Data Quality Model of ATandC Losses of Feeders in Power Sector – Case Study

Objective –

All energy supplied to a distribution utility does not reach to the end consumer. A substantial amount of energy is lost in the Transmission and Distribution (T&D) system by way of Technical and Non Technical Losses. The distribution system accounts for highest technical and non technical losses in the power system.

Aggregate Technical and Commercial Losses (AT & C) of individual feeders over a span period of twenty two months was available for WBSEDCL from National Power Portal. Objective was to see the quality of data of individual feeders compared to the loss minimized model of all the feeders' performance over the same period.

Definitions –

Billing Efficiency – It is an indicator of proportion of energy that has been billed (includes both metered and unmetered sales) to consumers w.r.t. energy supplied to an area. It can be computed using formula provided below :

Billing Efficiency = Total Energy Billed to Consumers (kWh) / Total Energy Input (kWh)

Collection Efficiency - All consumers are billed on the basis of energy consumed by them which is obtained from meter reading and assessment of unmetered energy of consumers. The billed amount is computed on the basis of tariff fixed by regulatory commission for applicable customer category. Collection efficiency is an indicator of proportion of amount that has been collected from consumers w.r.t. amount billed to them. Collection efficiency can be computed using formula provided below :

Collection Efficiency = Revenue Collected (In Rupees)* / Billed Amount (In Rs.)

*The revenue collected shall exclude the arrears. Collection efficiency to be capped at 100%.

AT&C Losses

The aggregate technical and commercial losses shall be measured using formula mentioned below:-

AT&C Losses = {1 - (Billing Efficiency X Collection Efficiency)} X 100

The concept of Aggregate Technical & Commercial losses provides a realistic picture of loss situation in the context it is measured. It is combination of energy loss (Technical loss + Theft + inefficiency in billing)& commercial loss (Default in payment + inefficiency in collection).

Solution –

Attempt was made to study AT&C losses of feeders of a sample organization WBSEDCL for 8960 records of Twentytwo months data from April 2015 till January 2017. The Data was Broken into two sets of 6000 records for training and 2960 records for Testing the model created.

When we run Linear Regression on Numerical variables there are two methods that can be used. 1) Using Normal Equation for Linear Regression and 2) Linear Regression using Gradient Descent for Cost Minimization of the Dependent Variable in this case ATandCLoss attribute.

Case I - Inputs used were Original Independent Variables: amount_billed, abr (average billing rate), billed_energy, net_energy, reporting_Period, ReportingMonth_Start, ReportYear_Start.

Under Normal Linear Regression Model obtained had the following parametric weights.

0.00017464 -0.0055126 * feeder_id +3.0204E-07 * amountBilled -4.5407E-06 * BilledEnergy +2.6462E-06 * NetEnergy -0.0026649 * AvgBillingRate +0.047059 * ReportingPeriod +0.18089 * ReportingMonthStart +0.36075 * ReportingYearStart

This can be used to predict the AT&C loss for the feeder under consideration for the next few months. Ofcourse as the data gets fed back into the system from all other feeders in the ecosystem the model may change as performance improve.

Linear Regression Model using Gradient Descent for Loss Minimisation obtained had the following parametric weights. In the table below Theta Training is obtained by initializing the Theta (parametric weights to Zero) and these weights get adjusted as there is gradient descent of this cost (Loss) Function over a number of iterations to get to the lowest value.

Then this model is tested against another set of data and overfitting/ underfitting of model is reduced by using the parametric weights obtained from training data to train the test set. The Final theta (parametric weights) obtained can then be used to depict the model for incoming data.

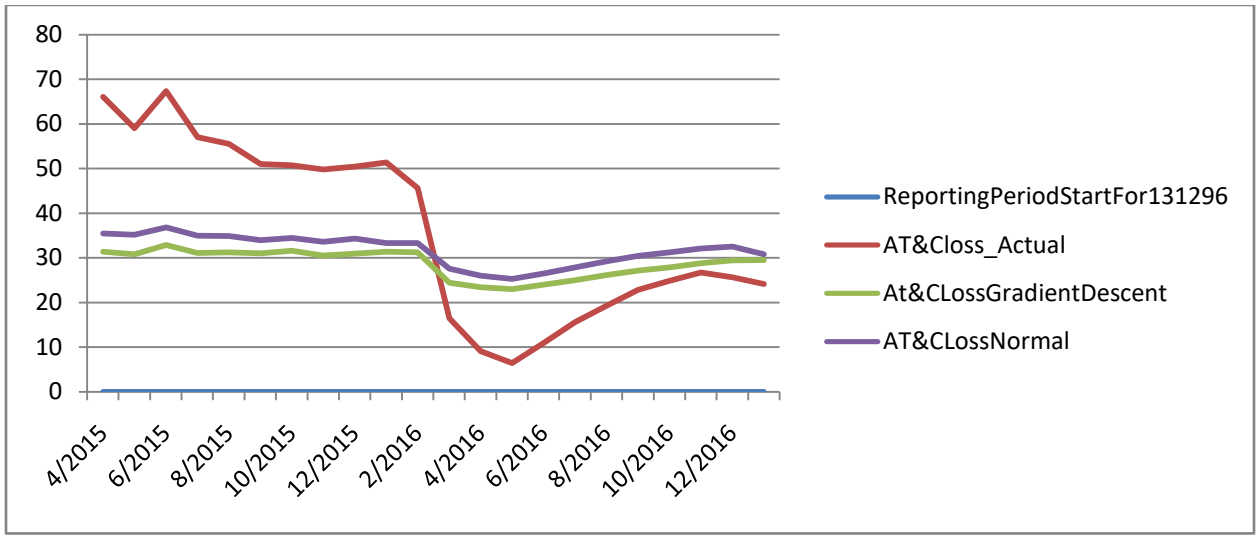
	theta_grad_descent_theta 1Original (Obtained with parametric weights initialized to Zero)	**theta_grad_descent_Theta -2Original** (Obtained with parametric weights initialized to Theta1)
Intercept	26.50358961	**27.03018263**
feeder_id	-1.01509536	**-0.384665399**
amountBilled	-4.034334026	**-7.558710439**
BilledEnergy	-4.506295047	**-8.919947418**
NetEnergy	5.485532778	**12.86473903**
AvgBillingRate	-1.239889991	**-2.34854916**
ReportingPeriod	0.801044116	**1.314110216**
ReportingMonthStart	0.610837784	**1.187159472**
ReportingYearStart	0.145955345	**2.253332468**

Hence cost minimized model becomes

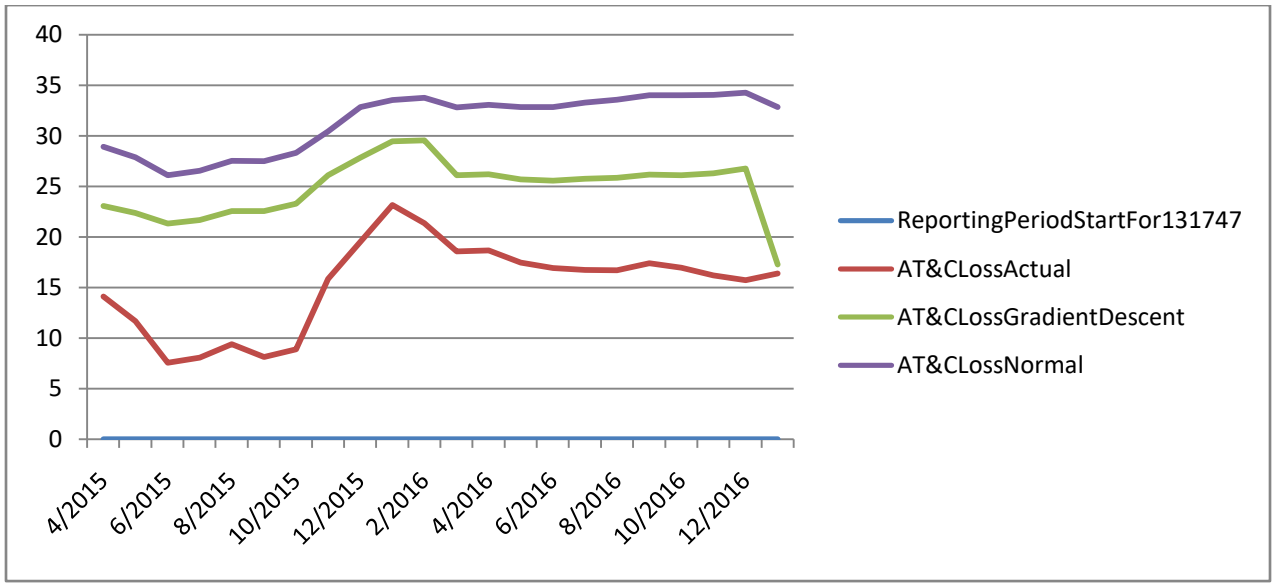
27.03018263 -0.384665399 * feeder_id -7.558710439 * amountBilled + 8.919947418 * BilledEnergy +12.86473903 * NetEnergy -2.34854916 * AvgBillingRate +1.314110216 * ReportingPeriod +1.187159472 * ReportingMonthStart +2.253332468 * ReportingYearStart

This model has been optimized for losses suffered by all feeders over the period under consideration for the dataset given.

A few sample graphs attached. For the Feeder No. 131296 it can be seen that AT&C Loss Actual has improved over the models AT&CLoss Normal as well as AT&CLoss CostMinimised over all the feeders.



Let us take another case for Feeder No. 131746. Similar results can be seen.



Let us take another case for Feeder No. 137844. Actual AT&C Loss of the feeder is more then AT&C Lost Cost minimized across all feeders.

Case II - Inputs used were variables: amount_billed, abr, billing_efficiency, collection_efficiency, reporting_Period, ReportingMonth_Start, ReportYear_Start.

Under Normal Linear Regression Model obtained had the following equation with parametric weights as obtained.

0.00043541 -0.0036092 * Feeder ID + 0.52837 * AvgBillingRate -0.50076 * BillingEfficiency -1.0856 * CollectionEfficiency + 0.97972 * ReportingPeriod + 0.020461 * ReportingMonthStart + 0.14021 * ReportingYearStart

Linear Regression Model using Gradient Descent for Loss Minimisation obtained had the following parametric weights.

	theta_grad_descent_Theta-1 (Obtained with parametric weights initialized to Zero)	**theta_grad_descent_Theta-2 (Obtained with parametric weights initialized to Theta1)**
Intercept	26.75434096	**26.54087322**
Feeder ID	-0.528690235	**0.15655912**
AvgBillingRate	-0.560958973	**-0.533449332**
BillingEfficiency	-13.21915618	**-14.5597808**
CollectionEfficiency	-8.273208881	**-9.594929446**
ReportingPeriod	0.28142881	**-0.170876693**
ReportingMonthStart	0.209784018	**-0.317578784**
ReportingYearStart	0.093845104	**-0.077563785**

Observations - It was seen from both models that there was a bias of 26.5 to 27% which is the intercept saying that Minimum AT&C Loss will be in this range across all feeders of this organization. It compares well with the figures released by CEA as shown below.

	2010-11	2011-12	2012-13	2013-14	2014-15
T&D Losses	23.97	23.65	23.04	21.46	22.77 (P)
AT and C Losses	26.35	26.63	25.48	22.56	24.62 (P)

Reference - http://www.cea.nic.in/reports/monthly/executivesummary/2018/exe_summary-03.pdf

WHY LOSS MINIMIZATION IS ESSENTIAL?

The Transmission and Distribution Losses in our country, which were around 15% up to 1966-67, increased gradually to 28.36 % by 2011-12. Total Transmission and Distribution Losses are about 30-50%. The major part of the loss is taking place only in distribution sector which accounts for 80-90% of total T&D losses. Cost of power theft is Rs 20,000 Crores / year and the total loss incurred by all State Electricity Boards is Rs 26,000 Crores per year in India.

The Transmission and Distribution losses in advanced countries of the world have been ranging between 6 to 11%. As per the T & D losses issued by CEA, taking into consideration the Indian conditions, it would be reasonable to aim for containing T & D losses within 10 to 15% in different States. The losses (on accrual basis) of all the utilities increased from Rs 24,796 crores in the year 2008-09 to Rs 30,466 crores in 2009-10. In the year 2010-11, the aggregate losses of all utilities decreased to Rs 29,701 crores .

Accurate estimation of T&D Losses has gained importance as the level of losses directly affects the sales and power purchase requirements and hence has a bearing on the determination of electricity tariff of a utility by the commission. Higher AT & C losses have severe impact on tariff as well as economy. So, minimization of this losses is utmost essential.