



Diagnostic Framework for Electricity Losses in Pakistan Using Data Visualization Techniques

Fatima Tariq¹, Dr. Atif Alvi²

^{1,2}Department of Computer Science, Lahore Garrison University, Lahore.

¹F18-mscs-016@lgu.edu.pk.

Abstract

To obtain the best results and leads to accuracy, data visualization techniques allows the big data and unprocessed data in a structured format. There are a lot of techniques that provides the best results for both high and low dimensional data preprocessing. Electricity crisis are becoming the part of day to day life and increasing rapidly. There are a lot of factors that affects the distribution and transmission losses of electricity. PEPCO, GENCO and other distribution companies are responsible for electricity distribution .There are 10-12% electricity losses during the distribution from main to other. At different sectors electricity production and installation from different resources of energy needs a proper monitoring system .Except of system losses a lot of electricity lost due to non-technical factors which leads to the shortage of electricity in a Country. .Grid stations should be planted according to the population of the province .Diagnostic framework includes the previous and present data comparison which predicts and forecasts the future consumption that helps to overcome losses by time series analysis algorithm (ARIMA)

Keywords: Data visualization, ARIMA model, time series analysis, distribution, transmission

1. INTRODUCTION

It is very important to know the significance of data which we are going to use. For this purpose data visualization is one of the important key feature which plays an important role to expose the data in an easy and precise way. Data visualization is the basically the data representation after processing in graphical from. Some of the data cannot be described and detected easily

in text-based data so it has to be used in patterns, correlations and charts with the help of data visualization techniques. For the advanced predictive analysis in big data, data visualization has its special significance over accuracy and faster decision making. Energy production is one of the crucial and important need of today's world but this energy needs, maintenance, performance evaluation and fault detection at every level

which is possible with data visualization [1]. Energy is using in houses and for other recreational activities which may be lost or used improperly by the time. Electricity optimization for data visualization helps to prevent the data loss over larger datasets. Data visualization tools plays a key role in decision making on the basis of collected datasets from different sources. At different time and different season of the year electricity consumed in many different ways but production depends on so many factors .Sometimes electricity produced cannot be enough and needs more installation from private companies or sometimes bought from other sources which results in a huge deficit of economy and rise in electricity bills. Natural disasters plays a vital role in losses and due to low maintenance of systems .Electricity production decreases and needs a lot of efforts to distribute it properly. In different situations electricity consumption changes and consume less or more by the need. This consumption should be monitored and maintained by making best decision support systems which lags in past [2].

The main purpose of this work is to look into detail how hyper-convergence technology improves the reliability of data centers and how hyper-converged systems operates.

1.1 Data Visualization for Monitoring:

Data visualization is the process of acquiring data from different areas and give results. This helps to overcome the complication of communication in data. The analysis to facilitate and identify the data is done by data visualization and data mining. There are a lot of data visualization techniques and tools to meet the users need but the main

function is to provide the data visualization techniques effectively which is possible with the developed services and platforms [3]. NEPRA and NTDC are the main sources that makes report every year that how much electricity is installed, generated and consumed. However they make this report at broader level but not compares the datasets at different levels and stations .There are a lot of electricity sectors which consumes energy more than their projected demand but cannot be monitored by sources due to lack of maintenance which results shortage of electricity and more generation from private companies like IPPS and that affects the economy .There are 42 IPPS (independent power stations) which contributes electricity in Pakistan but there is still shortage of electricity due to technical and non-technical losses.

1.1.1. Data Visualization Framework for Monitoring

Data Visualization characteristics comprise of different phases which determines the framework for energy consumption and distribution losses in proposed model.

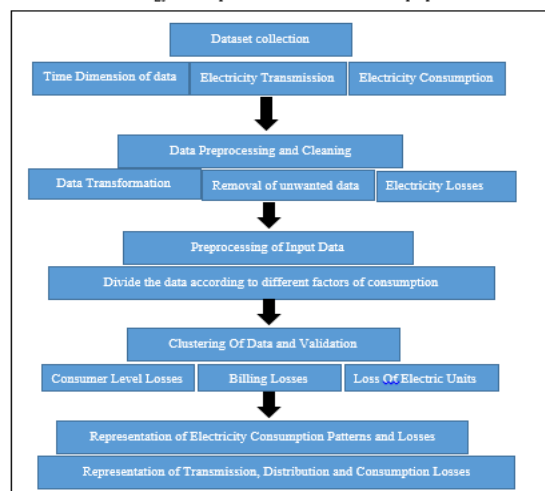


Figure 1: Data Visualization Framework for Monitoring Of Electricity

2. LITERATURE REVIEW

For the extraction of difficult data and transform it in an easy way we need data visualization techniques to handle the data more carefully and provide the opportunity to the user for better understanding and results for this data passed through different phases like collection and analysis and by the time for extraction of important factors data mining is used, because datasets needs to be sorted first to avoid the mess.

Electricity losses in a power system are either technical or non-technical in order to avoid and overcome the technical losses at grid station we use to normalize the data and inefficient use of electricity and its transmission .There are a lot of factors that happened during electricity losses. Sometimes the equipment has less capacity and more energy transmitted at high voltage so one of the major thing is we need to optimize the equipment and normalize the existing usage. Transformer uses frequently may not be advantageous because due to technical and non-technical factors. This needs normative value to monitor the power losses. To control the electricity in a transformer we need regulators in them to overcome the technical factors and avoid power losses because old transformers with low maintenance have more chance of power losses as compared to the new ones. Regulators change the power flows in a transformer [4-6]

It is the hour of need to overcome the power losses that helps to overcome the shortfall of electricity. One of the major causes of high consumption of electricity is due to poor wiring at primary and secondary level. Due to so many technical losses economy deficits

a lot and needs more money on new equipment's. That's why we need to improvise the equipment which results in less losses and not exceeds than the desired ones. In order to create a framework for the diagnostics of electricity losses in a country .Monitoring of electricity losses and their distribution must be concerned because technical and non-technical factors are ahead. There must be a deficit of 3000MWH electricity /Year as the production and distribution level also increases but still the consumption remains high.

Power cables used in a grid are the critical things to handle because unexpected failure lead to shortage of energy distribution and sometime complete failure which results in system failure and losses[7].Materials used in a transformer affected by overheating ,overloading and aging which causes supply system failures and power losses[8].

People are unaware of the system that how much energy is consumed in a country so it must be known that how electricity distributed and consumed with the help of some patterns and understanding that can be possible with data visualization techniques [9].Consumer often have no idea about their electric consumption pattern sometimes the appliances they use become obsolete and instead of giving advantage actually becomes the cause of loss and higher energy consumption .They visualization through time series analysis helps to figure out the losses and learns the electricity pattern for balanced amount of distribution and consumption[10].In order to introduce the smart things and implement them for efficient and accurate results ,different strategies and prerequisites of the issue

must be studied and practiced in detail[11]. Economic stability of the world for electricity consumption depending on different geographical and social factors as well as the human behaviors[12]. electrical efficiency should be maintained for

different domains to reduce the loss of electricity and increase the efficiency. There are a lot of factors that affect the energy efficiency in residential as well as in recreational

2.1. Possible Solutions

Table 1. Possible solutions and exploration of framework

Analysis	Description	Example
Descriptive	What is happening around	Previous datasets
Diagnostic	Reasons behind the scenario	Detection of losses
Predictive	What would happen	Manage losses
Prescriptive	Possible ways to manage	Mitigate losses
Context	Circumstances and consequences	Climate change /irregular precipitation /higher temperature

3. FRAMEWORK AND METHODOLOGY

Data visualization techniques help in filtering and abstraction of data can be easily described by the specific models which help in accurate decision making [12]. To extract the meaningful data from the given collected information benefits the organization to keep things assemble and update the data on time. To measure the efficiency and accuracy of electricity distribution and consumption in a country it is necessary to record and monitor the parameters at real time [13-15].

First of all data collected from different sources must be refined according to the key

variables. NEPRA(National Electric Power Regulatory Authority) and NTDC(National Transmission and Dispatch company) generates statistical report every year which provides the information that how much electricity generated, installed, and consumed according to the different sectors. Temperature varies from place to place so as the climate also affects the distribution and transmission of electricity.

At provincial level different distribution companies are responsible to transmit electricity from grid station to low level electricity usage at domestic, industrial

agriculture level and also for traction(Transport).System losses are monitored due to technical factors but non-technical factors may not be monitored or sometimes not known. Different types of consumers belongs to different sectors of the country at which losses percentage varies .In the proposed model diagnostic framework helps to overcome the electricity losses and analyze the factors that affects the electricity consumption and distribution. Electricity is transferred from low voltage to high voltage through different levels of transmission lines. Electricity distributed in different areas of the country through different grid stations with the help of transmission lines which are 750KV, 550KV, 350Kv, 240KVand 138KV.

The system have:

- a. The ability to identify the losses on the basis of electricity distribution and consumption at different levels.
- b. The ability to do time series analysis to forecast the electricity demand at peak hours.
- c. From previous losses, system is able to diagnose the deficit in electricity consumption on the basis of different factors.
- d. The diagnostic framework helps to implement the latest techniques which overcome the losses during electricity consumption.

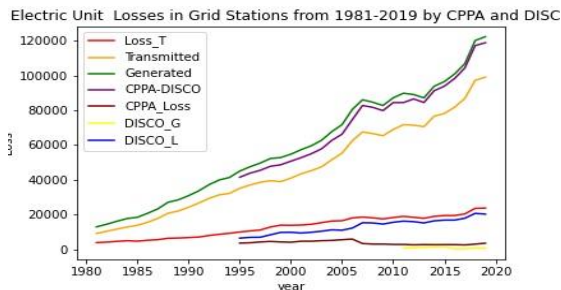


Figure 2: Electric Unit Losses in CPPA and DISCOs

3.1. Model for Diagnostics of Electricity Losses:

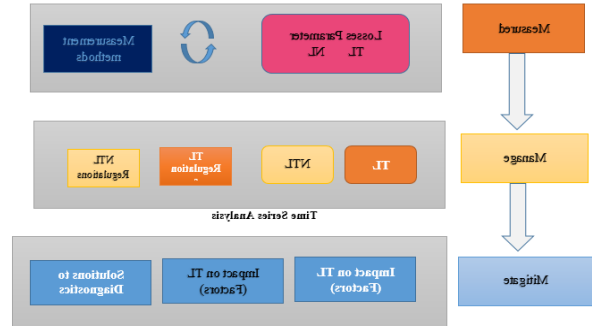


Figure3: Diagnostic framework for electricity losses

For supervised learning:

$$Y = F(X) \quad \dots\dots (i)$$

Where,

Y is a dependent function of X, This is helpful for time series analysis where time measured according to the different factors whereas factors vary from situation to situation. Here is an autoregressive model following time series analysis that helps to forecast the electricity consumption in different sectors.

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_n y_{t-n} + \epsilon_t \quad \dots\dots (ii)$$

The above equation helps to predict the electricity consumption on the basis of time series. In autoregressive model ACF is a coefficient of correlation function for time

$$Corr(y_t, y_{t-k}) \quad \dots\dots (iii)$$

3.2. Algorithm for Experiments

Algorithm: 1	:	Losses Diagnostics and Prediction
Input	:	TL/NTL parameters
Output	:	Forecasting/diagnostics of losses and electricity consumption
<i>ParaList</i> []	:	Electricity Units, Consumer level, billing, Consumed units
<i>Managecategory</i> []	:	Climate change, Temperature, inefficient consumption
Begin		
Step 1	:	Measured Loss()
Step 2	:	Extract parameters \leftarrow <i>ParaList</i> []
Step 3	:	for each category[i] monitor losses
Step 4	:	Call: Measure Loss()
Step 5	:	Manage Loss()
Step 6	:	for each <i>Managecategory</i> []
Step 7	:	Manage <i>ParaList</i> []
Step 8	:	Split datasets according to values
Step 9	:	Train datasets of <i>ParaList</i> []
Step 10	:	Test datasets of <i>ParaList</i> []
Step 11	:	Train and test auto regression(train ,test ,lags) $lag = [history[i] \text{ for } i \text{ in range}(length - window, length)]$
Step 12	:	Model = AR model (Autoregression model)
Step 13	:	Fit the training and testing datasets in model
step 14	:	Call: Manage loss()
Step 15	:	Mitigate Loss()
Step 16	:	Predict the possible solution from <i>ParaList</i> []
Step 17	:	for each
Step 18	:	If no.of predicted results=12 then it must be a year
	:	Else
Step 19	:	months[]
Step 20	:	Check the accuracy of the model from the predicted

- values
 - abc.predict(test)
 - Step 21 : Check the residual errors of ARIMA model
RMSE , MAE , MSE
 - Step 22 : Plot residuals and predicted results
 - Step 23 : Call: Mitigate Loss()
 - Step 24 : **End** ← if accuracy is appropriate
Else
 - Step 25 : Check the residuals and fit model
-

4. SIMULATIONS AND EXPERIMENTS

4.1. Losses in Billing at Provincial level:

Different companies are responsible to supply electricity in different provinces. From PEPCO to DSICO some losses takes place .Each company is able to provide electricity in different provinces and is responsible for its transmission and losses either at transmission level or billing level .Parameters to measure the losses are the units supplied and consumed by each province. Below the figure shows the billing losses in previous years from 1981-2019[16].

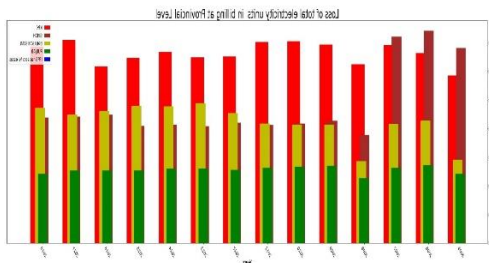


Figure 4: Data Visualization of provincial level losses

4.2. Consumption losses every year with respect to consumer:

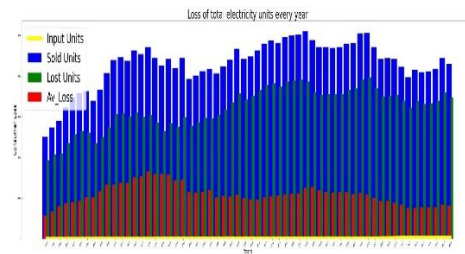


Figure5: Consumer level losses every year

The parameters taken to measure the losses at consumer level .By each distribution company every year a specific amount of energy is distributed and transmitted to each consumer by which losses can also takes place at consumer level also.

4.3. Effects of Climate Change on Loss of Electricity

Climate change is the biggest factor that affects the electricity distribution and transmission in a country. If the temperature goes high low medium wires gets melted and cannot be able to transmit specific amount of electricity which results in short

fall. Due to high temperature load on power stations increased and results in power losses and electricity deficit. Data collected from the meteorological reports of climate [17] where loss of electricity can be calculated and helps to overcome them. From the given simulations annual precipitation level can be calculated from the \maximum and minimum monsoon seasons and electricity consumed according to them in different cities.

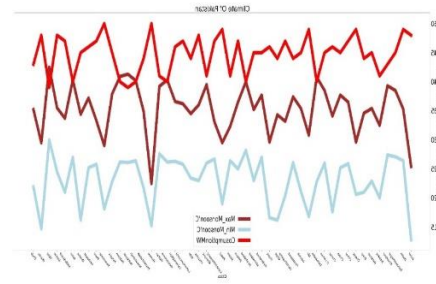


Figure 6: Effect of climate on electricity consumption

5. RESULTS AND CONCLUSION

Prediction of Consumer level losses

Input:

Parameters (Electricity units, consumers, Losses).

Target Values:

Loss Result

Output:

Prediction and accuracy

5.1. Division of Data Set for Storage:

The shows the percentage of data that is divided for training, testing and validation purpose.

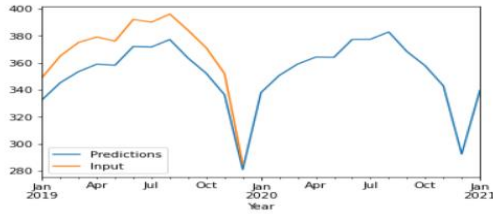
Table 2. Consumer level dataset division

Data Set Division	
Total Number of Samples	720
Training Data	70%
Validation Data	15%

Test Data	15%
-----------	-----

5.2. Result of Losses and Prediction with 70 – 30 Division of Dataset:

The graph shows the actual and predicted results on the basis of previous data .From time series analysis we used auto regression model which forecasts the values and also find the residual errors in a model. Model shows the 96% accuracy at consumer level predictions and 94% at Input Units. Data collected from the sources of NBEPR and NTDC (National Transmission and Distribution companies) which are input units delivered in the city every year to fulfill the desired needs. With the help of the previous calculation we are able to get the predicted results by using specified model.



Input units

Figure 7: Predicted results from ARIMA

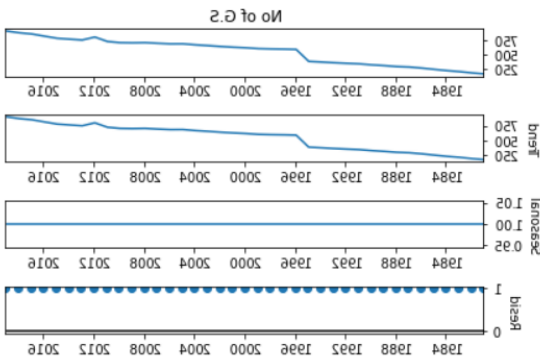


Figure 8: Directions towards future grid stations

REFERENCES

[1] X. Lü, T. Lu, C.J. Kibert, M. Viljanen, Modeling and forecasting energy consumption for heterogeneous buildings using a physical-statistical approach, *Appl. Energy* 144 (2015) 261–275. <https://doi.org/10.1016/j.apenergy.2014.12.019>.

[2] Cai, Y., G. Huang. An optimization-model-based interactive decision support system for regional energy management systems planning under uncertainty *Expert Systems with Applications*, 36 (2009), 3470–82

[3] Tan Guilong, Chen Yuttti. Application Research of information visualization method based on parallel coordinates [J]. *Journal of Beijing Technology and Business University: Natural Science Edition*, 2008,

Electricity distribution changes a lot in a smarter way but there are still more losses to measure and monitor .Even as the time passes the huge difference in losses occurs every year even the production growth increase as per 20-30% by the policy. In past years smart metering systems were introduced which also helps to reduce losses but they are very expensive and only in smaller number .How can losses be overcome at villages and other areas where billing units are also lost and a huge difference occurs between consumed and supplied units at consumer level.

In order to avoid all these losses we need to use the forecasting of data to predict some closely related data that helps to overcome the losses which results in less electricity consumption.

26 (2): 7579. *Roy. Soc. London*, vol. A247, pp. 529-551, April 1955

[4] O. Rubanenko, I. Hunko, O. Rubanenko and A. Rassölkin, "Influence of Solar Power Plants on 0.4 kV Consumers," 2019 IEEE 60th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), Riga, Latvia, 2019, pp. 1-5.

[5] S. L. Gundebommu, O. Rubanenko and I. Hunko, "Analysis of Threelevel Diode Clamped Inverter for Grid-connected Renewable Energy Sources," 2019 IEEE 20th International Conference on Computational Problems of Electrical Engineering (CPEE), Lviv-Slavske, Ukraine, 2019, pp. 1-6. doi: 10.1109/CPEE47179.2019.8949104

- [6] Reduction of Technical and Non-Technical Losses in Distribution Networks. Final report 20/11/2017 pp. 88–95, 2016. <https://doi.org/10.3311/PPee.8893>
- [7] C. Zhou, H. Yi, and X. Dong, “Review of recent research towards power cable life cycle management,” *High Voltage*, vol. 2, no. 3, pp. 179–187, 2017.
- [8] S. Charles and T. McGrail, “Sweep Frequency Response Analysis Transformer Applications: A Technical Paper from Doble Engineering,” Doble Engineering Company, MA, USA, Version 1.0, pp.1-47, Jan 2003.
- [9] Attari, S. Z., DeKay, M. L., Davidson, C. I., & De Bruin, W. B. (2010). Public perceptions of energy consumption and savings. *Proceedings of the National Academy of Sciences*, 107(37), 16054–16059. doi: 10.1073/pnas.1001509107
- [10] Herrmann, M. R., Brumby, D. P., & Oreszczyn, O. (forthcoming). Watts your usage? A field study of householders’ literacy for residential electricity data visualizations. *Energy Efficiency*.
- [11] Fan, C., Xiao, F., Wang, S.: Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Appl. Energy* 127, 1– 10 (2014)
- [12] Hsu, D.: Identifying key variables and interactions in statistical models of building energy consumption using regularization. *Energy* 83, 144–155 (2015)
- [13] Luo Jian. Research and implementation of visual data mining method [D]. Xi'an: University of Electronic Science and technology of China, 2009.
- [14] Alonso, M.; Amaris, H.; Alcalá, D.; Florez, D. M. Smart Sensors for Smart Grid Reliability. *Sensors* 2020, 20, 2187.
- [15] Lee, S.; Choi, D.H. Energy Management of Smart Home with Home Appliances, Energy Storage System and Electric Vehicle: A Hierarchical Deep Reinforcement Learning Approach. *Sensors* 2020, 20, 2157.
- [16] [https://media.npr.org/documents/about/annualreports/2019 Annual Report.pdf](https://media.npr.org/documents/about/annualreports/2019%20Annual%20Report.pdf)
- [17] <https://www.pmd.gov.pk/>