Paper 3944 -2019

Distribution Transformer Health Monitoring and Predictive Asset Maintenance

Prasenjit Shil, Ameren; Tom Anderson, SAS Institute Inc.

ABSTRACT

Distribution transformer is the most vital asset in any electrical distribution network. Hence, distribution transformer health monitoring and load management are critical aspects of smart grids. Transformer health monitoring becomes more challenging for smaller transformers where attaching expensive health monitoring devices to the transformer is not economically justified. The addition of Advanced Metering Infrastructure (AMI) in smart grids offers significant visibility to the status of distribution transformers. However, leveraging vast amount of AMI data can be daunting. This paper uses the hourly usage data collected from Ameren Illinois' AMI meters to determine distribution transformer outage, failure, and overload. The proposed methodology not only detects and visualizes outage and congested areas in near real time, but also detects transformers and distribution areas with a long history of outage and congestion. This paper also offers a predictive algorithm to enhance regular equipment maintenance schedules and reduce repair truck trips for unscheduled maintenance during unplanned incidents like storms. SAS® Enterprise Guide®, SAS® Enterprise Miner™, and SAS® Visual Analytics were used to efficiently produce the information necessary for operational decision-making from gigabytes of smart meter data.

INTRODUCTION

Electric power distribution circuits provide power to end users through a network of cables, transformers, switches, and other devices. The loading of each of these elements in the electric circuit is limited by the physical capacity of devices. When actual load exceeds certain design threshold, the circuit is shut down or damaged or becomes susceptible to future failure. It's necessary for the utilities to successfully identify the transformers, switches, devices etc that are of higher risk to avoid interrupting power supplies to end-users. However, given the quantity of the distribution transformers deployed in the distribution network, the line distribution transformers are not monitored by any automated processes. Distribution service transformers typically do not have sensors, instrument transformers, and other equipment to monitor its health. Typically, fault issues are captured on these devices when customers call in with a problem or when an outage is identified followed by investigation and perhaps, replacements.

The cost of distribution service transformers ranges from ~$1000 to more than $100,000+ based on its size whereas the power transformer costs can easily exceed $1 million. Preventing or reducing transformer failure will help in reducing transformer maintenance cost (O&M expense for the utilities) and increase system reliability, which should provide enough incentives for the utilities to proactive prediction of pending transformer failures. Table 1 summarizes the cost-benefit of transformer health monitoring to make a business proposition for the utilities (EPRI, 2016; EPRI, 2006).

With the adoption of smart meter or automated metering infrastructure (AMI) and big data platforms to collect and analyze AMI meter data, utilities are equipped to continuously monitor and proactively manage distribution transformer overloading. This paper utilizes hourly usage (or electric load) data collected from Ameren Illinois' AMI meters and other distribution network data such as distribution transformer outage, failure, and overload. The proposed methodology not only detects and visualizes outage and congested areas in near real time, but also detects transformers and distribution areas with a long history of outage and congestion. SAS Enterprise Guide, SAS Enterprise Miner™, and SAS Visual Analytics were used to efficiently produce the information necessary for operational decision-making from gigabytes of smart meter data.

Cost/Benefit Analysis Categories		Favorable/ Unfavorable In/Decrease	Comments
System Operations	Fuel		
	Purchased Power		
	Ancillary Services		
	Direct Emissions Costs		
	Operator Costs		
	Revenue from Enabled Sales		
Utility Operations (Non-Fuel O&M)	Generation	↓	Labor costs of crews to perform periodic inspections.
	Transmission		
	Distribution		
	Customer		
	A&G (Admin & General)		
Utility Operations (Non-Production Assets)	Trucks, A&G, Tools, Software	↑	Cost of algorithm and application development and implementation.
Capital Revenue Requirements	Generation	↑	Identification of pending failure would cause replacement/ maintenance sooner.
	Transmission		
	Distribution		
Reliability	Damage from Voltage Variations	↓	Avoided damage via early detection.
	Damage from Harmonics; PQ other		
	Interruption Costs Sustained/Normal		
	Interruption Costs Momentary		
	Interruption Costs Major Events		
Customer Direct	Value of Service (Comfort, Light, etc.)		
	Cost of Equipment (Devices)		
	Value of Information Provided		
Environment	ΔTons SO_2	↓	Reduction in or elimination of truck rolls to inspect asset.
	ΔTons NOx		
	ΔTons CO_2		
	ΔPounds Hg		
	ΔParticulates		
Security Impacts	Oil Saved	✓	Recognizes positive impacts of avoided miles/km driven.
	Major Blackouts Avoided		
Efficiency Impacts	ΔkWh System Losses		
	ΔkW System Losses		
	ΔkWh Consumed		
	ΔkW Consumed		
Equity Impact	Metering Accuracy		
Safety Impact	Public Safety	✓	Analytics reduces or eliminates the need for field inspection of asset.
	Employee Safety		

Notes: Green or Red text or symbol indicates a favorable or unfavorable change, respectively. A ✓ indicates a change that may be difficult to quantify with precision. Arrows (↑↓) indicate an increase or decrease.

Table 1: Cost-Benefit Case for Transformer Health Monitoring

Source: EPRI technical report "Data Analytics Cases for Asset Awareness," EPRI, Palo Alto, CA: 2016. 3002008799. Visit ***www.EPRI.com*** *for details.*

DATA AND DATA MODEL

Electric distribution transformer rarely breaks down suddenly (EPRI, 2006). Usually there is enough priori signs from the gradual breakdowns of transformer components that leads to eventual transformer failure (EPRI, 2006). Such breakdowns provide early warning signs of possible future transformer failure. This paper explores ways to utilize smart meter data and outage, failure, or customer information to detect such early detection signs for future possible transformer failures.

This paper utilizes data from various sources as shown in Figure 1. The primary data set is the AMI meter dataset from 13 different circuits in Ameren Illinois' distribution network. The AMI meter data were augmented by mapping them to the corresponding customer class, geo-location, distribution transformers, device, switch, and reclosers etc. The analysis also uses transformer ratings provided by the manufacturers, outage and known failure information (event data) by feeder/transformer and weather data for the nearest weather station. Figure 2 shows the analytical framework using the data presented in Figure 1. Analytical framework shown in Figure 2 is similar to the framework presented by Anderson and DeBose (2013) in a previous publication.

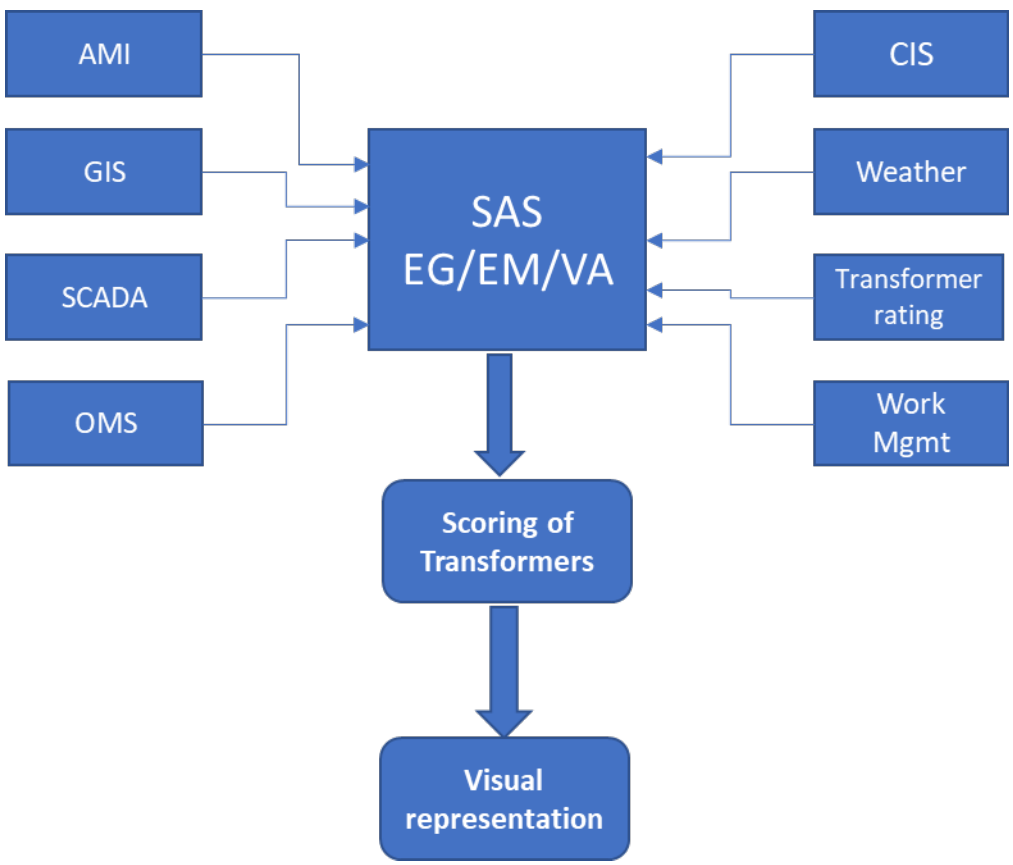

Figure 1: Input Data for Model Construction

The following section provides a brief description of various input data.

- AMI: AMI or Advanced Metering Infrastructure provides usage data (kW or kWh) from the customers' meters. The paper utilizes the hourly or sub-hourly interval usage (kW) data collected from the AMI system.
- GIS: The geospatial information system or GIS provides the geolocation of the elements in the feeder/circuit.
- SCADA: The Supervisory Control and Data Acquisition or SCADA data provides time stamped information about the distribution system devices.
- OMS: The Outage Management System or OMS acts as a repository for all the outage related information and metrics. It's an important parameter in the model presented in this paper.
- CIS: The Customer Information System or CIS contains all customer related information such as customer address, customer number, meter number, usage etc including upstream distribution transformer that provides power to the customer. It's essentially a database of historical information of customer load and other operational data.
- Weather: Various weather data elements from different weather stations across the service territory are sourced from outside vendor and stored within a repository.
- Transformer Rating: Transformer rating information helps in calculating transformer loading or busy factors at a given instance
- Work Management: Work management data comes from various internal systems and provides information about cause, description and duration of the outage and parts replaced etc.

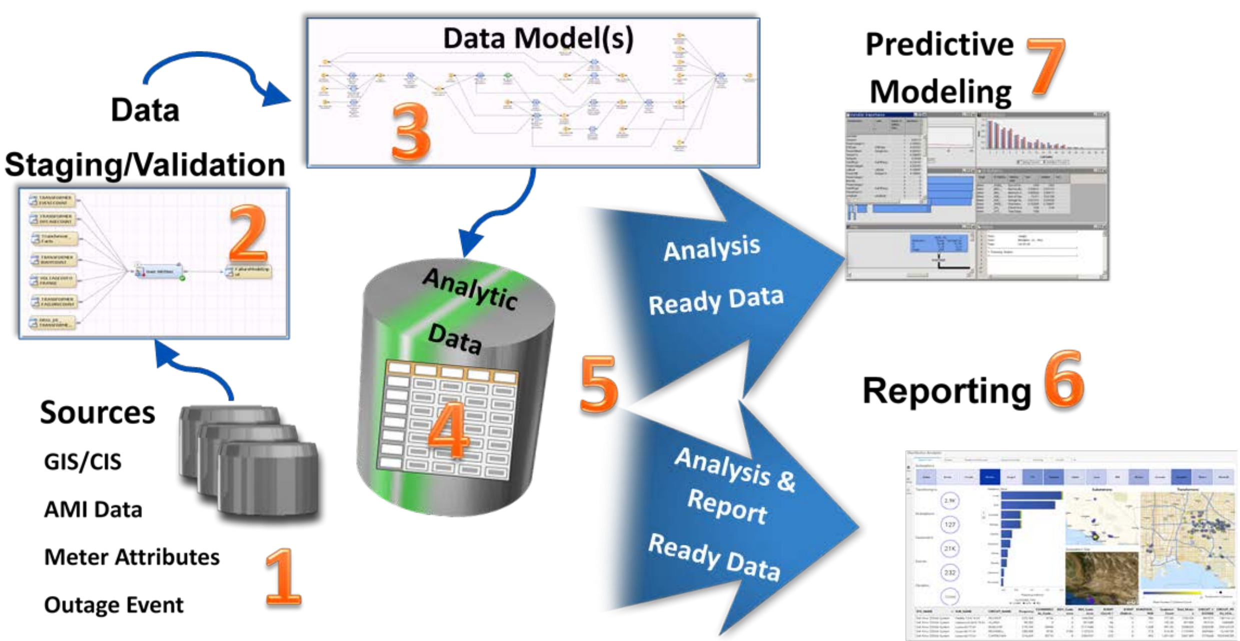

Figure 2: Analytic Framework for Transformer Health Index

The dataset from various sources, as shown in Figure 1, are brought into SAS EG to process and build data model for efficient and easy analytic model development and deployment. The final dataset is imported into SAS EM to construct predictive models and finally the results are represented in SAS VA. Figure 3 shows a highlight of the data models in SAS EG.

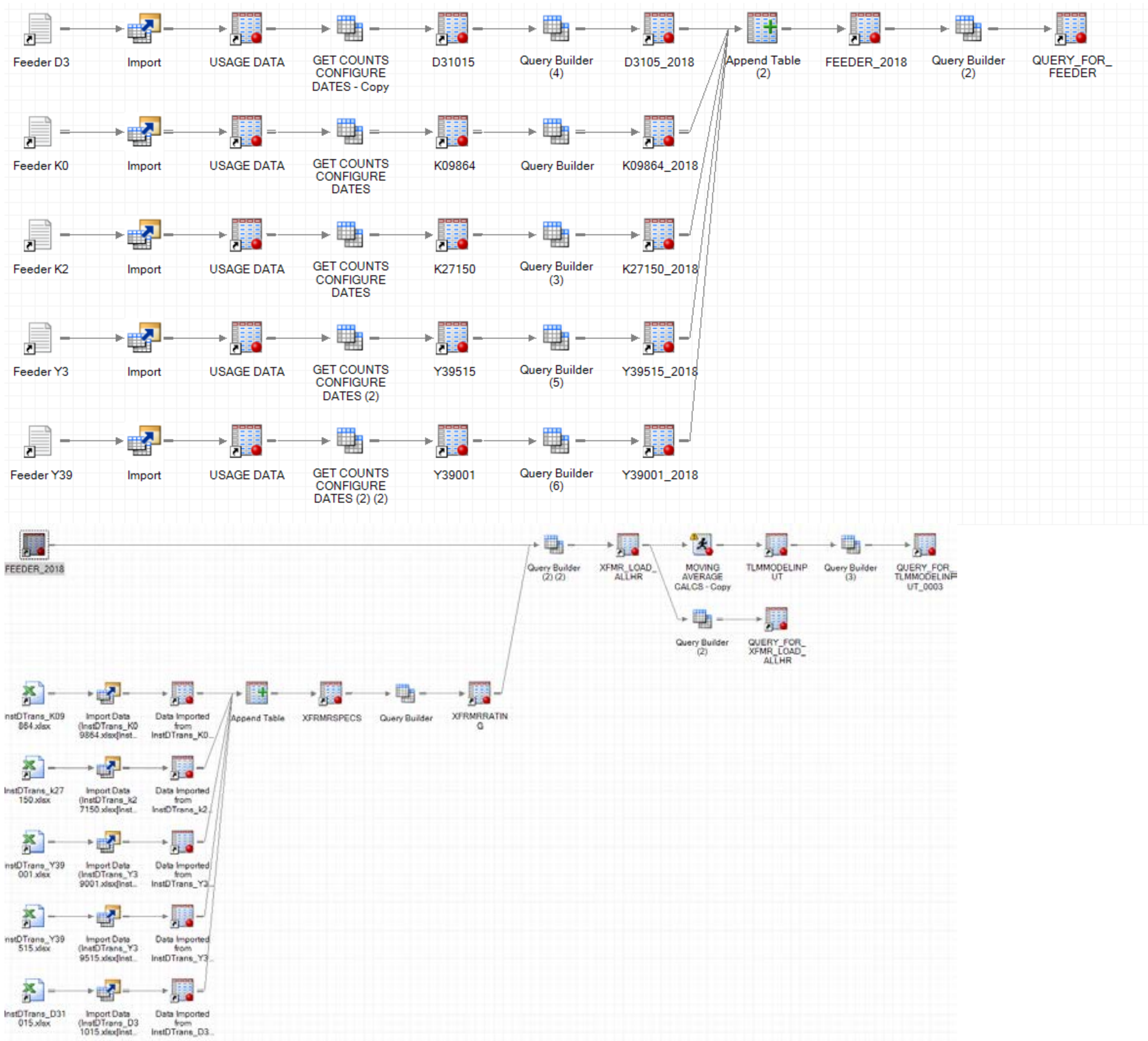

Figure 3: SAS Process Flow for Data Input Process

ANALYTICAL CONCEPTS AND BACKGROUND

Numerous studies have been published by both academia and industry to predict transformer failure or loss of life based on the aging of the isolation of the transformer and the operation temperature due to loading (EPRI, 2016; Rashid, 2011; Montsinger and Dann, 1934). Publications are available to construct mathematical models for the temperature of a transformer submerged in oil or the hot-spot temperature. This temperature is used for the evaluation of a thermal aging rate due to a cyclical load under ambient air temperature. Ambient temperature is the temperature of air conducting the radiators or heat exchangers. Research publications indicate that ambient temperature is an important factor affecting the life expectancy of the transformer (Rashid, 2011; Montsinger and Dann, 1934; Alquthami and Meliopoulos, 2014). The ambient temperature and the heat produced by the transformer windings are directly proportional to the load the transformer is bearing. Due to instrumentation limitation, utilities do not regularly measure the ambient temperature of the distribution transformers unlike the substation or bulk transformers. As load correlates to weather (temperature, humidity, wind speed, and dew point), this paper tests the significance of weather data in transformer failure.

Transformer loading is defined as the ratio of the load on the transformer to the actual nameplate rating of the transformer. Transformers experiencing loading above the nameplate rating will be prone to eventual failure. In electrical distribution systems, the transformers can be loaded above their nameplate capacity ratings due to emergency-situations or due to higher usage from the customers served by the transformer (Rashid, 2011; Jardini et al, 2000). Transformer overloading can be planned, short termed or long-termed. Risk of loss of life of the transformer is greater in case of long-term overloading. Overloading of transformers above its nameplate capacity or an ambient temperature higher than what it is designed for accelerates the aging of the transformer more than that in the case of normal loading (Rashid, 2011; Leal et al, 2002). The short-term overloading can also be significantly risky (Rashid, 2011). Sudden increase in load yields higher hot-spot temperature, which will breakdown of the thermal insulation system that will lead to transformer loss of life (Rashid, 2011; Jardini et al, 2000). Clearly, the key component of the transformer failure lies in higher than optimal loading. Therefore, this paper attempts to predict transformer failure using transformer loading or running average of busy factor as defined later in the paper.

Historical AMI meter data aggregated to the transformer level provide a loading (or, busy factor as noted in this paper) history on each transformer. The methodology proposed in this paper utilizes AMI interval data aggregated to the transformer and device level to calculate transformer utilization or transformer busy factor to predict possible loss-of-life in near real-time manner.

Ultimately, given the volume and complexity of the data, a good system will require an interactive system that can be applied to all transformers individually in near real time. This paper attempts to achieve the same.

Variable used for model construction are described below:

- Moving Values based on Average, Max and Variance
  - Time Periods
    - 6 hours
    - 12 hours
    - 24 hours
    - 72 hours
- Busy Factor: Busy Factor is defined as ratio between the load on the transformer and Transformer rating from the manufacturer.
- Weather: Weather variables that can spike up transformer load and subsequently busy factors are temperature, humidity, dew point and wind speed.

This predictive modeling step started with a Decision Tree, followed by a Forrest and Gradient Boost models. The results were compared before selecting the champion model. The following section discusses these models.

DECISION TREE

As described in literature and SAS manuals, a decision tree creates a hierarchical segmentation of the input data based on a series of rules applied to each observation. Each rule assigns an observation to a segment based on the value of one predictor. Rules are applied sequentially, which results in a hierarchy of segments within segments. The hierarchy is called a tree, and each segment is called a node. The original segment contains the entire data set and is called the root node. A node and all of its successors form a branch. The final nodes are called leaves. For each leaf, a decision is made about the response variable and applied to all observations in that leaf. The exact decision depends on the response variable. Figure 4 shows the Decision tree model presented for transformer failure.

Constructing the decision tree requires a measure, category, or date response variable and at least one predictor. This decision tree presented in this paper utilizes event data as categorical variable. This variable identifies an actual failure event on a transformer. The variable importance graph shows that the busy factor and certain weather conditions are most significant based on the relative importance of each variable. The average square error (ASE) for the model presented in this paper was 0.0893. ASC provides fit statistic for the decision tree and a better model would have lower ASC.

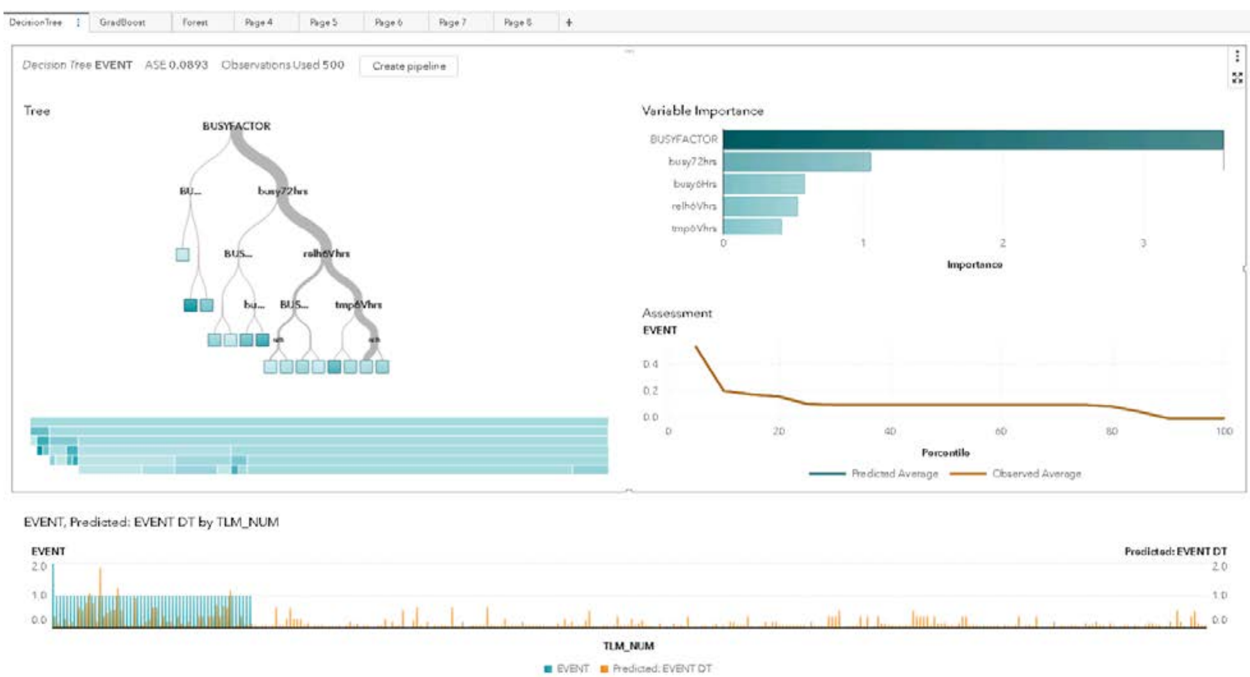

Figure 4: Predictive Model Output – Decision Tree

FOREST

As described in literature and SAS manuals, forest is an ensemble model containing a number of decision trees. Forest, a popular data science methodology, can be used to build predictive models for both classification and regression problems. To avoid any overfit the data, each tree in the forest is built on a different sample of the training data and when splitting each node, a set of candidate inputs for the split are selected at random, and the best split is selected from those. This ensures that the trees aren't correlated to each other.

Reviewing the Assessment Plot and the ASE score (0.0854), the Forest performed a little better than the Decision tree in predicting transformer failure. The variable of importance graph, as shown in the Figure 5, highlights that the Forest model used more data points in its evaluation than the Decision tree did.

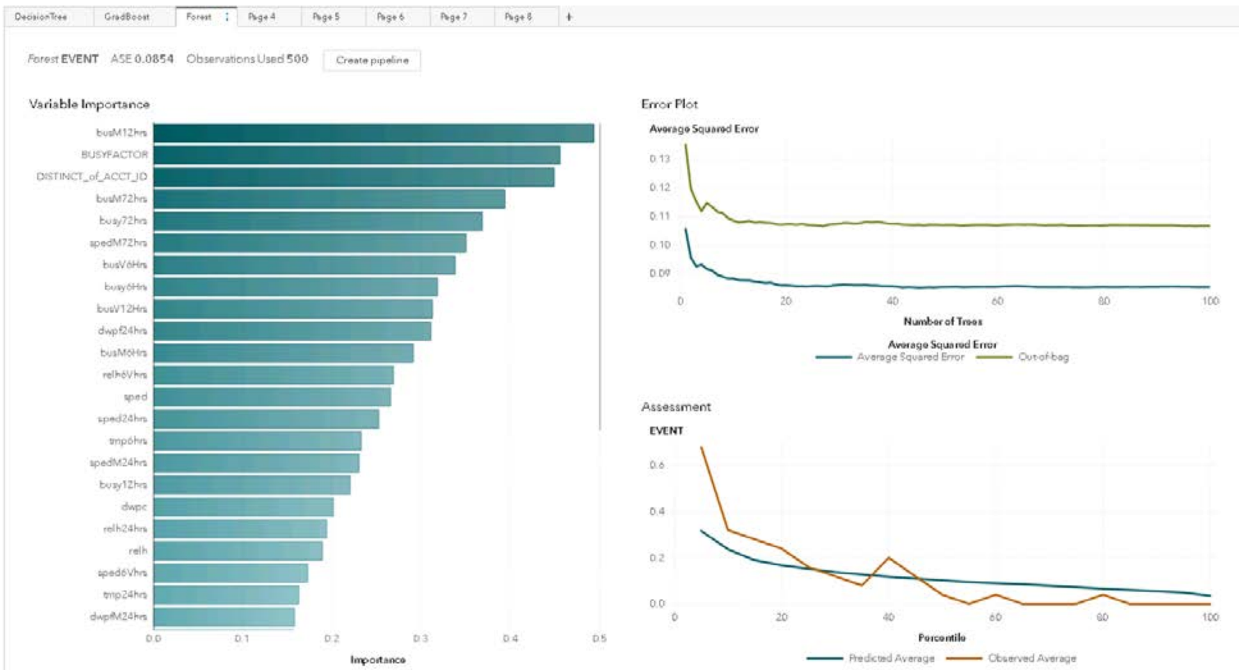

Figure 5: Predictive Model Output –Forest

GRADIENT BOOST

Gradient boosting is an iterative approach that creates multiple trees where each tree is typically based on an independent sample without replacement of the data. A gradient boosting model hones its predictions by minimizing a specified loss function, such as average square error. The first step creates a baseline tree. Each subsequent tree is fit to the residuals of the previous tree, and the loss function is minimized. This process is repeated a specific number of times. The final model is a single function, which is an aggregation of the series of trees that can be used to predict the target value of a new observation.

The Gradient Boost performed far better than the Decision Tree and Forest. With an ASE of 0.0110, the model predicted over 80% of the events. Noticeably, this model fares better when predicted event values are compared to actual event values (Figure 6).

Given low ASE and model fit, it was determined that Gradient Boost model had a better fit for the problem presented in this paper. Hence this model was used to score the distribution transformers in order to create the scorecard report displayed in Figure 8. Furthermore, this model can be implemented in real-time by leveraging SAS Event Stream Processing.

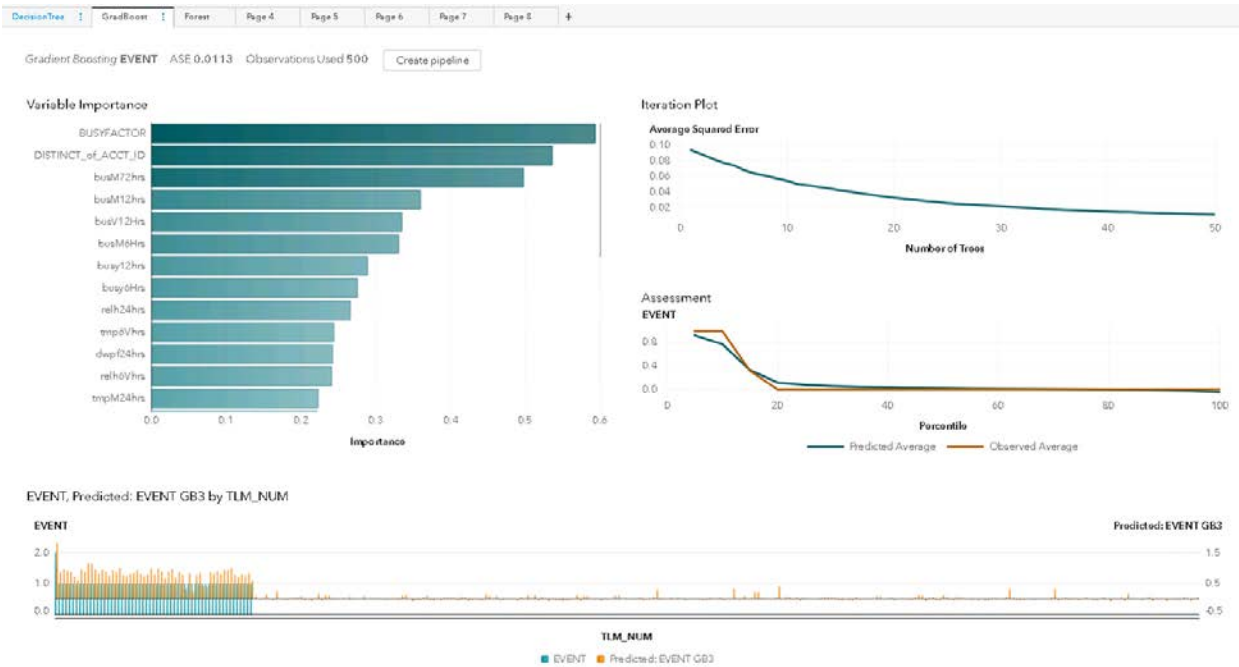

Figure 6: Predictive Model Output – Gradient Boost

Figure 8: Transformer Scorecard: Health Index with Historical Reference

It's important to note that the models developed and presented in this paper could be implemented in SAS Event Stream Processing (ESP). For the purpose of illustration, the gradient boost model was exported and implemented as an ASTORE model. ASTORE, indicative of Analytical Store in SAS systems, allows the analyst to save predictive models and subsequently use it to score new data for future prediction. In this paper, the gradient boost model was used to score the transformers for propensity of failure in real time. This provides a dashboard identifying key transformers with a score of 0.80 or greater, displays a map as well as identifies the Feeder (circuit) associated with the transformer (Figure 9).

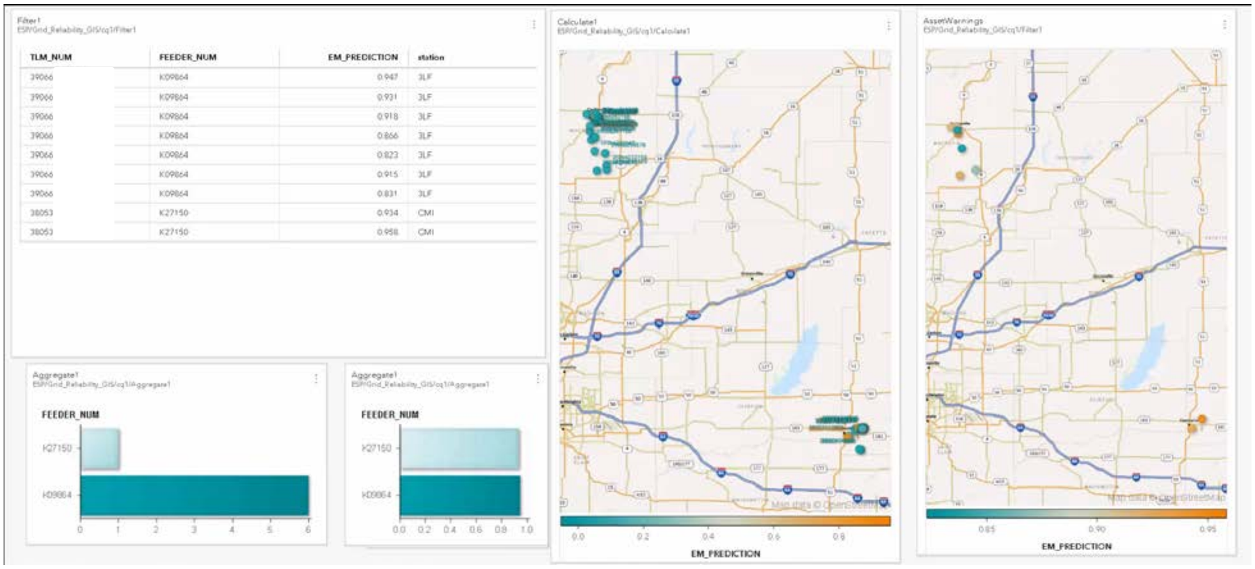

Figure 9: Event Stream Viewer Displaying Output from SAS ESP

CONCLUSION

The methodology presented in this paper not only detects and visualizes outage and congested areas in near real time, but also detects transformers and distribution areas with a long history of outage and congestion. This paper provided examples of predictive algorithms to enhance regular equipment maintenance schedules and reduce repair truck trips for unscheduled maintenance during unplanned incidents like storms. This solution leveraged SAS® Enterprise Guide®, SAS® Enterprise Miner™, SAS® Visual Analytics as well as SAS Event Stream Processing to efficiently produce the information necessary for operational decision-making.

REFERENCES

1. Electric Power Research Institute (EPRI) technical report. “Data Analytics Cases for Asset Awareness,” EPRI, Palo Alto, CA: 2016. 3002008799. Accessed March 25, 2019. www.EPRI.com

2. Electric Power Research Institute (EPRI) technical report. “Steady-State Trend Data Analysis for Application in Predictive Maintenance and Operations,” EPRI, Palo Alto, CA: 2006. 1010200. Accessed March 25, 2019. www.EPRI.com

3. Rashid, Nawzad. 2011. “Short-time Overloading of Power Transformers.” Master Thesis, Royal Institute of Technology, Department of Electrical Engineering. ABB Corporate Research.

4. Tenbohlen S., Stirl T. and Stach M. "Assessment of overload capacity of power transformers by on-line monitoring systems." 2001 IEEE Power Engineering Society

Winter Meeting. Conference Proceedings (Cat. No.01CH37194), Columbus, OH, USA, 2001, pp. 329-334 vol.1.

5. Montsinger V. M. and Dann W. M. "Overloading of power transformers." Electrical Engineering, vol. 53, no. 10, pp. 1353-1355, Oct. 1934.

6. Alquthami T. and Meliopoulos A. P. "Distribution transformer load management system." 2014 IEEE PES General Meeting, Conference & Exposition, National Harbor, MD, 2014, pp. 1-5.

7. Jardini, J. A., Schmidt H. P., Tahan, C. M. V., de Oliveira, C. B., and Ahn, S. U. "Distribution transformer loss of life evaluation: A novel approach based on daily load profiles." IEEE Transactions on Power Delivery, vol. 15, no. 1, pp 361-366, January 2000.

8. Jardini, J. A., Tahan, C. M. V. , Gouvea, M. R. , Ahn, S. U. and Figueiredo, F. M. "Daily load profiles for residential, commercial and industrial low voltage consumers." IEEE Transaction on Power Delivery, vol. 15, no. 1, pp 375-380, January 2000.

9. Leal, Adriano, Jardini, J.A., Magrini, Luiz, Ahn, S.U., and Battani, D. "Management system of distribution transformer loading." Proceeding of the Transmission and Distribution 2002, São Paulo, Brazil.

10. Su, C., Pu, Y., Ching, H., Kuo C. and Kuo, J. "Design of transformer load monitoring systems by utilizing smart meter data." 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), Edinburgh, 2017, pp. 88-92.

11. SAS Visual Analytics User's Guide. Accessed March 25, 2019. http://support.sas.com/documentation/cdl/en/vaug/68648/PDF/default/vaug.pdf

12. Anderson, T. and DeBose T. "Transformer Load Management", SAS and CenterPoint Energy, DistribuTECH 2013

13. SAS® Visual Data Mining and Machine Learning 8.3: User's Guide. Accessed March 25, 2019. http://documentation.sas.com/api/docsets/vdmmlug/8.3/content/vdmmlug.pdf?locale=en

14. Kabir, F., Foggo, B., and Yu, N. (2018). Data Driven Predictive Maintenance of Distribution Transformers. 312-316.