Paper 232-2013

# Weathering the Storm: Using Predictive Analytics to Minimize Utility Outages

Kathy Ball, SAS® Institute; Mark Konya, Ameren Missouri

## ABSTRACT

In order to meet ever-increasing customer service expectations, electric utilities must continuously improve the reliability of their electric distribution systems. Over the last decade, utilities have invested in digital technologies that give them near-real time readings on the health of their electric grid. This data is incredibly useful during major storm outages, but the flood of data pouring in from transformers and meters can quickly overwhelm even the most seasoned Distribution Engineer. Without analytical technologies, engineers cannot respond quickly enough to prevent additional outages and improve restoration times.

Distribution Optimization for storm management equips utility engineers and dispatchers to predict which assets will be affected by storms while optimizing the placement of crews, thus decreasing outage restoration times. Combining geospatial visualization with predictive analytics, the predictive enterprise utility can shorten outages from weather events and identify weak points in the electrical distribution system thus preventing future outages.

## INTRODUCTION

The days immediately preceding a storm are a critical decision period for an electric utility distribution company. The actions a company takes with respect to monitoring weather forecasts for an upcoming storm, predicting damage from that event according to its Emergency Response Plans (ERPs), and obtaining joint service agreements to obtain work crews will determine how well a utility can respond to customer outages when a storm hits. Then, a utility meets its second critical obstacle of determining the best method to restore power to its customers given safety, crew, and system constraints. As seen in Figures 1 and 2, storm restoration conditions can be perilous given weather and system conditions.

In this paper we discuss pre-storm and post-storm operations where predictive analytics can assist electrical distribution companies in restoring power to customers given constrained conditions. This paper uses simulated data for research purposes to improve storm outage and restoration methodologies. During pre-storm preparations this paper will discuss methods for analyzing weaknesses in the distribution grid via reliability analyses for potential equipment failure potential, and storm forecasting methodologies as it relates to Event Level Classifications for anticipated number of outages, crew requirements, and full system restoration time. For post-storm, this paper will focus on work crew optimization routing to restore power to customers in the least amount of time and trucks travelling the least distance based on optimization modeling techniques.

**Figure 1, Hurricane Irene Photo**

**(Source: Accuweather.com)**

**Figure2, Hurricane Irene Photo**

**(Source: www.processensors.com)**

For analysis purposes, this paper will be using storm track information from Hurricane Irene as it impacted the state of Massachusetts in August, 2011. Data has been simulated for customers in Massachusetts and made to correspond with the actual number of outages reported by day post emergency power restoration efforts (9/1/2011). This paper will outline predictive analytical strategies and processes a utility can use for storm optimization minimizing outages. Prior to the storm, minimizing outages can be accomplished by applying an analytical approach that takes into account the health of the assets in a system combined with upcoming weather events. In this paper, we analyzed transformer assets and its many characteristics to determine a scorecard as to the health of all the assets in the network and developed a potential failure model. Post storm-outages can be minimized via optimal crew routing using multiple optimization techniques that resulted in solutions being developed based on the skill sets of the crews, distances to be travelled, and work force limitations.

## PREPARING FOR THE STORM

Once a major storm has been identified, a utility starts analyzing the impact of the storm on its service territory immediately. A utility must determine an Event Level Classification which projects the expected severity of customer outages and expected damage to a distribution system. In determining this impact, utilities must review distribution information contained in multiple systems such as SCADA, OMS, and other customer and equipment related databases. In this paper, we added weather information such as rainfall potential and wind and combined this with system data to project potential service outages with information from sources such as Figures 3 and 4. Combining weather data with inventory data can be cumbersome in an organization where most data is contained in silos. Utilities operate with limited data regarding equipment health in the field. Distribution data contained in SCADA provides more efficient communications from field equipment; however it is still not enough in an asset-intensive industry faced with aging equipment and increasing damage from storm events.

When traditional analytics are combined with advanced visual analytics, a utility can see the expected assets impacted based on a given storm forecast path and develop a wider situational awareness to detect system weakness in real time across the grid. Combined with historical and projected system weakness areas, this holistic information can help a utility assess the projected damage assessment correctly and assigned the correct Event Level Classification for service restoration. As we delve further into the predictive side of storm preparation and assessment, we will be reviewing the impact of variables impacting equipment failures due to storms as well as the correlation of other variables relating to system and overall historical system performance.


**Figure 3, Storm projection for Hurricane Irene 2011**

**(Source: Accuweather.com)**

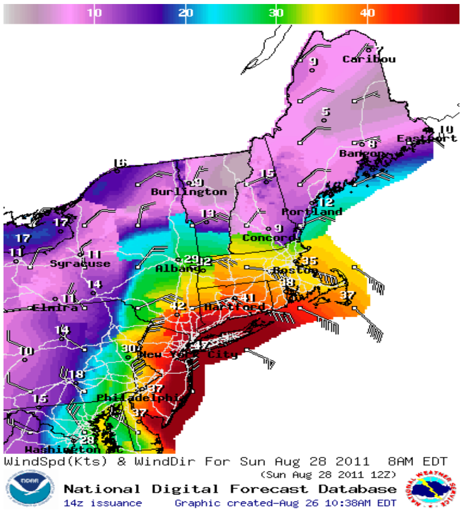



**Figure 4, Hurricane Irene Wind Speed**

**(Source: NOAA Archives)**

Utilities can use predictive and visual analytics to help determine the impact on critical assets in the distribution system. At a simplistic level, a utility can use traditional analytics to determine which assets will fail based on the coordinates of a storm and the number of assets such as transformers in the storm's path. Figure 5 shows the traditional methodology based on the assets within the storm cone of damage based on inventory and DMS information. However, this will not give you the complete picture on an assets health. Using SAS Visual Analytics Explorer, a distribution worker can quickly determine assets impacted in the storm from historical and predictive measures in real-time and can adjust ad-hoc to the constantly changing conditions associated with a storm as seen in Figure 6. Advanced analytics will produce results such as geographic outage information, variable correlations for transformer failure as it relates impact from storm conditions combined with other equipment information. The results from the correlation analysis and other data visualizations such as forecasting equipment failure can help a distribution engineer model the potential for failure by eliminating variables with no impact on failure prediction and build a richer model yielding greater reliability analysis.


**Figure 5, Traditional Asset Impact Method**

(Source: ESRI Storm Archives)

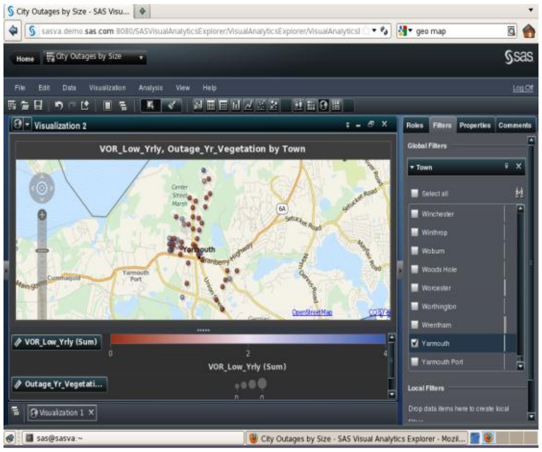



**Figure 6, Advanced Analytical Impact Method**

(Source: SAS® Visual Analytics)

## STORM PREPARATION ANALYTICS: PREDICTING OUTAGES

Storms are inevitable and difficult to predict. Utilities can improve their storm performance by using traditional analytics combined with predictive analytics forecasting outages and equipment failures. Based on storm track alone, the utility can analytically define a cone of storm impact and identify the assets affected. In addition, a utility can use predictive analytics to score the probability of asset failure based on equipment data contained in SCADA, OMS, and DMS systems. Then, by integrating data from social media, a utility can better pinpoint outages in real-time from sources such as Twitter and can incorporate this information into real time forecasting including optimization analytics. In Figure 7, real-time outage information is viewable in an analytical format.

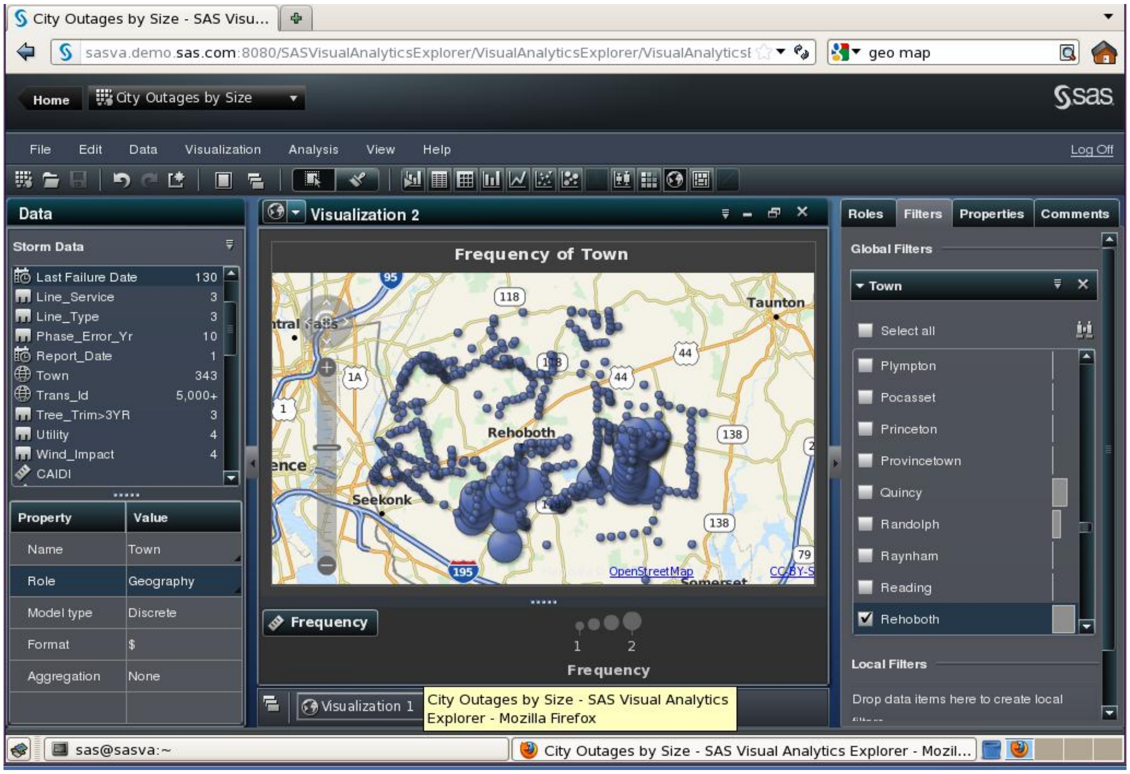



**Figure 7, Outages by City during Hurricane Irene 2011 (OMS and Social Media Information)**

(Source: SAS® Visual Analytics)

## METHODOLOGY

### STEP ONE: NARROWING DOWN THE VARIABLES AFFECTING ASSET FAILURE

Beginning with data extraction and integration, a base table (BT) is established combining asset information, previous outage data, storm track information including wind speed and direction, and system configuration information. The BT is a structure of data suitable for predictive modeling and consolidates information based on common data points. Once the data was available as a BT, we built predictive models using a powerful graphical interface that enabled us to perform variable correlation analyses simplifying the tasks associated with data mining. In-memory analytics tools within Visual analytics give a user the capability to do complex correlation matrices to show storm variable relationships interactively.

Starting with traditional analytics, the primary technique used in determining the assets affected by the storm is a SAS® graphing technique known as PROC GINSIDE. Results can be seen in Figure 8. This technique helps you determine which assets will be affected based on geography alone. The procedure determines which points fall outside of a predetermined location identified in this paper as the hurricane zone. Each individual transformer is mapped and only those transformers within the polygon will appear as a subset as being potentially affected by the storm. This narrows down the master set of data to only those likely to be impacted. In this particular example, four thousand transformers were predicted to fail as a result of the storm as opposed to six thousand transformers simply located within the storm cone.

Program ▾

Program* | Log | Output Data (3)

OUTPTS ▾

Filter and Sort | Query Builder | Data ▾ | Describe ▾ | Graph ▾ | Analyze ▾ | Export ▾ | Send To ▾

| | x | y | _ONBORDER_ | id | Utility | Town | Latitude | Longitude |
|---|---|---|---|---|---|---|---|---|
| 1 | -71.44040108 | 42.48513726 | 0 | InPath | Nstar | Acton | 42.48513726 | -71.44040108 |
| 2 | -71.43799782 | 42.48463089 | 0 | InPath | Nstar | Acton | 42.48463089 | -71.43799782 |
| 3 | -71.43799782 | 42.48554868 | 0 | InPath | Nstar | Acton | 42.48554868 | -71.43799782 |
| 4 | -71.43958569 | 42.48561198 | 0 | InPath | Nstar | Acton | 42.48561198 | -71.43958569 |
| 5 | -71.437397 | 42.48842857 | 0 | InPath | Nstar | Acton | 42.48842857 | -71.437397 |
| 6 | -71.43495083 | 42.48937795 | 0 | InPath | Nstar | Acton | 42.48937795 | -71.43495083 |
| 7 | -71.43748283 | 42.48944124 | 0 | InPath | Nstar | Acton | 42.48944124 | -71.43748283 |
| 8 | -71.43589497 | 42.48991593 | 0 | InPath | Nstar | Acton | 42.48991593 | -71.43589497 |
| 9 | -71.43190384 | 42.48906149 | 0 | InPath | Nstar | Acton | 42.48906149 | -71.43190384 |
| 10 | -71.43477917 | 42.48874503 | 0 | InPath | Nstar | Acton | 42.48874503 | -71.43477917 |
| 11 | -71.43834114 | 42.48592845 | 0 | InPath | Nstar | Acton | 42.48592845 | -71.43834114 |
| 12 | -71.43589497 | 42.4881121 | 0 | InPath | Nstar | Acton | 42.4881121 | -71.43589497 |
| 13 | -71.45396233 | 42.49323863 | 0 | InPath | Nstar | Acton | 42.49323863 | -71.45396233 |
| 14 | -71.43924236 | 42.48874503 | 0 | InPath | Nstar | Acton | 42.48874503 | -71.43924236 |
| 15 | -71.43280506 | 42.48991593 | 0 | InPath | Nstar | Acton | 42.48991593 | -71.43280506 |

**Figure 8, Predicted Transformer Failure Points in Path of Storm using PROC GINSIDE**
**(Source: SAS®)**

Next step is to add the system health information relating to transformer failure and weather conditions. Combining data from multiple systems creates a cumbersome dataset. Oftentimes utilities have a deluge of data and need to narrow down data while not excluding some of the less obvious variables leading to failure. SAS Visual Analytics provides a quick way to do this while sorting through billions of rows of information with hundreds of variables (Figure 9). SAS Visual Analytics speeds up the analytical lifecycle by quickly identifying significant correlations in data.

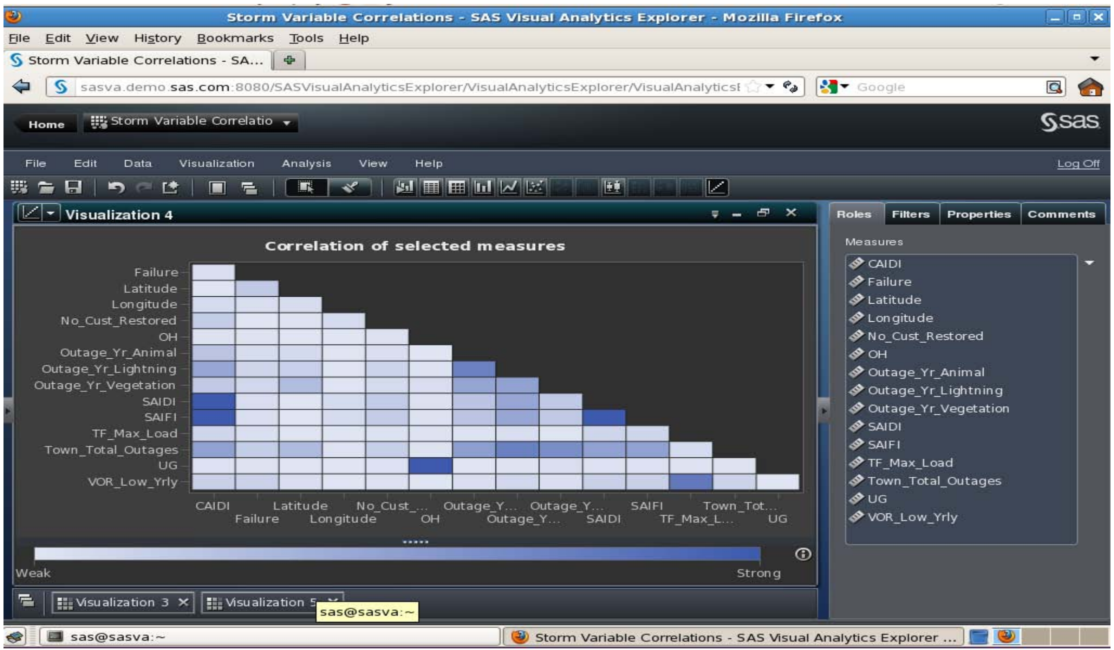



**Figure 9, Correlation Analysis of Equipment Health Information for Outage Prediction**

**(Source: SAS® Visual Analytics)**

Further explorations of the data were conducted in conjunction with the correlation matrix to examine the interplay of the variables selected. Regression analyses were then conducted on multiple variables from the DMS system as seen in Figure 10.

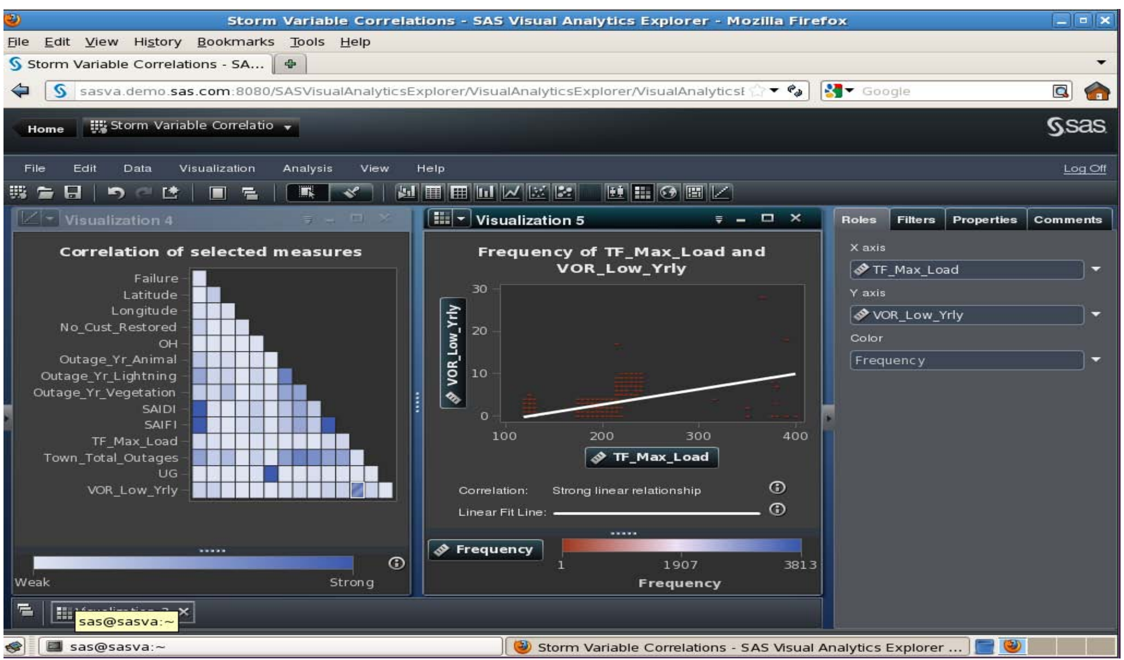



**Figure 10, Correlation Analysis combined with Variable-Specific Analysis**

**(Source: SAS® Visual Analytics)**

**STEP TWO: SURVIVAL ANALYSES MODELING ASSET POTENTIAL STORM FAILURE**

In this paper, we were able to quickly sort through variable correlations from the matrix that ten variables had the greatest correlations for modeling equipment failure for storm impact. From here, we took this information and developed a Reliability Analysis in SAS® Enterprise Guide using Rapid Predictive Modeler. A model competition was created using the advanced feature. These models were then fed into SAS® Enterprise Miner for scoring and deeper analyses. Survival analyses, also known as Reliability Analyses in engineering, apply supervised techniques like Regression Analyses, Decision Trees, and Neural Networks. From this stage of the analysis, we combined this information with the predictive analytics associated with asset failure using survival analysis modeling. Survival analysis involves the modeling of time to event data; in this context, transformer failure. The failure is considered an "event" in the survival analysis whereby a single event occurs and the asset is broken. Figure 11 displays output from the advanced models in Rapid Predictive Modeler predicting variables contributing to an outage.

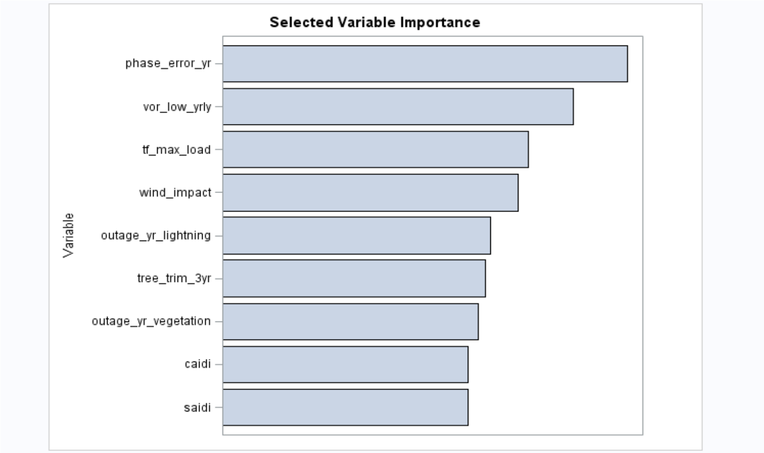



**Figure 11, Survival Analysis using Advanced Modeling Techniques**

**(Source: SAS® Enterprise Guide)**

**STEP THREE: OUTAGE PREDICTION MODEL SCORING**

The process of applying a predictive model to a set of data is referred to as scoring the data. The scoring process consists of two basic steps with the outcomes displayed below:

- Build the model and save the file using the dataset for the desired outcome which is failure known in this example as a transformer outage.
- Apply that model to validate and test dataset for the same transformer outage to obtain predicted outcomes. Results of this task are seen in Figures 12 and 13.

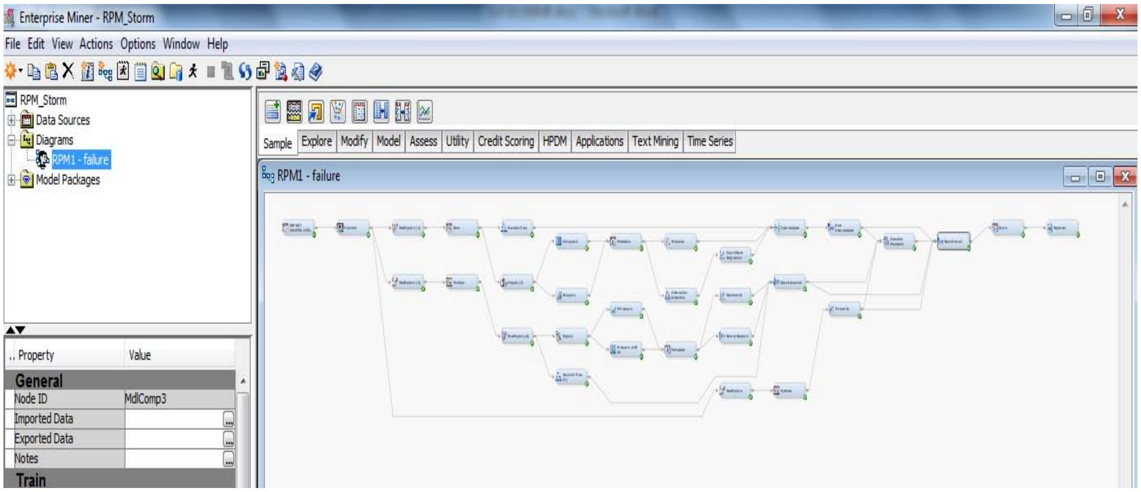
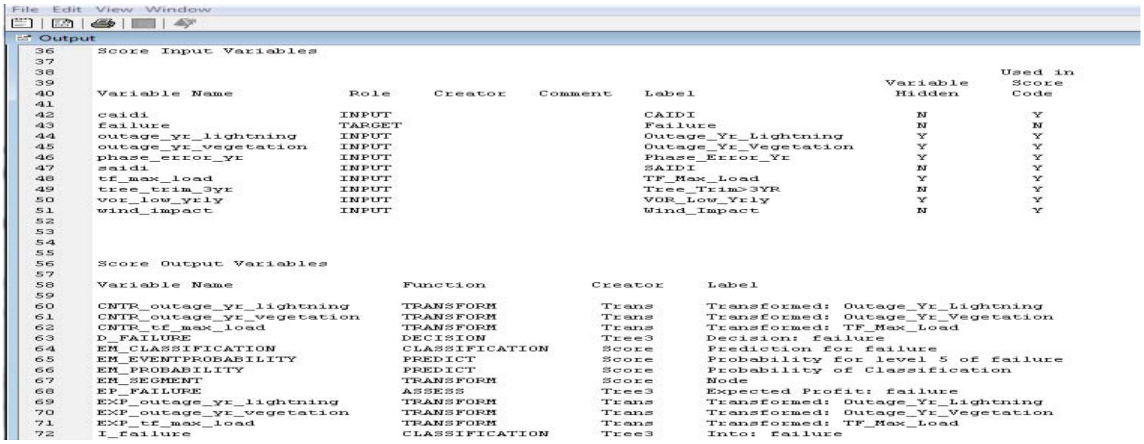



**Figure 12, Advanced Modeling Techniques using Enterprise Miner**

**(Source: SAS® Enterprise Miner)**



Score Input Variables

| Variable Name | Role | Creator | Comment | Label | Variable Hidden | Used in Score Code |
|---|---|---|---|---|---|---|
| caidi | INPUT | | | CAIDI | N | Y |
| failure | TARGET | | | Failure | N | N |
| outage_yr_lightning | INPUT | | | Outage_Yr_Lightning | Y | Y |
| outage_yr_vegetation | INPUT | | | Outage_Yr_Vegetation | Y | Y |
| phase_error_yr | INPUT | | | Phase_Error_Yr | Y | Y |
| saidi | INPUT | | | SAIDI | N | Y |
| tf_max_load | INPUT | | | TF_Max_Load | Y | Y |
| tree_trim_3yr | INPUT | | | Tree_Trim>3YR | N | Y |
| vor_low_yrly | INPUT | | | VOR_Low_Yrly | Y | Y |
| wind_impact | INPUT | | | Wind_Impact | N | Y |

Score Output Variables

| Variable Name | Function | Creator | Label |
|---|---|---|---|
| CNTR_outage_yr_lightning | TRANSFORM | Trans | Transformed: Outage_Yr_Lightning |
| CNTR_outage_yr_vegetation | TRANSFORM | Trans | Transformed: Outage_Yr_Vegetation |
| CNTR_tf_max_load | TRANSFORM | Trans | Transformed: TF_Max_Load |
| D_FAILURE | DECISION | Tree3 | Decision: failure |
| EM_CLASSIFICATION | CLASSIFICATION | Score | Prediction for failure |
| EM_EVENTPROBABILITY | PREDICT | Score | Probability for level 5 of failure |
| EM_PROBABILITY | PREDICT | Score | Probability of Classification |
| EM_SEGMENT | TRANSFORM | Score | Node |
| EP_FAILURE | ASSESS | Tree3 | Expected Profit: failure |
| EXP_outage_yr_lightning | TRANSFORM | Trans | Transformed: Outage_Yr_Lightning |
| EXP_outage_yr_vegetation | TRANSFORM | Trans | Transformed: Outage_Yr_Vegetation |
| EXP_tf_max_load | TRANSFORM | Trans | Transformed: TF_Max_Load |
| I_failure | CLASSIFICATION | Tree3 | Into: failure |

**Figure 13, Model Scoring**

**(Source: SAS® Enterprise Miner)**

**RESULTS FROM PREDICTIVE METHODOLOGY**

Catastrophic failures are common during major storm events. Predictive and visual analytics play a key role in pinpointing transformer outages in a proactive versus reactive manner. The methods used in this paper such as storm transformer outage prediction combined with geospatial modeling helps a utility better plan for assistance, materials, and potential storm impact. While the accuracy results in this simulated dataset were higher than should be expected at 98% over existing techniques, additional studies with real storm data could be beneficial for all. This could lead to timelier outage restoration, improved system reliability, and better communications with governmental and regulatory bodies.

## WEATHERING THE STORM: AFTER THE STORM ANALYTICS

After a storm has passed and the first few days of emergency repairs are complete, a utility is faced with the daunting task of restoring power to multiple customers given work crew and system constraints. In this paper, we will examine the optimal method for restoring power using an enhancement to an optimization technique known as Capacitated Vehicle Routing Problem using SAS® OR.

The objective is to restore the most customers while the work crews are covering the least distance given skill set, labor, and vehicle constraints. Each work crew must start and finish at an Operations Center and can only work a fixed number of hours a day. Additionally, we have added skill set constraints such that Underground repair crews can restore overhead and underground outages while overhead crews can only restore overhead outages. The final algorithm creates a new customer power restoration technique for utilities. The new modeling technique can accommodate changes to constraints regarding outage restoration times to meet restoration goals.

Each crew will restore one transformer outage before moving to the next and can only visit the outage site one time. Repair time can vary per transformer. The objective of the vehicle routing is to restore the most number of customers while minimizing drive time and accommodating other constraints. The route is the total number of outage restorations made by a single Utility work crew. Crews are divided into Overhead and Underground crews and each crew is limited to 16 work hours per day inclusive of drive time. These are the primary constraints used in the enhanced capacitated vehicle routing problem whose objective is to restore power to the most customers the quickest while minimizing customer outage time reflected in CAIDI measures.

Three days after Hurricane Irene, after the initial emergency restorations, the state of Massachusetts was faced with over 65,000 customer outages and over 10,000 transformer failures. At this stage of the storm restorations efforts, the simple repairs restoring thousands of customers have been completed as well as those to critical assets such as hospitals, jails, emergency call centers and other critical services. Outages are now more difficult than getting a feeder back to service which restores multiple customers quickly. This stage of the storm can be improved heavily through analytics via Optimization algorithms including but not limited to Mixed Integer Linear Programming (MILP) using a modified Capacitated Vehicle Routing Problem with Simultaneous Pickup and Delivery (CVRPSDP). Distance-based clustering techniques are ineffective in this problem since the goal is to first optimize customers restored and then distance which most clustering techniques begin with.

### STEP ONE: TRAVEL TIME CALCULATION FOR MODEL PREPARATION

Reducing the number of potential solutions is one of the key drivers for any Optimization engine else unlimited solutions could stymy or prevent model convergence. Distances must be calculated between all transformer outage locations to calculate drive times for the purposes of this paper. Actual driving times were calculated for the base case but the distance variance between actual and Euclidian distances was negligible. Distances were calculated and later put into a customer restored minutes per mile metric.

### STEP TWO: MODIFY MILP FRAMEWORK FOR CUSTOMER RESTORATION CONSTRAINTS

Since the majority of the VRP problems are distance focused, changes to the basic VSP Sub problem must be made so that the Repair Crews are targeting the highest number of customers restored at an outage location first. In order to account for the skill set constraint, the skill set available is greater to or equal to the skill set needed. Underground Repair Crews have the ability to service Overhead and Underground Outage locations. Conversely, Overhead Repair Crews can only restore Overhead outage locations.

### STEP THREE: MILP SOLVER TO CREATE OPTIMAL SOLUTIONS VERSUS STANDARD UTILITY ROUTING

When using solvers such as Mixed Integer Linear Programming (MILP) s, dynamic work crew routing is determined and graphed. The MILP solver contains the OPTMODEL procedure, implements an LP-based branch-and-bound algorithm. The model solver attempts to solve the original problem by solving linear programming relaxations of a sequence of smaller sub problems. The MILP solver also implements advanced techniques such as generating cutting planes and applying heuristics to improve the efficiency of the overall algorithm. Results from an optimization routine for outages covered in one day in a trial city are shown in in Figure 14. The outage not covered exceeds shift limitations.

The optimal paths from OPTMODEL were compared with the current utility routing method of routing via highest number of customer restored per outage site. The results were varied across the simulated outage and transformer locations in our Massachusetts model. Cities that contained multiple outages in close proximity had fewer differences that the standard utility model. However, in cities where outages were dispersed and skill constraints were a factor, outage restoration times were improved for this dataset.

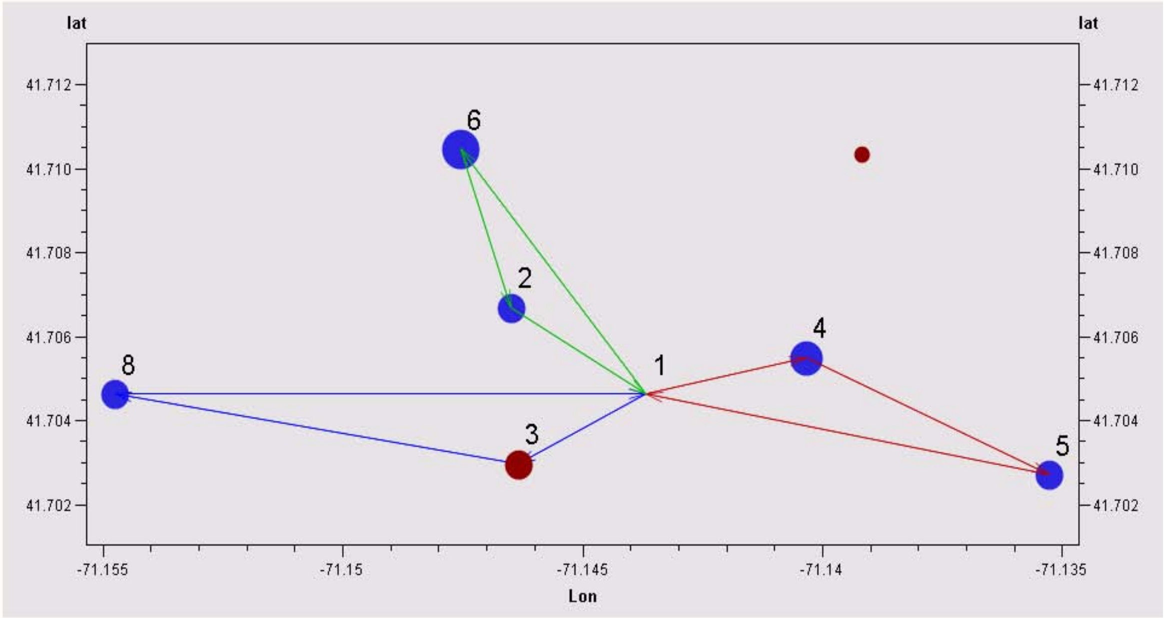



**Figure 14, Constrained Customer Restoration Problem – Repair Crews and Logistics Problem**

**(Source: SAS® OR)**

### RESULTS

Results from the Constrained Customer Restoration Optimization varied in this simulated dataset from no change to a 22% improvement in outage restoration time for optimized routes from non-optimized routes. On average, towns were able to have customers restored 13.8% faster than the present heuristic restoration technique. Variables contributing to the greatest gains include distances between outages, number of underground outages, with crew and vehicle capabilities playing a role as well. This technique could certainly add value to any utility faced with storm and emergency response efforts to restore power and services to its customers.

## CONCLUSION

A utility has many analytical tools it can use to improve its storm performance before and after a storm hits. By better predicting which assets are likely to fail, Resource Planners in Distribution Engineering can optimize the use of preparatory resources, better estimate the number of crews needed and secure Joint Agreement crews earlier if necessary. The planner may also be able to pre-stage materials and Vegetation Management crews in advance of the storm impact.

Predictive and visual analytics are great tools to assess how many people will likely be without power as a result of one piece of equipment damage such as a transformer. Combined with locational information and real-time visual analytics, utilities can focus on restoring power faster by making data-driven decisions. Finally, being able to dynamically route crews or schedule parallel activities enables the next most important job to be prioritized. In a dynamic environment where storm paths often change, advanced analytics can make sense of the data deluge to keep a utility aligned with its restoration goals.

## REFERENCES

Massachusetts Government information from the Department of Public Utilities, Storm Orders regarding Hurricane Irene. Available at http://www.mass.gov/eea/grants-and-tech-assistance/guidance-technical-assistance/agencies-and-divisions/dpu/storm-orders.html

"S.P. Anbuudayasankar, Amrita University, K. Ganesh, K. Mohandas, Amrita University, " Mixed Integer Linear Programming for Vehicle Routing Problem with Simultaneous Delivery and Pick-Up with Maximum Route-Length", International Journal of Applied Management and Technology, Vol. 6, Number 1.

## ACKNOWLEDGMENTS

Albert Hopping, SAS® Institute, Operations Research Optimization Modeling

Jim Duarte, SAS® Institute, Survival Analysis Modeling

## RECOMMENDED READING

- *SAS® OR ® 9.22 User's Guide: Local Search Optimization (Genetic Algorithms)*
- *SAS® OR ® 9.22 User's Guide: Mathematical Programming (PROC OPTMODEL)*
- *SAS® Mixed-Integer Linear Programming Solver*
- *Batch Production of Driving Distances and Times Using SAS® and Web Map APIs, Ash Roy and Yingbo Na, SAS Global Forum 2012, Coder's Corner*