21, rue d'Artois, F-75008 PARIS
http : //www.cigre.org

CIGRE US National Committee
2012 Grid of the Future Symposium

Industry-wide Database For Transformers

BHAVIN DESAI, MICHAEL LEBOW
Electric Power Research Institute, Coplanar Consulting LLC., United States

SUMMARY

Managing fleets of aging power transformers is a critical issue for utility companies striving to maintain reliability and control costs. Predicting the failure rate of transformers nearing the end of their design life is challenging, making typical asset management activities such as repair and replacement decisions more difficult. Generic transformer reliability data may not be adequate to inform such decisions, and data from a single company may not be extensive or diverse enough for useful statistical analysis. To help utilities better predict transformer service life, EPRI created the Transformer Industry-Wide Database (IDB). The IDB pools transformer operating and failure data from supporting utilities in order to assemble a statistically valid population that includes a variety of power transformers.

To acquire additional performance data to support risk-based transformer management, EPRI and member utilities continue developing an industry-wide database (IDB) for power transformers. The transformer IDB captures in-service and failure data from many sources in a common format for data mining and statistical analysis. Data is obtained on in-service transformers and those removed from service due to failure or retirement. The data are historically accurate and include failure mode, operational and repair history, and equipment design information.

An important application of IDB data is to develop hazard functions through parametric analysis. This involves fitting a model to the data in order to mathematically describe the transformer aging and wear-out process over time. Because transformers have different designs, different components, and fail for different reasons, analysis groups of similar transformers must be properly selected to develop failure models appropriate for each group. Data must also be assigned to the correct group.

An innovative approach is used to separate transformer populations into subsets of similar units with shared characteristics and behavior. Statistical analysis focusing on those subsets yields insights into transformer aging, failures, and projections of future performance that support maintenance repair and replacement decisions, and asset management decisions to minimize lifecycle costs of equipment replacement and maintenance, including failure costs.

KEYWORDS

Substations
Transformers
Asset Management

bdesai@epri.com

Introduction

Managing fleets of aging power transformers is a critical issue for utility companies striving to maintain reliability and control costs. Effective asset management decisions that maximize performance and minimize lifecycle costs are based on risks associated with actual transformer condition and performance. Predicting the failure rate of transformers nearing the end of their design life is challenging, making typical asset management activities such as repair and replacement decisions more difficult. Generic transformer reliability data may not be adequate to inform such decisions, and data from a single company may not be extensive or diverse enough for useful statistical analysis. To help utilities better predict transformer service life, EPRI created the Transformer Industry-Wide Database (IDB). The IDB pools transformer operating and failure data from utilities to assemble a statistically valid population that includes a variety of power transformers.

An advanced risk-based approach to managing transformer assets uses historical performance data to group and rank transformers. This decision support foundation performs four key steps: 1) assessing existing performance, 2) specifying required performance, 3) projecting future performance, and 4) understanding how to bridge gaps

This approach requires equipment failure models and hazard rates based on the best available data. Generic reliability data is inadequate to meet decision support requirements. A given utility's transformer data is unlikely to be sufficiently extensive or represent statistically diverse transformer population subsets. To acquire additional performance data, EPRI and member utilities are developing an industry-wide database (IDB) for power transformers. The transformer IDB captures in-service and failure data from many utilities and transforms it into a common format for data mining and statistical analysis. Data is obtained on in-service transformers and those removed from service due to failure or retirement. The data are historically accurate and include failure mode, operational and repair history, and equipment design information.

The IDB provides a broad-based repository of transformer performance data as a key part of a decision support foundation. With proper care and analysis, this data can provide information about the past performance of equipment groups (e.g. substation transformers) and subgroups (e.g. 345 kV auto transformers) and the factors that influence that performance (e.g. voltage, manufacturer). With enough data, projections can be made about future fleet performance (e.g. expected number of failures), and both past and projected future performance information can be useful for operations, maintenance, and asset management decisions.

Transformer Demographics and Life Expectancy

Many substation transformers were installed in the 1960s and 1970s and are approaching the end of their nominal design lives. Figure 1 shows the age profile for over 7,000 units in a particular subset of in-service transformers contained in the transformer IDB. Clearly depicted is the "asset wall" in the 35 to 45 year age bracket. This IDB data is aggregated from eight utilities with a variety of sizes and service territories and is thought to be representative of the general industry in North America.

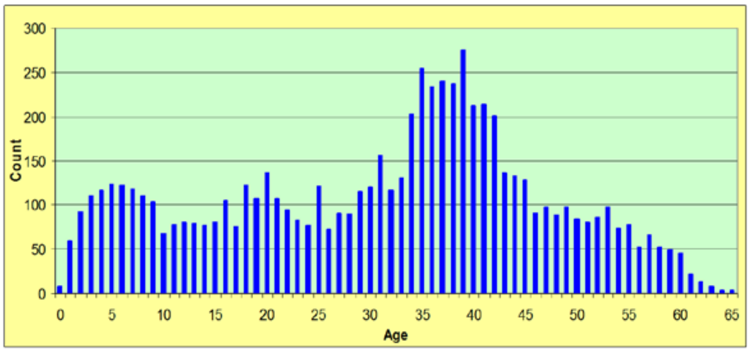

Figure 1: *Typical Age Profile – In-service Industry-wide Database for Transmission Substation Power Transformers*

Like other equipment, transformers may follow a familiar bathtub curve—an initially high rate of infant mortality failures, followed by a relatively low and constant failure rate during a long service life, then an increase in wear-out failures with impending end of life.

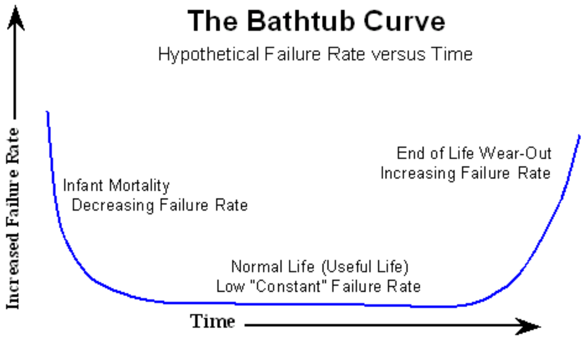

Figure 2: *Failure rate over time*

One application of the IDB is to assess whether this curve accurately describes historical transformer performance. If the bathtub curve applies to transformer life, what are the parameters of the curve—especially the wear-out portion of the curve? Do the curve parameters change with different transformer makes, models, vintages, and applications? Answering these questions is more important than ever as transformer fleets age and high replacement costs and uncertain lead times put more pressure on asset managers striving to meet high reliability standards.

Transformer Data Categories
The IDB includes population and failure data sets based on transformer type, make, model, application, and age. There are two broad classes of transformer data included: in service units (population data) and units permanently removed from service due to failure or retirement. For the purpose of IDB, failure is the termination of a transformer's ability to perform its functions with acceptable risk without major repair. This includes failure with a forced outage—requiring the transformer's immediate removal from service by relay operation or emergency switching—and failure with a scheduled outage in which the transformer is removed from service at a selected time (for example, due to unacceptably high dissolved gas levels).

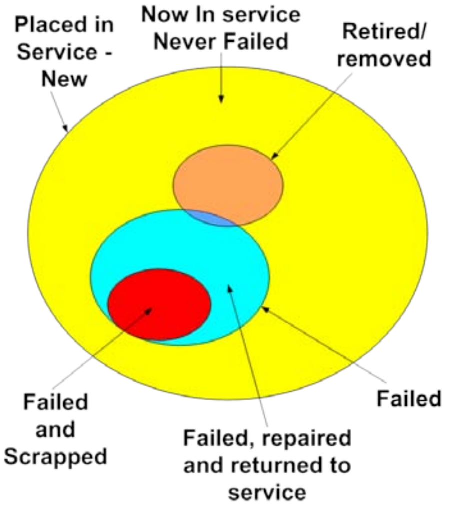

Figure 3: *IDB data categories (not to scale)*

In-service population data includes

- Transformer name plate information
- Serial number or unique identifier

- In-service or manufactured date
- Application
- Type, e.g., auto or non-auto transformer
- Previous rewind (yes/no, date)
- Service location (utility, substation, transformer position)

Removed from service data includes the preceding information, plus the date and the reason the transformer was removed from service.

Data Processing

An important application of IDB data is to develop hazard functions (hazard rates) through parametric analysis. This involves fitting a model to the data in order to mathematically describe the transformer aging and wear-out process over time. Since transformers have different designs, different components, and fail for different reasons, analysis groups of similar transformers must be properly selected to develop failure models appropriate for each group. Data must also be assigned to the correct group—auto transformers must be grouped with auto transformers and rewound units grouped with rewound units. Grouping transformer populations into subsets of similar units with shared characteristics and behavior is essential to enable an "apples to apples" analysis. The need for such groupings is illustrated by the notable differences in survival performance between new and repaired units and between auto and non-auto units (Figure 4) demonstrated by a particular subset of the IDB.

Failure data details are also desirable in defining subgroupings, including failure location (main body, load tap changer, bushing) and cause (internal or external to the transformer system). In general, utility-supplied data requires review and cleansing before it is suitable for inclusion in the IDB.

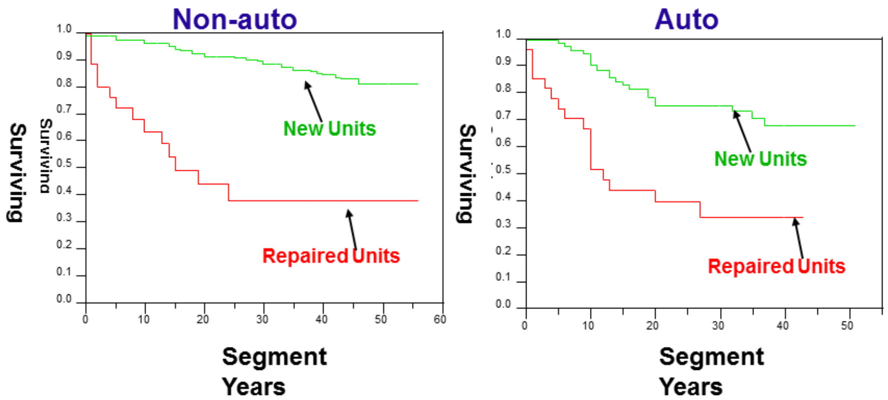

Figure 4: *Survival plots for* new *versus repaired units—notable differences in performance*

Modeling Objectives and Data Analysis

Because the models under initial development focus on the wear-out portion of the bathtub curve, we are primarily concerned only with failures that could be wear-out related. Therefore, main body failures initiated by external events such as a stuck breaker or mis-operating relay are identified and analyzed separately. To model wear-out, only failures beyond the expected age for inception of wear-out are used for analysis. Load tap changers and bushings can be expected to wear-out at different rates from each other and from the main body and therefore also are analyzed separately, even if their failure results in failure of the transformer.

The analysis begins with a search for the best wear-out models to fit one utility's data. A second utility's data is analyzed to find its best models. The two sets of data are compared to assess similarity. If the two utilities' models are similar enough based on standard statistical tests, the two utility data sets are combined and the best models are found for the aggregated data. A third utility is then analyzed and compare to the aggregated models. The process continues, building a larger aggregated data set. As the set gets larger, the models' confidence bounds can be expected to decrease.

To estimate parameters of a model, the challenge is to find the parameters that make that model fit the data best. To compare among models, we find models which fit the data best, and decide if one or more models fit sufficiently better than the rest. Maximum likelihood estimation (MLE) and likelihood ratio (LRT) tests are used. MLE reflects goodness-of-fit metrics based on the likelihood (probability) of seeing the data given a particular model. In this case this means both the general form of the model and specific model parameter values. The maximum likelihood estimates of the parameters are those values that make the observed data most likely to have happened. LRT compares two nested models, testing whether the nesting parameters of the more complex model differ significantly from their null values. LRT tests whether the extra goodness of fit to the data is worth the added complexity of the additional parameters. IDB statistical analysis is complicated by the large degree of truncation in the aggregated data and the different censor dates associated with the different utility data sets.

Project Status and Preliminary Results

Data collection for the IDB began in 2006, and the database now contains records on more than 20,000 power transformers. Several thousand more records are in various stages of review for inclusion. Figure 5 presents an overview of the IDB process. Extensive interaction with the supplying utility is often required to assure sufficient data quality and proper descriptive classification.

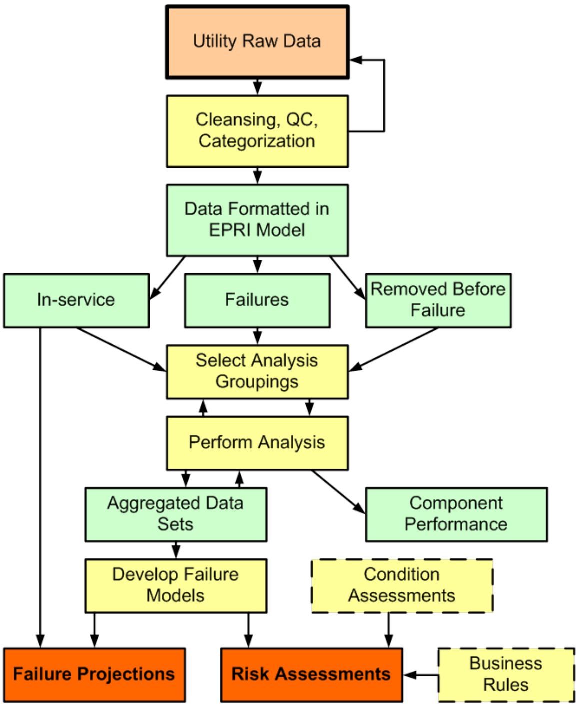

Figure 5: *The IDB process*

One goal of the IDB work is to develop appropriate hazard rates for transformer subsets of interest. The hazard functions can be convoluted with the corresponding in-service population to provide forecasts of anticipated failures. In Figure 6, an application example for a set of transformers from a particular utility provides the probability distribution of the number of failures in the next year based on a hazard rate determined from IDB analysis. Also provided are 95% confidence bounds on these probabilities. These results were computed using the appropriate hazard function and the transformer set demographic data. For example, the probability of having two failures in the next year is about 0.27. The black bars are the upper and lower 95% confidence bounds on the individual probabilities. There is essentially 0% chance of having greater than nine failures. Such calculations can provide information useful for asset management and planning.

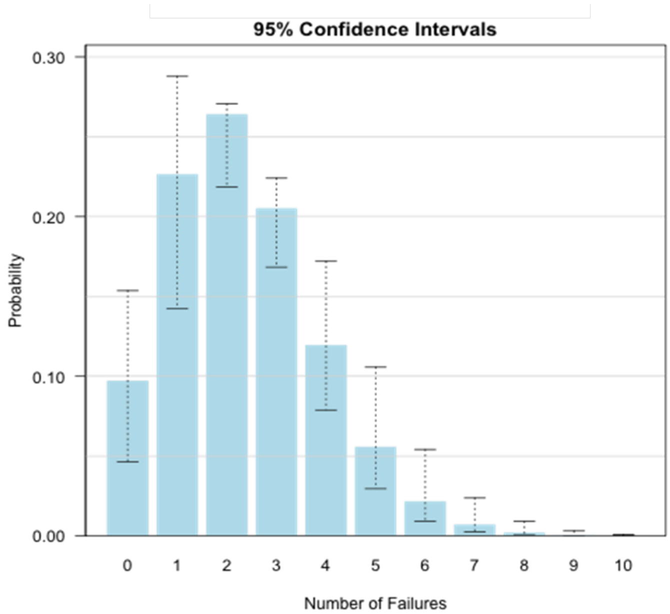

Figure 6: *Application example: yearly failures*

The development of an industry-wide database for power transformers is an ongoing effort. Through the work undertaken to date, we have positively addressed several fundamental project issues:

1. Does the potential value of an IDB justify the project?
2. Is there sufficient, useable data available?
3. Can data be collected, organized and reviewed with a reasonable level of effort?
4. Are available statistical techniques appropriate for analysis of the collected data?
5. Can hazard models be constructed with necessary confidence limits from the data?

EPRI's transformer IDB is an ongoing development and the insights, underlying methodology, approach and findings continue to be fine-tuned, enhanced and evolve as new data-sets are added and existing data reviewed. Nonetheless there are some valuable accomplishments based on work to-date:

1. Established a value consensus and assembled critical data mass
2. Developed practical data model
3. Developed data validation processes
4. Tested and identified applicable analysis techniques and methodologies
5. Produced promising preliminary results
    a. Descriptive failure statistics
    b. Aggregated data from different utilities
    c. Hazard rates that show age dependent wear-out
    d. Application of results to project anticipated number of failures

The Transformer IDB will provide utilities valuable insights and information to support maintenance repair and replacement decisions, and asset management decisions to minimize lifecycle costs of equipment replacement and maintenance, including failure costs.

BIBLIOGRAPHY

[1] Industry-wide Equipment Failure Database (Power Transformers): Data Subsets, 1023119, (EPRI, Palo Alto, CA, 2011) [Electronic media: spreadsheet-based database]

[2] Industry-wide Equipment Performance Database (Power Transformers): Data Analysis, 1023047. (EPRI, Palo Alto, CA, 2011) [Electronic Media—Power Point Presentation]

[3] Industry-wide Equipment Failure & Performance Database: Enhanced Data Model.1020010. (EPRI, Palo Alto, CA, 2010)

[4] Transformer Performance Database: A Value Proposition for an Industry-wide Equipment Performance Database (IDB) of Substation Transformers.102358. (EPRI, Palo Alto, CA, 2006)