Mr. Jake Langthorn, Chair, TADSWG  
Director of Transmission Policy  
Oklahoma Gas & Electric Co.  
PO Box 321 MC 208  
Oklahoma City, OK 73101-0321

Dear Mr. Langthorn:

Based on the conversations that took place during EIA’s November 2011 visit to the NERC offices in Atlanta and follow-up conference calls, EIA has prepared this memo to help provide a possible statistical framework for thinking about the significance of some of the data that NERC collects. EIA is not necessarily recommending that the assumptions and methods in this memo are appropriate ones to make; EIA is simply trying to point out the types of analyses that could be implemented if certain assumptions are made. Other assumptions might be more appropriate to make. These other assumptions might require different statistical techniques to be applied. Unfortunately the constraints on EIA’s time do not allow EIA to explore every possible analysis.

Scope of Memo

For each circuit under NERC’s jurisdiction NERC collects data about the transmission outages that occur in a given year. NERC attributes these outages to various causes and has developed reliability metrics based on these attributions. These reliability metrics include:

1. ALR 6-11: Automatic AC Transmission Outages per circuit Initiated by Failed Protection System Equipment,
2. ALR 6-12: Automatic AC Transmission Outages per circuit Initiated by Human Error,
3. ALR 6-13: Automatic AC Transmission Outages per circuit Initiated by Failed AC Substation Equipment, and
4. ALR 6-14: Automatic AC Transmission Outages per circuit Initiated by Failed AC Circuit Equipment.

EIA was asked to help provide a statistical framework for thinking about the statistical significance of changes in these metrics from year to year.

Based on follow up discussions that took place during December, 2011 EIA recommends that the frameworks outlined in this memo are not appropriate to use for analyzing common-mode outages and dependent-mode outages. The frameworks outlined in this memo are only appropriate when the outages at any circuit are independent of the outages at any another circuit.

Such an assumption is not appropriate when analyzing common-mode outages and dependent mode outages. NERC staff is better suited than EIA staff for determining whether such an assumption is reasonable when analyzing single-mode outages.
Variables Affecting Transmission Outages

When analyzing reliability metrics involving transmission outages it is helpful to recognize that many variables may influence the number of outages a circuit experiences in a given year. When enough data are collected about a circuit and its conditions, it is sometimes possible to identify correlations between certain variables and the observed outages the circuit experiences in a given year. Such studies might be beneficial in the future, but they would require additional time and possibly additional data collection.

Challenges with Measuring Reliability

Both NERC and the public are interested in knowing whether the reliability of transmission is improving. Unfortunately the influence of some volatile variables such as the number of hurricanes that make landfall on the United States can have a very significant influence on the actual reliability of transmission from year to year. These volatile variables can sometimes mask the improvements (or deterioration) in transmission infrastructure. So it may be possible to significantly improve the transmission infrastructure from year A to year B but still observe an increase in the number of outages over the same period. Whether the reliability has improved from year A to year B may lead to arguments over the exact definition of reliability.

By focusing on reliability metrics that may be somewhat less susceptible to highly volatile random fluctuations, NERC may be able to focus the reliability discussion on aspects of transmission reliability that are more controllable by policy decisions. Metrics ALR6-11 through ALR6-14 may be metrics that fall into this category.

Assumptions to Consider

It is relatively straightforward to calculate the observed rate of failure for each circuit from year to year. However, it is much more challenging to determine whether the changes in these rates are due to improvements in reliability, if the changes are due to other factors that are not the target of interest, or if they are simply due to random fluctuations. Disagreements over the definition of reliability contribute to this challenge.

By making some assumptions about the number of outages that occur each year it may be possible to draw some conclusions about whether observed changes are signs of actual changing reliability. Obviously the validity of the assumptions will influence the validity of the conclusions.

Paired Normal

By making the following four assumptions it is possible to apply a paired analysis method to assess whether the observed rates of outages for the circuits under NERC’s jurisdiction in different time periods are statistically different.

1. Assume the change in outages for each circuit is independent of the change in outages for every other circuit.
2. Assume the change in number of outages between two time periods for each circuit is described by a normal distribution (a bell-shaped and symmetric curve).
3. Assume the means and the standard deviations of the above normal distributions are identical for every circuit.
4*. Assume that any circuit that had an outage in either time period was operating during both time periods.

* The last assumption above has been included in order to accommodate NERC’s lack of a circuit inventory. When working with the change in outages between two time periods, only circuits which were operating in both time periods should be analyzed.

Unpaired Poisson

However, by making different assumptions an unpaired statistical method may be applied.

1. Assume the outages for each circuit in year 1 comes from identical but independent Poisson distributions (a Poisson distribution is a commonly used distribution to describe discrete rare events).
2. Assume the outages for each circuit in year 2 comes from identical but independent Poisson distributions.
3. Assume the number of outages in year 1 and the number of outages in year 2 are independent.

Significance Tests

Based on the assumptions in the previous section standard statistical methods may be applied to the TADS dataset. Table 1 summarizes some results based on the Paired Normal assumptions and method. Table 2 summarizes some results based on the Unpaired Poisson assumptions and method. However as has been discussed at the start of this memo these assumptions are not appropriate to apply to the existing TADS datasets that include common-mode outages and dependent-mode outages. Consequently the results in this section are shown for illustrative purposes only. No conclusions should be made using these actual results.
Table 1 – Automatic AC Transmission Outages per circuit Initiated by Failed Protection System Equipment (ALR 6-11) based on the Paired Normal assumptions and method

<table>
<thead>
<tr>
<th>For Illustrative Purposes Only (Results are not meaningful when method is applied to datasets containing dependent mode outages or common-mode outages)</th>
<th>2009-2008</th>
<th>2010-2009</th>
<th>2010-2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Estimated Mean <strong>Change</strong> in Number of Outages per Year per Circuit</td>
<td>-0.012</td>
<td>0.001</td>
<td>-0.011</td>
</tr>
<tr>
<td>2. Standard Error of Mean <strong>Change</strong> in Number of Outages per Year per Circuit</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>3. Significance of Mean <strong>Change</strong> in Number of Outages per Year per Circuit</td>
<td>***</td>
<td>NS</td>
<td>**</td>
</tr>
<tr>
<td>4. Average Number of Outages per Year per Circuit</td>
<td>0.055</td>
<td>0.042</td>
<td>0.055</td>
</tr>
<tr>
<td>5. Estimated Mean Percent <strong>Change</strong> in Number of Outages per Year per Circuit</td>
<td>-22%</td>
<td>3%</td>
<td>-20%</td>
</tr>
<tr>
<td>6. Upper Bound of 90.0% Confidence Interval</td>
<td>-0.005</td>
<td>0.009</td>
<td>-0.003</td>
</tr>
<tr>
<td>7. Estimated Mean <strong>Change</strong> in Number of Outages per Year per Circuit</td>
<td>-0.012</td>
<td>0.001</td>
<td>-0.011</td>
</tr>
<tr>
<td>8. Lower Bound of 90.0% Confidence Interval</td>
<td>-0.020</td>
<td>-0.006</td>
<td>-0.019</td>
</tr>
</tbody>
</table>

Key to Significance Labels

<table>
<thead>
<tr>
<th>NS</th>
<th>Not Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>90 Percent Confident</td>
</tr>
<tr>
<td>**</td>
<td>95 Percent Confident</td>
</tr>
<tr>
<td>***</td>
<td>99 Percent Confident</td>
</tr>
</tbody>
</table>

Technical Notes

A key to understanding the numbers in the table above is to realize that the estimated mean change in number of outages per circuit (#1) is actually a random variable. The standard error (#2) in Table 1 conceptually represents how much uncertainty there is in this estimate of the mean (#1). More technically it is an estimate of the expected standard deviation of the mean if it had been possible to estimate the mean many times.

The 90-percent confidence interval around the mean should be understood to state that there is a 90 percent chance that the “real” mean lays between the upper bound (#6) and the lower bound (#8). If one were to randomly estimate a 90-percent confidence interval 100 times then one would expect the “real” mean to fall between the upper and lower bounds 90 times. However one would expect the “real” mean to fall outside of the confidence interval for 10 of the estimates. There is no way to know whether the estimate in Table 1 represents one of the 90 “good” estimates or one of the 10 “bad” estimates.

Since NERC and the public are interested in knowing whether the reliability of transmission is statistically improving EIA has also included a significance flag (#3). Conceptually the significance level (#3) represents how confident one may be that there is a “real” difference between the two time periods.
Table 2 – Automatic AC Transmission Outages per circuit Initiated by Failed Protection System Equipment (ALR 6-11) based on the **Unpaired Poisson** assumptions and method

<table>
<thead>
<tr>
<th>For Illustrative Purposes Only (Results are not meaningful when method is applied to datasets containing dependent mode outages or common-mode outages)</th>
<th>2009-2008</th>
<th>2010-2009</th>
<th>2010-2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Estimated Mean Percent <strong>Change</strong> in Number of Outages per Year per Circuit</td>
<td>-24%</td>
<td>0%</td>
<td>-24%</td>
</tr>
<tr>
<td>2. Significance of Mean <strong>Change</strong> in Number of Outages per Year per Circuit</td>
<td>***</td>
<td>NS</td>
<td>***</td>
</tr>
</tbody>
</table>

**Key to Significance Labels**

- **NS** Not Significant
- * 90 Percent Confident
- ** 95 Percent Confident
- *** 99 Percent Confident

**Confidence Interval of Means**

Graph 1 summarizes confidence intervals around the observed means based on an assumption that the number of outages in each year is a Poisson distribution. This graph is not used for the significance tests above.
Graph1 – Automatic AC Transmission Outages per circuit Initiated by Failed Protection System Equipment (ALR 6-11) assuming that the number of outages in each year is a Poisson distribution.

**Limitations and Criticisms**

It is instructive to comment on some limitations of these methods and to point out some possible criticisms of these assumptions.

One limitation of both methods is that multiple time periods cannot be simultaneously compared. Only two time periods can be compared against one another using these methods.

Another limitation of the methods is that if you reject the assumptions then you may not be comfortable with the conclusions.

As has already been pointed out the assumption of independence is not appropriate for common-mode outages and dependent-mode outages. Consequently these methods are not appropriate for analyzing these rates of outages. Alternative techniques that take into consideration outage dependencies between circuits will need to be pursued to analyze these types of outages.

A possible criticism of the Unpaired Poisson assumption set is to argue that the number of events in year 1 and the number of events in year 2 are not independent.
Concluding Remarks

During EIA’s visit to NERC’s office in Atlanta EIA tried to emphasize that there is seldom a single “proper” statistical technique to analyze any particular situation. Rather statistical analysis is a way of thinking about data that carefully identifies assumptions and precisely asks questions. The details of the assumptions can alter which methods are appropriate to use and this in turn may potentially suggest different results.

In this memo EIA has identified two possible sets of assumptions and corresponding methods that might be applied when analyzing certain subsets of the transmission outage data for which outage independence can be reasonably assumed. There may be other sets of assumptions and methods that would be more appropriate to make. However, time constraints have limited EIA’s ability to investigate and comment on other possible assumptions and methods.

A deeper understanding of transmission reliability will likely require an in-depth study of common-mode outages and dependent-mode outages. Such a study will require more extensive collaboration with subject matter experts so that the factors that contribute to common-mode outages and dependent mode outages can be better understood. However before embarking on such a study, resources should be properly allocated. Statistical analyses become much more complicated and time consuming when assumptions of independence do not hold.

Sincerely,

Jonathan DeVilbiss
Mathematical Statistician
Office of Electricity, Renewables, and Uranium Statistics
Office of Energy Statistics
U.S. Energy Information Administration

cc:
Mr. Mark Lauby, NERC
Mr. Jeff Schaller, TADSWG (Hydro One)
Ms. Jessica Bian, NERC
Mr. James Robinson, NERC
Mr. Andrew Slone, NERC
Mr. Clyde Melton, NERC
Mr. James Powell, NERC
Mr. Matthew Varghese, NERC
Mr. Michael Curley, NERC