov-Smirnov test statistic will not need cumulative distribution function test. It only applies to continuous distributions. The test is calculated as

$$D = \max_{1 \leq i \leq N} \left\| F(Y_i) - \frac{i}{N} \right\| \tag{6.14}$$

Where F is the theoretical cumulative distribution for what we are conducting this test which must be a continuous distribution.

6.3.2.2. Anderson- Darling test

The Anderson-Darling test is used to test the given data, whether this lies to the given probability distribution. The test is most often used to test of a given family of distribution, in which case parameters of that family needs to be estimated. Anderson-Darling test is mainly the modified adaptation of the Kolmogorov-Smirnov test, and it gives more weight to the tails, where Kolmogorov-Smirnov test gives more to the center of the distribution. This has the advantage of allowing a more sensitive test.

Mathematical formulation of the Anderson - Darling test statistic is (Anderson and Darling, 1952)

$$A^2 = -n-S \tag{6.15}$$

Where

$$S = \sum_{i=1}^{n} \frac{2i-1}{n} [\ln F(X_i) + \ln(1 - F(X_{n+1-i}))] \tag{6.16}$$

Here F is the cumulative distribution function of the given family of distribution. Note that the values of X_i are the ordered data.

6.3.2.3. Cramer-von Mises test

The Cramer-von Mises test is used to test the fitness of distribution of a given family of distribution. In statistics the Cramér–von Mises test is used for testing the goodness-of-fit for a cumulative distribution function F compared to an empirical distribution function F_n for the given data, or comparing of two empirical distributions. It is also used in part of another algorithm as minimum distance estimation. This is represented as (Anderson, 1962)

$$w^2 = \int_{-\infty}^{\infty} \{F_n(x) - F(x)\}^2 dF(x) \tag{6.17}$$

In this study, here F is the theoretical cumulative distribution function for the wind speed data of BHU region, and F_n is the empirically observed distribution. Two distributions can be empirically estimated ones in a two-sample case. This Cramér–von Mises test is an alternative to the Kolmogorov–Smirnov test.

Let $x_1, x_2, \ldots, x_n$ are the observed values, in ascending order. The test statistic is

$$T = nw^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left[\frac{2i-1}{2n} - F(x_i)\right]^2 \tag{6.18}$$

6.4. Results and discussions

In this work, regular and complete measured data for the year 2014 aiming hourly load demand inside the BHU campus has been taken from the EWSS centre, BHU. We are investigating the feasibility of integrated renewable distributed generation with the smart grid at BHU campus. In the demand-side management of BHU campus, we considered stochastic renewable energy sources like solar and wind energy to setup renewable distributed generation unit with decentralized generation. For this, we have collected hourly load demand data for the year 2014 at BHU campus. As per annual demand pattern, we can setup optimal generation capacities to fulfill our demand without any interruption. There are two feeders of power supply at BHU campus for better control of power distribution. In case anyone of these two feeders fails, the second one will be over utilized and supply the required power demand. If both will fail, then there will be blackout like situation. EWSS center is using 133 kVA capacities of both the feeders. We measure power in either kW (kilo watt) or kVA (kilo Volt Ampere) unit. In this study, we have the hourly load demand data in kVA unit. As we all know that a total number of hours in a year (except leap year) become eight thousand seven hundred and sixty. For the year 2014, we have collected the data, which shows that the total eight thousand seven hundred and eleven hours powers have been supplied. Remaining forty-nine hours, there was a blackout in the year 2014. For the statistical analysis via R programming language software, we considered 8760 hours for evaluation with 8711 hours available data and remaining 49 hours interpolated data.

After computing, results show in Figure-6.3 the graphical representation of a histogram of annual hourly load demand distribution data of BHU campus. Here we have drawn the histogram with the help of R programming language software. Hines et al. (2008) stated that choosing the number of class intervals approximately equals to the square root of the sample size often works well in practice. Here, we used this theory to find out the class intervals of the Histogram. Because, if we have very few intervals for the larger data then this will give us the blocky or coarse figure or if we have very large intervals compared to not much larger data then this will show a ragged figure. For the selection of best fit, we should choose appropriate class intervals for the given sample size. In our case, we have total sample size equal to eight thousand seven hundred and eleven, which gives the value of ninety-four with approximation after the square root of sample size. We have drawn the histogram in ninety-four class intervals for the given three Lognormal, Gamma and Weibull distributions.

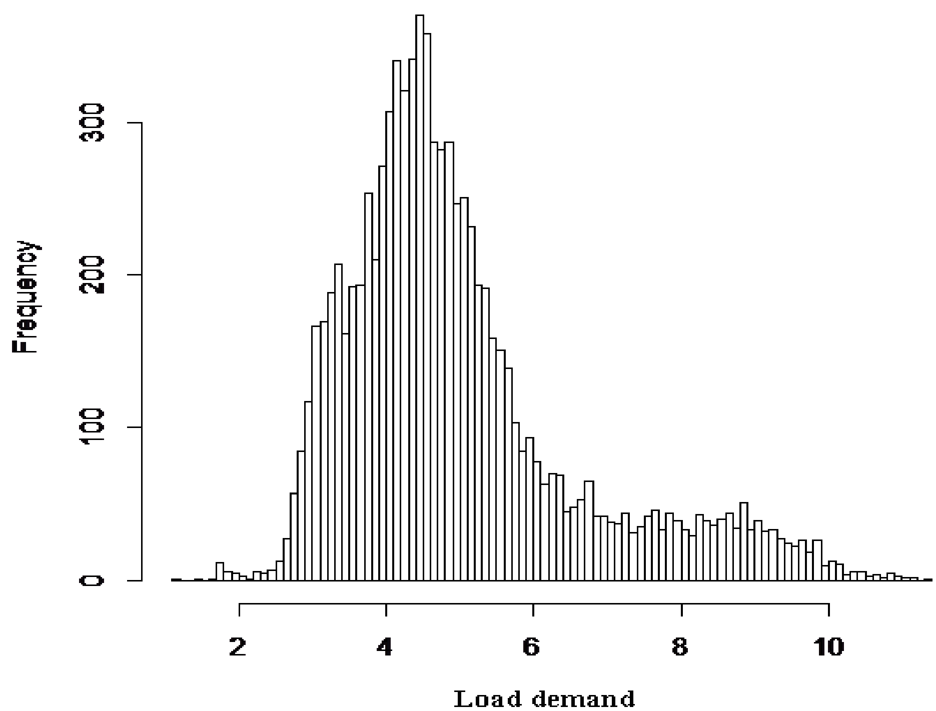

Figure 6.3: Histogram of hourly load demand data

Results are showing graphical representation of theoretical probability density functions (pdf), empirical and theoretical cumulative distribution functions (CDF), Q-Q plot and P-P plot, table of ranked distributions with goodness-of-fit tests results and parameters values. Figure-6.4 is showing PDF of Lognormal, Gamma and Weibull distributions respectively. Also, Figure-6.4 is showing PDF of Lognormal is best fitted compared to other two distributions. The fitness of PDF of Gamma distribution in Figure-6.4 is showing poorer than the Lognormal distribution and better than the Weibull distribution.

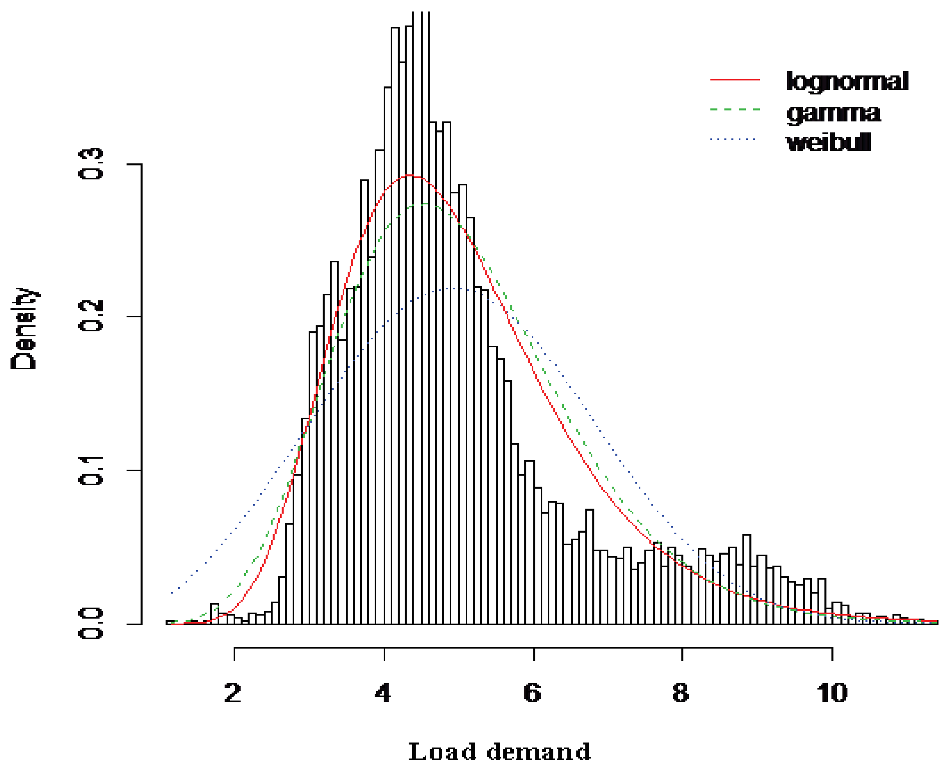

Figure 6.4: Density plots of theoretical distributions with load demand data

Where CDF of lognormal distribution in Figure-6.5 is showing the good fit with 95% lower confidence band and 95% upper confidence band in percentile. On the other side, CDF of Gamma and Weibull distributions in Figure-6.5 respectively are showing lacks fitness. With the comparative study of these three distributions, we have the best-fitted distribution as Lognormal distribution. From the zero to hundred percentile with mean is 4.991912 and the standard deviation is 1.623279 CDF for specified distribution function is showing the Lognormal distribution as the best one.

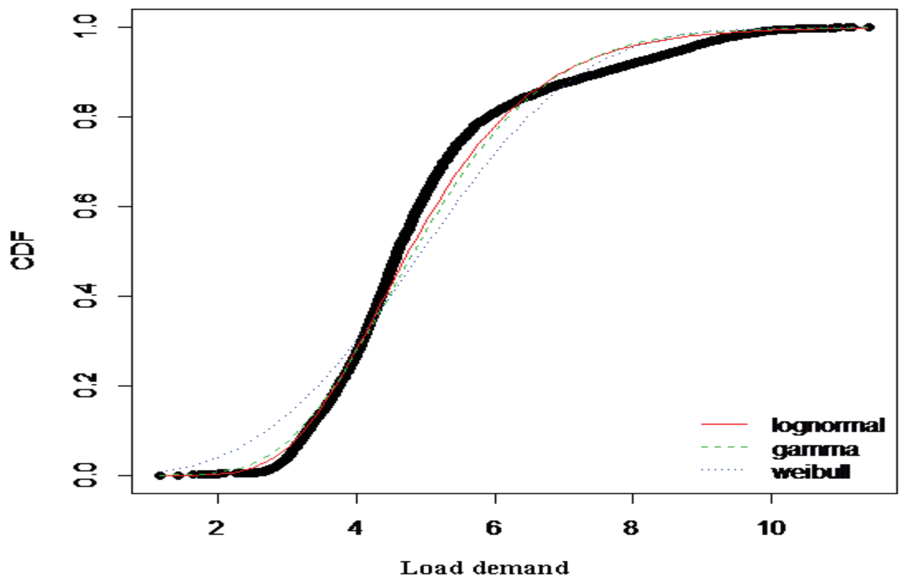

Figure 6.5: Empirical and theoretical CDFs

Figure-6.6 shows P-P (probability-probability) plot of the given data. In a P–P plot, this assesses how closely two data sets agree for the two cumulative distribution functions against each other. Here in Figure-6.6, Lognormal distribution is showing a close relation with the given dataset for the cumulative distribution function compared to other two Gamma and Weibull distributions. With the help of P-P plots, we can evaluate the skewness of the distribution. Figure-6.7 is showing the Q-Q (quantile-quantile) plot. In a Q–Q plot, this also assessed how closely two data sets agree for the two cumulative distribution functions against each other with their quantile plots. Q–Q plot also helps to compare the shapes of distributions with the graphical presentation to view how properties like location, scale, and skewness are similar or different in the two distributions. Q–Q plots compared to collected data and theoretical distributions. Q-Q provides an assessment of "goodness of fit" graphically, rather than having a numerical summary. With the interpretation of Figure-6.7, we can find the best-fitted Lognormal distribution in this study compared to other two Gamma and Weibull distributions.

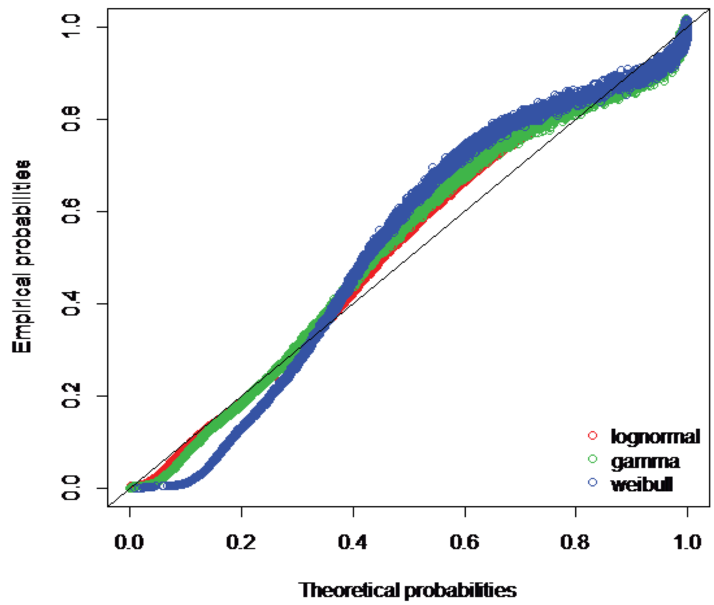

Figure 6.6: P-P plot

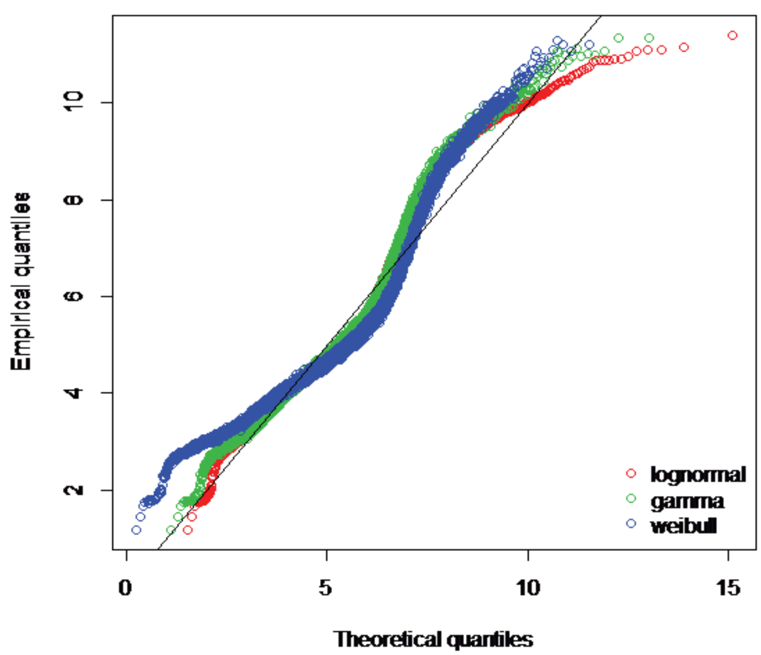

Figure 6.7: Q-Q plot

Table-6.2 is showing the results of parameters for the proposed distributions. Mean and standard deviation of one-year period hourly wind speed data of BHU campus is 3.53

and 1.87 respectively. These parameters have been calculated by the maximum likelihood methods because of its more accurate values. For Lognormal distribution location parameter $\mu = 1.5611$ and scale parameter $\sigma = 0.2993$. Values of scale and shape parameters of two parametric Gamma distribution are respectively $c = 5.5678$ and $k = 3.1278$. Two-parameter Weibull distribution scale parameter $c = 3.130725$ and $k= 5.568298$ has been given.

Table 6.2: Parameter values of proposed distributions

Probability distributions	Parameters	Results
Lognormal	Meanlog	1.561264
	Sdlog	0.299232
Gamma	Shape	10.87202
	Rate	2.17755
Weibull	Shape	3.130725
	Scale	5.568298

In having goodness-of-fit tests, Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises tests are calculated the distance between the theoretical cumulative distribution function with empirical distribution function. With goodness-of-fit tests results in Table-6.3 we have ranked these three distributions separately by findings of Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises goodness-of-fit tests. The Kolmogorov-Smirnov test results in Table-6.3 ranked them in the order of Lognormal, Gamma and Weibull distributions. From the Table-6.3, we can see the Anderson-Darling test results, which also ordered in Lognormal, Gamma and Weibull distributions. The third one, Cramer-von Mises test results in Table-6.3 also ranked them in the order of Lognormal, Gamma and Weibull distributions. From this, we can say that the Lognormal distribution is the best one among the proposed distributions.

Table 6.3: Goodness-of-fit test results

Probability distributions	Kolmogorov-Smirnov	Anderson-Darling	Cramer-von Mises
Lognormal	0.06886589	74.14607392	11.85236535
Gamma	0.09059829	131.54815513	21.74483485
Weibull	0.1314375	306.0348499	51.9529980

6.5. Optimal power flow

Due to the large-scale integration of renewable distributed generation, traditional economic power distribution system, which relies on an accurate forecast of power generation and load demand, will not be directly applied in the future smart grid. As discussed in the previous section, the randomness in renewable power generation is characterized based on stochastic models. Without taking into account the randomness, traditional economic power distribution system may schedule more conventional energy sources such as coal-fired or gas generators and underutilize the renewable energy sources, which increase power generation cost and decrease power system reliability. Stochastic models need to be developed for the economic dispatch of power to address the randomness in the renewable distributed generation. Optimal dispatch problem in power distribution was first introduced by Carpentier in 1962, and later it used as the optimal power flow which interchangeably used (Azami et al., 2011). Sasaki et al. (2017) developed robust stochastic dynamic load dispatch in their work against uncertainties of the renewable energy sources and load demand. They focused on a new dynamic load dispatch method for mitigating the irregularity associated with RES. The developed load dispatch method is to schedule the committed generating units' outputs to meet required irregular load demand estimation which is frequently updated in real-time operation circumstance.

Renewable distributed generation scenarios can be generated via micro-generation certification scheme (MCS) and incorporated in a stochastic locational marginal price (LMP) electricity market model to examine the impact of renewable distributed generation on price settlement, load dispatch, and reserve requirements (Ahmed et al., 2011). Scenario reduction can be used to reduce the computational complexity of MCS by classifying the renewable distributed generation into specific levels based on the uncertainty of renewable energy sources.

Another way to reduce the computational complexity of MCS is to use the moment estimation technique. System demand can be modeled as a random vector with correlated variables such that the dependency between load type and the location can be characterized (Madrigal et al., 1998). Then, a probabilistic optimal power flow problem can be formulated, and a maximum likelihood estimation method can be applied to evaluate the stochastic properties of a specific solution of the probabilistic optimal power flow problem. Maximum likelihood estimation is a method of estimating the parameters of a statistical model given observations, by finding the parameter values that maximize the likelihood of making the observations given the parameters. In general, for a fixed set of data and underlying statistical model, the method of maximum likelihood selects the set of values of the model parameters that maximize the likelihood function. Intuitively, this maximizes the "agreement" of the selected model with the observed data, and for discrete random variables, it indeed maximizes the probability of the observed data under the resulting distribution. Maximum likelihood estimation gives a unified approach to estimation, which is well-defined in the case of the normal distribution and many other problems. Suppose there is a sample $x_1, x_2, \ldots, x_n$ of n independent and identically distributed observations, coming from a distribution with an unknown probability density function $f_0()$. It is however surmised that the function f_0 belongs to a certain family of distributions. This represents a fundamental assumption underlying classical results on the properties of the maximum likelihood estimator (e.g., Le Cam, 1953; Wald, 1949) is that the stochastic law which determines the behaviour of the phenomena investigated (the "true" structure) is known to lie within a specified parametric family of probability distribution models, or we can say, the probability model is assumed to be "correctly specified (White, 1982).

The procurement of energy supply from conventional base load generation and renewable distributed generation can be investigated based on a multi-timescale scheduling in a dynamic programming framework (He et al., 2013). Specifically, the optimal procurement of energy supply from base load generation and a day-ahead price is determined by day-ahead scheduling given the distribution of renewable distributed generation and demand. On the other hand, the optimal real-time price to manage opportunistic demand for system efficiency and reliability is determined via real-time scheduling given the realizations of renewable distributed generation. The smart grid helps in real-time price management with the two-way communication system and showing actual consumption pattern in demand side management. This gives better managerial decision-making process to reduce losses and control over optimal power flow.

6.6. Conclusion

Based on the proposed architecture, the demand side management system receives the theoretical load curve as an input and calculates the required load control actions to fulfill the desired load consumption. Set up of the smart grid for renewable distributed generation technology at any location necessitates good load demand estimation. With good estimation, we can optimize investment cost of the renewable distributed generation technology and deliver continuous power supply to the desired location with a minimum per unit cost. For this, there is a requirement of statistical analysis to find out the best-fitted probability distribution of previous load demand data. Therefore, distribution of the electricity load demand at BHU campus is statistically analyzed for the year 2014 on an hourly basis. Because of randomness of the hourly load demand, it is expected to fit a probability distribution which will give better fitness. In this analysis, three probability distribution functions have been proposed, which are Lognormal, Gamma and Weibull distributions. For the reason, the significance level alone cannot help to decide for selection of any distribution without going for any statistical goodness-of-fit test. For showing intelligence, the reliability of the information is very important factor in estimation and making decisions. Since decisions made with the help of statistical analysis are in certain confidence level, this case analysis was done at 95% confidence level. We used three goodness-of-fit tests named Kolmogorov-Smirnov, Anderson-Darling and Cramer-von Mises, and graphical analysis to identify the fittest distribution. With the comparative analysis of proposed distributions, Lognormal has been determined to be the best-fitted distribution to representing given load demand data of the BHU campus. Best fitted distribution can reduce chances of error and risk in making the investment decision for setting up of renewable distributed generation units and optimal power flow at BHU campus. For future work, one can find fitness of previous data with theoretically derived new probability density function.