Abstract
This paper presents an approach that uses statistical techniques to analyze and interpret Time of Outage Restoration (TOR) data in distribution systems with respect to several selected factors. The factors are classified into three main categories: (1) Time, (2) Consequence, and (3) Environment. Each category is further broken down into groups and classes for more detailed analysis. Both statistical, quantitative results and field engineering insights will be correlated and discussed. The approach taken in this paper can be easily modified for other utilities' TOR analysis. These analyses can provide a better understanding as to how the restoration time is affected by different influential factors. The methodology and general discussion in the paper will be useful for future research and development of the Time of Outage Restoration issues in distribution systems, such as model development for customer service improvement, distribution operations, and planning.

Key Words: distribution systems, outages, restoration time, statistical analyses.

I. Introduction
One of the primary missions of power companies is to provide a reliable, economical and safe supply of electricity to the customers. With the growing usage of electricity by the public, the complexity of power system operations is increasing at a fast rate. In order to achieve its mission with a high standard, power systems nowadays rely heavily on the refinement of conventional models, techniques and experience, extensive usage of computers, and the application of high-tech, emerging technologies, along with research in various other areas, in order to meet its current and future challenges [1, 2].

Power distribution system operations require a major amount of man power and resources to properly maintain the systems' daily functions. These operations are highly unstructured, nonlinear, stochastic and time varying (compared to generation systems and transmission systems). In addition, distribution systems cover larger physical areas, thus increasing the chances and the number of unexpected events that can occur to these systems. The large exposure of the distribution systems to outdoor environments makes them difficult to analyze, model and, at the same time, more vulnerable to natural disturbances.

Distribution faults have a large impact on power systems' reliability, security and quality, among other important factors.

Most utility companies experience a large number of distribution faults in their systems. Customers (industrial, commercial, and residential) are very concerned about the duration of the outages because outages can directly affect their economic revenue, production rate, safety, and convenience. Therefore, when an outage occurs, one of the most common questions that customers may ask is: How long will it take to restore the power? We term this as the Time of Outage Restoration (TOR) problem in this article. Currently, when an outage occurs and customers call in to report the outage, they usually request for an estimate of the TOR. In general, operators will make an educated guess about the restoration time based on their previous experiences and other available information. This approach does not objectively and efficiently utilize available information and data, and may lead to a large error variance between the estimate and the actual outage duration.

If we are able to provide a reasonable estimate of the TOR, then customers will be much happier because they can make contingency plans to minimize the negative effects of the outage on their side. With a better understanding of the effects of different influential factors on TOR, we certainly can provide a better estimate of the TOR. However, to provide an objective and accurate TOR estimate is a challenging task because it depends on so many different factors, such as weather conditions, time of occurrence, the fault causes, the availability of manpower, the number of outages need to be restored at that time, the location of the outage, etc. To further complicate the TOR analysis, different influential factors may have direct and indirect effects on the restoration time. For example, a stormy weather may cause more lightning strike faults. It is very difficult to isolate only direct influential factors for TOR analysis. It is unrealistic for this paper to investigate and report all possible factors that can influence TOR. However, we can develop and present a general approach to correctly extract factors that influence TOR for further analysis, so that the TOR estimate can be improved.

In addition to providing better customer services, the utility can also be benefited with better understanding of Time of Outage Restoration under different scenarios. For example, the crew assignment and outage restoration process can be scheduled more efficiently and effectively, thus increasing the reliability of the distribution systems and reducing the operation costs.

Some utility companies have been collecting data related with outage restoration time. Outage data is recorded on the Duke Power Distribution System each time that an outage occurs as a result of the operation of a breaker relay, fuse, or another protective device. The recorded data contains information such as the geographical location, circuit ID number, weather condition, time of the fault occurrence, and the time of power restoration. Twenty-two information items...
are recorded for each outage at Duke Power Company. Some of
the information is based on a coding system, designed by
Duke Power engineers and summarized in charts, that describe
the circumstances under which the outages occurred (e.g.,
weather condition, season of the year, time of the day). Other
information items are based on data collected by equipment
which monitor the operation of the distribution system (e.g.,
number of phases affected, protective devices used, etc.).
Once an outage is detected, crewmen are dispatched to restore/repair
the system as well as to find the fault causes, if possible [3-6]. Duke Power has accumulated a huge database
on outage information [7]. However, data is meaningless
until they are interpreted correctly. There are many statistical
tools that can be used to analyze this data. However, the
statistics may lead to misleading results if interpreted
carelessly because the actual problem may not be as simple as
it appears to be.

In this paper, we will present an approach that uses
statistical techniques to analyze and interpret the TOR data
with respect to several factors such as time, weather
condition, etc., and correlate the results with actual field
experience in distribution systems. Outage data for the
Durham Region in Duke Power System will be used in this
paper for illustration purposes. Knowledge about distribution
systems can help explain the statistical results, and the
statistical results can provide quantitative information about
the particular distribution system of interest. This type of
quantitative information is very useful for improvement in
customer service and different distribution system operations
such as maintenance, construction and design, priority
setting, and resource allocation. The approach taken in this
paper can be easily modified for other utilities' TOR analysis.
These analyses can provide a better understanding as to how
the outage restoration time is affected by different influential
factors. This knowledge can be very helpful for future
improvements in customer services, distribution operations
and planning.

II. Selected Factors and Statistical Tools for
the Time of Outage Restoration Analysis

This paper focuses on several factors for TOR analysis and
the factors are categorized into three main categories: (1)
Time, (2) Consequence, and (3) External. Each category is
further broken down into groups and classes for more detailed
analysis. The general categories under investigation in this
paper are schematically shown in Figure 1. (Note that there
are other factors, such as fault location effects on ETOR, can
be analyzed using similar approaches proposed in this paper).

Many statistical tools are available for data analyses. We
use graphical displays of TOR data to facilitate our discussion
because "one picture is worth a thousand words". In order
to illustrate the statistical measures used in this paper, Figure 2
shows a typical histogram of TOR, with the quantile box
drawn above the histogram. The median is shown as a line
across the middle of the box, and the 25th and 75th quantiles
are at the ends. The quantile boxes provide a summary of the
distribution's shape. A mean diamond is also shown as a
schematic that illustrates a sample mean and a 95% confidence
interval. The line across each diamond represents
the group mean. The width of each diamond represents the
95% confidence interval for each group. The diamond is
very useful to examine whether the mean for each class under
comparison is statistically different.

The histogram in Figure 2 suggests that TOR is not
normally distributed, as confirmed by a test for normality.
The Kolmogorov-Smirnov statistic used for the normality
test has a value of $D = 0.128$ [8], and the corresponding $p$-
value is less than 0.01. (Note that $p$-value is the probability
of obtaining a more extreme value of $D$ based on factors such
as the $\chi^2$ value and degree of freedom (DF) of the problem.)
The $p$-value obtained here indicates that TOR does not follow
a normal distribution. We thus cannot use the conventional
ANOVA technique to test for the equality of the mean or
median of the TOR at different classes for the selected factors
due to the violation of the normality assumption. In this
paper, we use the Nonparametric One Way Analysis of
Variance technique to analyze the TOR with respect to each of
the selected factors. This procedure performs analysis of
variance on ranks, is distribution free, and the assumption
of normality is not made. In particular, the Kruskal-Wallis test
is conducted here to test whether the Group medians are the
same across Groups of the selected factors. For a more
detailed description of Kruskal-Wallis method, please refer to
[8].

The $\chi^2$ approximation to Kruskal-Wallis test, together
with its $p$-value, will be given in the following sections
when we discuss the inferences about TOR for each specific
factor. A small $p$-value for the Kruskal-Wallis test indicates
that the responses are different across the Group for the factor.
On the other hand, a large $p$-value indicates that the responses
are the same across the Group for the factor. The usual cutoff
value for $p$-values is either 0.05 or 0.1. The results of the $p$
values for our data for each factor are clearly much larger or
smaller than the usual cutoff, thus the p-values provide a clear indication for our conclusions.

Besides determining the influential factors of TOR based on the p-values of the Kruskal-Wallis test, we also rank the importance of each factor for the TOR by $R^2$ measure, which ranges between 0 and 1. $R^2$ measures the proportionate reduction of total variation in the response variable associated with the use of independent variables, which represent each specific cause we consider in this paper. Higher $R^2$ indicates that the cause is more influential.

III. Non-Parametric One-Way Variance Analysis of the Time of Outage Restoration

In order to illustrate the approach of using statistical analyses for power distribution's outage restoration time analysis, the data set containing 4,323 outage records collected from Durham, North Carolina between 1992 and 1993 is used in this paper. The same methodology can be easily modified and applied to different distribution outage data. Note that the maximum TOR shown is 500 minutes (8 hours and 20 minutes), because outages which take more than 500 minutes are treated as abnormal situations (outliers in statistical terms) in this analysis and are disregarded in this TOR analysis.

A. Time Factors

The effect of the time of outage occurrence is first investigated. For a more detailed investigation, we further break down the time factor into three qualitative descriptions (groups): (1) time of the day, (2) day of the week, and (3) month.

1. TOR by Time of the Day

In this study, the time of the day was divided into twelve time zones (classes): (1) midnight - 2am, (2) 2am - 4am, (3) 4am - 6am, (4) 6am - 8am, (5) 8am - 10am, (6) 10am - noon, (7) noon - 2pm, (8) 2pm - 4pm, (9) 4pm - 6pm, (10) 6pm - 8pm, (11) 8pm - 10pm, (12) 10pm - midnight. The TOR distribution with respect to the twelve classes is graphically depicted in Figure 3. The plot is similar to the histogram shown in Figure 2 except that the count of the TOR observation is now shown as the density of the dot in the figure. In addition to the graphical display of the data, the means and variances are graphically displayed by the mean diamond boxes, as described in the previous paragraph. To test whether the median for TOR are the same for all different time zones, the Kruskal-Wallis test is performed. Its $\chi^2$ approximation has a value of 302.7535 with $DF = 11$ and the $p$-value = 0.0000, which indicates that time of the day is statistically significant to the TOR.

Figure 3 clearly indicates that time of the day is an important factor for the TOR estimate. The statistical results agree with our intuition. For example, during normal working hours, there are more crewmen on duty and they are stationed either in the operation centers or in the field. They can respond faster to outage calls (outages usually have the first priority in distribution operations). During off-hour, there are less crewmen on duty and they are stationed in their homes on-call, so the response time will certainly be longer than during normal working hours. In summary, the TOR is inversely proportional to the availability and readiness of crewmen.
in order to explain the discrepancy between the statistical results and our intuition.

3. **TOR by Month**

The Time category is also grouped into Months: (1) January, (2) February, (3) March, (4) April, (5) May, (6) June, (7) July, (8) August, (9) September, (10) October, (11) November, and (12) December. The TOR with respect to Month is graphically depicted in Figure 5. The $\chi^2$ value is 42.5849 with DF = 11 and the $p$-value = 0.0000, which indicates that the month of the outage occurrence is influential to TOR.

**B. Consequence Factors**

1. **TOR by Number of Phase Affected**

Number of Phases Affected is one of the consequences of outage occurrences. The number of phases affected is coded as: (0) single-customer, (1) single-phase, (2) two-phases, and (3) three-phases. The single-customer can be on a single-phase, two-phase, or three-phase line. The outage is classified as a single-customer outage as long as the outage occurs on a service line rather than on the primary line. In this paper, we show the TOR for two regions, Durham and Chapel Hill, to demonstrate the local properties of different regions. Durham, which has approximately 90,000 customers, is a much larger region than Chapel Hill, which has approximately 35,000 customers. The TOR with respect to the number of phases affected in Durham is graphically shown in Figure 6a. The $\chi^2$ value of the number of phases affected with respect to TOR in Durham region is 89.3097 with DF = 3 and the $p$-value = 0.0000. The TOR with respect to the number of phase affected in Chapel Hill region is graphically shown in Figure 6b. The $\chi^2$ value of time of the day with respect to TOR in Chapel Hill region is 88.7600 with DF = 3 and the $p$-value = 0.0000, which indicates that the number of phases affected is also influential for the TOR.
too many two-phase outages in Chapel Hill, thus the statistical estimate of the TOR is not so representative, as indicated by the wide diamond box confidence interval.

In addition, Figure 6b indicates that single-customer outage takes longer time to restore than other types of outages, which is a little surprising. After further investigation, the statistical results actually agree with the actual situations. The TOR can be thought of as both the time required to find the outage location and the time to fix the outage. For single-customer outages, the outage locations can be pin-pointed quickly, thus taking much less time to find the outage location. However, in the Chapel Hill region, many of the single-customer services are underground, while the single-phase, two-phase, and three-phase lines are in general overhead. By and large, it takes much longer time to fix an underground outage than an overhead outage. The TOR for single-customer outages is mainly spent on fixing the outages instead of finding them.

In summary, the TOR is proportional to the area covered by the circuits in terms of search time, and also to the type of distribution lines, whether underground or overhead, in terms of fixing time.

2. TOR by Activated Protection Devices

An outage will activate protection devices such as blow fuses and circuit breakers. This section discusses the TOR with respect to different protection devices which are activated by the outages. The protection devices are coded as: (0) Transmission Device, (1) Station Circuit Breaker, (2) Line Reclosure, (3) Primary Fuse, (4) Transformer Fuse, (5) Transformer CSP/Other, (6) Panel/Meter Base, (7) SEC/SVC Self Clearing, and (8) Manual Device. The TOR with respect to these protection device codes is graphically depicted in Figure 7. The \( \chi^2 \) value is 359.6365 with DF = 8 and the \( p \)-value = 0.0000, which indicates that activated protection devices provide statistically significant information for TOR.

Figure 7 clearly indicates that TOR is a function of the activated protection devices. However, the TOR is not only a function of how easy it is to restore the protection devices. More importantly, TOR is a function of how significant the protection device is to the system. Suppose a thunderstorm caused many protection devices to activate, including Static Circuit Breakers (code 1) and primary fuses (code 3). Physically, it may take less time to restore a primary fuse than a Static Circuit Breaker (such as OCB). However, from the distribution system reliability viewpoint, we want to restore power to most of our customers as soon as possible. Thus we need to set priorities for protection device restorations. The activation of an OCB will affect an entire circuit, say 5,000 customers, while a primary fuse may affect only a section of a feeder, say 10 customers. Therefore, the OCB protection device has a much higher priority than the primary fuses, thus the statistics showed that OCB takes less time to restore than other protection devices. Other than the priority issue of restoration, the actual time required to restore the protection device will also come into the picture. For example, SEC/SVC Self Clearing protection devices (code 7) are usually buried underground, which takes much longer to repair than an overhead fuse failure. In summary, the TOR is proportional to the importance of the protection device in terms of the system reliability as the first concern; then other factors such as to the actual time required to re-install the protection devices will be considered.

C. External Factors

External factors such as weather conditions, outage causes, etc. can affect the outage restoration time. This section discusses the TOR with respect to weather conditions and outage causes.

1. TOR by Weather Conditions

In this study, the weather conditions are categorized into: (0) Fair, (1) Cold, (2) Wind, (3) Wind & Lightning, (4) Wind & Snow, (5) Lightning, (6) Snow, (7) Ice, (8) Hot, and (10) Fog. The TOR with respect to weather conditions is graphically depicted in Figure 8. The \( \chi^2 \) value is 125.5354 with DF = 9 and the \( p \)-value = 0.0000, which indicates that the weather condition provides useful information for TOR.

Figure 8 indicates that the weather condition is a strongly influential factor for the TOR. Under bad weather conditions, e.g. stormy weather with high winds, lightning, etc., the outage restoration time takes much longer than in fair weather conditions. There are many reasons that explain why the weather condition can significantly affect the outage restoration time. To name a few: (1) It takes longer for crewman to restore an outage in bad weather than in fair weather; (2) stormy weather may cause many different outages.
in a short period of time and there is a long waiting period before the crewman can start restoring some of the outages. The mean values of TOR of snow (class 7) and ice (class 8) conditions in Fig. 8 indicate that snow causes a longer outage restoration time than ice. Intuitively, ice conditions is in general cause longer TOR than snow conditions due to reasons such as slippery roads. However, the data under analyses is in the Duke Power territory, which is located in the southern part of United States. We do not experience too much snow and ice conditions. The width of the diamond boxes, which indicates the variance of the mean estimate, for snow and ice conditions in Fig. 8 is very wide and has a large overlap on TOR for each others. This indicates that we cannot draw statistical conclusions of TOR in terms of the mean estimate to compare the snow and ice conditions. In summary, TOR is proportional to the severity of weather conditions — similar arguments as for month factor.

2. Types of Outage Causes

This section analyze the TOR with respect to different outage causes. The causes under consideration are coded as: (0) Customer Problem, (1) DPOC Accident, (2) Public Accidents, (3) Danger Trees, (4) Trees, Vines, (5) Animals, (6) Overload, (7) Lightning Strikes, BIL, (8) UG Services, (9) UG Primary, (10) Transformer Failure, (11) HPP/MHP Breaker, (12) Meter Socket/HPP/MHP, (13) Bad Secondary Connector, (14) Improper installation, (15) Natural Deterioration, (16) Vandalism, and (17) Other code. The TOR with respect to the cause codes is graphically depicted in Figure 9. The \( R^2 \) value is 1087.9372 with DF = 17 and \( p \)-value = 0.0000, which indicates that outage causes are influential to TOR.

![Graph](image)

Fig. 9. Time of Outage Restoration by Cause Code.

As indicated in Figure 9, the cause of an outage correlates strongly with the TOR. Actually, the outage cause is the most correlated factor among all the factors discussed previously. We will discuss the reason behind this strong correlation in the next section, where we discuss about the ranking of the factors’ influences for estimating TOR.

D. \( R^2 \) Analysis

Among all the influential factors under investigation, most of them are found statistically significant for the TOR. In this section, we use the \( R^2 \) value as described in Section II to evaluate and rank the degree of influence of the selected factors. The \( R^2 \) values with respect to the different factors for TOR estimate are listed in Table 1 in descending order.

As indicated previously, outage causes is the most influential factor for TOR. In [3], the authors have presented their work on estimating the outage causes based on the other factors considered in this paper, achieving satisfactory results. This implies that the factors under consideration may be correlated with each other as well. For example, an animal outage usually activate a single-phase primary fuse in a nice morning [9]. The crewman can drive to the fault location, change the fuse and restore the outage efficiently. However, if a distribution transformer fails, the crewman needs to take down the transformer, and install a new transformer. Apparently, this process takes more time and effort to restore the outage.

It is not surprising that the Day of the Week category has the lowest \( R^2 \) value because the classes under these category do not really provide information for the TOR, as indicated in the \( x^2 \) statistics discussed previously.

Table 1. \( R^2 \) values of chosen factors with respect to the TOR.

<table>
<thead>
<tr>
<th>Factors</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outage Cause</td>
<td>0.2130</td>
</tr>
<tr>
<td>Time of the Day</td>
<td>0.0567</td>
</tr>
<tr>
<td>Protection Device Activated</td>
<td>0.0550</td>
</tr>
<tr>
<td>Weather Conditions</td>
<td>0.0295</td>
</tr>
<tr>
<td>Number of Phase Affected</td>
<td>0.0137</td>
</tr>
<tr>
<td>Month</td>
<td>0.0110</td>
</tr>
<tr>
<td>Day of the Week</td>
<td>0.0023</td>
</tr>
</tbody>
</table>

IV. Conclusion

In this paper, we have analyzed the time of outage restoration with respect to several selected factors. Both statistical, quantitative results and field engineering insights have been discussed. We used the quantitative results to support our field experience insights to this problem (which is qualitative information); likewise, we used our field experience to interpret the statistical results so that the statistical results will not be mis-interpreted due to the lack of field experience. The methodology and general discussion in the paper will be useful for research and development on the Time of Outage Restoration issues in distribution systems, such as modeling development for customer service improvement.

Acknowledgment

The authors of this paper would like to acknowledge the support of Electric Power Research Center project grant 5-31326, North Carolina State University and Duke Power Company project grant 5-30585.

Reference


**Biography**

**Mo-yuen Chow**

Mo-yuen Chow earned his B.S. degree at the University of Wisconsin-Madison (1982); M. Eng. degree (1983) and Ph.D. degrees at Cornell University (1987), all in Electrical Engineering. Upon completion of his Ph.D., Dr. Chow joined the faculty of North Carolina State University in Raleigh, NC, where he is presently an Associate Professor in the Department of Electrical and Computer Engineering. Dr. Chow has also worked as a consultant for Taiwan Power Company, J.W. Harley Company, and as a faculty intern at Duke Power Company.

Since 1987, Dr. Chow has been working as a Principal Investigator in several projects in the areas of system monitoring, fault detection and control, applications of artificial neural network and fuzzy logic to power engineering. Dr. Chow served as a guest editor of *IEEE Transactions on Industrial Electronics* Special Issue on Application of Intelligent Systems to Industrial Electronics. He is currently an Associate Editor of the *IEEE Transactions on Industrial Electronics*. He is also listed in *Who's Who in Asian Americans*.

**Leroy S. Taylor**

Leroy S. Taylor is a Senior Distribution Engineer for Duke Power Company, and is a registered professional engineer in North Carolina. He was born in 1949 in Greenville, N.C. and received a B.A. degree (Physics) from the University of North Carolina in 1971. Joining Duke Power in 1977, he acquired extensive experience in distribution system engineering, operation, and construction. Since 1987 he has conducted intensive investigations on the cause of power quality disturbances which originate in the distribution system. He has also redesigned several Duke Power mainframe reporting systems used to evaluate and improve distribution system reliability and power quality.

**Mo-suk Chow**

Mosuk Chow got her Ph.D. degree in the field of Statistics from Cornell University 1983. During her graduate study, her expertise was in the area of decision theory and statistical computing. Since then her research interest branched out and encompassed a diverse applications of statistical method to various problems such as efficient estimation of superimposed exponential signals, capture and recapture problems and regression problems. Dr. Chow had taught at Northeastern University and Penn State University. She Currently works for C M Soft Company as a senior research associate and is an adjunct faculty of the Department of Management Science and Information System of the Penn State University.