

**MANAGING ENVIRONMENTALLY STRESSED AGING ASSETS
IN ELECTRIC POWER UTILITIES**

A Dissertation
Presented to
The Academic Faculty

by

Urenna Onyewuchi

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical & Computer Engineering

Georgia Institute of Technology
December 2012

Copyright 2012 by Urenna Onyewuchi

MANAGING ENVIRONMENTALLY STRESSED AGING ASSETS IN ELECTRIC POWER UTILITIES

Approved by:

Dr. Miroslav Begovic, Advisor
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Dr. Reginald DesRoches
School of Civil and Environmental
Engineering
Georgia Institute of Technology

Dr. Thomas Habetler
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Dr. Ronald Harley
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Dr. Saibal Mukhopadhyay
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Date Approved: November 12, 2012

To my family and loved ones

ACKNOWLEDGEMENTS

My decision to pursue a PhD began with my advisor, Prof. Begovic, when an opportunity opened up with the National Electric Energy Testing Research and Applications Center (NEETRAC). I appreciate the opportunity he gave me to stretch my mental capacity and learn several unique things only possible through a PhD path. He taught me how one can reach near-perfection by working towards it in patience rather than hurriedly. I have learned to look at problems from multiple perspectives. I have developed a new kind of resilience and strength through this academic refinement. I thank him for my maturity and growth over the past few years.

It is imperative that I acknowledge Dr. Okorie of Fair Lakes Urgent Care Center, VA and the faith he had in me from the first day we met. He, to this day, believes I am some kind of genius. Whenever I had doubts and fears about the conclusion of my research, his belief in me strengthened me. I am grateful for family and friends I made in Virginia while at George Mason University, especially Tim, Gbenga, Chinedu and Preye. I appreciate the advice and encouragement of Profs. Zietsman (who is now at Virginia Tech), Beale (now retired) and Gertler (my Signals and Systems professor) to pursue a PhD degree. I also appreciate the friends I made in Georgia.

I am grateful my path crossed that of Mrs. Milliner in the ECE Development Department, who introduced me to Ms. Pittman (Ms. P. as she is fondly called). These women provided me with probably the only social “life” I had at Georgia Tech. They believed in me, encouraged me, and could see a very bright future for me even when I couldn’t. They connected me with people they felt would provide me with guidance in my career. I have grown, just knowing them. Because of my friendship with them, Mrs.

Milliner thought of me in a meeting with Corning. That meeting changed my life, with regards to research and my career, and has presented me with a new chapter in my life.

Doyin Oyelaran and Ian Jefferson found my resume interesting. Doyin did everything possible to get me an interview with Corning Inc., which is a great company, and stay connected with the company. He advised me about how to succeed in corporate America and stayed in touch with me during my career decision process. Because of my interest in electric power, I was matched with Gerard Cliteur, a multi-talented and open-minded Corning employee, whom I respect, look up to, and admire. Gerard and Doyin have from the start, inculcated in me, the importance of staying true to myself. Gerard believes in me and is excellent about making time to listen to questions or ideas concerning my research or Corning work. His belief greatly increased my confidence in my latter semesters. Sometime in there, I became a new person (for the better). So, I can say I am indebted to Gerard and to Doyin. I hope I make them proud at Corning.

I would not be the woman I am today without the influence of my mother. I have had the honor of living with her the past five or so years as an adult, and many years as a child of course. She bore the brunt of my frustrations and took them well. She provided food for me; so I didn't have to spend time cooking but could focus on research instead. She ensured I did not leave my apartment looking haggard. She prayed for me constantly and told me I would do great. She labeled me as a "PhD" before I had even *proposed*. I could write about her for days but would like to summarize by saying that I hope I can be as great a mother as she is to me and my siblings.

My siblings have always encouraged me, and never doubted that would excel. Every time I share with them that I excelled in some exam, test or interview, they make it

a point to say “What’s new?” I am grateful to have them as my best friends: to chat with them at any time. I am also grateful for Osa’s friendship for almost two years now. He helped me review my PhD proposal, listened to and critiqued my presentation. Whenever he was not pleased with my delivery of my work, he said so and told me to remember why my research is important to the world: for public safety and management of key assets that drive the American lifestyle (and those of other nations). Keeping this point in mind gave me passion during my presentations. The discipline I have shown during my entire academic life can be attributed partly to that instilled in my siblings and me by my father. I am grateful for his lessons. My uncle Morris and his wife have always been proud of me and my work. They have kept an open door for me while I was in Atlanta and I appreciate them tremendously.

In conclusion, saving the best for last, I give all glory to God and my Christian faith. Somehow, God has kept me away from taking the wrong paths all these years. He guided me through this PhD and let me know that I was not alone. If, as I mentioned previously, my family, my advisor, Gerard and everyone believed in me, then God Himself has been my confidence, my pillar, and has heard every single prayer. I hope this dissertation is to His liking because I would not have been able to do it without Him. I am grateful to Georgia Tech as a whole for this great opportunity. I also thank Ms. Doria Moore, Siri, Chris, Ms. Mycko, Dr. Ferri and Tasha for lots of help; also, Dr. Gary May for signing my first teaching assistantship. I was opportune to attend school without need for a loan. There is no greater blessing than that, where school is concerned.

I consider myself blessed to have met everyone I mentioned in this journey of life, and so many more that who have been close to me.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
LIST OF TABLES	ix
LIST OF FIGURES	xiii
SUMMARY	xxii
1 Introduction	1
1.1 Problem Statement	3
1.2 Literature review	5
1.2.1 Failure prediction	5
1.2.2 Diagnostics	9
1.2.3 Economics of Management of Power Utility Components	11
1.3 A new approach to component management	14
1.3.1 Failure prediction	14
1.3.2 Diagnostics and preventive replacements	15
1.3.3 Economics	16
1.4 Navigating the dissertation	18
1.5 Summary	18
1.6 REFERENCES	20
2 Reliability-type Modeling of Components in Electric Power Utilities Using Preventive Replacement Data	24
Introduction	24
2.1 Reliability-Type Analysis of Preventive Replacements of Components from Utility Inspection Databases	25
2.1.1 Applying the Kaplan-Meier Method to Preventive Replacement Data	26

2.2	Comparing Reliability Models Generated from Power Utility Inspection	
	Databases to Models from Failure Databases	32
2.2.1	Metric for Quantifying Difference in Reliability Estimates of Components	
	using Corrective or Preventive Replacement Data	34
2.3	An Ideal Failure Database for Components in an Electric Power Utility	39
2.4	Summary	40
3	Modeling Failure Rates	41
	Introduction	41
3.1	Age-Based Failure Estimation and Prediction	42
3.1.1	Exponential Case	43
3.1.2	Non-Exponential Case	59
3.2	Stress-Based Failure Estimation and Prediction	80
3.2.1	Fragility Assessment of Wind-Stressed Components	82
3.2.2	A Second Predictor of Stress-Based Failure Risk: Age	90
3.2.3	A Strategy for Prioritizing the Geographic Scheduling of Component	
	Inspections using Environmental and Fragility Risk: Development of the Inspect	
	Index	97
3.2.4	Economic Effects of Component Replacements	101
	Sample Results: Wood Poles	103
3.3	Summary	107
3.4	REFERENCES	110
4	Predictive Maintenance: Diagnostics and Preventive Replacement of High-Risk	
	Components	112
	Introduction	112
	Navigating this Chapter	113
4.1	Analyzing the Accuracy of Individual Diagnostic Tests	114
4.1.1	Measures of Diagnostic Accuracy	115

4.1.2	Sampling Power System Components for Destructive Tests	118
4.1.3	Ranking Accuracies of Diagnostic Tests using Statistical Tools and Economics of Diagnostic Decisions	134
4.2	Improving Diagnostic Accuracy and Validity by Combining Diagnostic Test Results	143
4.3	Analyzing the Accuracy of Combined Diagnostic Tests	149
4.3.1	Applying Logistic Regression Analysis to Multiple Diagnostic Test Predictions	149
4.3.2	Applying Receiver Operating Characteristics (ROC) Analysis to Multiple Diagnostic Test Predictions	155
4.3.3	Applying Combinatorial Logic Rules	156
4.4	Effects of Cost-Relationships on Decisions of Preventive Replacements from Multiple Diagnostic Testing	166
4.5	Summary	170
4.6	REFERENCES	173
5	Stochastic Optimization of Preventive Maintenance Programs for Power System Components	174
	Introduction	174
5.1	Optimizing the Component-Management Process	176
5.1.1	Costs that affect Decision-Making in Management of Environmentally Stressed Aging Components	178
5.1.2	Defining the Decision Parameters	179
5.1.3	Defining the Random Variables	182
5.1.4	The Objective Function	197
5.2	Comparisons between Stochastic Sampling Approaches	201
5.2.1	Simple Random Sampling	202
5.2.2	Stratified Sampling	203

5.2.3	Latin Hypercube Sampling	205
5.2.4	Simulation Results: Application on wood pole data	207
5.3	Single-Year Optimization	213
5.3.1	Simulation Results	215
5.4	Multi-Year Optimization	226
5.4.1	Simulation Results: Multi-Year Optimization	228
5.5	Summary	244
5.6	REFERENCES	247
6	Future Work and Contributions	250
6.1	Future Work	250
6.2	Contributions	251
6.3	Publications	254
VITA		256

LIST OF TABLES

	Page
Table 1 Population of wood poles inspected, treated, and preventively replaced in Utility D during years 2003-2007.	29
Table 2 Complete set of failure times of $k = 50$ items with $T = 0.0155$ time units.	55
Table 3 Table of deviations of predicted expected times to failure from the true MTTF for different values of k , N and ratio k/N given different levels of missing failure data.	58
Table 4 Table of attributes of component failure risk at the time instants that an onset alert is triggered: time-varying hazard rates at those instants and the binomial probability of observing an onset (a run length of one) within three trials.	75
Table 5 Wind load statistics for ASCE 7 load criteria (Ellingwood and Tekie, 1999).	87
Table 6 Proportion of Florida Power and Light wood pole restorations after five real hurricanes [6].	89
Table 7 The average difference of the fragility curves are shown as $MD(t)$ for $t = 25, 50, 75, 100$ years using 100 simulated wind speeds . The recommended minimum age for inspections is highlighted in bold.	96
Table 8 Mean of fragilities per age and p-values of the t-statistics of the difference between fragilities with the null hypothesis being that the fragilities are equal and $\alpha = 0.05$..	97
Table 9 Number of wood poles inspected in years 2000 and 2001 from an inspection database and costs of excluding poles less than 20 or 25 years using an inspection cost C_i of \$20 per pole. The cost of failures less than either minimum age is used with the inspection cost to determine the benefits of recommending either 20 or 25 years.	106
Table 10 2x2 contingency table for classification of conditions and diagnostic test results of components.	116
Table 11 Parameters of simulated test for estimating the positive predictive value (PPV) of a diagnostic test over a 10-year inspection cycle. MC size is the Monte Carlo size, PPV_true is the simulated PPV, Cdt is the cost of a destructive test, C_{pr} is the cost of a preventive replacement, U represents a uniform distribution.	127
Table 12 Table of means (E_{PPV}) and standard deviations of PPV estimates for a Monte Carlo sampling size of 10, 000 showing 4 proportions of preventive replacements that were destructively tested.	130

Table 13	Table of the length of period in years it takes to reach minimum coefficient of variation of 5%.	132
Table 14	Table of costs of a destructive testing program when 5% - 50% of annual preventive replacements are destructively tested in each year. The number replaced is assumed to be 500 per year, the cost of one destructive test is \$200, one preventive replacement is \$2,000, the initial PPV is 0.6 and later 0.8 for proportions 10% - 50% after n_d years.	132
Table 15	Means and standard deviations of strengths of wood poles predicted by diagnostic tests one through six. They show the separations between predicted strengths for poles that failed the destructive test and those that passed, for significance level $\alpha = 5\%$.	136
Table 16	Sensitivity and specificity of six diagnostic testing techniques applied to at most fifty distribution wood poles. The tests are arranged in the order that NEETRAC considered best performance.	137
Table 17	Statistics of coefficient estimates of logistic fits of distinct diagnostic scores. The response variable is the classification of poles as weak or strong by destructive tests. The diagnostic scores are the predictor variables. Lower deviances are better than higher deviances.	138
Table 18	Area under the curve (AUC) for logistic regression model of each diagnostic test and combined prediction of tests one and three.	140
Table 19	Error probabilities of diagnostic decisions taken by six testing procedures. The tests were compared to destructive test results for evaluation of accuracy. The performance ranking in ascending order of costs is 5, 3, 4, 2, 1, 6.	143
Table 20	Correlation coefficients between predicted strengths of wood poles from six different diagnostic procedures.	148
Table 21	Correlation coefficients between predicted classifications of wood poles from six different diagnostic procedures using logic values zero for poles that tested good and one for poles that tested bad.	149
Table 22	Progression of deviance of a logistic regression model using a forward selection algorithm. The final intercepts and slopes of the included predictions are in the last rows of the table.	152
Table 23	Statistics and deviance of a forward selection logistic regression model that starts by including diagnostic test one in the initial model. The deviance increases to about 38 with inclusion of predictor three.	154
Table 24	Area under the curve (AUC) for logistic regression model of each diagnostic test and combined prediction of tests one and three.	156

Table 25 “OR” and “AND” combinatorial rules between predictions of two diagnostic tests. Logic one stands for components that test positive and logic zero stands for components that test negative.	157
Table 26 Selections of two or three diagnostic tests combined using the “OR” logic rule for making decisions on preventive replacements of wood poles. The decision costs of each combination are shown in the column labeled “Total.” “CNPV” is $P[\text{bad} \text{tests good}]$ and “CPPV” is $P[\text{good} \text{tests bad}]$.	161
Table 27 Selections and decision-costs of four diagnostic tests combined using the “OR” logic rule for making decisions on preventive replacements of wood poles.	162
Table 28 Selections of five or six diagnostic tests combined using the “OR” logic rule for making decisions on preventive replacements of wood poles.	162
Table 29 Accuracies and resulting inspection-cost ratios for six diagnostic testing procedures. The base accuracy (for test six) is 0.755.	169
Table 30 Derived failure probabilities of wood poles for distinct ages 0 – 100 years.	188
Table 31 Mean and one standard-deviation limits for age-specific failure proportions $F(t)$ estimated for distinct ages 0 – 125 years.	193
Table 32 Computation time, expected value and variance of differential cost S between a run-to-failure program and a preventive maintenance program for increasing sample sizes.	207
Table 33 Statistics and properties of three sampling plans: simple random sampling (SRS), stratified sampling (SS) and Latin hypercube sampling (LHS) applied to the management cost problem and implemented in MATLAB, where “ \wedge ”: power-exponent.	210
Table 34 Statistics of escalated costs of inspection, preventive replacements and corrective replacements as a product of base costs and ECI percentage increases using ECIs for total compensation for private industry workers. “*”: These costs are re-scaled by discount rates.	218
Table 35 Parameters of population properties for single-year optimization.	220
Table 36 Values of the random variables for the mean-value problem, where “a,” “b” and “c” are modes, or upper or lower bounds of the distributions.	225
Table 37 Ranks of 15 feasible solutions of the deterministic optimization problem. The last column (σ) is the 3-digit form of the proportion shown more fully under the “Constraints” column.	225
Table 38 Parameters of the random variables, excluding net present value scaling.	230

Table 39 Expected values, variances and standard deviations of five feasible solutions of combinations of the decision parameters: inspection cycle c and proportion of inspections to replace preventively σ . 232

Table 40 Goodness-of-fit statistics for probability plots of the distribution of $F(80)$. 239

LIST OF FIGURES

	Page
Figure 1 Black box showing inputs (left) and some outputs (right) of a reliability function algorithm.	28
Figure 2 Population distribution of inspected Utility D poles.	30
Figure 3 Population distribution of inspected Utility D poles, which received remedial chemical treatment (left) and did not receive treatment (right).	30
Figure 4 Population distribution of Utility D poles scheduled for preventive replacement, which received remedial chemical treatment (left) and did not receive treatment (right).	31
Figure 5 Survival Function of inspected Utility D poles, which received remedial chemical treatment (top) and did not receive treatment (bottom).	31
Figure 6 Age distribution of over 150,000 wood poles in a real power utility.	36
Figure 7 Hypothetical hazard rate and probability density function of real wood pole data under an assumed normal weather condition. Here, a Weibull density is used with shape parameter 5.2 and scale parameter 107. The pdf decreases for larger ages > 107 years.	36
Figure 8 An empirical survivor function of wood poles estimated using a hypothetical Weibull failure density function to represent real failure data and a corresponding function estimated using preventive replacement data obtained from statistics of inaccuracy of diagnostic testing on the poles.	38
Figure 9 Bayesian updates of posterior distributions for $\beta=1$ and $\alpha=2$ (top subplot), and $\alpha=3$ (bottom subplot). Peaks of the 10 posteriors are observed to shift to the right for each additional failing item. The prior distribution is the innermost curve.	46
Figure 10 Bayesian updates of posterior distributions for $\beta=1$ and $\alpha=4$ (top subplot), and $\alpha=5$ (bottom subplot). Peaks of the 10 posteriors are observed to shift to the right for each additional failing item. The prior distribution is the innermost curve.	47
Figure 11 Prior (large variance) and posterior (small variance) distributions of θ for $N = 10$ items, $k = 5$ items, with true $\theta = 1/3$ and α and β equal to 3 and 1, respectively.	48

- Figure 12 Plot of $F(t|t>T)$ versus time—probability distribution function of times to failure of devices still surviving at $T = 1.3094$ time units, showing the true distribution (bottom curve) for $\theta = 1/3$, and the distribution predicted using Bayesian updates of posterior of θ (top curve). Here, $N = 10$ items and $k = 5$ items. 49
- Figure 13 Prior (left) and posterior (right) distributions of θ for $N = 1000$ items, $k = 5$ items, $\theta = 3$, α and β equal to 3 and 1, respectively. 50
- Figure 14 The conditional probabilities of time to failure $F(t|t>T)$ for $N = 1000$ items, $k = 5$ items, where the bottom curve is the true distribution obtained using $\theta = 3$ of the generated times to failure, the top curve is the predictive distribution using the prior and posterior of θ and $T = 0.0025$. 51
- Figure 15 Prior (left) and posterior (right) distributions of θ for $N = 1000$ items, $k = 20$ items, $\theta = 3$, α and β equal to 3 and 1, respectively. 51
- Figure 16 The conditional probabilities of time to failure $F(t|t>T)$ for $N = 1000$ items, $k = 20$ items, $\theta = 3$ and α and β equal to 3 and 1, respectively. True distribution is top curve and predicted is bottom curve. 52
- Figure 17 The conditional probabilities of time to failure $F(t|t>T)$ for $N = 1,000$ items, $k = 50$ items, $\theta = 3$ and α and β equal to 3 and 1, respectively. For the complete failure dataset, the true function is the bottom curve and the predicted is the top curve. 56
- Figure 18 Prior (largest variance) and 3 posterior densities (small variances) of θ are shown in the top subplot. In the clockwise direction, the true distribution $F(t|t>T)$ and Bayesian-predicted distributions for a complete failure dataset, 40%, and 50% missing failure data are shown in the bottom subplot for $k = 50$ and $N = 1000$ items. 57
- Figure 19 Histograms of the mean time to failure $E_A(t)$ for failure data missing 40% and 50% of the failure data(left subplots), and the deviations from predictions with complete data (right subplots). 58
- Figure 20 Graph showing the relationship between percent differences of predicted expected times to failure from the true MTTF for different values of k , N and ratio k/N ($k = 50$ and $N = 1000$, $k = 500$ and $N = 10000$, $k = 100$ and $N = 10000$, $k = 50$ and $N = 10000$, respectively) with respect to different levels of missing ratio l/k (left subplot) and failure data l (right subplot). 59
- Figure 21 Illustration of sliding time windows of length t_i where approximately exponential failure rates h_{tv} are estimated. The difference between consecutive sliding windows is defined to result in a large overlap between the windows. 61
- Figure 22 Hazard curve used for generating random variates with $t_0 = 30$ time units, $t_n = 100$, $\alpha = 5$, $\beta = 50$. 63

- Figure 23 Histograms of pseudo-random numbers distributed on $U(0,1)$, the uniform distribution with minimum and maximum 0 and 1 respectively. They show limitations of computer programs in generating perfect uniform distributions. 64
- Figure 24 Probability density functions (top sub-plots) and hazard functions (bottom subplots) of Weibull densities with varying shape α and scale parameters β . The legends in the top subplots apply to the bottom subplots. The probability density functions (pdfs) shift to the right for $\alpha = 5$ as β increases from 20 to 70. The pdfs reduce in variance and increase in height for $\beta = 50$ as α increases from 2 to 7. Non-linearity reduces in the hazard functions for $\beta = 50$ as α increases from 2 to 7. The hazard rates stretch to the right for $\alpha = 5$ as β increases from 20 to 70. 71
- Figure 25 Histogram of generated lifetimes for $N = 10,000$ components for Weibull parameters $\alpha_2 = 5$ and $\beta_2 = 70$. 71
- Figure 26 Theoretical cumulative distribution function (cdf) of the simulated times-to-failure (line with triangular marker) and the empirical cdf of the generated times-to-failure (stairs directly behind the theoretical cdf). 72
- Figure 27 Theoretical hazard curve (dashed line) and time-varying exponential hazard rates (solid line), showing slight overestimation of the failure risk with increasing time. 73
- Figure 28 Smoothed estimate of the time-varying exponential hazard rates using a Kalman filter (solid line in top subplot) overlaid with the theoretical Weibull hazard curve (dashed line in top subplot) and the error covariance of the Kalman filter (bottom subplot), showing convergence of the error covariance around 10 years. 74
- Figure 29 Plot of the probability that the most recent failure rate estimate at some age is the maximum point among prior estimates. 75
- Figure 30 Histogram of wood poles replaced in a power utility. The replacements show a partial bathtub hazard behavior up to about 45 years as seen in the figures following. 77
- Figure 31 Nonparametric hazard plot of replaced wood poles using the actuarial method in MINITAB. Here, $N = 27,098$. 77
- Figure 32 Time-varying exponential hazard rates for a real power utility wood pole replacement dataset. 78
- Figure 33 Smoothed estimate of the time-varying exponential hazard rates using a Kalman filter (top subplot) and the error covariance of the Kalman filter (bottom subplot), showing convergence of the error covariance around 12 years. 79

Figure 34 Plot of the probability that the most recent failure rate estimate at some age is the maximum point among prior estimates.	80
Figure 35 Number of poles per class for one year of inspection data in a real power utility, showing that there are significantly more wood poles of classes three and five than any other class, of the 5,792 poles analyzed.	84
Figure 36 Distribution of heights of wood poles for classes three and five.	85
Figure 37 Empirical and fitted cumulative distribution functions of heights of class three and five poles. The fit is for a lognormal distribution function.	85
Figure 38 Cumulative distribution function (cdf) of heights, in feet, of poles sampled from the fitted lognormal distribution using the Latin Hypercube Sampling technique.	86
Figure 39 Distributions of moment capacity of class three and five wood poles.	87
Figure 40 A plot of probability versus moment: The distribution of moment demands of wind loads and moment capacities of wood poles of classes three and five for $V=200$ mph in power distribution networks.	88
Figure 41 Combined mean fragility curve of Class 3 and 5 wood poles (solid line) ± 1 standard deviation (dashed lines), excluding wind loads on overhead lines.	89
Figure 42 Expected residual strength of the poles as a function of age.	93
Figure 44 Illustration of derivation of the “inspect index” for a component in a county or district, signifying the fragility-risk of components in that region given historical hurricane information.	99
Figure 45 Flowchart of strategy for scheduling inspections of environmentally stressed aging components in electric power utilities, where MNE stands for maximum no-effect winds, City [i] is an ordered list of cities, $n_{c[i]}$ is the number of components over the inspection start-age of components for the ordered list of cities, N is the total number of components owned by the electric utility and c is the inspection interval or schedule.	100
Figure 46 Regions of a power utility, where wood poles were inspected in 2000 and 2001, showing the numbers in parentheses (#, #) as the historical averages of wind speeds from 1900 to present in mph and one standard deviation of the average ages (in years) of the poles inspected in the regions respectively. While Region 2 has the highest inspect index, Region 1 has the least.	104
Figure 47 Illustration of the classification of components as good and bad based on destructive test results (lower sets), and predictions of conditions of the components using diagnostic testing (dashed sets). The circled markers are inaccurate diagnoses of the conditions marked by squares.	116

- Figure 48 Illustration of annual destructive tests performed on some proportion of unique components (with label “R”) that test positive over n years of inspections for estimation of the positive predictive value (PPV). 120
- Figure 50 A plot of mean positive predictive value (PPV) versus number of years. Expected PPV of a diagnostic test per year of analysis for different sizes of destructive tests. The standard deviations of 5% and 50% PPV estimates for years 1 and 10 are also shown. 130
- Figure 51 Coefficients of variation for PPV estimates for 1 through 10 years of destructive testing simulations, showing that the least coefficients were found when 50% of preventive replacements were destructively tested. 131
- Figure 52 Receiver operating characteristics curves of six diagnostic tests using the results of logistic regression analysis for classification of components. 140
- Figure 53 A tree diagram showing costs and decisions of diagnostics and replacements of a component with costs in parentheses. 142
- Figure 54 Normal probability plots of six diagnostic testing scores of strengths of utility wood poles. The plots show that a normal distribution may not be the best fit for the scores. 151
- Figure 55 Change in model deviance from an initial null model to inclusion of the first predictor (diagnostic test three) and then, inclusion of the second predictor (diagnostic test four). The other predictors were not found to significantly improve the model deviance, as per hypothesis tests. A forward-selection algorithm is used. 153
- Figure 56 Receiving operating characteristics (ROC) curves for classification of condition of distribution wood poles by logistic regression modeling of diagnostic test scores one through six. The logistic model that combines predictions from diagnostic test one and three is shown to have a generally higher ROC curve than the model of the individual tests. 155
- Figure 57 Plot of the minimum costs that result from preventive replacements of wood poles using single (“Diagnostic 1 alone” or “Diagnostic 5 alone”) or logic OR-combined diagnostic tests. The x-axis shows the number of tests combined (2, 3, 4, 5 or 6). The selections yielding each minimum cost are shown in text boxes in the figure. 163
- Figure 58 Plot of error rate of missing a bad wood pole after a compensatory combination of diagnostic tests against the correlations between the combined tests. Initial error rate when diagnostic test one is used alone is 0.2273; two, 0.1389; three, 0.2143; four, 0.1622. 164

Figure 59 Plot of error rate of prematurely replacing a good wood pole after a compensatory combination of diagnostic tests against the correlations between the combined tests.	164
Figure 60 Plot of the minimum costs that result from preventive replacements of wood poles using single (“Diagnostic 1 alone” or “Diagnostic 5 alone”) or logic AND-combined diagnostic tests. The x-axis shows the number of tests combined (2, 3, 4, 5 or 6). The selections yielding each minimum cost are shown in text boxes in the figure.	165
Figure 61 Decision-costs generated by combining diagnostic tests using logic “OR” rules, “AND” rules or some combination of the rules. The third and fourth data-points are derived by combining the first two tests using the “AND” rule or the “OR” rule. Then, the results are combined with the last third test using the “OR” or “AND” rule correspondingly.	166
Figure 62 A plot of minimum costs of different combinations of diagnostic tests used in making decisions on preventive replacement of wood poles. Here, the ratio of corrective to preventive replacement costs is two.	167
Figure 63 A plot of minimum costs of logic “OR” combinations of diagnostic tests used in making decisions on preventive replacement of wood poles. Here, the cost of individual inspections is twice the initial cost.	168
Figure 64 A plot of minimum costs of logic “OR” combinations of diagnostic tests used in making decisions on preventive replacement of wood poles. Here, the costs of diagnostic procedures are directly proportional to the accuracies of the procedures.	169
Figure 64 Wind zone map of the United States by Federal Emergency Management Agency (FEMA), showing relative risks of high winds among states [2].	176
Figure 66 Lifecycle of a typical inspected component up to its end-of-life either from failure or diagnostic decision for preventive replacement.	178
Figure 67 Normal probability plot of residuals of a simple regression fit of a log-transformed $F(t)$.	189
Figure 68 Simple regression fitted plot of the log-transformed failure probability and its confidence bounds.	190
Figure 69 Mean and 95% confidence interval of age-specific proportions of exposed wood poles that may fail under possible hurricane wind speeds.	191
Figure 70 A closer view of $F(t)$ estimate for lower ages.	191
Figure 71 Age distribution of wood poles in a real power utility used in comparing three stochastic sampling approaches.	192

Figure 72	Age distribution of expected failures during a hurricane event for a southern Florida hurricane probability distribution.	193
Figure 73	Simple regression fit of one of 729 possible combinations of age-specific failure proportions.	194
Figure 74	Densities of annual expected failures using simple regression fits of enumeration of means and one standard deviation limits of the age-specific failure proportions.	195
Figure 75	Histogram and lognormal regression fit of the expected failures.	196
Figure 76	Histogram and lognormal regression fit of the expected failures for ages ≥ 20 years.	196
Figure 77	Annual costs of competing strategies (run-to-failure RTF and preventive maintenance) that play a role in component management decision-making, ignoring the cost of penalties.	198
Figure 78	A plot of the estimate of the mean and 95% confidence intervals of the mean for increasing sample sizes of the random variables for the simple random sampling scheme.	208
Figure 79	Expected annual differential cost of competing pole-management approaches by sampling using three approaches: simple random sampling (SRS), stratified sampling (SS) and Latin hypercube sampling (LHS).	211
Figure 80	Difference in ranges of 95% confidence limits between the stratified sampling (SS) and Latin hypercube sampling (LHS) methods, and the simple random sampling (SRS) method. The SS provide provides consistently smaller and relatively significant confidence ranges while the LHS range improves after 3,125 samples.	212
Figure 81	Comparisons of memory requirements for three stochastic sampling methods.	212
Figure 82	Comparisons of computation time requirements for three stochastic sampling methods.	213
Figure 83	Cumulative distribution function of optimum expected differential cost between two competing maintenance approaches, showing that excluding lost revenue, negative public perception from a run-to-failure (RTF) program, penalties for failing to run a preventive maintenance program, the RTF appears to be more financially viable than a preventive program.	221

Figure 84 Sensitivity of expected differential cost of five random variables: cost of unit corrective replacement Ccr, cost of unit preventive replacement Cpr, cost of unit inspection Ci, expected annual number of failures out of inspected components Ya, and sensitivity s1, using a correlation approach.	222
Figure 85 Cumulative distribution function of differential costs between the present utility preventive maintenance approach and a run-to-failure (RTF) approach, showing both a large difference favoring RTF and a large variance.	223
Figure 86 Cumulative distribution function of expected management differential cost for a highly sensitive diagnostic test.	224
Figure 87 Distribution of differential costs for the first year of optimization excluding net present value scaling.	231
Figure 88 Sensitivity of differential cost to five random variables: number of failures for poles at least 20 years old (Ya), accuracy of diagnostic test (s1), cost of inspection (Ci), cost of preventive replacements (Cpr) and cost of corrective replacement (Ccr), using a correlation approach.	231
Figure 89 Expectation and variances of feasible solutions of the first year for the optimization problem, showing an inspection cycle of 20 years as the solution.	233
Figure 90 Overlay chart of annual differential costs (left distributions) between competing management approaches for two years of optimization including the distribution of their sums (right).	234
Figure 91 Sensitivity of the sum of two annual differential costs to nine random variables using a correlation approach.	235
Figure 92 Sensitivity of the sum of three annual differential costs to seven shown random variables using a correlation approach. “U13” is for the sensitivity s1 in the 3 rd year.	236
Figure 93 Overlay chart of annual differential costs (left distributions) between competing management approaches for three years of optimization including the distribution of their sums (right).	236
Figure 94 Overlay chart of annual differential costs (left distributions) between competing management approaches for three years of optimization including the distribution of their sums (right).	237
Figure 95 Sensitivity of the sum of three annual differential costs to seven shown random variables using a correlation approach.	238
Figure 96 Probability plot of distribution of failure proportions for 80-year old poles.	239

Figure 97 Four instances of distributions of failure proportions for ages 59 – 62.	240
Figure 98 Overlay chart of annual differential costs (left distributions) between competing management approaches for two years of optimization and the distribution of their sums (right).	240
Figure 99 Sensitivity charts of the two-year sum of differential costs, the differential cost in the first year ($p=0$) and the cost in the second year ($p=1$).	241
Figure 100 Overlay of distributions of differential costs for a highly sensitive diagnostic test.	243
Figure 101 Sensitivity chart of the four-year sum of differential costs.	243

SUMMARY

Electric power delivery is an old and well-established sector of service providing in the United States. The power utilities, which provide the service, are under continual scrutiny from both the public and the government. Power reliability indices are published periodically showing trends in performance improvements or declines of power services. While severe environmental hazards are considered “acts of God” and do not affect utility indices, customers still expect a significant level of hazard preparedness from the utilities to avoid interruption of power and life in the midst of hazard seasons. Environmental hazards include wind storms (tropical storms and hurricanes), ice storms, and earthquakes, among others. The performance of utility assets during severe weather is related to the management of the assets before the hazard.

The traditional approach to managing electric utility assets involves running them to failure: a corrective and reactive approach that could cost the utilities dearly during severe weather. Public perception and reputation of the utilities turns negative with even a slight appearance of poor asset management. The run-to-failure approach allows assets to sometimes get as old as the power utility: a disaster waiting to happen. In the wake of fatalities and customer complaints as storms appear to increase in severity over the years, some power utilities are now expected or even mandated to conduct strict regular management programs on their assets.

Preventive maintenance approaches can be very costly. Power utilities currently spend millions of dollars annually to discover a relatively small number of assets that may fail only under stress of environmental hazards. When the frequency or intensity of the hazards the assets are exposed to is small enough, the millions could be viewed as

wasted investments. To make informed decisions on the management of these decisions, the power utility requires useful information on the failure of the assets and statistical and probabilistic expertise to process the information.

Findings in this research suggest a number of improvements that could save power utilities millions of dollars annually from the expenses they currently make in the management of their assets. Real-time decision-making on inspections of assets could also be improved by combining models developed in the course of this research with good data acquisition processes on asset failures.

In the first chapter of the research work, available data on replacement of assets from power utilities were analyzed statistically to infer asset reliability as reflected by the diagnostic company serving the utility. Using prior knowledge on the inaccuracy of the diagnostic procedure, an experiment was designed to test the effect of both diagnostic inaccuracy and an incomplete database of a population of identical components on reliability estimates of the components.

A metric was developed to quantify the disparity between the reliability estimate using diagnostic *replacement* data and corresponding hypothetical *failure* data for the experimental results. The results showed that diagnostic tests that fail to detect a significant number of components at high risks of failure (low sensitivity) overestimate reliability estimates. On the other hand, tests that wrongly classify low risk components as high risk components (low specificity) underestimate reliability estimates. The experiment simulated an inspection database consisting of about 10% of a population of components, and diagnostic sensitivity of 0.38 and specificity of 1.00, the deviation between true reliability and diagnostic-measured reliability was about 80%. Thus, for

accurate reliability estimation of utility assets, the power utility should embark on recording and processing failure data rather than diagnostic replacement data.

In the next chapter of research, methods that applied failure data in predicting future failure occurrences or estimating annual failure numbers were developed: first for non-stress failures and then for stressed failures. The non-stress failures were assumed to be age-based or time-based, which is a general assumption by electrical reliability engineers. Exponential distributions are very common in reliability studies because their hazard rates are constant and easily differentiable or integrable. They are also related to other density functions like Weibull and gamma, which represent more realistic hazard or instantaneous failure rates than exponential distributions.

A Bayesian approach was developed in the course of this research to obtain the predictive distribution function of lifetimes of surviving components from a dataset of recorded lifetimes assumed to be sampled from an exponential distribution. The contribution of the work was both in estimating the function when the rate parameter is unknown and its convergence even when the failure dataset was small relative to a larger population of surviving components. The distribution function was found to depend on the population size, the number of failed components, the times-to-failure of the observed failures and the distribution parameters of a Gamma density function assumed to be the prior density of the rate parameter. The Gamma density function was chosen because it is a conjugate prior of an exponential likelihood function, making the product of the functions easy to integrate. While the procedure is strong in predicting the distribution of future lifetimes under uncertainty in the rate parameter, it relies on the choice of the posterior density function of the parameter. This is a limitation.

The exponential assumption was found to be useful for gradually building a bathtub hazard curve of a population of components as failure information is acquired over small intervals of time: one year, for instance. Maximum likelihood estimation was applied to the lifetimes of components that failed within the period of say one year, the number of failures that occurred before the time interval and the number surviving to obtain estimates of time-varying rate parameters. The estimates were then processed through a Kalman filter and a model that uses extreme value theory to detect increase in failure risk. The work serves as a tool to prompt an asset manager of the onset of increasing failure risk and possibly an appropriate age to begin preventive maintenance on the component.

Historically, some sets of components have been reported to fail more significantly from environmental hazards than aging. These components require stress-based reliability estimation. Fragility assessment techniques are often applied in structural engineering to estimate probabilities of loss from the damage of structures like residential buildings, bridges and the like from environmental hazards. However, for the sake of research on aging and aged utility assets, the loss of strength with age was incorporated into fragility assessment.

Fragility curves were generated for distinct ages of components. Analytical experiments were conducted for utility wood poles under hurricane stress. The results showed that under the same intensity of wind speeds, newly installed wood poles carrying overhead lines and 25-year old poles performed approximately the same. Thus, replacing a 25-year old pole preventively would not decrease its probability of failing under a hurricane. Modifying utility inspection strategies to incorporate this finding for

the current utility inspection cycle led to 25% savings in annual inspection investments alone for a power utility studied, or approximately \$1 million.

To improve diagnostic recommendations, results of destructive tests used in evaluating the accuracy of different diagnostic techniques for one set of utility assets were analyzed. A technique popularly used in at least the eastern part of the United States was found to have a low sensitivity to detecting high risks of failure in a small sample of components. Misclassification costs were found for this technique and compared to those of other techniques. The technique was found to be inferior to some others. However, when the classification pairs of the popular technique and that of a more superior technique were combined using an “OR” logic rule, savings were obtained of over \$200 per component per year for an expected component failure rate. The study showed strong performance in combining two diagnostic techniques. The costs began to escalate with every other additional combination of diagnostic techniques.

The question of the benefit of preventive maintenance in light of diagnostic inaccuracy and uncertainties in failure probabilities from unpredictable environmental hazards arises. This was handled by defining a cost-benefit function between a preventive maintenance approach and its competing run-to-failure (RTF) counterpart. A stochastic optimization framework was developed for decision-making over a planning or budgeting horizon. Two decision or choice variables were chosen for component management: an inspection cycle of the components of interest and the proportion of the components to replace preventively per year. The latter is important because it provides room for the deliberate reduction (replacement) of aged components.

Excluding potential penalties for failing to run a stringent maintenance program, the benefit of the preventive maintenance program to the utility may be observed in the number of failures that can be averted through accurate diagnostic detection. The simulations run using utility wooden poles as a case study showed the power utility may have the potential to save over 75% of their current expenditures on inspections and preventive replacements by targeting only poles at high risk of failure under environmental stress. For some power utilities, this results in over \$5 million in savings annually. By focusing on the higher risk poles and saving the utility millions, out of every 10,000 poles, the probability of discovering a high-risk pole improves by about 30%. The results show an improvement to current asset management strategies.

In all, the objective of this PhD research is in optimizing electric utility decisions on the inspection and maintenance of environmentally-stressed aging power system components, accounting for uncertainty in diagnostics, failures and unit cost variables that influence life cycle cost model of the components. The current utility strategy was found to be inefficient in both cost and discovery potential of high-risk components. It lacks risk-based analysis of the failures and consequences of inaccurate diagnostic actions on the components studied. As part of the research, useful data on diagnostic recommendations were statistically analyzed, novel failure prediction and estimation approaches were developed, and management decisions were stochastically optimized. The optimum decision ensures as much economic and reliability benefits as realistically possible from adopting a preventive maintenance program within some specified planning horizon.

1 INTRODUCTION

The deregulation of the electric power industry increased competition between power utilities to sell electricity to consumers at relatively low rates, causing the utilities to seek cost reductions in other areas of their business including in the management of their assets. A significant number of electric utility assets are run to failure: that is, are utilized in service until they fail. Only then are they replaced. With relatively slow rates of failure in normal environments, the industry has seen a growing amount of aging infrastructure that some reliability engineers believe will fail to meet power demands in the coming years. Fear of the outlook of power infrastructure around the United States in the next few decades has prompted much discussion as seen in [1], [2].

However, fear and criticism alone are unlikely to change the traditional electric utility behavior on managing their assets. An understanding of the maintenance investment-payback, on the other hand, might. (It should be noted that the terms utility, power utility and electric utility will be used interchangeably in this dissertation.) Utility managers assume that the cost of preventive maintenance far outweighs run-to-failure cost, especially when the majority of failures of some groups of the utility components occur only during natural disasters like earthquakes, hurricanes, ice storms, and other similar events. Nonetheless, below-par performance of utility assets during these disasters sometimes attracts the scrutiny of public service commissions (PSC's). In the event that performance is lower than expectations, asset management audits are performed. The outcome may include mandatory regulations that could prove highly cost-ineffective for utilities. Thus, pre-emptive cost-effective strategies should be sought to avoid any hints of negligence in management of assets. Such strategies would also promote public safety and judicious use of annual budgets, even following inevitable natural disasters.

The year 2005 was a tragic year for power utilities in the wake of two powerful hurricanes: Katrina and Wilma. Florida Power & Light (FPL) was one such utility. It faced both public and government criticism in light of the damages to their assets and the resulting loss of power to millions of customers, which took weeks to restore. Customer disapproval and general concern over performance caused the utility to hire an external consulting company to evaluate its management methods and prove that it had performed within electric reliability standards and codes. While FPL showed it performed reasonably, the utility now faces the mandate to tighten its preventive maintenance program. Among other things, FPL is now required to inspect each of its wood poles, a subset of its infrastructure that suffered greatly, every eight years. Meanwhile, most other utilities, which apply regular inspection schedules, inspect their wood poles on an average of about 10 - 15 year schedules.

Concern over the impact of severe weather to power utilities did not start in 2005. Prior to 2005, damages caused by storms to power utilities have been recorded and analyzed by researchers and consultants in the power field. The Edison Electric Institute published a report in 2004 with a listing of utility losses and outage durations caused by past storms. For instance, according to the report, a 2002 North Carolina ice storm cost Duke Power about \$87 million, affected over 1.3 million customers, and caused a 9-day outage. Hurricane Fran of 1996 destroyed about 5,500 poles supporting overhead lines; 2,800 transformers; and about 3,000 miles of wire [3]. Though utilities replace failed assets and lose revenue during these extreme weather events, they are typically fined only if found negligent in observing reliability and safety codes. Environmental hazards (storms for instance) events are normally treated as “acts of God.” Such events are currently not included in reliability indices used to measure trends in quality of service of the utilities.

Given that management practices on components are relatively flexible within reasonable limits: inspection schedules in most states are unregulated, it is in the best

interest of electric utilities to define a cost-effective approach to strengthening their infrastructure and reducing asset loss and failure risk even under severe weather conditions. As the infrastructure ages, asset failure under naturally hazardous conditions will shift from the safe net of “acts of God” event classifications to debates on negligence. The likelihood of hefty fines and penalties will increase with the latter. Major areas of component management that require substantial analysis to increase effectiveness of maintenance operations against fundamentally corrective approaches include:

- Failure prediction,
- Failure prevention, and
- Scheduling of preventive measures affecting lifecycle costs.

The problem statement is defined in the following section.

1.1 Problem Statement

The three areas listed constitute interesting research in component management literature. They affect expenditures, performance and potentially public safety. It is important to mention that the field of asset management is very vast and complicated: there are many assets, maintenance operations and lots of uncertainty, with often little data. The next step in this research was naturally to define the asset management problem in a manageable way.

Electric power utilities are very capital-intensive, and the bulk of their capital is in their assets. The Edison Foundation puts the total value of generation, transmission, and distribution infrastructure in the United States to about \$440 billion, with capital expenditures expected to exceed \$70 billion as of 2007 [4]. Diagnostic and replacement procedures for each type of utility equipment are different; failure assessment also differs by equipment. For this reason, this work focuses on generating *frameworks* for effective management of certain groups of utility equipment rather than procedure for specific

equipment. Qualifying equipment include those that are vulnerable to natural disasters, *may* not be individually critical to power delivery when they fail, but are in such large use in service that their simultaneous failures would cost the utility millions of dollars in replacement and lost revenue. Such equipment includes supporting infrastructure like some transmission structures, wood poles, and overhead lines, among others.

In tackling issues of failure prediction, failure prevention and scheduling in power utilities, it became obvious, though not surprisingly so, that key data required for analysis from the utilities were incomplete. The utility is not in the habit of recording and storing data. This may partly be because utility managers are unsure what constitutes relevant data for decision-making, and lack expertise, labor and time to deal with the data. Researchers sometimes assume that the data is complete. However, by assuming completeness, results of analysis are misinterpreted. As part of this work, the effects of making decisions with incomplete data in failure prediction and budgeting are analyzed. The uncertainty in the data and developed failure models represents some of the stochastics applied in the work.

The use of algorithms, especially complex algorithms that integrate stochastics, in decision-making is very common in the financial sector: trading, investments, stock broking. In the past few decades, stochastic optimization algorithms have also been applied in the power sector: hydrothermal capacity scheduling, energy planning, distributed generation, energy trading, among others. The algorithms have been used in these fields because the underlying issues are too complicated to be handled by the human brain. Also, the decisions are so critical that they could result in hundreds of thousands and possibly millions of dollars lost daily or weekly. While the issue of component management may be less financially critical than say energy trading, the drudgery of complex analytics could still result in annual savings of millions of dollars to the utility from their current practices. These savings could be applied towards other

utility needs. So, the “complex” algorithms are worth the investment. This dissertation work shows that.

The objective of this PhD research is to optimize power utility decisions on the management and maintenance of environmentally-stressed aging power system components, accounting for uncertainty in variables that influence annual utility budgets on the components. Useful data from utility inspection databases are extracted and statistically analyzed. The databases provide information on the current practices of the utility and reflect the potential for improvement. Novel failure prediction and estimation approaches are developed for power system components using the risk of exposure to damage and the consequence of aging. Also, utility annual investments on component management are stochastically optimized relative to a traditional run-to-failure management approach by solving for optimum decisions on inspection scheduling and preventive replacements of inspected components. Optimizations are carried out for one-year and then multi-year planning horizons. The benefits of the strategies developed will be observed both in economics and reliability (measured in the loss of assets using probabilistic models of environmental hazards in specific geographical locations of the components).

1.2 Literature review

1.2.1 Failure prediction

The plethora of work in reliability studies involving electric utility or power delivery components focus on time dependency: how the probability of failure changes with time. This is logical since the failure risk of most components increases with time. Time-dependent reliability functions are survivor functions, hazard functions, failure probability density functions, mean-residual life functions, and their derivatives. The functions are explained in great detail in [5].

Survivor functions represent the conditional probability that a component that has survived up to some instant in time will survive past that time. The failure probability density function (pdf) gives the probability that a component will fail at some time instant. Hazard functions are not probabilities but are defined as the ratio of the probability density function to survivor function of components. They show the evolution of failure risk of the components. Let the hazard function of a component with time be $h(t)$, the pdf be $f(t)$ and the survivor function be $S(t)$. Then, the hazard function is

$$h(t) = \frac{f(t)}{S(t)}.$$

Say the failure risk (hazard) of a component does not start to increase until time T . Then, utility maintenance actions on the components before time T will not be very helpful towards the failure risk of the population of the components. After time T , however, the hazard could be significantly reduced if the maintenance actions are effective. Obtaining the time T for some component constitutes some of the relevance of failure modeling.

Each reliability function: $f(t)$, $h(t)$, $S(t)$, can be derived from the other. Thus, if one function is known, others can be easily estimated. The survivor function is a complement of the cumulative probability function of failure (or time-to-failure, to be more precise). In reliability studies, the term “complete datasets” is typically used to describe datasets that consist of all times-to-failure of components, when all the components in the study have been observed to fail. This should not be confused with whether information is missing from a dataset. Censored datasets, on the other hand, refer to when a proportion of studied components have either not failed or have fallen out of observation by the time of analysis. The majority of failure datasets of groups of identical components are censored. This is because a large percentage of most types of utility components have neither failed nor been replaced since installation; hence, the problem and controversy of “aging infrastructure.” Reliability modeling can be used in analyzing, understanding, estimating and predicting failures of the components.

In [6], Begovic *et al.* discusses the use of data on times-to-failure of equipment to predict future failures. It was found that for the components analyzed, failure risk became evident (or started increasing) several years after installation of the components. Parametric modeling was used, as it often is, for failure prediction. A Weibull distribution was applied to the dataset and found to be a good fit. The Weibull function is actually a distribution frequently used to model failures of most components because its hazard function is flexible: can be constant, increasing or decreasing.

The two-parameter Weibull function is characterized by a scale β and a shape α parameter. The hazard function is constant when α is equal to one (becomes an exponential distribution), is increasing when α is greater than one, and is decreasing for α less than one. The scale parameter affects the width of a plot of the pdf of the Weibull distribution. The pdf $f(t)$ is shown below.

$$f(t) = \left(\frac{\alpha}{t}\right) \left(\frac{t}{\beta}\right)^{\alpha} e^{-\left(\frac{t}{\beta}\right)^{\alpha}}.$$

For α greater than one, an increase in β implies a right-shift of the peak of the density function.

In [7], Ng constructs prediction intervals for mean lifetime and intervals for few missing times-to-failure of components when the failure dataset is complete. The work was done for a dataset that follows a two-parameter exponential distribution. That is, the dataset could be left-, right-, or doubly-censored dataset. The results presented by Ng were a generalization of a prediction distribution obtained by Leemis in [8], which finds predictions for a right-censored dataset. Knowledge of expected failure times of components helps in planning for maintenance, possible outages, and even budgeting.

The modeling of failures of power utility components under severe weather has been studied by a number of researchers. In [9], Zhou *et al.* apply Poisson regression and Bayesian network models to actual overhead line failure data for different weather conditions: thunderstorm and wind, in Kansas. The purpose of the paper included an

evaluation of weather impact on distribution lines and an estimation of failure rates under the conditions. The effects of lightning and winds on line failure were analyzed simultaneously by assuming that the counts of weather failure events followed a Poisson distribution.

Wind speeds of magnitudes 45 mph and higher were analyzed as a group, and prediction bounds for the estimates were obtained using a Monte Carlo simulation in [9]. Similar analysis was performed for medium voltage overhead lines and underground cables in [10], which uses a simple regression model for estimating outage rates for observed wind speeds. In [11], dynamic traveling wave techniques were applied in simulating lightning events on unshielded transmission lines, and then used to predict outage rates of the lines depending on the location of the lightning strike and the type of line.

Outage predictions of distribution and transmission networks under hurricane loads are obtained in [12]. Fragility curves: functions that relate probability of damage to the intensity of stress (hurricanes, in this case), are found for components of the power networks. The components include substations, utility wood poles, and transmission and distribution lines. Using the interconnectivity between substations and distribution load points provided by Texas electric power utilities, Monte Carlo methods were used to simulate hurricane fields and measure impact on spatially distributed outages for the entire system. The approach utilized led to a 15.59% outage prediction model error and a small standard deviation of 0.02%. The results were validated against historical observed results from Hurricane Ike of 2008.

The papers and works listed in this subsection provide some prior research into applications and approaches for failure modeling in reliability theory. Also, the use of fragility analysis is frequent in civil engineering for structural design, reliability assessment, or loss predictions from natural hazards. Novel approaches in parametric

modeling and interesting results on reliability for aging infrastructure are provided in this dissertation.

1.2.2 Diagnostics

Power system components undergo different types of condition monitoring during their lifetimes, depending on how critical the components are, regulations, whether the components are repairable, and the costs of diagnostics and maintenance. The accuracy of applied diagnostics affects the benefits of the resulting maintenance actions. For instance, say a component that is healthy and still has a long time-to-failure is replaced prematurely due to inaccurate diagnostic recommendations, the replacement is a waste of diagnostic investments. When there are several such components, the effect on the power utility becomes more severe. Avoiding such inaccurate diagnostic recommendations is key to managing components well. The concept of diagnostic accuracy is well-studied especially in the medical or clinical field (epidemiology). This is discussed in the following paragraphs.

In [13], the principles of receiver operating characteristics (ROC) are explained. ROC curves are often plotted to determine the accuracy of a diagnostic test in terms of its sensitivity (how well a diagnostic test detects a “disease”) and specificity (how well a diagnostic test does not misclassify a “non-disease” or “good health”). The term disease is used here because the majority of literature in this field is clinical. Disease implies a state of non-normalcy in a patient; it is equivalent to a component that is below reliability or safety standards for the purpose of this dissertation. In other words, a component below standards may be substituted for a patient with some disease in the medical research. An ROC curve is a plot of the true positive rate (TPR) against the false positive rate (FPR) of a diagnostic test, where “positive” is a disease or some damage state (below standard). Because TPR and FPR are within the range of zero to one, the area under a

perfect ROC curve is one. A perfectly accurate test has a TPR of one and an FPR of zero. Curves of real tests appear below the ROC of a perfect test.

Comparisons between samples of several patients known to be at low or high risk of some disease, disorder or illness may help show the characteristics of disease in a person. A gold standard is required to determine the accuracy of a diagnostic test; that is, to find out where the diagnostic decision or finding is the same as the findings of the gold standard. Realistically, there are often no obvious or perfect “gold standards.” The gold standard is determined based on knowledge and understanding of a good amount of historical data or evidence.

The accuracy of a diagnostic test can be determined from a long-term assessment of decisions of the test. Here is a simple illustration. The error rate of an over-the-counter pregnancy test may be determined by collecting data over time on the number of women who tested positive for pregnancy but showed no eventual signs of pregnancy. This would yield a false positive rate. Likewise, women who test positive and were confirmed pregnant through growth of egg(s), miscarriage or delivery, among other obstetric signs could be used to determine the true positive rate of the pregnancy test.

In the case of components that fail mechanically, the “gold standard” may be found through destructive testing. That is, assuming a component will fail under hurricane load because it is below standard, a destructive test simulating intensity of the environmental load would cause the component to break, become damaged or fail. Destructive tests are expensive and render components useless. This is why diagnostic tests are important. If a diagnostic test is accurate, it eliminates the need for a destructive test. However, a destructive test may be used in the front end to determine the accuracy of the diagnostic test rather than wait long periods of time to assess accuracy. In such cases, the diagnostic test is as accurate as the number of component health classifications that match those of a highly accurate destructive test. Destructive tests were performed on fiberglass distribution poles in [14] to determine the strengths of the poles. In [15], the

variability and bias that may come from using a few samples or components for assessing diagnostic test accuracy are discussed.

The accuracy and validity of diagnostic testing may be increased by combining classifications from multiple tests. In [16], by assuming that the costs of misclassification of diagnostic tests are equal, binary regression, linear (LDA) and non-linear discriminant analyses, decision trees, Bayesian and neural networks are mentioned as possible statistical approaches used to determine the level of accuracy achievable by combining diagnostic tests. LDA results in a new scale of analyzing diagnostic test scores; linear combinations of the scores may be difficult to interpret in the field. Other papers on combining diagnostic tests to improve accuracy include [17] and [18].

1.2.3 Economics of Management of Power Utility Components

The lifecycle cost assessment of components involves purchase, installation in service (into the power network), operation, possible inspections (diagnostics) and maintenance, and then end-of-life. End-of-life typically implies replacement of the component, either correctively or preventively. “Correctively” means a failure occurred while “preventively” means replacement is based on diagnostic recommendations. Research on lifecycle costs (LCC) of power utility components constitute the estimation of expected cost rates: the ratio of expected cost during the life of the component to the expected lifetime of the component. The latter is estimated using reliability or survival analysis. The expected cost is partly based on choice, like the frequency or probability of inspection of a component, and partly on prevalence of failure risks or causes.

Age replacement policies are common for LCC analysis of components. Sometimes, the cost of purchase is neglected since it is a sunk cost and will not affect decisions on component management. Other times, it is not. In [19], cost models are developed for replacements of components with warranty and those without. Optimal replacement ages are found for the components using a calculus-based approach on the

lifecycle cost function. The results show that the optimal replacement age policy for the preventive approach should be adjusted based on warranty to exploit possible cost rate reductions.

In [20], the recommendation is to compare optimal replacement decisions based on these cost criteria: using the expected average cost per unit time (averaging costs over an unbounded horizon applying limit theory), or using expected discounted costs (cumulative net present values of lifecycle costs). The decisions of the optimal replacement age may differ depending on the cost criterion of choice. There is no clear explanation of which criterion is preferable. An illustration of the application of the models is done for the maintenance of a cylinder on an existing swing bridge. Given the formulation of the failure and cost models, the plots of the cost functions were convex showing a clear age replacement intervals, within which the costs were minimum. The optimum interval differed for each cost criterion.

Heuristic optimization approaches have been applied in decision-making concerning power utility components. In [21], a genetic algorithm (GA) was developed for selecting the appropriate maintenance strategy for a substation. The approach was used rather than a traditional LCC analysis because of the large and diverse number and ways of making decisions affecting the substation. One decision is the choice of an air-insulated substation (AIS) over a gas-insulated substation (GIS) since the AIS is more easily maintainable than a GIS or vice versa. The GIS is of course more reliable. However, the the expected lifetimes of substations can be assumed to be a function of the interval of maintenance. Thus, the choice of maintenance, inspection cycles, expected lifetimes, type of substation, interconnectivity between substations, costs of maintenance are just a few parameters influencing the maintenance decisions for the GA approach. The mutations and crossovers search for combinations of the choices that give sufficiently small LCC's, where the LCC is the fitness function.

Since the parameters that cause the lifecycle cost of a maintainable power utility component are subject to change in each year and are often times uncertain, stochastic optimization approaches are better suited to obtaining risk-informed decisions. For instance, the occurrence of natural hazards in a geographical location is unpredictable. Likewise, the damage caused to vulnerable components is also uncertain and likely dependent on the intensity of the hazard. Applying stochastic optimization approaches rather than deterministic ones provides information of the statistics of objective cost functions. One such piece of information would be the probability that the objective function would exceed a certain value of interest: the amount of risk a manager is willing to take in implementing a decision.

Stochastic programming methods are often applied in the field of finance as seen in [22]. In [23], the methods were applied for optimization purposes in control systems, signal processing, telecommunications (fields of electrical engineering), and stochastic finance. When the decisions in future periods or years are dependent on the present or past periods, multi-stage optimization techniques are applied. This is as seen in [24] for decision-making in hydrothermal capacity generation.

A stochastic optimization approach is applied in decision-making for the management of environmentally stressed aging utility components since the parameters influencing annual costs of management are largely stochastic in nature. These parameters include vendor service costs, component failure probabilities and accuracies of diagnostic classifications. To ensure that optimum decisions on preventive maintenance are carried out with regards to reliability and cost, numerous possible scenarios of combinations of parameters are simulated using stochastic sampling approaches.

1.3 A new approach to component management

Novel approaches developed in this dissertation work for the three major sectors described in the prior subsection are summarized here. The approaches and frameworks are of intellectual merit to the field of power utility asset management should provide motivation for electric utilities to modify their current management strategies.

1.3.1 Failure prediction

Maximum likelihood estimation is one popular technique for parametric modeling of failure rates of components using observed failures. In this dissertation, a Bayesian approach is applied in predicting the distribution of future times-to-failures of surviving components even when the rate parameter of an exponential failure dataset is unknown. The prior density function is however assumed to be a known Gamma density, which is the conjugate prior of the exponential likelihood function. Simulations show the approach performs well even for a small proportion of failures from a large population of identical components. However, it is limited by the choice of the prior density function of the unknown rate parameter. That is, the approach becomes inconvenient when the prior density function is not a conjugate prior of the exponential likelihood function.

The relative simplicity of the exponential assumption lays some groundwork for generating the more realistic bathtub hazard curve of most components, even without using a Bayesian approach on observed data. The failure rates of components were estimated gradually per period (say, years) as though the failure dataset was exponential. The approach is applied in real-time detection of an effective time in the life of a population of components to begin inspections and maintenance operations. The appropriate time is logically at the onset of increasing failure risk in the population. The approach uses both Kalman filtering and extreme value theory for decision-making. Another Kalman filter application is discussed in [25] and the theory of extreme values for detecting novelties in signals is discussed in [26].

While conventional research on reliability of components uses time (or age) as the dominant factor for observing the evolution of failure risk of the component, history has shown that environmental stress plays a strong role in the failure rate of components. Fragility assessment adapted from civil engineering is employed in modeling failure rates of power utility components vulnerable to natural hazards like hurricanes. Because the power utility infrastructure is aging, for the first time, at least based on literature searches, a framework is developed for estimating age-dependent fragility of vulnerable utility components. This is extended to formulating information-driven strategies for scheduling and prioritizing the inspections of environmentally affected utility components.

1.3.2 Diagnostics and preventive replacements

While the cost of misclassification of diseases by diagnostic tests in clinical trials may cause death (false negatives) or raise false feelings of fear (false positives), the costs are more financial in the testing of electric utility components. Depending on the kinds of components, the costs of corrective replacements for false negatives may be upwards of three times the cost of a preventive replacement for a false positive. The use of economics rather than just statistical approaches for ranking the effectiveness of competing diagnostic tests or combining tests to increase accuracy is more beneficial for interpretation and applicability. For instance, the use of linear discriminant analysis, a statistical approach used in medical research for combining competing diagnostic scores generates a new scale, different from that of each diagnostic test. This is often difficult to implement in the field or explain in simple terms.

In addition, a different turn is taken in this dissertation on the justification of acquiring information on the accuracy of diagnostic tests. Here, analysis of how information gained on accuracy affects maintenance-investment planning is performed. The costs and benefits of short-term and long-term assessment of diagnostic accuracy are

quantified in terms of whether a power utility would save by switching diagnostic testing companies based on accuracy. If the diagnostic accuracy of the current testing company is high, then assessment of diagnostic accuracy may appear unnecessary but the confidence in the choice of the diagnostic test would be high. On the other hand, if the accuracy is found to be low, the power utility would be equipped with information to seek a more accurate diagnostic testing company.

Economics is highlighted here as a strong decision-making tool for selection of diagnostic testing methods used in determining whether to perform preventive replacements of components diagnosed as below standard. It is a very practical tool for decision-making in business.

1.3.3 Economics

Penalties, reliability indices and management costs play roles in choosing maintenance strategies for components. Utilities try to avoid penalties, and United States reliability indices are generally high but vary with utilities. Thus, in literature, cost functions are defined as objectives to obtain the optimal replacement, warranty, and maintenance age in age-based. It should be noted that utility reliability indices exclude storm days considered “acts of God.” Utilities are often not penalized for energy not served during periods of natural hazards. The cost functions minimized are usually the expected lifecycle costs of components with respect to their expected lifetimes. However, lifecycle costs may not be practical to asset managers, who may only serve in their positions for a shorter amount of time than the expected lifetime of the class of components. It is therefore important to analyze optimum inspection and preventive replacement decisions based on *annual* costs to the utility for managing a population of components rather than lifecycle costs.

A few lessons can still be learnt from lifecycle cost (LCC) analysis. For one, as discussed in [27], acquisition costs can be left out of LCC analysis since they do not

affect future costs or decisions. For the same reason, they will be left out of the analysis for annual utility costs resulting from component management. Other costs that are candidates for analysis include:

- Annual lost revenue from energy not served due to natural hazards,
- Annual corrective replacement costs,
- Annual preventive replacement costs, and
- Annual inspection costs.

You may observe that switching costs are neglected. This is because realistically, when severe storms occur, it is possible for all lines to and around feeders to be damaged, making switching impossible. The analysis in this dissertation neglects lost revenue because it is complicated and difficult to isolate lost revenue caused by one class of components from another.

A segment of the cost effectiveness analysis presented in this dissertation is based on the comparison between the annual cost of a run-to-failure program and a preventive maintenance program. Lost revenue from severe weather for both programs may not be substantially different. This provides another argument for neglecting lost revenue in the cost-effectiveness analysis between the two management programs and optimization of the cost function defined in the work.

In this dissertation, Monte Carlo (MC) methods are chosen to simulate numerous scenarios of varying parameters in costs, diagnostic accuracies and component failure risks. The probability density functions of the parameters are defined to be discrete. The use of MC is also much simpler to understand than the more complicated decision mechanisms of say genetic algorithms. Because the objective cost functions are non-convex in the decision parameters (either an inspection cycle or the proportion of components to replace preventively annually or both), random search is used as the relatively simple stochastic optimization algorithm in the work.

1.4 Navigating the dissertation

The dissertation begins, after the introduction, with a statistical analysis of preventive replacement data records on components kept by electric power utilities. It then proceeds in the next chapter to deal with modeling and predictions of failure rates of components. Diagnostics and preventive replacements are studied in the following chapter, before the stochastic optimization of the maintenance process. The dissertation ends with listings of intellectual and practical contributions of this work to the field of asset management, and finally references for works cited in the body of the dissertation.

1.5 Summary

Failure prediction, failure prevention through diagnostics and preventive maintenance operations, and optimization of management decisions continue to be strong subjects of research in asset management. The increasing average ages of power infrastructure are a cause for concern on future electric safety and corrective replacement costs of utility equipment especially following severe environmental hazards. If it is assumed that the hazards become more severe with time based on the concept of global warming and observing worsening trends of time-to-severe storms, the average annual failures of exposed aging components may also increase significantly. Pre-emptive actions are necessary to mitigate and control these likely failures.

The models and frameworks developed in this dissertation are important in learning conditions of systems of identical components and making cost-effective decisions on managing the components. Novel failure estimation, failure prediction, diagnostic decision and annual cost optimization models are developed. Simulations of the proposed method shows a superiority of at least \$5 million annually for the

environmentally exposed utility components studied, over the current utility strategy on maintenance.

In the conclusion of the research, information gained and processing completed on diagnostics and failure data are combined to present electric utilities with an optimum practical solution to managing the set of infrastructure that is aging more rapidly in recent years and are under environmental stress.

1.6 REFERENCES

- [1] H. L. Willis et al., *Aging Power Delivery Infrastructures*. New York, U.S.A.: Marcel Dekker, Inc., 2001.
- [2] L. Philipson and H.L. Willis, *Understanding electric utilities and deregulation*. Florida, U.S.A.: CRC Press, 2006.
- [3] B. Johnson, "Utility storm restoration response," Edison Electric Institute, U.S.A., Jan. 2004.
- [4] M.W. Chupka and G. Basheda, "Rising utility construction costs: Sources and impacts," The Edison Foundation, Washington, D.C., U.S.A., Sept. 2007. <
http://www.eei.org/ourissues/finance/Documents/Rising_Utility_Construction_Costs.pdf>
- [5] L. M. Leemis, *Reliability: Probabilistic Models and Statistical Methods*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1995.
- [6] M. Begovic, P. Djuric, J. Perkel, B. Vidakovic, and D. Novosel, "New probabilistic method for estimation of equipment failures and development of replacement strategies," in *Proceedings of the 39th Annual Hawaii International Conference on System Sciences, 2006. HICSS '06.*, Hawaii, 2006.

- [7] V. Ng, "Prediction intervals for the 2-parameter exponential distribution using incomplete data," *IEEE Transactions on Reliability*, vol. R-33, no. 2, pp. 188-191, 1984.
- [8] J. F. Lawless, "Prediction intervals for the two parameter exponential distribution," *Technometrics*, vol. 16, pp. 341-344, May 1974.
- [9] Y. Zhou *et al.*, "Modeling weather-related failures of overhead distribution lines," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1683-1690, Nov. 2006.
- [10] P. Carer and C. Briend, "Weather impact on components reliability: a model for MC electrical networks," in *Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems*, Puerto Rico, pp. 1-7, 2008.
- [11] A.C. Liew and M. Darveniza, "Calculation of the lightning performance of unshielded transmission lines," in *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, no. 6, pp. 1471-1477, June 1982.
- [12] J. Winkler *et al.* "Performance assessment of topologically diverse power systems subjected to hurricane events," *Reliability Engineering and System Safety*, vol. 95, no. 4, pp. 323-336, Apr. 2010.
- [13] C.E. Metz, "Basic principles of ROC analysis," in *Seminars in Nuclear Medicine*, vol. 8, no. 4, pp. 283-298, Oct. 1978.
- [14] M.F. Miller, "Fiberglass Distribution Poles-A Case Study," *IEEE Transactions on Power Delivery*, vol. 10, no. 1, pp. 497-503, Jan. 1995.

- [15] A. Alperovitch, "Assessment of diagnostic tests," in *Bulletin du Cancer*, vol. 67, no. 4, pp. 378-383, 1980.
- [16] M. Pepe and M.L. Thompson, "Combining diagnostic test results to increase accuracy," *Biostatistics*, vol. 1, no. 2, pp. 123-140, 2000.
- [17] M. McIntosh and M. Pepe, "Combining several screening tests: optimality of the risk score," vol. 58, no. 3, pp. 657-664, Sept. 2002.
- [18] J.Q. Su and J.S. Liu, "Linear Combinations of Multiple Diagnostic Markers," *Journal of American Statistical Association*, vol. 88, no. 424, pp. 1350-1355, Dec. 1993.
- [19] R.H. Yeh *et al.* "Optimal age-replacement policy for nonrepairable products under renewing free-replacement warranty," *IEEE Transactions on Reliability*, vol. 54, no. 1, pp. 92-97, Mar. 2005.
- [20] J.M. van Noortwijk, "Optimal replacement decisions for structures under stochastic deterioration," in *Proceedings of the Eighth IFIP WG 7.5 Working Conference on Reliability and Optimization of Structural Systems*, Poland, pp. 273-280, 1998.
- [21] M. Hinow, "Substation maintenance strategy adaptation for life-cycle cost reduction using genetic algorithm," *IEEE Transactions on Power Delivery*, vol. 26, no. 1, pp. 197-204, Jan. 2011.
- [22] A. Collomb, "Dynamic asset allocation by stochastic programming methods," Ph.D Dissertation, Dept. Mngt. Sci. and Eng., Stanford Univ., Stanford, CA, 2004.

- [23] M. Milisavljevic, “Information Driven Optimization Methods in Control Systems, Signal Processing, Telecommunications and Stochastic Finance,” Ph.D Dissertation, Dept. Elect. Eng., Georgia Inst. of Tech., Atlanta, GA, 2001.

- [24] M.V.F. Pereira and L.M.V.G. Pinto, “Multi-stage stochastic optimization applied to energy planning,” *Mathematical Programming*, vol. 52, no. 1-3, pp. 359-375, 1991.

- [25] H. Lee and S. J. Roberts, “On-line novelty detecting using the Kalman filter and extreme value theory,” in *19th International Conference on Pattern Recognition 2008. ICPR 2008*, Tampa, FL, pp. 1-4, Dec. 2008.

- [26] S. J. Roberts, “Novelty detection using extreme value statistics,” *IEE Proceedings on Visual Image Signal Processing*, vol. 146, no. 3, pp. 124-129, June 1999.

- [27] *Life-Cycle Costing Manual for the Federal Energy Management Program*, Handbook 135, National Institute of Standards and Technology (NIST), Gaithersburg, MD, 1996.

2 RELIABILITY-TYPE MODELING OF COMPONENTS IN ELECTRIC POWER UTILITIES USING PREVENTIVE REPLACEMENT DATA

Introduction

Power system components can be removed from service for different reasons. For instance, components installed on rights of way may be replaced for reasons of construction or vehicular accidents. In terms of the management of components, components can be said to be replaced correctively if they fail abruptly while in service, or preventively based on some diagnostic recommendation. The ability to model replacements of components for any of the listed (or other) reasons depends on availability and access to replacement data.

A thorough dataset for modeling reliability consists of an inventory of groups of identical components owned by a power utility, times (years) of installation of the components, times of replacement, vivid descriptions of the reason(s) for replacement, among others. The reason for replacement, especially whether preventive or corrective, is a good attribute to use for classification in replacement modeling of components. Reliability modeling is the term used when replacements are corrective, describing failures of components.

In the absence of failure data, reliability-type functions may still be found for the components using age information from diagnostic recommendations. The information may be stored in what some utilities call inspection databases of the components. It is important that the functions found for preventive replacements not be substituted for corrective replacements as has been done by some researchers in the past. The functions may instead show hazards of age-based diagnostic replacements of the components or deterioration trends of the components.

In this chapter, preventive replacement data for components in power utilities are analyzed using conventional reliability models (survivor functions, hazard functions etc). These data were provided by the utilities in the hope that the resulting models would be conclusive and informative about the failures of the components. The models, if accurate, would then be useful for failure prevention and prediction. Investigations showed that the datasets were incomplete, that is, that information on only a small percentage of identical components owned by the utilities (those inspected) was recorded. The larger percentage of the components has not yet been inspected and thus, did not provide any replacement information. Out of the components inventoried, several of them lacked the year of installation or reason for preventive replacement. The incomplete data would of course create a lack of confidence in generated models. Unless otherwise specified, the reliability functions in the “Sample Results” subsections of the chapter are not from failures, which are ideal, but from preventive replacements, since age-failure data were unavailable from the utilities.

As a part of this chapter, comparisons are made between reliability functions obtained from a component inspection database and those obtained from a hypothetical ideal failure database. The latter are generated by using realistic accuracy-parameters of diagnostic recommendations on the components. The comparison will be used to show the extent to which wrong decisions can be made when preventive replacement data are used in place of corrective replacement (or failure) data.

2.1 Reliability-Type Analysis of Preventive Replacements of Components from Utility Inspection Databases

A complete failure dataset would contain ages of all failed components as well as ages of surviving components. Then, by grouping the ages and applying the Kaplan-Meier product limit estimator, approximate survivor functions (and related lifetime functions like hazard and failure probability functions) are obtained. Identical

components are grouped for analysis. This means conductors are analyzed differently from transformers or generators, for instance. The times-to-failure of the components are assumed to be independently and identically distributed (i.i.d).

Here, because preventive replacements are the parameters of interest, the term “times-to-failure” in this section will be changed to “times-to-replacement,” where replacement means preventive replacement. The terms are used interchangeably in this chapter unless otherwise mentioned. Then, using the Kaplan-Meier product-limit estimator method explained below for the reader’s ease, the survivor function is obtained from component inspection databases.

2.1.1 Applying the Kaplan-Meier Method to Preventive Replacement Data

The ages of components inspected in some year I can be found as the positive difference between the year of inspection of each component and the year of inspection. If the component is scheduled for replacement during the year of inspection, the age at replacement can be easily extracted using any data-extraction software. We assume that the risk of diagnostic replacement is similar across all vintages (years of manufacture or installation) of the components. This prevents sparseness or thinness of datasets, especially since the data that will be extracted is really not of the component but of diagnostic decisions. Assuming there are k distinct ages at replacement from a discrete distribution, $y_1 < y_2 < \dots < y_k$, the conditional probability or hazard function $h(y_j)$ can be interpreted as

$$h(y_j) = P[T = y_j | T \geq y_j],$$

where T is the lifetime of the component (as caused by preventive replacement and not failure). Let $R(t)$ be a risk set for all components at risk just before some time t. Then, for the complement $R(t)^c$, the survivor function S(t) can be written as

$$S(t) = \prod_{j \in R(t)^c} [1 - h(y_j)].$$

Since only a small percentage of the components are typically replaced preventively, the dataset of replacement ages may be called a right-censored dataset for reliability-modeling purposes. The likelihood function L of the lifetimes for a right-censored dataset, where only a proportion of components have failed or are replaced, is

$$L = \prod_{i=1}^N \{f_T(t_i)\}^{d_i} \{S_T(t_i)\}^{n_i - d_i},$$

where f_T is the probability density function of lifetimes; n_i is the number of components in each distinct age group t_i , d_i is the number of observed replacements per age group, $S_T(t_i)$ is the survival probability of components of age group, and N is the total number of identical components in the inspection database.

By maximizing the log-likelihood function, an estimate of the hazard function is the ratio of d_j at y_j to the number of components still surviving just before time y_j . Using the estimate, the estimate of the survivor function $S(t)$ becomes

$$\hat{S}(t) = \prod_{j \in R(t)'} \left[1 - \frac{d_j}{n_j}\right].$$

An estimate for the variance of the estimate of the survivor function $V[S(t)]$, often known as Greenwood's formula, is

$$\hat{V}[\hat{S}(t)] = [\hat{S}(t)]^2 \sum_{j \in R(t)'} \frac{d_j}{n_j(n_j - d_j)}.$$

The asymptotically valid confidence interval for the survivor function is then

$$\hat{S}(t) - z_{\alpha/2} \sqrt{\hat{V}[\hat{S}(t)]} < S(t) < \hat{S}(t) + z_{\alpha/2} \sqrt{\hat{V}[\hat{S}(t)]},$$

where $z_{\alpha/2}$ represents the number of standard deviations from the estimate at the $(1-\alpha)$ confidence level.

The survivor function of populations of power system components found using the non-parametric method is terminated at the last observed replacement age, and does not have a value of zero at that point. Note that the survivor function of a set of

components that have all been replaced by the time of analysis has a value of zero at the last observed replacement age. The probability density of replacement per age is then estimated as a product of the hazard and survivor estimates at that age. An illustration of an algorithm that estimates the reliability-type functions of a power system component is in Figure 1. The inputs to the algorithm are the ages of the components, the number of components per age group and the number of components replaced per age group. The outputs are the estimates of the reliability-type functions: hazard, survivor and probability density function of replacements.

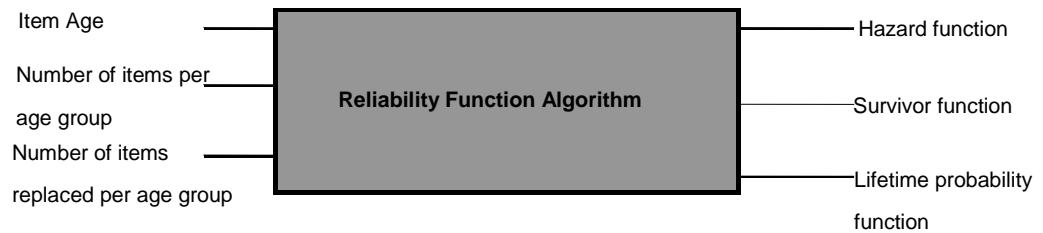


Figure 1 Black box showing inputs (left) and some outputs (right) of a reliability function algorithm.

The shape of the resulting hazard function gives an idea of how the perceived risk associated with the replacement (preventive or diagnostic replacement) of the component changes with time. Since non-parametric methods can only estimate lifetime functions up to the last replacement or failure age, parametric distribution-fitting methods are used in predicting those of incremental ages. Exponential and Weibull distributions are typical distributions used in lifetime modeling. Q-Q plots and goodness-of-fit tests such as Anderson Darling and Cramer von Mises tests can be used for fitting parametric models and estimating optimum distribution parameters for observed data. The expectation of the lifetime of the components may then be derived from the reliability-type functions.

Sample Results: Wood Poles

One set of power system equipment for which inspection data are kept and updated is utility wood poles. Utility wood poles support overhead distribution lines. Some utilities have wood poles that have been in service since the 1900's, though inspection databases may date back only to a decade or two in a number of utilities. While wood poles in some utilities receive remedial (in-service) chemical treatment to reduce fungal decay, those in other utilities do not. A power utility termed Utility D in this work for privacy purposes, which provided inspection data on its wood poles for analysis, applies remedial treatment to a portion of its population. Sample data from the populations of remedially treated and untreated poles are separated for analysis to illustrate interpretations of comparisons between survivor functions. The populations of the poles inspected and preventively replaced in Utility D over five years are given in Table 1.

Table 1 Population of wood poles inspected, treated, and preventively replaced in Utility D during years 2003-2007.

Description	Number of Poles
Number of poles inspected	222,284
Number that received remedial chemical treatment	47,894
Number that did not receive remedial chemical treatment	174,385
Number with unknown treatment history	5
Number of poles preventively replaced	6369

According to the table, only 2.87% of wood poles inspected in those years were diagnosed as weak or preventively replaced. This is a very small proportion of the nearly quarter-million wood poles that incurred diagnostic costs in those years. The age distribution of the inspected dataset is in Figure 2; separated into treatment populations in Figure 3.

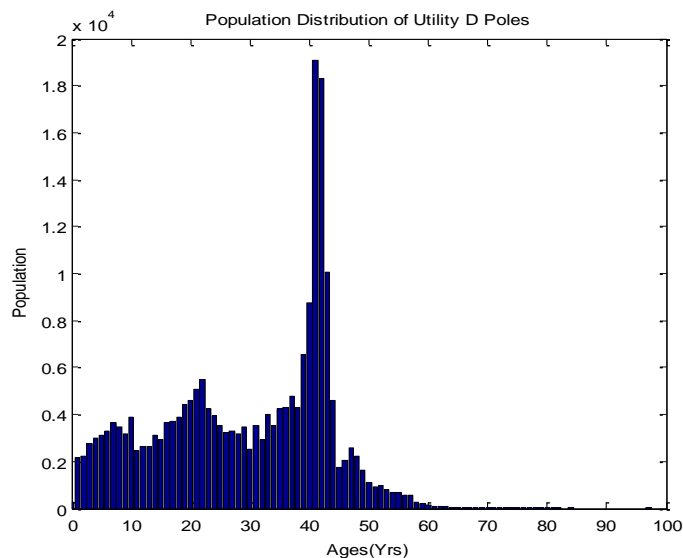


Figure 2 Population distribution of inspected Utility D poles.

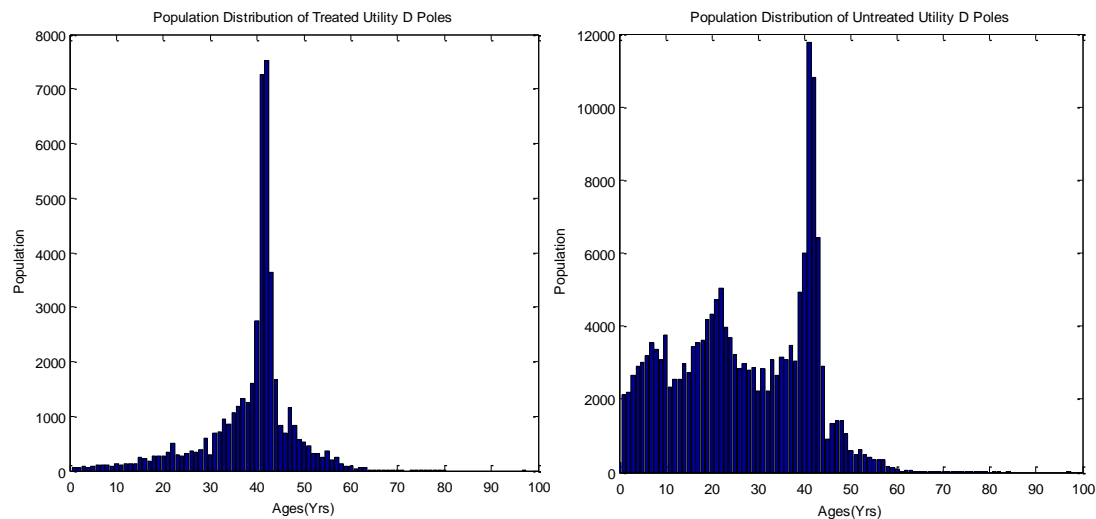


Figure 3 Population distribution of inspected Utility D poles, which received remedial chemical treatment (left) and did not receive treatment (right).

It can be speculated from the plots that either a large number of wood poles were installed in the 1960s' and/or a large number of those poles were inspected (with regards to the spike around 40 years). The range of populations in Figure 3 also shows that a larger proportion of wood poles have not received chemical treatment up to the current inspection year than those that have received treatment. The treated or untreated poles scheduled for preventive replacement are in Figure 4.

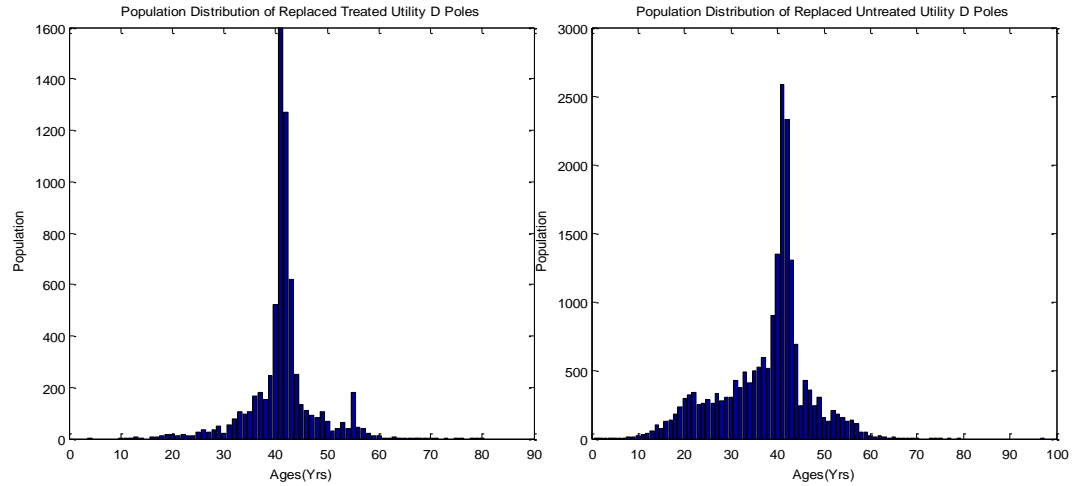


Figure 4 Population distribution of Utility D poles scheduled for preventive replacement, which received remedial chemical treatment (left) and did not receive treatment (right).

The survival functions of the population of wood poles are shown in Figure 5, and make it easier to compare the per-age replacement or survival of the poles. One survivor function appears above the other. The higher survival function is of the treated poles and shows that the risk of preventive replacement of treated poles is lower than the risk for poles that were not chemically treated.

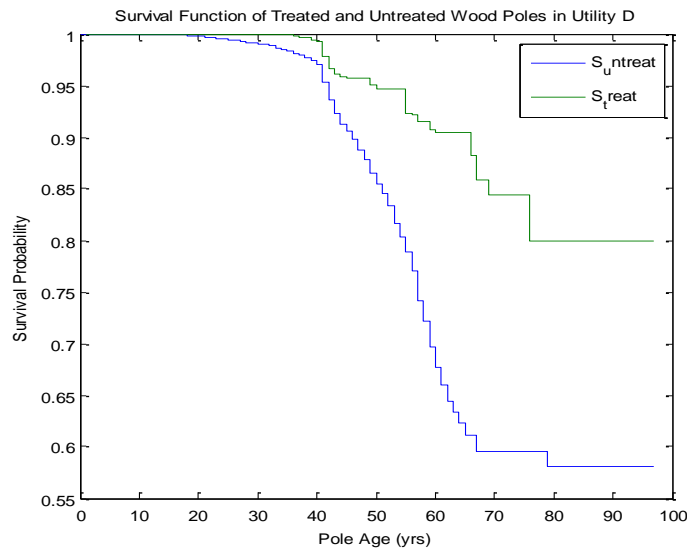


Figure 5 Survival Function of inspected Utility D poles, which received remedial chemical treatment (top) and did not receive treatment (bottom).

It is important to mention that the inspection company in this illustration schedules wood poles for replacement when the poles show extensive decay at the ground-line. Depending on the level of decay, approximate strengths of the poles are found. Poles are then replaced if the *computed* strengths fall below two-third of the original required strength of the poles. Research from National Electric Energy Testing Research and Applications Center (NEETRAC) has shown that such computations of strength are often inaccurate. This is why this chapter finds reliability-type models and does not assume that preventive replacement data can be used in place of corrective replacement data.

2.2 Comparing Reliability Models Generated from Power Utility Inspection Databases to Models from Failure Databases

As explained in the introduction, reliability-type models generated from utility inspection data are for preventive replacements; whereas, ideal reliability models would be from failure or corrective replacement databases. There are two major characteristics that are expected to cause discrepancies between reliability models obtained from either type of data. They are listed below.

- Missing data
- Diagnostic inaccuracy

“Missing data” means that the data corresponding to the inputs of the reliability algorithm of Figure 1 are not complete. That is, records of ages of the components in the inspection databases are incomplete, even of those scheduled for replacement. This encompasses loss of information and is expected to affect results of reliability-type analysis. The term “missing” is used instead of “completeness” since the latter is also used in reliability modeling to describe a dataset where all components studied have already been observed to fail. “Missing” here implies “not available.”

“Diagnostic inaccuracy” describes the false classifications of inspected components. Components that are good or reliable but are diagnosed as bad or unreliable are called false positives. Similarly, components that are bad or unreliable but are diagnosed as good are called false negatives. In the event of the maximum normal environmental loading in the neighborhood of the components, the false negatives will fail and incur high corrective replacement costs. The false positives receive premature replacements since they otherwise would not have failed if left in service. True negatives are good or reliable components accurately diagnosed as such. Likewise, true positives are bad or unreliable components accurately diagnosed as such.

Diagnostic inaccuracy can be measured in terms of diagnostic classifications of components (as good: “don’t replace” or bad: “replace”). It can also be measured in terms of some gold standard of the true condition of the component. An example of such gold standard could be determining whether a component fails a destructive test. The measures of inaccuracy using diagnostic classification are

- Negative predictive value (NPV) and
- Positive predictive value (PPV).

Using the gold standard of the true condition of a component, the measures of diagnostic inaccuracy are sensitivity and specificity.

Specificity is the ratio of true negatives to the number of good components. Diagnostic negatives represent inspected components left in service. These are components found to be above set reliability or safety standards. A high specificity implies that the number of good components inaccurately diagnosed as bad is relatively small. Sensitivity, on the other hand, is the ratio of true positives to the number of bad components. Diagnostic positives are scheduled for replacement. These components are those that are found inadequate in meeting set standards based on diagnostic procedures or measurements. A high sensitivity implies that a diagnostic test shows relative accuracy in detecting components that are really below par.

2.2.1 Metric for Quantifying Difference in Reliability Estimates of Components using Corrective or Preventive Replacement Data

One significant purpose of reliability models of components is for failure prediction and prevention. By finding the age at which the risk of failure of components starts to increase, an effective age-based strategy for running diagnostics and conducting preventive maintenance on components can be found. The inability of power utilities to provide complete failure databases but instead providing inspection records with information on preventive replacements is unfortunate. The results of reliability-type modeling of diagnostic replacements can be misleading in learning the evolution of failure risk of components. Regrettably, inspection companies and some researchers have used results from diagnostic replacements to describe failures of components. In this subsection, missing data and diagnostic inaccuracy will be simulated to quantify misinformation that may arise from using diagnostic replacement models.

Given known or expected accuracies (sensitivity or specificity) of a diagnostic test, the following metrics are developed for quantifying the dissimilarity between the true survivor function of a complete population of a vintage of components and that found using an inspection database (diagnostic replacement data from power utilities). Let Pop. 1 be a label for a complete population of components with available *failure* data. Also, let Pop.2 be a label for a subset of the population representing incompleteness alone (the small percentage inspected) and Pop.2.1 be the label that accounts for both incompleteness and inaccuracy in diagnostic records. Then, for each (sub-)set, Pop.1 through Pop.2.1, define a variable S_diff_j as follows.

$$S_diff_j = \sum_{i=1}^{length(S_1)} (1 - S_j(t_i))^2,$$

(1)

where j is one, two, or two point one for any of the three sample-labels; t_i is each distinct age group of either failure or diagnostic replacement in a population; value one represents the maximum value of a survivor function; and S_j is the survivor probability of the j -label sample set at age t_i .

Since the true survivor function of a component is that of Pop.1, the survivor function of Pop.1 is the baseline survivor function. The deviation $D_{S2.1}$ of the survivor function $S_{diff2.1}$ using a hypothetical inspection database from that of a failure database S_{diff1} can be found using the following formula.

$$D_{S2.1} = \frac{S_{diff2.1}}{S_{diff1}}. \quad (2)$$

The closer $D_{S2.1}$ is to value ‘one,’ the better the reliability-accuracy of the inspection database. If $D_{S2.1}$ is greater than one, it can be interpreted that the survivor function of the inspection database appears below the true survivor function—underestimates it. Thus, the survivor function is pessimistic, giving the impression that components are more hazardous than is true. This may lead electric utilities to replace more components than they should. Likewise, if $D_{S2.1}$ is less than one, the true survivor function is overestimated. This would cause utilities to replace less components than they should. In calculating the metric, when the last observed replacement age according to an inspection database is less than the last observed *failure* age, the expression $[1 - S_j(t_i)]^2$ in Equation (1) in the remaining age groups t_i is assumed to be one.

Sample Results: Hypothetical Failure Dataset

A population of wood poles gotten from a real power utility is analyzed. The age distribution of the poles is shown in Figure 6. The largest number of wood poles is of vintage (installation year) between 1965 and 1970. This may be a result of either large power-plant expansion during that period or the acquisition of a power plant that had

been operating for that long. The population is assumed to fail according to a Weibull density function with shape parameter 5.2 and scale parameter 107. The scaling was used because it is the age of the oldest wood poles in the population. The shape parameter was chosen to ensure an increasing hazard rate in the poles. This is shown in Figure 7.

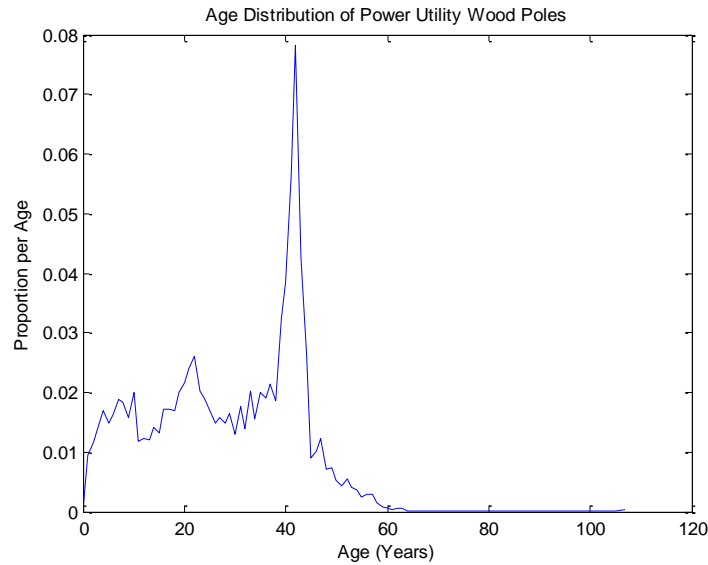


Figure 6 Age distribution of over 150,000 wood poles in a real power utility.

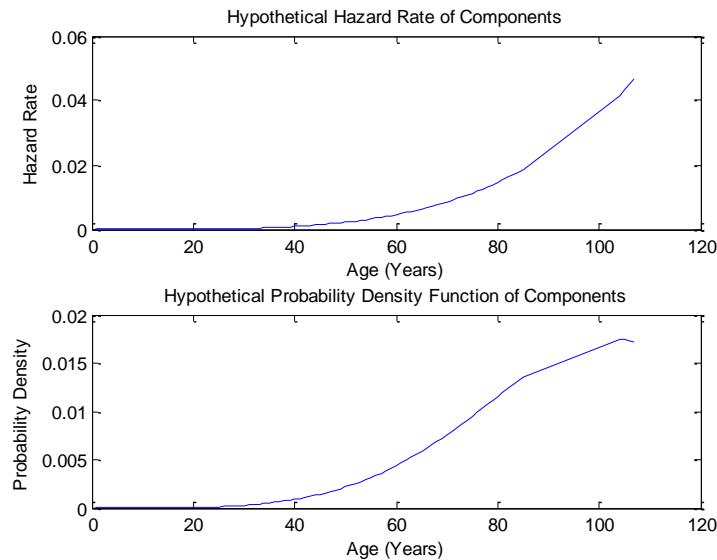


Figure 7 Hypothetical hazard rate and probability density function of real wood pole data under an assumed normal weather condition. Here, a Weibull density is used with shape parameter 5.2 and scale parameter 107. The pdf decreases for larger ages > 107 years.

The oldest 10% of the population is simulated to be inspected to coincide with a logistical methodology for inspecting wood poles in the country. This results in simulated inspections of age-44-and-up wood poles. In reality, there may be a sparse population of younger wood poles in the vicinity of the older poles. However, for ease of illustration and because such populations are likely relatively small, the younger population is ignored in this experiment. The sensitivity and specificity of a popular wood-pole diagnostic technique are used to estimate the number of wood poles diagnosed to be under specification. They are assumed to be the same for all ages of these poles were simulated to be replaced preventively. The sensitivity and specificity of the technique are 0.38 and 1.00 respectively.

Using the Kaplan-Meier approach, the number of poles estimated to fail $d(t)$ per age t for the Weibull hazard $h(t)$ and number of poles per age $N(t)$ is

$$d(t) = f(t) \times N(t).$$

Then, with test sensitivity as s_1 and specificity as s_2 , the number of poles, out of those inspected, estimated to be replaced $d_1(t)$ and those not replaced $d_2(t)$ is the inspected proportion of

$$\begin{aligned} d_1(t) &= [s_1 \times d(t)] + [N(t) - d(t)](1 - s_2), \\ d_2(t) &= N(t) - d_1(t). \end{aligned}$$

The empirical survivor functions of the poles using the failure counts and replacement counts from the Kaplan-Meier approach are estimated and the plots shown in Figure 8. Given the selection of inspected poles, the result of the reliability analysis shows that the survivor curve generated from preventive replacement data is optimistic relative to the true survivor function (*lower curve*). It appears above that of the failure dataset giving the impression that the poles behave better than they actually do.

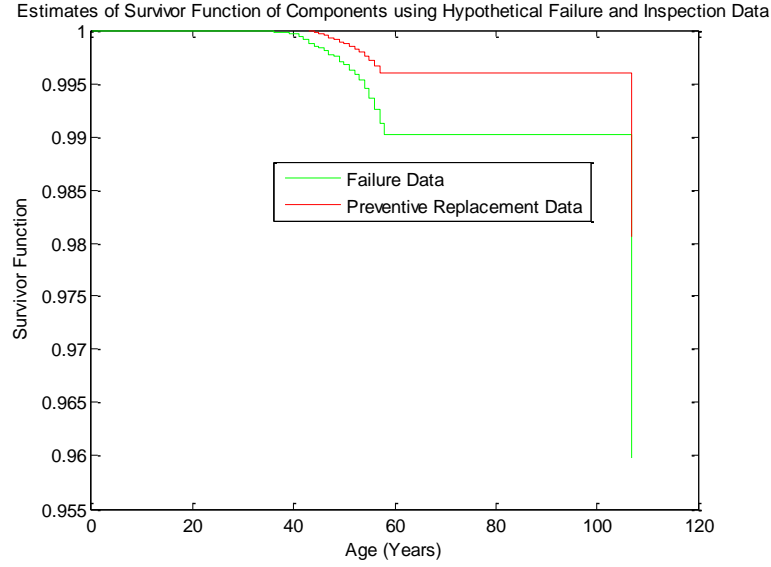


Figure 8 An empirical survivor function of wood poles estimated using a hypothetical Weibull failure density function to represent real failure data and a corresponding function estimated using preventive replacement data obtained from statistics of inaccuracy of diagnostic testing on the poles.

The metric developed in this section is used to quantify the difference between the two survivor functions. The parameter S_{diff_1} in Equation (1) is calculated to be 2.00^{-3} and $S_{diff_{2,1}}$ is 4.31^{-4} . The ratio $D_{S_{2,1}}$ in Equation (2) is found to be 0.21. It can be said that the survivor function generated using preventive replacement data deviates by over four-fifth of the survivor function generated using failure data. It represents almost five times the number of preventive replacements that utilities should have replaced to avert component failures. Assuming the cost of corrective replacements is three times that of preventive replacements, it translates into approximately fifteen times the cost of preventive replacements that would retroactively would spent in replacing failed poles.

The results of this experiment show why preventive replacement data from inspection databases should not be used to learn failure patterns of components.

2.3 An Ideal Failure Database for Components in an Electric Power Utility

Analysis of the hypothetical databases in the illustration of the previous section showed the insufficiency of reliability models generated from inspection databases for learning failure risks of power utility components. In the first section, an application of Kaplan-Meier or non-parametric methods to replacement datasets was illustrated. The work shows that for reliability analysis where age is used as the independent variable, the following properties need to be recorded for every component.

- Installation year,
- Boolean or other indicator variable identifying whether the component is still in service or has been replaced,
- The reason or mode for replacement or failure (corrective measure: broke unsuspectingly without environmental impact or broke as a result of extreme weather conditions; preventive measure: reason for preventive measure; other damage: vehicular accident, construction purposes), and
- Year or other time-unit of replacement.

Other properties of the components, like physical properties accounting for dimensions, geometry and the like, may also be recorded for classification purposes. However, for age-based reliability modeling, the listed properties are required.

Given these properties, the Kaplan-Meier method may be applied to the datasets using the algorithm of Figure 1. A right-censored dataset is expected, where a portion of the components is likely to be surviving at the point of analysis. Though provision of such an ideal database will be useful for modeling, it does not provide direct savings to the utility. It will instead help with making better decisions in failure prediction and prevention of components. The economics of inventory should be analyzed to find

adequate timing or scheduling of keeping records of all components, whether failed or not.

In the next chapter, innovative methods of failure prediction and estimation methodologies will be explored and used for practical decision-making strategies.

2.4 Summary

Electric power utilities rarely keep or update failure databases on their components. Databases generated by inspection companies are however more available and accessible. Performing reliability-type modeling using diagnostic replacement data from the inspection databases may show evolution of the risk of replacement but not necessarily failure. Missing data and inaccuracy of diagnostic recommendations (premature replacement of good components or undetected bad components) were used in simulations in this chapter to show erroneous decisions that can result from modeling with inspection data.

For an inspection database consisting of about 10% of a population of components, and diagnostic sensitivity of 0.38 and specificity of 1.00, the deviation between survivor functions generated from real failure data and diagnostic replacement data was about 80% using the developed metric S_{diff} . This study serves as a motivation to discourage electric power utilities from relying strictly on diagnostic data for making appropriate decisions on inspections of their assets.

3 MODELING FAILURE RATES

Introduction

Failures of power system components may lead to disruption of power-service to residential, commercial or industrial customers, depending on contingency design and whether the components are connected in series, parallel or some combination of either. The management of components affects the lifetimes of the components. As paraphrased from Leemis (1995), reliability is a term used to describe the probability that a component will perform satisfactorily whatever purpose it was installed for, for some expected amount of time and environmental condition. Electrical engineers typically model reliability of power system components with age as the independent variable. However, because some power system components fail mechanically and not electrically, stress-based reliability models are needed in estimation of their failure rates.

The simplest age-based parametric model assumes that the times-to-failure of a component are exponentially distributed. This implies a constant failure rate and is not common among components, which tend to have increasing risk of failure after a certain amount of time. Times-to-failure are random variables. Given well-known statistical tools, the distribution of times-to-failure of components that are still surviving by some present time instant can be obtained using already observed failures. However, the tools are often complicated and sometimes unsolvable analytically. It takes intricate design of the model to reach simpler analytical results.

In this chapter, Bayesian learning will be used to predict future failures probabilistically for an exponential rate assumption. Since it has been established that power utilities rarely keep or update failure data on their components, the effect of incomplete data on failure prediction is explored.

A growing risk of failures with age is expected for most components, implying a less simple failure model than an exponential distribution. The bathtub hazard curve is sometimes used to generalize the failure risk of components. It has three regions of risk: an early decreasing failure region (burn-in or early mortality), a constant failure region (random failures) and an increasing failure region (aging effects). The Bayesian process utilized for the exponential assumption is restricted to an assumed prior density function, and may not be effective or optimum for failure prediction in the non-exponential case. Instead, a new interval-based estimation method is developed in this chapter for obtaining the failure rate of components by assuming the failure rate is constant (exponential) for a small enough interval of time. This avoids cumbersome calculations of Fisher's information matrices and multiple parameters that must be found for parametric modeling of failures using mixtures of Weibull, lognormal, and other such densities.

The second major section of the chapter uses a mix of environmental stress and age rather than solely age as the independent variables in estimating reliability or the probability of failure. It involves analytical modeling of the age-specific failure of a component given the profile of the identified environmental stress. While environmental stress is aptly studied in civil engineering, it is the inclusion of aging in the fragility process that makes the approach very interesting, unique and applicable to the utilities' aging infrastructure. Together, the resulting age-specific fragility curves of components are used in developing a framework for geographical risk-informed prioritizing of annual inspections of sub-populations of components.

3.1 Age-Based Failure Estimation and Prediction

Under normal environmental conditions, failures may be highly attributed to aging effects of a component including accumulation of loading (mechanical, electrical, environmental, or otherwise). Data for normal or extreme weather failures or replacements from a power utility may be contrasted to show this. Discussions with

utility managers show that age-based failures in power utilities have been somewhat low in normal weather, but are expected to rise over the next few decades as the components installed at the beginning or middle of the twentieth century start approaching 70 years and above. With records of age-failure data, failure predictions can be carried out. We look first at an exponential case and then migrate to the more realistic non-exponential case of failure rates.

3.1.1 Exponential Case

Given k uncensored (observed) lifetimes $t_1, t_2, t_3 \dots t_k$ of a proportion of an entire population N of identical components by some present time instant T , their times-to-failure are assumed to be exponentially distributed. The rate parameter θ is unknown, but the times-to-failure of the remaining $N - k$ components are to be predicted from the known failure times. The probability distribution function (pdf) of the times-to-failure may be written as

$$p(t|\theta) = \begin{cases} \theta e^{-\theta t}, & t \geq 0 \\ 0, & \text{otherwise} \end{cases}.$$

Assuming independence and identical distribution (i.i.d) of the times-to-failure, the distribution of times-to-failure of the remaining components is the predictive distribution $p(t | D, t > T)$, where

$$D = \{t_1, t_2, \dots, t_k, t_{k+1} > T, t_{k+2} > T, \dots, t_N > T\} \text{ and}$$

$$p(t|D) = \int_{\theta} p(t|\theta, t > T) p(\theta|D) d\theta.$$

(3)

The conditional probability $p(t|\theta, t > T)$ is the exponential pdf of time t given the parameter θ and the condition of remaining lifetimes $t > T$, where T is the current time instant; $p(\theta|D)$ is the a posteriori density of θ given the failure information.

The first term in the integrand in (3) can be written as:

$$p(t|\theta, t > T) = \theta e^{-(t-T)\theta}, t \geq T.$$

The second factor of the integrand is defined by first specifying a prior of θ . Taking the prior $p(\theta)$ to be a Gamma density, which is a conjugate prior of the exponential likelihood with parameters α and β , the posterior of θ can be easily found.

$$p(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \theta \geq 0.$$

$$p(D|\theta) = \left(\prod_{i=1}^k \theta e^{-\theta t_i} \right) e^{-(N-k)T\theta}.$$

(4)

The posterior of θ can be written as

$$p(\theta|D) \propto p(D|\theta)p(\theta) = \theta^{k+\alpha-1} e^{-(\sum_{i=1}^k t_i + (N-k)T + \beta)\theta}.$$

As is expected, the posterior of θ is another Gamma density, with parameters

$$\alpha_p = \alpha + k,$$

$$\beta_p = \beta + \sum_{i=1}^k t_i + (N-k)T.$$

Having defined the factors in the integrand in Equation (3), the predictive distribution

$$p(t|D, t > T)$$

is found by integrating for θ from $0 \dots \infty$ to yield

$$p(t|D, t > T) = \frac{\alpha_p \beta_p^{\alpha_p}}{(t - T + \beta_p)^{\alpha_p+1}}.$$

(5)

The cumulative distribution of t is in turn given by

$$F(t|D, t > T) = \begin{cases} 1 - \frac{\beta_p^{\alpha_p}}{(t - T + \beta_p)^{\alpha_p}}, & t > T \\ 0, & t \leq T \end{cases} \quad (6)$$

From (5), we can find the estimate of the expected time-to-failure $E(t)$. With T as the current time instant, the estimate is

$$E(t) = T + \frac{\beta_p}{\alpha_p - 1} = T + \frac{\beta + \sum_{i=1}^k t_i + (N - k)T}{\alpha + k - 1}. \quad (7)$$

As is observed in Equations (5) to (7), the predictive density, the cumulative distribution and the expected time-to-failure are not directly dependent on the unknown parameter θ . Instead, they rely on the current time instant T , the parameters α and β from the prior of θ , the total number of items N , the number of failed items k at time T , and the sum of all known failure times. In other words, knowledge of the rate parameter of the exponential distribution of observed times-to-failure is not necessary for prediction of future failures given the assumptions made. The prior of θ and the observed failure times are however, required.

With subsequent failures, the parameters of the posterior density change, and the cumulative distribution function of the surviving items is Bayesian-updated. Prediction is sharpened with an increasing number of failing items. Since the Gamma distribution is the distribution of choice which makes for closed-form solutions for the predictive functions above, it is of great interest to observe how much impact the parameters of the prior has on Bayesian learning.

When the Gamma shape parameter is one, the Gamma distribution resembles an exponential distribution; its posterior may not converge for a long period of time. On the other hand, Gamma distributions with shape parameters greater than one are uni-modal so that the resulting posterior distributions are expected to converge relatively fast. The

experiments below test for the convergence of the posterior of θ using Bayesian learning for different values of the shape parameter α and then illustrate the developed Bayesian process for predicting future failure times with an unknown rate parameter for the selected shape parameter α .

Sample Results: Influence of Prior Parameters on Convergence of Posterior

Distribution of θ

Let the population of the items N be 1,000 and the number of failed items k be 10. The prior distribution of θ is assumed to have Gamma density parameter β equal to one; distinct values of α are tested. The times-to-failure of the N items are randomly generated following an exponential distribution with θ equal to three (three failures in one time unit). The times to failure may then be sorted so that the failure time of the $k = 10^{\text{th}}$ item is selected as the current time instant T . The posterior of θ calculated using Bayesian theorem, is observed after each subsequent failure in Figure 9.

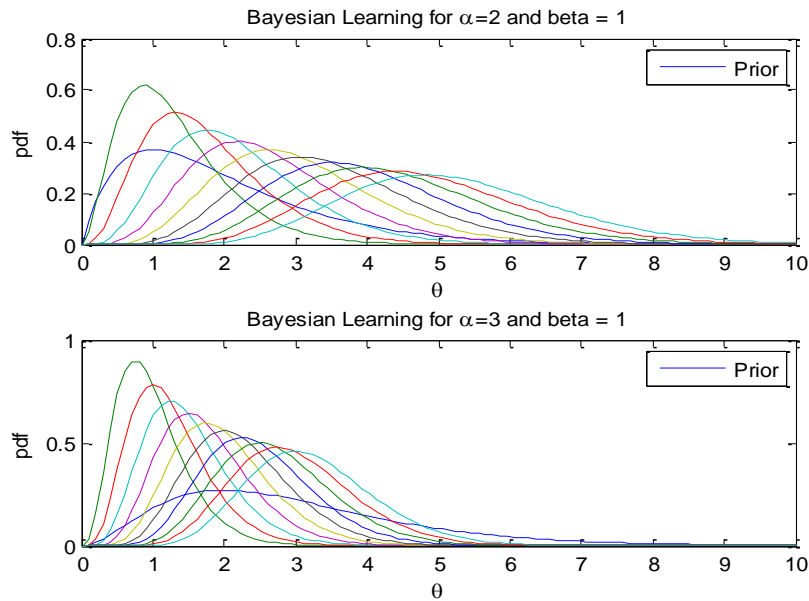


Figure 9 Bayesian updates of posterior distributions for $\beta=1$ and $\alpha=2$ (top subplot), and $\alpha=3$ (bottom subplot). Peaks of the 10 posteriors are observed to shift to the right for each additional failing item. The prior distribution is the innermost curve.

The posterior distributions shift to the right, in the direction of the true value of θ equal to three, as failure data are updated. After $k = 10$ failures, the posterior distributions from prior shape parameters α equal to three and four (in Figure 10) converge very close to $\theta =$ three. On the other hand, for the same failure data, the posterior distributions for prior α equal to two and five shoot way past θ equal three. They may require more information for a better prediction of θ . The results show that using α equal to three or four should be sufficient for future predictions.

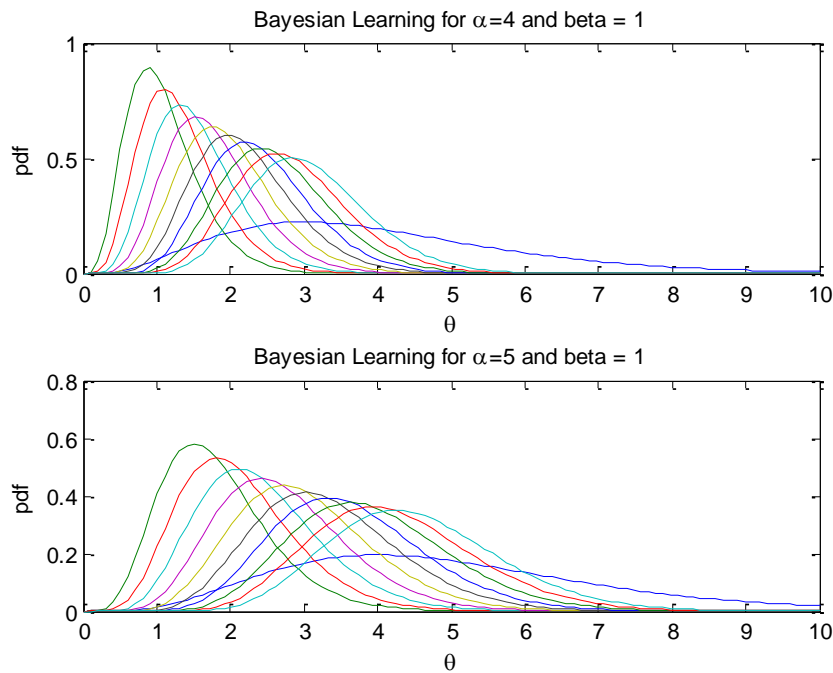


Figure 10 Bayesian updates of posterior distributions for $\beta=1$ and $\alpha=4$ (top subplot), and $\alpha=5$ (bottom subplot). Peaks of the 10 posteriors are observed to shift to the right for each additional failing item. The prior distribution is the innermost curve.

A similar test may be conducted to test the scale parameter β . However, this is unnecessary since the posterior distributions for α equal to three or four and β equal to one, as seen in Figure 9 and Figure 10, converge towards the true value of θ . Shape parameter α equal to three and scale parameter β equal to one are used in illustrating predictions of future failure times in the following illustration.

Sample Results: Prediction of Future Failures

The times-to-failure of a population of $N = 10$ components are randomly generated, following an exponential distribution with $\theta = 1/3$ (one failure in 3 time units) and $k = 5$ failed items. The prior Gamma density parameters α and β are chosen to be three and one respectively for the prior distribution of θ . The times-to-failure are then sorted so that the failure time of the $k = 5^{\text{th}}$ item is selected as the current time instant $T = 1.3094$ time units. The prior and posterior distributions of θ are plotted in Figure 11, showing a fair approximation of θ when information is gained from the failure of 50% of a very small population.

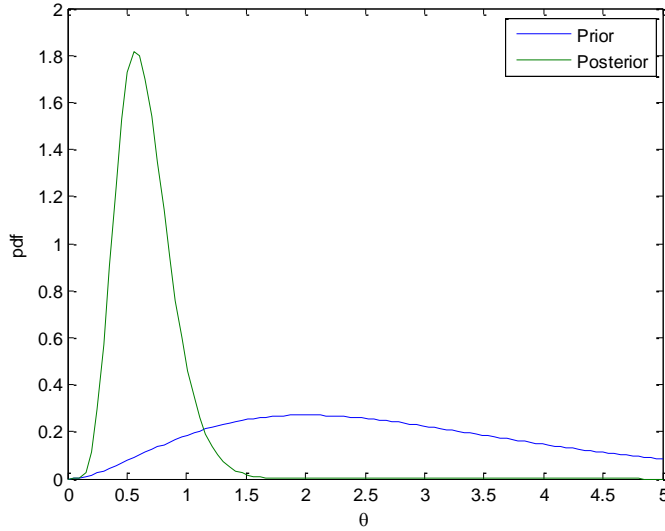


Figure 11 Prior (large variance) and posterior (small variance) distributions of θ for $N = 10$ items, $k = 5$ items, with true $\theta = 1/3$ and α and β equal to 3 and 1, respectively.

Using the developed Bayesian process, the predictive distribution function of the times-to-failure of the five items that had not failed by $T = 1.3094$ time units is shown in Figure 12, overlaid with the true distribution function of the remaining simulated times-to-failure. The expected time-to-failure is estimated to be 3.0551 time units using Equation (7), while the actual expected time to failure is 3. This is within a 95% confidence interval of the true value.

Though the distribution functions are visually dissimilar, their Pearson correlation is calculated to be 0.980 at a p-value of 0. Also, a two-sample T-test of the two distributions shows that they are equal based on a 95% confidence interval, where the test found them equal with a T-value of 2.31 at a p-value of 0.022. A low p-value means there is not enough evidence to reject the hypothesis (equality of the distributions).

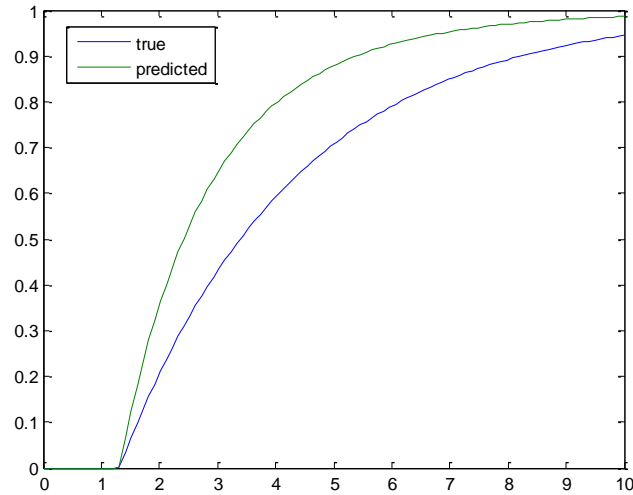


Figure 12 Plot of $F(t|t>T)$ versus time—probability distribution function of times to failure of devices still surviving at $T = 1.3094$ time units, showing the true distribution (bottom curve) for $\theta = 1/3$, and the distribution predicted using Bayesian updates of posterior of θ (top curve). Here, $N = 10$ items and $k = 5$ items.

By increasing N to 1,000 components; θ to 3; and leaving the number of failed items k as 5; the resulting posterior density is shown in Figure 13. The posterior has a mode that is very close to three. It shows that even when only 0.5% of a large number of components have failed, a very good estimate of θ can still be found.

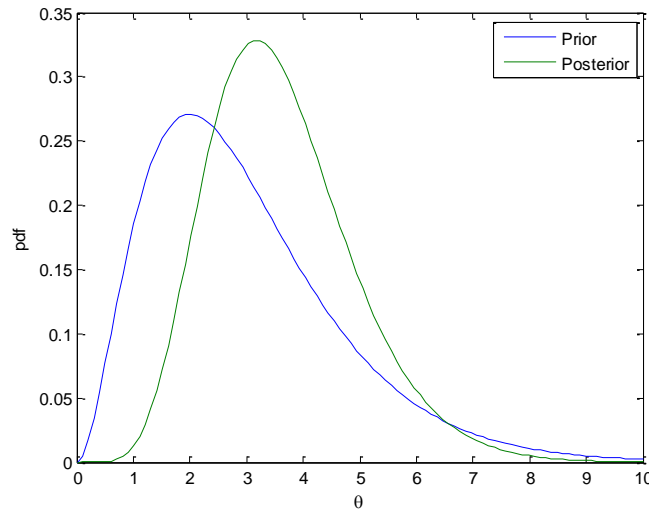


Figure 13 Prior (left) and posterior (right) distributions of θ for $N = 1000$ items, $k = 5$ items, $\theta = 3$, α and β equal to 3 and 1, respectively.

Figure 14 shows the distribution functions from the Bayesian-prediction methodology and the real data. The closeness of the estimation of θ for k equal to five compared to one thousand components is a very strong observation since most components in the power network have small failure rates (sometimes less than one per cent).

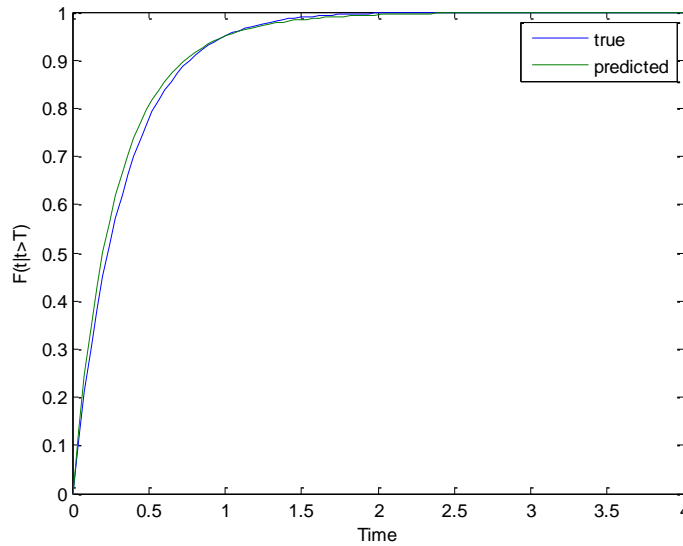


Figure 14 The conditional probabilities of time to failure $F(t|t>T)$ for $N = 1000$ items, $k = 5$ items, where the bottom curve is the true distribution obtained using $\theta = 3$ of the generated times to failure, the top curve is the predictive distribution using the prior and posterior of θ and $T = 0.0025$.

The times to failure of the five items that have failed by $T = 0.0012$ time units are 0.0008, 0.0008, 0.0010, 0.0012 and 0.0012 time units. Figure 15 shows that with the failure of 15 more items by $T = 0.0096$ time units, the prediction of θ is improved relative to the prior distribution. The predicted cumulative probability distribution, though underestimated, does not deviate much from the true distribution. The p-values for two-sample T-tests of the distributions show the failure times may come from the same distribution. Figure 16 is the corresponding cumulative distribution pair.

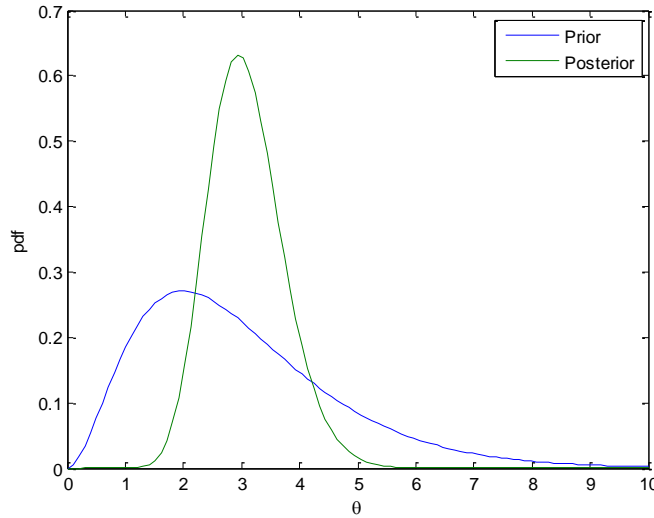


Figure 15 Prior (left) and posterior (right) distributions of θ for $N = 1000$ items, $k = 20$ items, $\theta = 3$, α and β equal to 3 and 1, respectively.

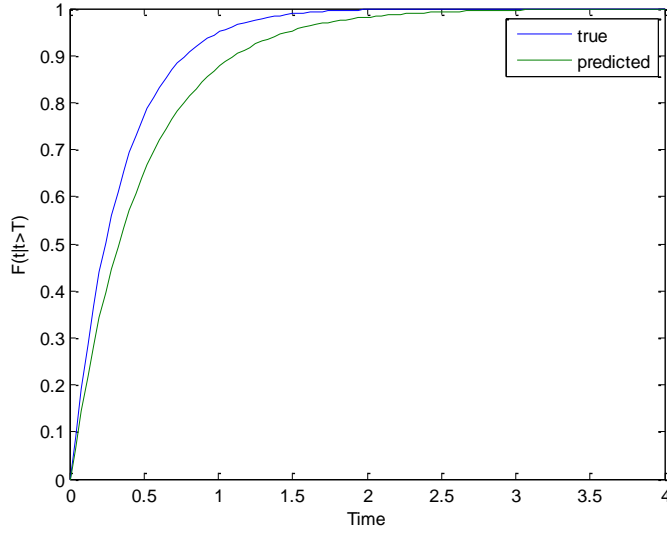


Figure 16 The conditional probabilities of time to failure $F(t|t>T)$ for $N = 1000$ items, $k = 20$ items, $\theta = 3$ and α and β equal to 3 and 1, respectively. True distribution is top curve and predicted is bottom curve.

3.1.1.1 The Effect of Missing Data on Failure Predictions

Failure prediction is expected to suffer with incompleteness in failure information; it will in turn affect the Bayesian updates, which rely on observed failure times. An experiment is developed to simulate the exclusion of proportions of observed failure times from a failure database using a Monte Carlo method. Then, its effect on failure predictions is analyzed. Estimates of expected lifetimes $E(t)$ for a complete dataset and expected lifetimes for an incomplete dataset are compared to find how much missing data excludes the true expected lifetimes of components.

Let D_1 be a completely random subset of D from Equation (3), such that D_1 contains only a proportion of the observed failure times and all the unknown failure times of remaining components. Let k be the actual number of failed items by the present time instant T , t_A be the available failure times, and $(k - l)$ be the number of recorded times-to-failure; that is l is the number of completely random missing failure times. Then,

$$D_1 = \{t_A, t_{k+1} > T, t_{k+2} > T, \dots, t_N > T | A = 1 \dots k - l\}.$$

Then the predictive density function $p(t|D_1, t > T)$ of the times-to-failure of remaining components given the incomplete recorded times-to-failure is obtained as

$$p(t|D_1) = \int_{\theta} p(t|\theta, t > T)p(\theta|D_1) d\theta. \quad (8)$$

The first factor of the integrand remains unchanged. As before, the second integrand requires the probability of the time-to-failure given the unknown parameter θ .

However, the set of failure times is now D_1 , so that $p(\theta|D_1)$ is given by

$$p(\theta|D_1) \propto p(D_1|\theta)p(\theta) = \theta^{k-l+\alpha-1} e^{-(\sum_{i=1}^{k-l} t_i + (N-k)T + \beta)\theta}.$$

The posterior of θ is again a Gamma density but now with new shape and scale parameters as defined respectively below,

$$\alpha_{pA} = \alpha + k - l,$$

$$\beta_{pA} = \beta + \sum_{i=1}^{k-l} t_i + (N - k)T.$$

Note that the number of future failures is still the difference between the entire population N and the number of components that have actually failed k , since we assume that the number of remaining components is known. Having defined the factors in the

integrand, the predictive distribution $p(t|D_1, t > T)$ is found by integrating for $\theta, 0 \dots \infty$ to yield

$$p(t|D_1, t > T) = \frac{\alpha_{pA} \beta_{pA}^{\alpha_{pA}}}{(t - T + \beta_{pA})^{\alpha_{pA} + 1}}. \quad (9)$$

The cumulative distribution of t is in turn given by

$$F(t|D_1, t > T) = \begin{cases} 1 - \frac{\beta_{pA}^{\alpha_{pA}}}{(t - T + \beta_{pA})^{\alpha_{pA}}}, & t > T \\ 0, & t \leq T \end{cases} \quad (10)$$

The new estimate of the expected time-to-failure $E_A(t)$ can be found, as shown in Equation (11), utilizing the less than actual number of failure times $(k - l)$. With T as the current time instant, the estimate is given by

$$E_A(t) = T + \frac{\beta_{pA}}{\alpha_{pA} - 1} = T + \frac{\beta + \sum_{i=1}^{k-l} t_i + (N - k)T}{\alpha + k - l - 1}. \quad (11)$$

The new estimates are expected to be different from the estimates obtained using a complete failure database. An illustration of the new results is given in the following subsection.

Sample Results: Hypothetical Dataset

A population of $N = 1,000$ components is generated in MATLAB assuming $k = 50$ failed components. The simulated times-to-failure of the 50 components are shown in Table 2.

Table 2 Complete set of failure times of $k = 50$ items with $T = 0.0155$ time units.

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
0.0002	0.0008	0.0010	0.0011	0.0013	0.0015	0.0016	0.0019
t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}	t_{16}
0.0026	0.0028	0.0036	0.0039	0.0039	0.0039	0.0040	0.0044
t_{17}	t_{18}	t_{19}	t_{20}	t_{21}	t_{22}	t_{23}	t_{24}
0.0047	0.0051	0.0053	0.0064	0.0064	0.0066	0.0068	0.0070
t_{25}	t_{26}	t_{27}	t_{28}	t_{29}	t_{30}	t_{31}	t_{32}
0.0071	0.0085	0.0085	0.0087	0.0091	0.0092	0.0092	0.0095
t_{33}	t_{34}	t_{35}	t_{36}	t_{37}	t_{38}	t_{39}	t_{40}
0.0096	0.0097	0.0098	0.0105	0.0115	0.0123	0.0124	0.0124
t_{41}	t_{42}	t_{43}	t_{44}	t_{45}	t_{46}	t_{47}	t_{48}
0.0128	0.0128	0.0129	0.0131	0.0145	0.0145	0.0145	0.0149
t_{49}	t_{50}						
0.0149	0.0155						

The true and Bayesian-estimated predictive distributions at $T = 0.0155$ time units are shown in Figure 17, assuming all k failure times $\leq T$ are known and accurately recorded. The true distribution is gotten directly from the simulated data. The 95% confidence interval of $E(t)$ that accounts for sampling errors in MATLAB is (0.2725, 0.4569) for a thousand-trial Monte Carlo simulation that generates a sorted list of $N = 1,000$ exponentially distributed times-to-failure with $\theta = 3$. The interval also accounts for errors in the convergence of the posterior density of θ , and is used in evaluating the deviation of $E_A(t)$ from $E(t)$ estimated from a complete dataset.

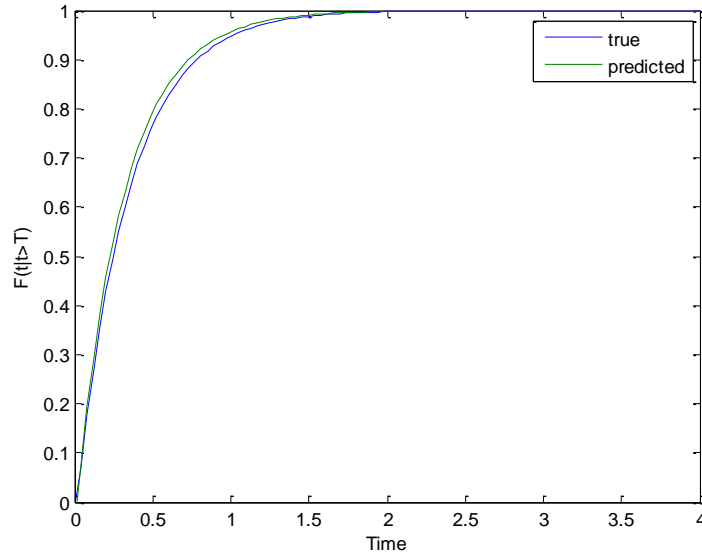


Figure 17 The conditional probabilities of time to failure $F(t|t>T)$ for $N = 1,000$ items, $k = 50$ items, $\theta = 3$ and α and β equal to 3 and 1, respectively. For the complete failure dataset, the true function is the bottom curve and the predicted is the top curve.

Randomly sampling distinct proportions of the failure times in Table 2 using bootstrapping, the predictive distributions and expected time-to-failure $E_A(t)$ of the incomplete dataset are generated. The graph of the new posterior density when 40-50% of failure data is missing is shown in Figure 18. The distribution functions found for the incomplete failure datasets get progressively farther away from the distribution function of the real failure data, and the Bayesian estimates from a complete dataset.

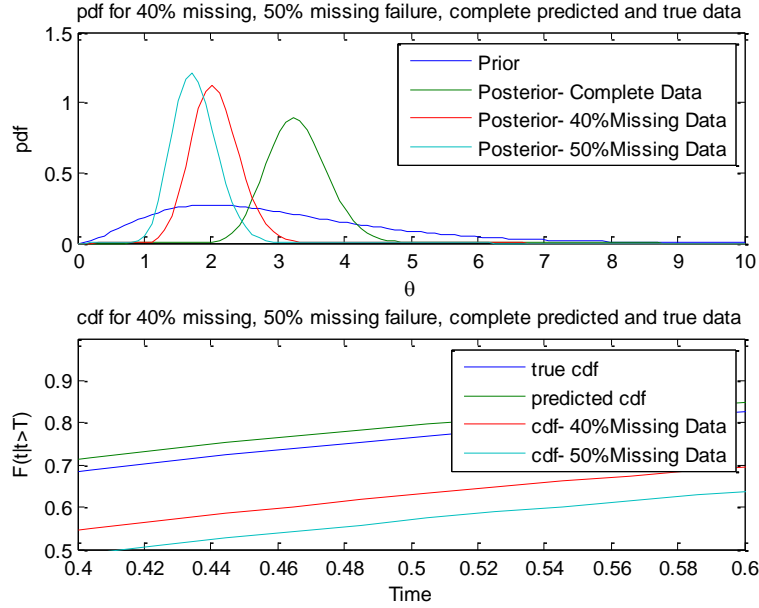


Figure 18 Prior (largest variance) and 3 posterior densities (small variances) of theta θ are shown in the top subplot. In the clockwise direction, the true distribution $F(t|t>T)$ and Bayesian-predicted distributions for a complete failure dataset, 40%, and 50% missing failure data are shown in the bottom subplot for $k = 50$ and $N = 1000$ items.

The distribution of $E_A(t)$: that is, the expected time-to-failure found from the incomplete dataset, is shown in Figure 19. It is obtained by selecting 40% or 50% of failure times completely at random from the failure dataset 1,000 times. The distributions fail to include the true $E(t)$, which is one-third.

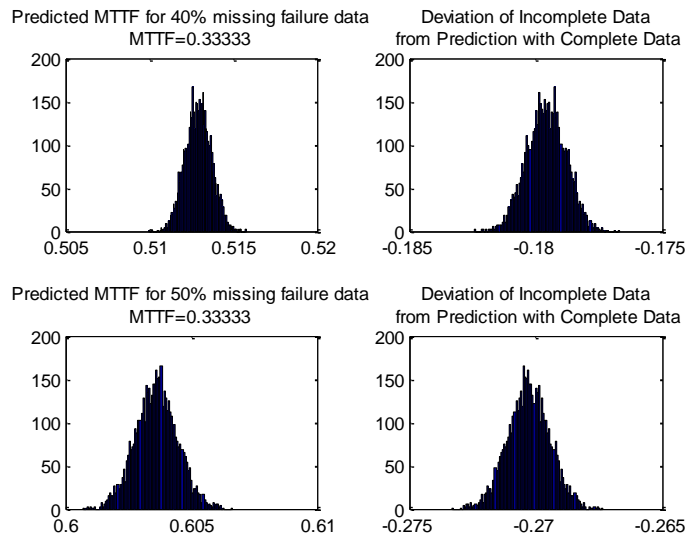


Figure 19 Histograms of the mean time to failure $E_A(t)$ for failure data missing 40% and 50% of the failure data(left subplots), and the deviations from predictions with complete data (right subplots).

Similar graphs for other levels of missing failure data can be obtained but are not shown here. The summary of the results is shown in Table 3 for varying population sizes N and number of failed devices k . As was expected, increasing the amount of missing failure data resulted in increasing deviations from the true expected time-to-failure.

A right shift is observed; that is, predicted failure times are larger than the true values, and estimates are optimistic. On the field, this means that using estimates of the predictive distribution from an incomplete dataset, power utilities will expect components to fail much later than they actually would. Thus, assuming a preventive replacement policy is being adopted, a good number of components may fail before even being inspected or preventively replaced.

Table 3 Table of deviations of predicted expected times to failure from the true MTTF for different values of k , N and ratio k/N given different levels of missing failure data.

N	1000		10000		10000		10000	
k	50		50		100		500	
Ratio l/k	Deviation from MTTF	Number of missing failure times	Deviation from MTTF	Number of missing failure times	Deviation from MTTF	Number of missing failure times	Deviation from MTTF	Number of missing failure times
2.00%	0.30%	1	8.71%	1	1.85%	2	3.03%	10
5.00%	0.90%	3	10.41%	3	2.93%	5	4.09%	25
10.00%	2.30%	5	12.26%	5	4.87%	10	6.03%	50
15.00%	4.60%	8	15.33%	8	7.03%	15	8.17%	75
20.00%	6.30%	10	17.65%	10	9.47%	20	10.60%	100
25.00%	9.20%	13	21.55%	13	12.23%	25	13.35%	125
30.00%	11.30%	15	24.48%	15	15.33%	30	16.45%	150
35.00%	15.10%	18	29.57%	18	18.93%	35	20.05%	175
40.00%	17.90%	20	33.47%	20	23.20%	40	24.25%	200
50.00%	27.10%	25	45.78%	25	33.88%	50	35.15%	250

The relationship between the Monte-Carlo-average estimated expected lifetimes for an incomplete dataset and the proportion or number of missing data is shown in Figure 20. Even for the same k/N ratio of missing information, but with different values of k and N , there is some variation in percent difference from the true expectation of the lifetime. For instance, the pairs ($k=50$, $N = 1,000$) and ($k = 500$, $N = 10,000$) have the

same ratio of missing failure data, that is 5%. Yet the curves that define their predictions of the mean or expected time-to-failure (MTTF) vary in both figures.

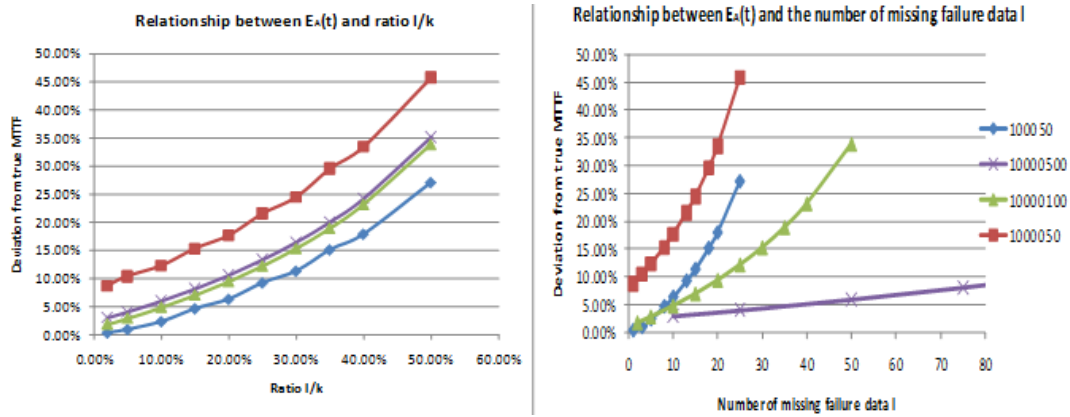


Figure 20 Graph showing the relationship between percent differences of predicted expected times to failure from the true MTTF for different values of k , N and ratio k/N ($k = 50$ and $N = 1000$, $k = 500$ and $N = 10000$, $k = 100$ and $N = 10000$, $k = 50$ and $N = 10000$, respectively) with respect to different levels of missing ratio l/k (left subplot) and failure data l (right subplot).

The deviation of the predicted expected time-to-failure $E_A(t)$ from $E(t)$ for the larger population ($N = 10,000$ items) is higher than for $N = 1,000$ in Figure 20. The prediction error of the smaller population ($N = 1000$ items) is much larger than for $N = 10,000$ items. Both plots in Figure 20 have this in common: increasing missing failure times leads to more inaccuracy in predictions. Also, the worst prediction errors were for the combination of the smallest value of k and the smallest value of ratio k/N , that is $k = 50$, $N = 10,000$, and $k/N = 0.5\%$.

Based on the illustrations presented in this section, it can be inferred that storing over 80% of failure times of components that rarely fail (so that k is small) is critical to predicting the expected failure time to within 95% of its confidence interval.

3.1.2 Non-Exponential Case

The hazard or failure rates of most components are not exponential over all times. The rates may decrease in early life and increase in older life. This describes a bathtub hazard rate. However, not all components show a burn-in or early mortality region of

failure. For the components that do, the time interval within which burn-in occurs is relatively short compared to the entire expected lifetime of the component. So, the burn-in region may be excluded leaving only two regions of concern for the bathtub curve.

Non-exponential parametric modeling has historically been used to estimate the hazard rates of components that indicate a bathtub behavior by obtaining multiple scale and shape parameters of assumed time-to-failure distributions from failure data. These distributions include exponential power distributions, double exponential power distributions, and mixtures of Weibull distributions. Other methods include the application of Bayesian statistics to some assumed prior density of the scale or shape parameter as done in the previous section. Drawbacks to the techniques include tediousness in finding multiple parameters for the distribution of times-to-failure and the non-existence of closed-form solutions to the integrations of posteriors, among others. Though numerical methodologies such as exploiting Monte Carlo may present a solution to the closed-form problem, computational complexity and sampling size of the techniques complicate the estimation. A simpler, more tractable method of approximating bathtub-shaped hazard curves is therefore beneficial.

Rather than estimate the hazard rates in the two identified regions by finding several parameters for cumbersome probability distribution functions from a large failure dataset, the hazard curve is generated gradually as failure data is acquired over time. Processing of the data is done as though the observed lifetimes were sampled from an exponential dataset. The gradual process is like sliding a time window through the life of a population of components. The rate parameters found within each time window are thus allowed to vary. An illustration is shown in Figure 21. The probability density function pdf is then a time-varying exponential density function.

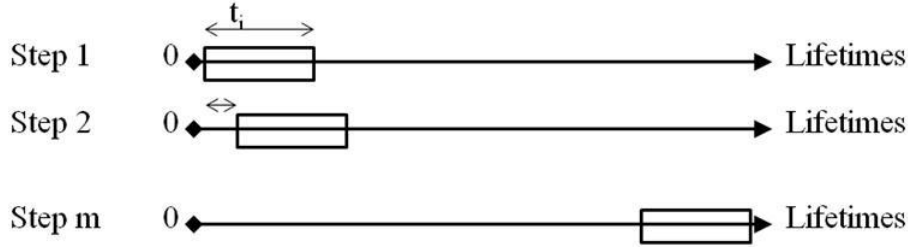


Figure 21 Illustration of sliding time windows of length t_i where approximately exponential failure rates h_{iv} are estimated. The difference between consecutive sliding windows is defined to result in a large overlap between the windows.

While the maximum likelihood estimate (MLE) of the rate parameter of the exponential distribution is simply the ratio of the number of observed failures to the sum of the observed lifetimes, the estimate required for analysis in this work is not as obvious. Because the small time window will truncate failure information to its left and right, another estimate is needed. The truncation results in the need for an MLE of a doubly-censored exponential dataset.

An algorithm is developed in this work to estimate a bathtub hazard curve using a new methodology. The uniqueness of the algorithm is that the onset of a growing failure hazard is simultaneously investigated. This is described in detail in the following subsections. Design variables for the algorithms include the following.

- Length of the sliding time window for counting the number of components failing within the window,
- Time interval between consecutive shifting windows, and
- Initial length of an evolving nested window of analysis for testing the development of the hazard rate within the window.

The strategy for estimating the hazard curve assumes that the hazard rates are constant (an exponential failure rate) within small-enough time intervals of observation.

The algorithm is tested both with synthesized and real datasets of replacements or failures of components. For the hypothetical dataset, either the thinning algorithm or inverse transformation of the cumulative distribution function may be used for random

variate generation. A mixture of Weibull density functions is used in forming the cumulative distribution function from which the random variates are generated.

3.1.2.1 Simulating failures using random variate generation from a mixture of Weibull densities

The bathtub hazard curve is, in general, defined as a piecewise function on time, with time intervals signifying regions of burn-in (neglected in this work), random failures and aging (increasing failure rate). The flexibility of a Weibull distribution function in representing constant, increasing or decreasing failure rates, depending on the magnitude of its shape parameter α , makes the distribution good for modeling bathtub curves. The bi-regional bathtub hazard curve discussed in this paper can be defined as a piecewise function in time.

The shape parameter α of a Weibull function can be chosen to produce a constant or increasing hazard function. When α is equal to one, the hazard function is constant or exponential. When it is equal to two, an increasing linear hazard function is obtained. A value greater than two results in a non-linear increasing hazard function, which is a typical shape for hazard function of aging components. With this in mind, let aging of a set of identical components begin at time t_0 . Thus, before t_0 , the hazard function is constant, and afterwards, it is increasing non-linearly. See Equation (12).

$$h(t) = \begin{cases} \frac{\alpha_1}{\beta_1} \left(\frac{t}{\beta_1} \right)^{\alpha_1-1}, & \alpha_1 = 1, t \leq t_0 \\ \frac{\alpha_2}{\beta_2} \left(\frac{t}{\beta_2} \right)^{\alpha_2-1}, & t_0 \leq t \leq t_n \end{cases}, \quad (12)$$

where t is the time-to-failure (or age of failure) of the component, t_0 is the knee of the hazard curve: where the failure rate starts to increase, t_n is a design variable: the maximum observed lifetime simulated, α_1 is the shape parameter of the Weibull density

function in the constant hazard region, α_2 is the shape parameter of the Weibull density function in the aging region, β_1 is the scale parameter of the Weibull density function in the constant region, and β_2 is the scale parameter of the Weibull density function in the aging region.

To ensure that the hazard function is continuous for all time $[0, t_n)$, the same parameter β_1 is defined to have the minimum value of the hazard function in the aging region, that is at t_0 .

$$\beta_1 = \frac{\alpha_2}{\beta_2} \left(\frac{t_0}{\beta_2} \right)^{\alpha_2 - 1}. \quad (13)$$

An illustration of a hazard curve designed using the equation is in Figure 22.

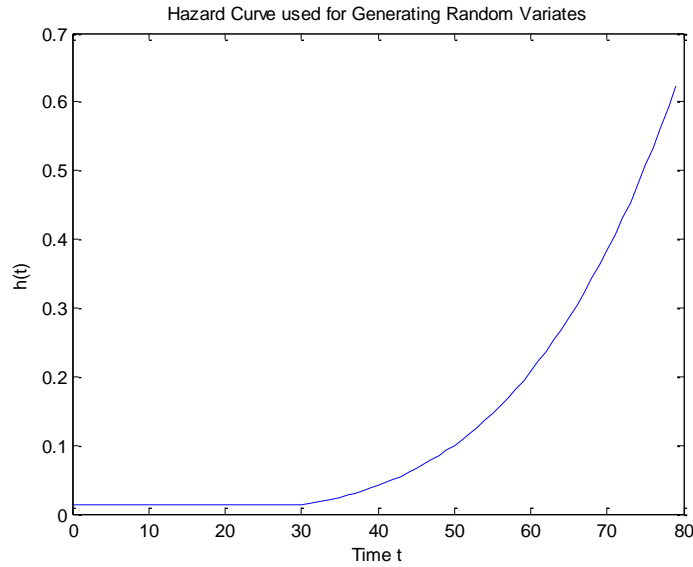


Figure 22 Hazard curve used for generating random variates with $t_0 = 30$ time units, $t_n = 100$, $\alpha = 5$, $\beta = 50$.

The cumulative distribution function $F(t)$ is easily found from the hazard function.

The relationship is given here for ease of understanding.

$$F(t) = 1 - e^{-\int_0^t h(\tau) d\tau}.$$

Since $h(t)$ is both integrable and continuous, $F(t)$ can be found in closed form. The inverse transformation method can then be applied on $F(t)$ to generate random lifetime variates. Implementing this using a Monte Carlo method implies that lifetimes $T \leftarrow F^{-1}(U)$, where $U \sim U(0,1)$. Alternatively, the thinning algorithm can be used. A Kolmogorov-Smirnov test can be used in comparing the distributions of generated lifetimes to the true distribution $F(t)$. The test returns a statistic and a p-value used to decide the probability that the distributions are equal at a specified confidence level α .

Due to the inexactness of randomization in generating the variates, a non-smooth function of the hazard h_{tv} is expected. That is, computer programs such as MATLAB use pseudo-random numbers usually dependent on clocks in generating uniformly distributed random numbers. Thus, a plot of several random numbers distributed on $U(0,1)$ generated on a computer program will not show a perfectly flat box distribution. See Figure 23 for two sizes of vectors generated: 1,000 and 10,000 data points.

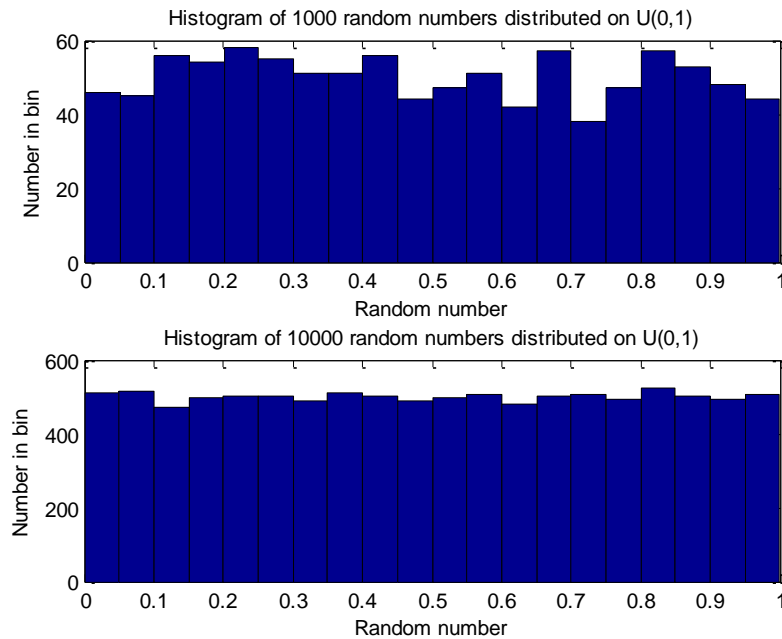


Figure 23 Histograms of pseudo-random numbers distributed on $U(0,1)$, the uniform distribution with minimum and maximum 0 and 1 respectively. They show limitations of computer programs in generating perfect uniform distributions.

3.1.2.2 Estimation of time-varying exponential rate parameters from failure data

Numerical integration may be required to obtain maximum likelihood estimates of parameters of multi-parameter distribution functions. Such computations can be tedious, and the accuracy and variance of the estimates would depend on sampling size of the Monte Carlo methods. Recognizing that the hazard curve of failing components can be built in portions of time simplifies the problem of estimation.

Let the time interval of the time window t_i , as shown in Figure 21, be the difference between t_{m1} and t_{m2} the minimum and maximum time instants of the time window respectively. So, m_1 and m_2 represent the order indices of lifetimes captured in the window. Then, the time window will start from $t = t_{m1} = 0$ to estimate the exponential rate parameter within the time window and end close to $t = t_{m2} = t_n$.

The failure rate at time instant t_{m1} , assuming an exponential distribution, is

$$h_{tv}(t_{m1}) = \frac{1 + n_{wd}}{\sum_{i=m_1}^{m_2} t_i + (N - m_2) \times t_{m2} - (N - m_1 + 1) \times t_{m1}}, \quad (14)$$

where h_{tv} is the time-varying exponential hazard rate, and n_{wd} is the number of lifetimes observed or captured in the time window that extends from t_{m1} to t_{m2} . While t_{m1} starts at zero, consecutive time instants t_{m1} are chosen so that overlaps between regions covered by preceding windows exist.

The algorithm for obtaining the time-varying exponential hazard rate from the generated random variates follows.

1. Set t_i to some design length.
2. Initialize t_{m1} and t_{m2} to zero and t_i respectively.
3. Set the sliding length v to some design value.
4. Initialize a variable j to one for the length of estimated h_{tv} .

5. While $t_{m2} \leq t_{(N)}$,
 - a. Find the number of lifetimes between t_{m1} and t_{m2} as n_{wd} .
 - b. Find $h_{tv}(t_{m1})$ using Equation (14).
 - c. Increment j by one.
 - d. Increment t_{m1} and t_{m2} by v .

One application of estimating the bathtub-shaped hazard curve of a component, like the one shown in Figure 22 is in finding the time at which the hazard/failure rate changes from being approximately constant to increasing t_0 . This time instant—the knee of the distribution, is usually a start-age for preventive maintenance on components. It is when the components are defined as “aging.” The goal would be flattening the hazard curve and reducing the rate of failure from t_0 onward.

3.1.2.3 Detecting the Onset of Increase in Failure Risk of Components

A plot of the estimated hazard curve for a complete lifetime dataset, where all components have failed, will graphically show where the failure rate starts to increase. This is where the aging region begins, and may be called the knee of the failure distribution. For purposes of proactive maintenance, it would be beneficial to detect the approximate instant when the failure rate starts to increase. Proactive and preventive maintenance operations, when initiated in a timely format, are expected to maintain the approximately flat level of the constant region of failure.

Segmented regression techniques are applied in estimating “break-points,” “change points” or “switch points” in relationships between some response variable and explanatory variable. Algorithmic methods that have been used in dealing with segmented regression include simple grid-searches, modifying a probability distribution of the response, differentiating the response function over the entire range of the explanatory variable, using regression splines, and Bayesian Markov-chain Monte Carlo (MCMC) methods. These methods will not be efficient in dynamically estimating the

onset of an increase in the hazard curve since they might require a large number of samples to the right of the breakpoint to identify a change in the slope.

An algorithm is developed to estimate occurrence of a knee (change-point) while approximating the actuarial failure rate of the components. This algorithm is embedded within that of the failure rate estimation. Since the bathtub curve studied in this work ignores the burn-in region, only two regions (the region of constant or random failures and the region of increasing failures) are possible. So, the algorithm should find only one knee.

Ideally, the hazard curve modeled here is perfectly flat in the random failure region before it starts to increase. In this sense, the onset of risk increase would be detected easily using a derivative approach, assuming the curve were differentiable. (The minimum point at which the derivative of the function was positive would be the onset.) However, in the practical sense, especially given the partitioning of failure rates in this paper, the estimated hazard function is a discrete function. Also, the number of failures in each time window of failure observations is unlikely to be uniform. Thus, the derivative approach is an unlikely solution to detecting the onset. Instead, a Kalman filter is used to determine the state of the hazard process and a procedure that relies on extreme value theory is used to identify the onset.

The dynamics of the state space of the failure rates can be modeled using a simple Kalman filter. Given the data set of failure rates $\left\{ \hat{\theta} \right\}_{j=0}^T$ $\{\bar{\theta}_j\}_{j=0}^T$, consisting of the one-dimensional estimates of the time-varying failure rates. The Kalman model, which relies on a normal distribution assumption, is considered as

$$a_j = a_{j-1} + w_j, \quad w_j \sim N(0, W_j),$$

$$y_j = a_j + v_j, \quad v_j \sim N(0, V_j),$$

where a_j is the state variable and follows a Gaussian diffusion process with Markovian dynamics; y_j is the observation variable, which is essentially each estimated failure rate; and w_j and v_j are the noise variables in the state and observation spaces respectively. The noise variables in a Kalman filter are also Gaussian and distributed with a mean of zero and variances of W_j and V_j , as shown in the equations. Control variables have been neglected since the failure process is not expected to be controlled by any external systems. Also, the state transition matrix is assumed to be identity meaning that no systematic trends in the state are expected. Finally, the observation matrix is also identity; in this model, the state relies perfectly on the observation and vice versa in the time update and observation update evaluations.

Though the failure process is not Gaussian-distributed, we assume that the variances of the noise in updating or predicting a future hazard state or measurement does not change with time but may be modeled as a Gaussian distribution. Then, W_j is some constant Q and V_j some constant R . If the estimated error covariance is P_j , the time updates of the state and covariance are

$$\hat{a}_j = \hat{a}_{j-1} + K_j(y_j - \hat{a}_{j-1}),$$

$$P_j = (1 - K_j)(P_{j-1} + Q),$$

$$K_j = \frac{(P_{j-1} + Q)}{(P_{j-1} + Q + R)},$$

where K_j is called the Kalman gain, and y_j is the j -th estimated failure rate obtained using the developed hazard algorithm of Equation (14).

The state estimate in the equations is used for detecting onset of failure increase since it is a smoother estimate of the time-varying failure rate from Equation (14). Before the convergence of the Kalman error covariance, the estimate of the state is not confident. When the covariance reaches steady state at a fast rate, the training set of normalcy

(random failure region) will consist of useful data early. Else, some errors may occur from using the state estimate before convergence. A trigger may be set to check for convergence. For instance, the following may be used as a test.

$$P_j - P_{j-1} < \varepsilon. \quad (15)$$

Around the transition of the failure rates of a component to increasing failure rates, almost every new state estimate will be more extreme (greater) than the prior estimates. Let $A_m = \{a_1, a_2, \dots, a_m\}$ be the set of m independent and identically distributed random variables of the Kalman state estimates. Define $x_m = \max(A_m)$ as the largest element observed in the set of m samples. Then, by the extreme value theorem as discussed in [11], the probability of observing some extremum $x_m \leq x$ is

$$P(x_m \leq x | \mu_m, \sigma_m) = e^{-e^{-y_m}}, \quad (16)$$

$$y_m = \frac{x_m - \mu_m}{\sigma_m},$$

where y_m is a reduced variate of the extrema, μ_m and σ_m are as norming parameters. The generalized extreme value (GEV) distribution is used as the limit distribution of normalized maxima of a sequence of i.i.d. random variables. In the case of Equation (16), the GEV reduces to a Gumbel distribution. While the asymptotic forms of the norming parameters in [11] are given for A_m drawn from a one-sided normal distribution, the parameters are defined somewhat differently in this paper.

For a set X consisting of a growing number of maxima of m -sized consecutive Kalman states x_{mi} , μ_{mi} and σ_{mi} are defined as the expected value of the elements of set X and the corresponding standard deviation at each point of estimation i . In [11], a “novelty” or onset for the bathtub case is detected when the probability in Equation (16) exceeds about 0.95. This threshold is loosened to 0.9 in this work.

Sample Results: Simulation

Experimentation begins with a large-sized population of components, $N = 10,000$, simulated to fail within at most 100 years. The average age of power system components like substation transformers, wood poles and transmission structures in some regions of the United States is greater than half a century. A proportion of such components may be close to 100 years old. An identical set of the N components is simulated to have a constant failure risk up to $t_0 = 30$ years.

The parameters of the Weibull and exponential density functions used in the illustration were based on the shapes and scaling of the hazard functions. In the probability density plot of Figure 24, for small values of the shape parameter ($\alpha < 5$), the population of components appears to fail in relatively large volumes in the early life. As the shape parameter increases for $\beta = 50$, the likelihood of failing around the scale parameter β increases. On the other hand, by increasing β for one value of α , the probability density flattens. Assuming the majority of the simulated components fail around 50 years and with moderate transition from the constant region of failure risks to increasing risk, the parameters of the Weibull distribution are set at $\alpha_2 = 5$ and $\beta_2 = 70$ for Equations (12) and (13).

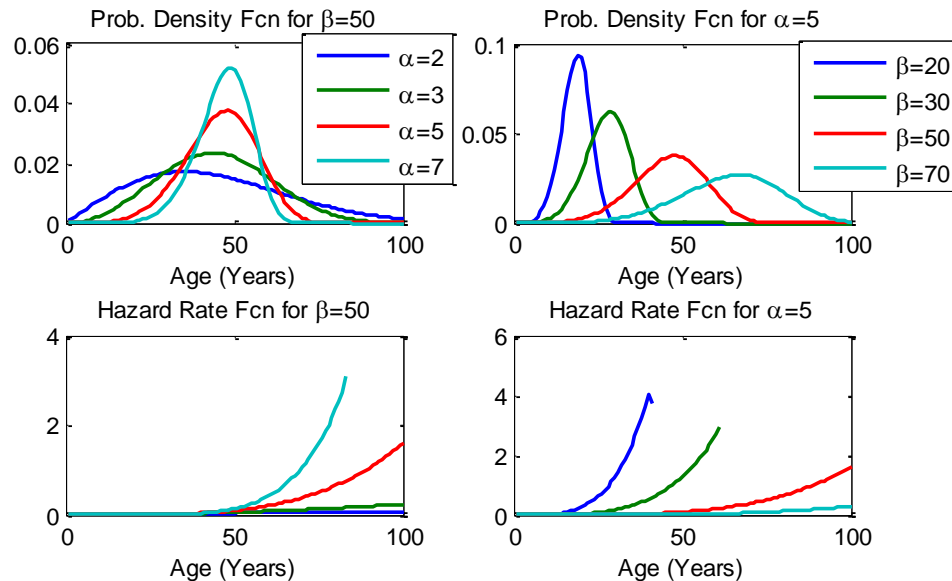


Figure 24 Probability density functions (top sub-plots) and hazard functions (bottom subplots) of Weibull densities with varying shape α and scale parameters β . The legends in the top subplots apply to the bottom subplots. The probability density functions (pdfs) shift to the right for $\alpha = 5$ as β increases from 20 to 70. The pdfs reduce in variance and increase in height for $\beta = 50$ as α increases from 2 to 7. Non-linearity reduces in the hazard functions for $\beta = 50$ as α increases from 2 to 7. The hazard rates stretch to the right for $\alpha = 5$ as β increases from 20 to 70.

The histogram of times-to-failure of the components is shown in Figure 25. Given the Weibull parameters, all the components are observed to fail within 70 years.

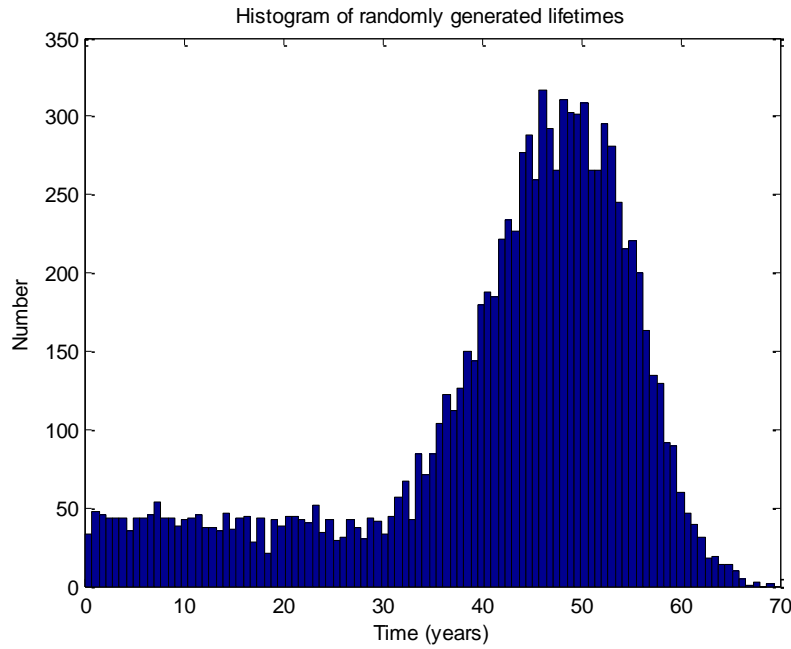


Figure 25 Histogram of generated lifetimes for $N = 10,000$ components for Weibull parameters $\alpha_2 = 5$ and $\beta_2 = 70$.

The true and theoretical distribution functions of the times-to-failure of the components are shown in Figure 26.

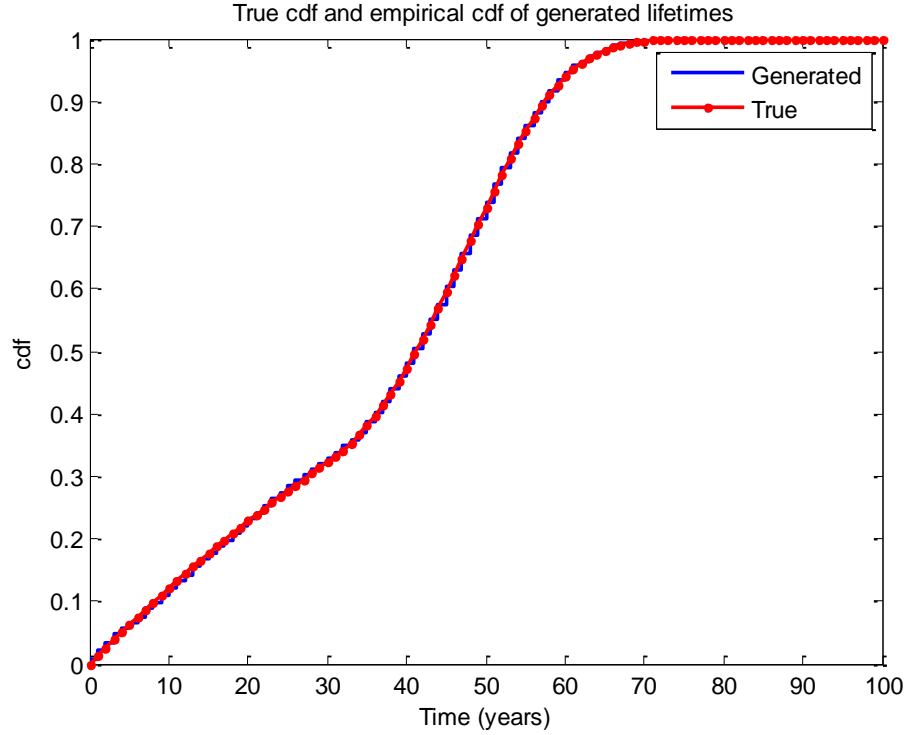


Figure 26 Theoretical cumulative distribution function (cdf) of the simulated times-to-failure (line with triangular marker) and the empirical cdf of the generated times-to-failure (stairs directly behind the theoretical cdf).

The time-varying hazard rate estimates θ_{hat} are calculated from Equation (14) within time windows of length equal to two years. The resulting approximate hazard curve using the developed time-varying technique and the theoretical Weibull hazard curve used for simulation are shown in Figure 27. The time-varying hazard curve is observed to slightly overestimate the theoretical hazard curve, increasing for longer stretches of time.

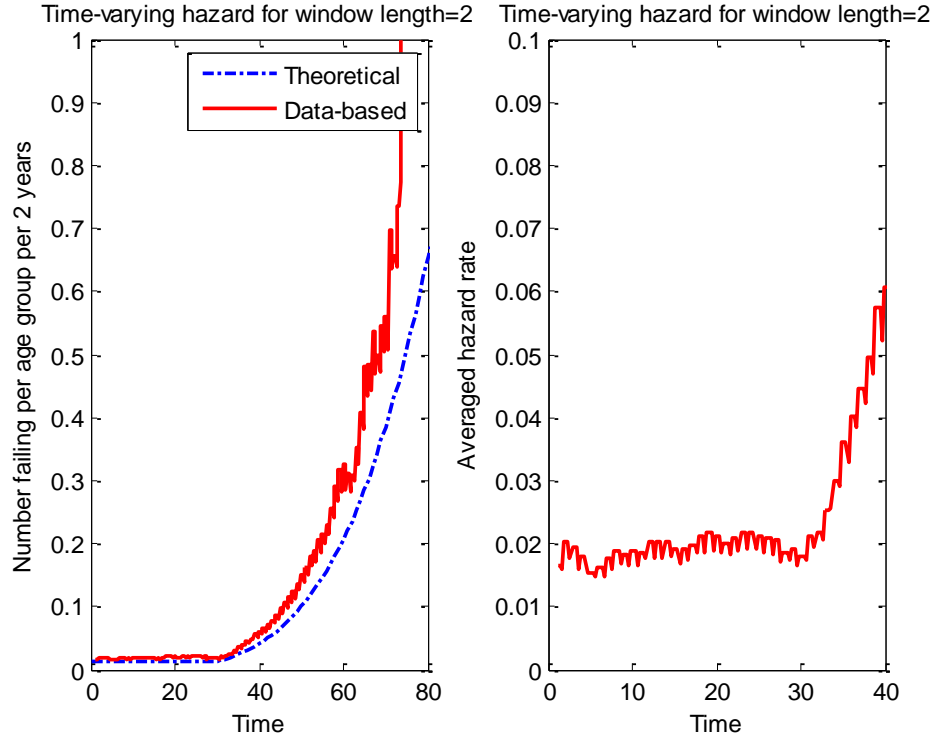


Figure 27 Theoretical hazard curve (dashed line) and time-varying exponential hazard rates (solid line), showing slight overestimation of the failure risk with increasing time.

The Kalman filter is designed with constant, low noise variances of $R = 1e-3$ and $Q = 1e-5$. The sensitivity of a Kalman filter to the values of R and Q can be found in literature. The state error covariance is initialized with $P_0 = 1e-5$ and the initial state of the failure process (failure risk) $f_0 = 0$. The resulting Kalman filter, which smooths the time-varying estimate, is seen in Figure 28, with the error covariance P_i . Note that P_i is not a probability but an error covariance. Convergence of the error covariance is observed around 10 years. Thus, the first trigger P-trigger is set around 10 years. Consequently, the Kalman filter starts to approximate the true state of the failure risk more accurately around the same age of 10 years and closely follows the theoretical hazard curve up to age 50. Then, it starts to overestimate the hazard risk slightly.

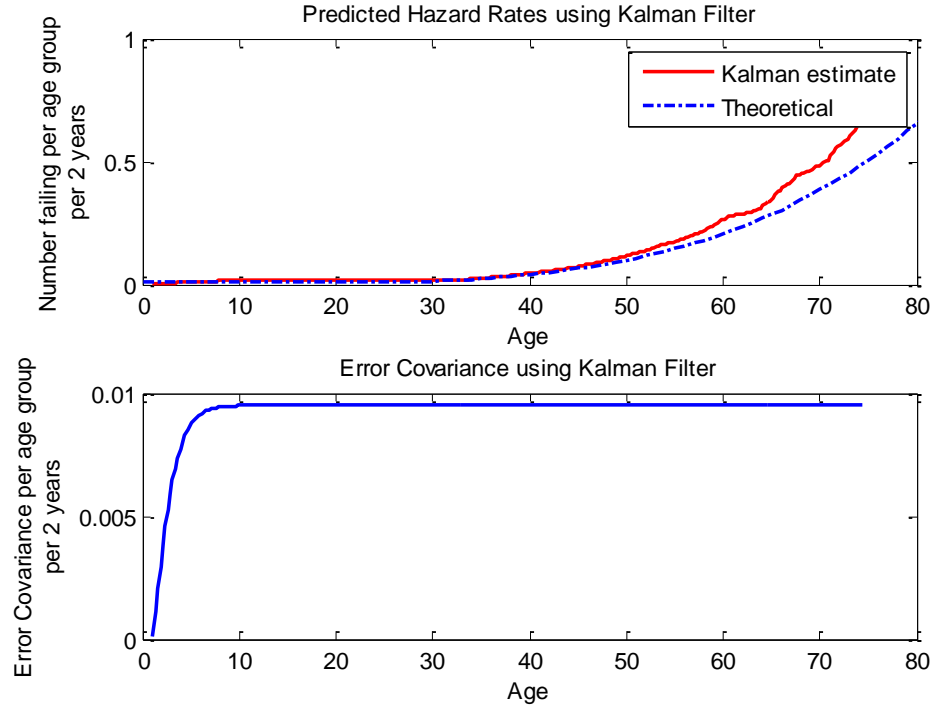


Figure 28 Smoothed estimate of the time-varying exponential hazard rates using a Kalman filter (solid line in top subplot) overlayed with the theoretical Weibull hazard curve (dashed line in top subplot) and the error covariance of the Kalman filter (bottom subplot), showing convergence of the error covariance around 10 years.

Using the concept of the extreme value theory, the probability of getting a maximum value over time is shown in Figure 29, and does not exceed the threshold of 0.9 till 38 years. The number of data points used for finding the maximum is 10. Roberts mentions in [11] that most data point sizes behave well. The increase in the probability before the age of 10 coincides with the increase in the Kalman state during this period, before the error covariance settles. Each point of increase in the probability other than at age 30 represents points that may have been false alarms in a decision function that used the mean or variance of the states directly.

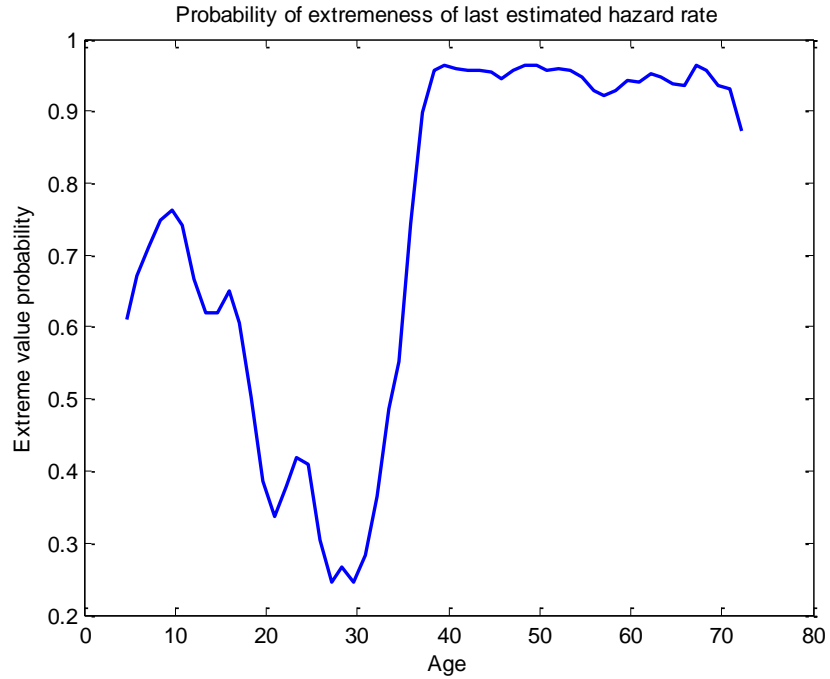


Figure 29 Plot of the probability that the most recent failure rate estimate at some age is the maximum point among prior estimates.

The attributes of the failure risk displayed as an output of the algorithm are shown in Table 4. The hazard rates of the components are observed to increase more consistently after 38 years.

Table 4 Table of attributes of component failure risk at the time instants that an onset alert is triggered: time-varying hazard rates at those instants and the binomial probability of observing an onset (a run length of one) within three trials.

Onset flag raised at	Time-varying hazard rate (θ_{hat_i})	Probability of a three-time onset occurrence up to flagged time
38 years	0.040	0.958
39 years	0.045	0.964
40 years	0.049	0.960
...
72 years	0.709	0.930

Sample Results: Real Wood Pole Data

The algorithm is applied to a real wood pole dataset for evaluation. There are over 150 million wood poles in service in the United States used to support overhead lines and other loads (like service transformers) in electric power utilities. Neglecting failures from natural disasters such as hurricanes and ice storms, wood poles are suspected to fail when their strengths decline over time in contrast to the loads they support. Unfortunately, power utilities do not routinely record the times or causes of non-environment-related failures of the poles. Influence of public service commissions has resulted in regular inspections of the poles in most states. Consequently, the replacements of the wood poles relying on the recommendations of the diagnostic companies are recorded. If we assume that a preventive replacement of a wood pole accurately reflects averted failures of the poles, such replacement databases may be analyzed for observing failure (more correctly, replacement) risks of the poles.

The dataset extracted and analyzed here consists of 27, 098 replaced power utility wood poles. The histogram of the replaced poles is shown in Figure 30.

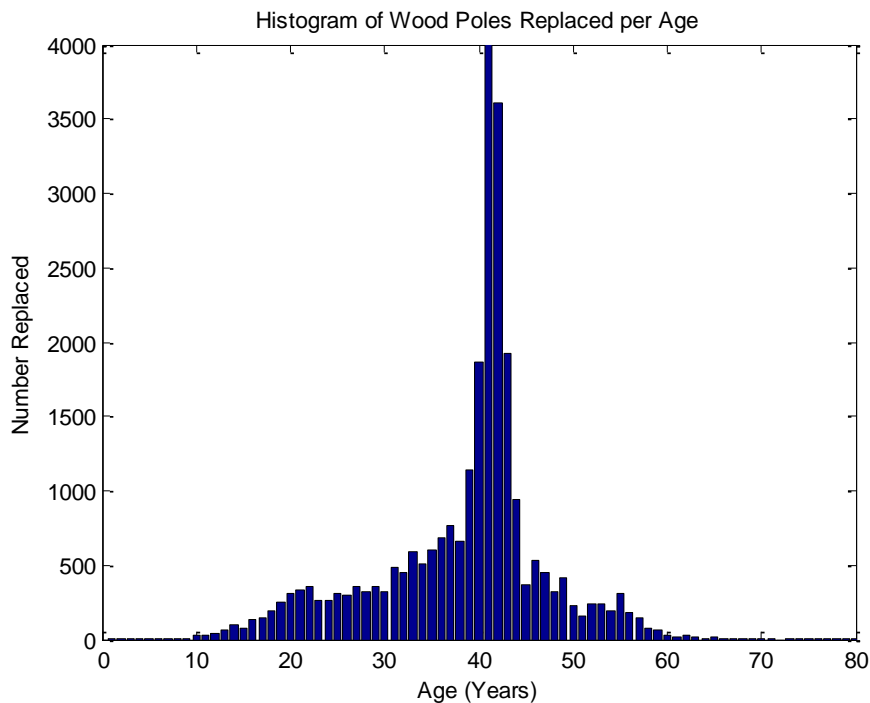


Figure 30 Histogram of wood poles replaced in a power utility. The replacements show a partial bathtub hazard behavior up to about 45 years as seen in the figures following.

A constant failure rate is observed within the first ten or so years after which the failures increase. The hazard curve of the wood poles was found using an actuarial method in MINITAB and shown in Figure 31. The curve reflects a partial bathtub behavior (constant and increasing failure risk regions) up to about 40 years. Other lifetimes are ignored with regards to this work since they are not consistent with the bathtub assumption. It is important to note that the sudden increase in the hazard plot at 40 years might be a result of a large proportion of 40-year old poles in the inspected region and may not be directly connected to replacements or failures.

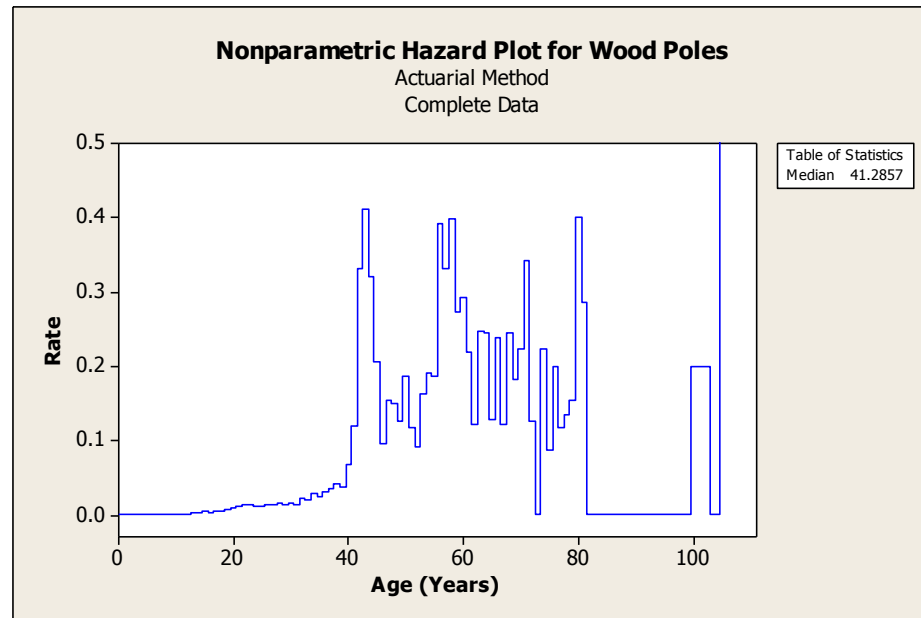


Figure 31 Nonparametric hazard plot of replaced wood poles using the actuarial method in MINITAB. Here, $N = 27,098$.

The estimated time-varying exponential hazard rates using the developed algorithm are shown in Figure 32. This closely resembles the hazard plot estimated using the actuarial method in MINITAB.

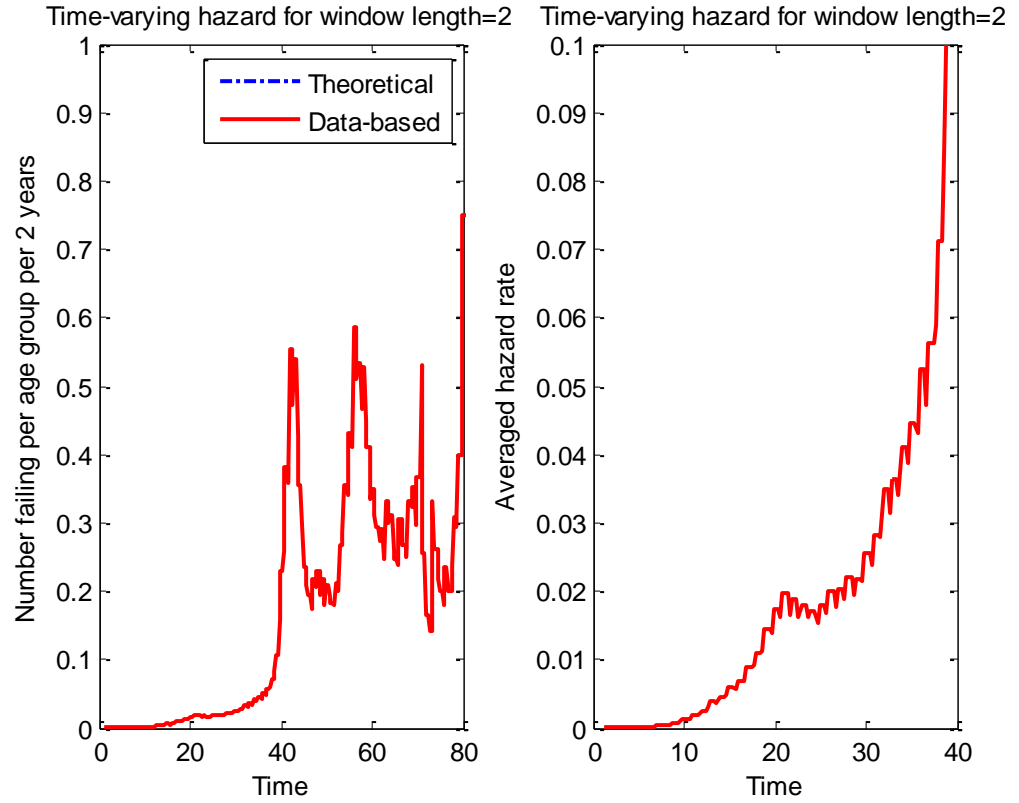


Figure 32 Time-varying exponential hazard rates for a real power utility wood pole replacement dataset.

The estimate of the failure risk using the same initial and noise parameters for the Kalman filter as in the synthesized dataset is shown in Figure 33.

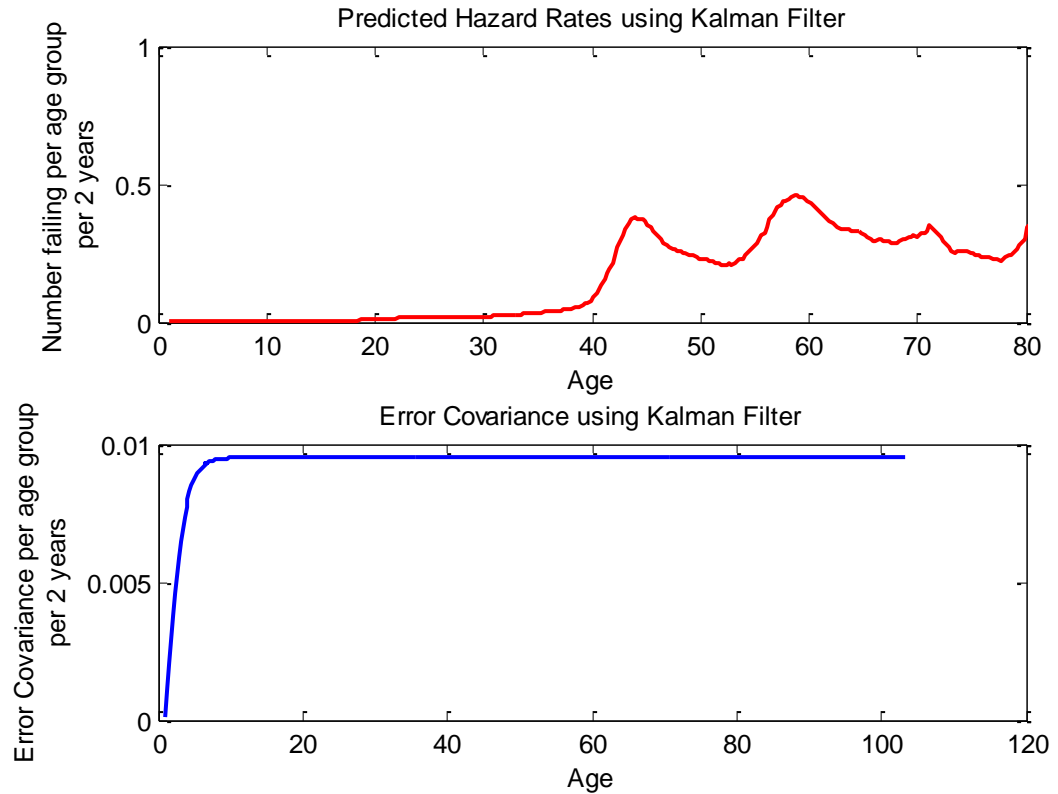


Figure 33 Smoothed estimate of the time-varying exponential hazard rates using a Kalman filter (top subplot) and the error covariance of the Kalman filter (bottom subplot), showing convergence of the error covariance around 12 years.

The onset is detected at around 18 years as shown in Figure 34. This can be seen as the point at which the failure risk in the early state becomes very pronounced in Figure 32. Though the threshold missed the initial onset around 10 years, the detection around 18 through 22 and 35 through 44 years are very beneficial since the hazard risk increases very sharply around these points. The decrease in the probability around 44 years is also consistent with the observed risk using the actuarial method, where the bathtub behavior is no longer observed.

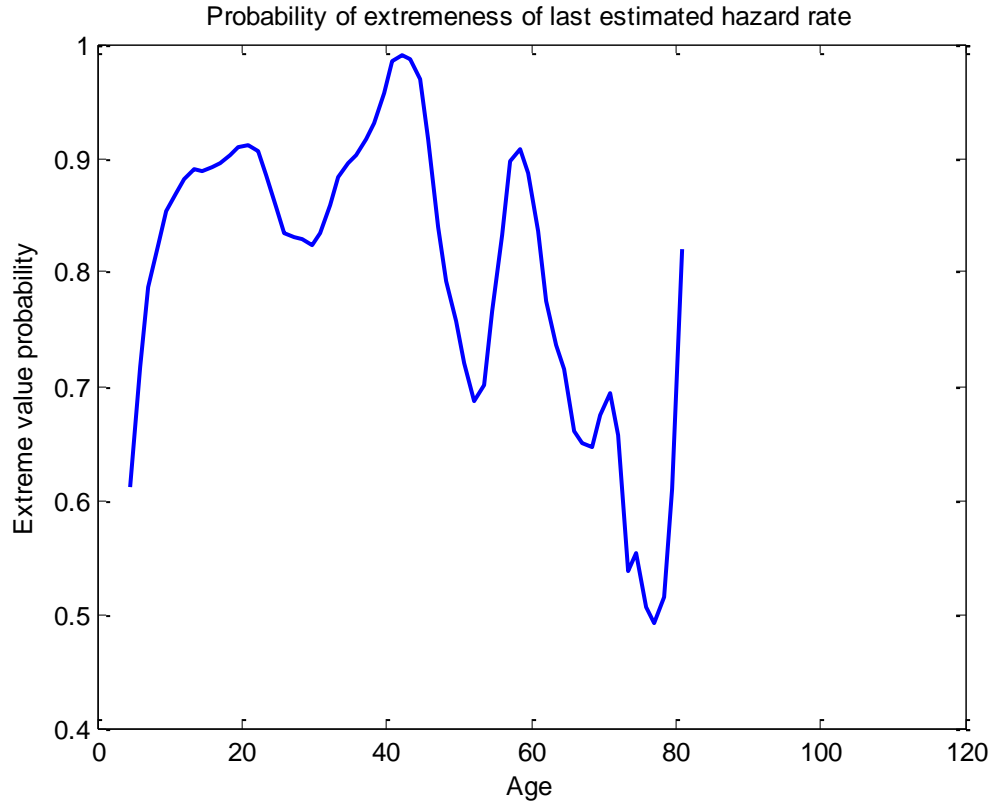


Figure 34 Plot of the probability that the most recent failure rate estimate at some age is the maximum point among prior estimates.

The algorithm was applied on a synthesized dataset of $N = 500$ and $1,000$ components. For the dataset consisting of $N = 500$ components, two false alarms were raised at 12 years and 19 years but a steady increase in the probabilities of onset occurred at 33 years. For the dataset consisting of $N = 1,000$ components, one false alarm was raised at 12 years and the true alarm was raised at 32 years.

3.2 Stress-Based Failure Estimation and Prediction

Power system components that are not housed within buildings face environmental stress that may cause them to fail. Over the past few decades, a number of hurricanes have caused the failure and subsequent replacement of such exposed components as transformers, overhead lines, utility poles and the like. A simple

regression analysis of failures in past hurricanes has been used to show that the failure rate of wood poles increases exponentially with the category of hurricanes as set by the Saffir-Simpson scale [1]. Since scaling is derived from wind speeds, an analytical model of failures related to wind speeds is more appropriate for failure prediction and making risk-informed decisions on the inspections of power system components exposed to the environment.

Since the components discussed here are assumed to fail mechanically more frequently than electrically, structural analysis is used. Modeling the failure of a component by obtaining the probability that stress on the component exceeds its strength is called fragility assessment in civil engineering. Fragility curves have been used in assessing seismic vulnerabilities of bridges and residential buildings to earthquakes, hurricanes and other hazards. It may also be seen as reliability modeling of the components with environmental stress rather than age as the independent variable. In fragility assessment, the damage state of the modeled component and the intensity of stress on the component must be specified.

In this chapter, only one damage state of a component is identified: failure that leads to replacement of the component. Its complement is the survival of the component under loading. For conductors, the damage is when they snap; for poles, when they break. The failure patterns of components and measures of each environmental condition are different. Earthquakes are measured on the Richter scale depending on the magnitude of ground shaking and seismic wave generated. Lightning strikes are defined by the number in a certain location and the charge carried. To avoid generalization, the intensity measure of stress here is focused on wind speeds since some data is available on failures of some electric utility components from historical hurricanes.

3.2.1 Fragility Assessment of Wind-Stressed Components

Using the provisions of the American Society of Civil Engineers 7 (ASCE7) standard, an analytical model for wind load profiles along an exposed component is considered [2]. Physical and material properties of the component that affect its moment of capacity are assessed. To represent the variability in material properties of the components and wind profile, uncertainties in applied wind loads and the structural response of the components are accounted for in this framework to improve the accuracy of the derived fragility curves. The wind load is called the demand on the component.

The cumulative distribution function of the probability that the demand on the component (wind load) exceeds its capacity (strength) can be computed as follows.

$$P(C < D | IM) = 1 - \Phi \left(\frac{\ln(m_C) - \ln(m_D)}{\sqrt{\beta_C^2 + \beta_D^2}} \right), \quad (17)$$

where C is the component capacity, D is the demand of the wind, $\Phi[.]$ is the normal standard distribution function, m_C is the moment of capacity, m_D is the moment of demand, β_C is the logarithmic standard deviation of the capacity, and β_D is the logarithmic standard deviation of the demand.

The log-normal distribution has been found to be effective in modeling fragility. The composite logarithmic standard deviation, which is the denominator in Equation (17), represents dispersion. Dispersion in the demand and capacity can be estimated from historical analysis of wind loads in a specified region of interest and analysis of the properties of the component.

The moment of capacity of a component can be estimated using a physics approach. The moment of a force at any point, for instance, is often calculated as a cross-product of the force on a component and the distance-vector between that point and the point where the force is applied. The moment is found specific to a component.

Wind pressure F acting in the lateral direction on the face of an exposed component can be calculated from the following relationship that is provided in the 2010 ASCE 7 standard [2].

$$F = q_z G C_f A_f,$$

$$q_z = 0.613 K_z K_{zt} K_d V^2 I,$$

where q_z is the velocity pressure evaluated at height z on the pole, G is the gust-effect factor, C_f is the force coefficient, A_f is the projected area normal to the wind, K_z is the velocity pressure exposure coefficient, K_{zt} is the topographic factor, K_d is the wind directionality factor, V is the basic wind speed which corresponds to 3-sec gust speed at 33 ft above ground in open ground. The values can be found in literature.

A Latin hypercube sampling technique can be used to generate and randomly pair samples of parameters of wind load D and component or structural response C . The realization of the wind demand and component capacities may then be compared in a Monte Carlo simulation across a wide range of wind speeds, which serve as the intensity measure of the demand. The failure fragility curve is developed from the comparisons as the ration of the number of times D exceeds C per wind speed to the number of samples. Theoretical results from the assessment are weighed against real failure observations to evaluate how closely the analytical method predicts failures.

Sample Results: Wood Poles

The capacity of distribution wood poles is characterized in terms of their fiber strengths. Experimental studies have shown that the following factors affect the fiber strength of a wood pole:

- Species of the wood pole,
- Geometric form (dimensions),
- Moisture content,

- Pretreatment conditioning, and
- Load sharing.

Data from one year of inspections of wood poles in a real power utility are analyzed. The distributions of classes and heights of the 5,792 poles are shown in Figure 35. It can be observed from the histogram that classes three and five are significantly denser than other classes represented. No wood pole of class eight, nine or ten was inspected in the year analyzed. Seventy poles were class two poles while there was fewer than fifty of each of the other four classes. Due to the class distributions, fragility assessment is conducted for classes three and five poles alone.

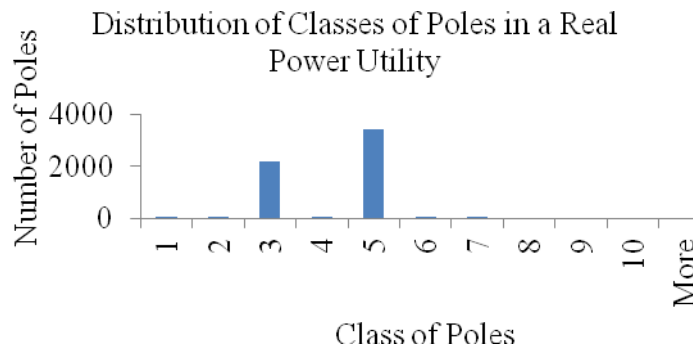


Figure 35 Number of poles per class for one year of inspection data in a real power utility, showing that there are significantly more wood poles of classes three and five than any other class, of the 5,792 poles analyzed.

The distribution of heights of the poles per class three or five is shown in Figure 36, with their respective cumulative distribution function (cdf) shown in Figure 37. Empirical and fitted cdfs are found. A lognormal distribution effectively fits the sample of heights, as confirmed using a non-parametric Kolmogorov-Smirnov test of equality.

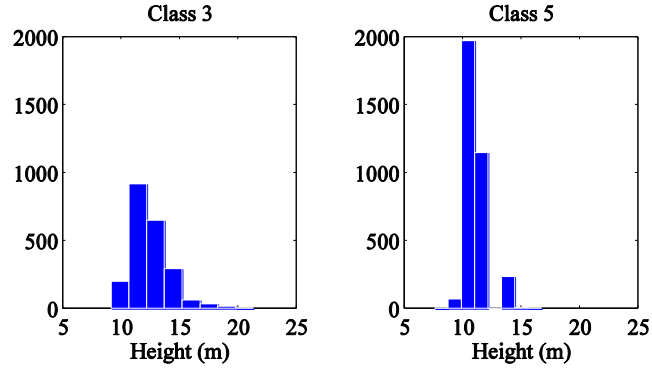


Figure 36 Distribution of heights of wood poles for classes three and five.

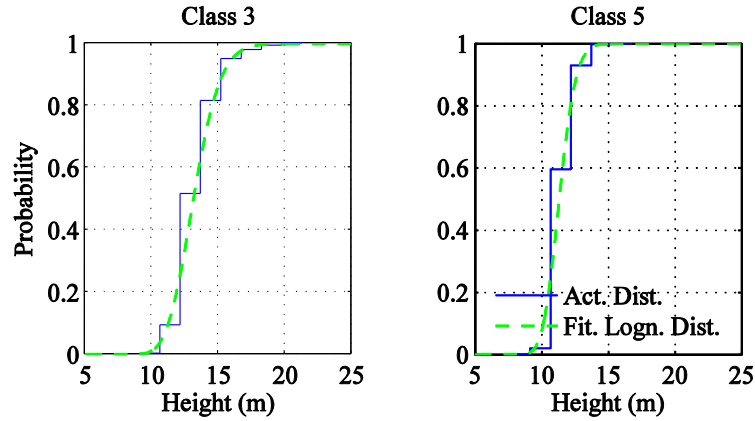


Figure 37 Empirical and fitted cumulative distribution functions of heights of class three and five poles. The fit is for a lognormal distribution function.

Using the parameters of the lognormal fit, 5,000 samples of heights of each class were generated with the Latin Hypercube sampling (LHS) technique. LHS is a constrained Monte Carlo sampling scheme used to obtain a smaller variance and faster convergence than that of the same sampling size of a conventional Monte Carlo technique. When used efficiently, it reduces the computational time of a stochastic sampling problem. The resulting cumulative distribution function (cdf) found using the LHS-sample is shown in Figure 38.

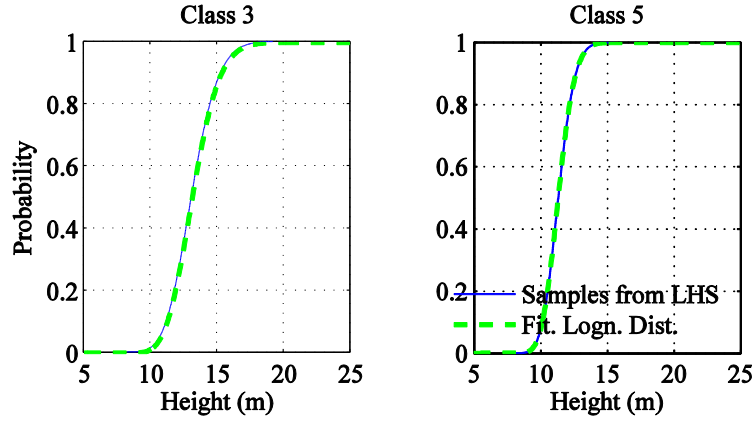


Figure 38 Cumulative distribution function (cdf) of heights, in feet, of poles sampled from the fitted lognormal distribution using the Latin Hypercube Sampling technique.

The heights are required for generating the minimum groundline circumferences of wood poles from the American National Standards Institute (ANSI O5) standard on wood poles [3]. The dimensions of each sample of the utility poles including the minimum circumference at top of the poles, ground-line distance from bottom of the poles, minimum circumference at 1.8 m from bottom, and minimum circumference at the ground-line are thus obtained for each class dimension. These dimensions are required to estimate the moment capacity of the wood poles and determine the induced moments in the poles as a result of lateral wind pressures.

The fiber strengths of the wood poles are derived as the product of some constant A and the ground-line circumference raised to a power B. The constants A and B are 8,480 and 0 respectively for Southern pine poles [4]. Other than species and classes, another factor that affects the fiber strengths of wood poles is the mode of pre-treatment of the poles before installation into the distribution network. Poles that are dried by heating in a non-aqueous solution under vacuum are said to be boultonized. This process reduces the strength of poles. Steaming may also reduce the strength of poles.

Assume the following proportions of modes of pretreatments of poles. The assumed percentage reduction in strength per pretreatment type is also given.

- 10% of the poles boultonized → 10% reduction in strength

- 10% of the poles steamed→15% reduction in strength
- 80% of the poles air-dried →0% reduction in strength

The accepted calculation for the moment capacity of a pole at the ground line M_{gl} as given in [4] is

$$M_{gl} = \frac{AC_{gl}^{B+3}}{32\pi^2}.$$

The distributions of moment capacities for the class three and five poles, with ground-line circumferences derived from wood pole databases, are shown in Figure 39.

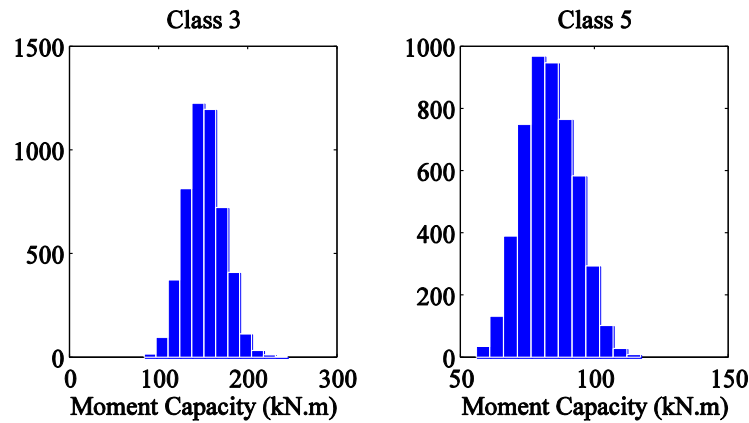


Figure 39 Distributions of moment capacity of class three and five wood poles.

Five thousand samples of the moment demand of wind loads were generated using the LHS technique with the following assumptions on the variables in wind pressure F , for a given wind speed. The probability distributions and associated parameters for the wind pressure model provided in [5] is used in this study and presented in Table 5.

Table 5 Wind load statistics for ASCE 7 load criteria (Ellingwood and Tekie, 1999).

Variables	Distribution	Coefficient of Variation (COV)
G	Normal	0.11

C_f	Normal	0.12
K_z	Normal	0.16
K_d	Normal	0.08

Using LHS, 5,000 random samples of wind pressure parameters are generated. These samples are randomly paired with the 5,000 random samples generated for the dimensions of the utility poles. The applied wind lateral forces on the poles for a given wind velocity are then calculated by determining the wind pressure on the poles using the sampled parameters and multiplying the resulting pressure by the sampled width of the poles. The induced moment in the poles at the ground level is then equal to the summation over height of the lateral forces times their distance from the ground.

Combining the capacity and demand models, the assessment of the probability of failure of wood poles can be performed at a given wind velocity. Figure 40, for instance, shows the probability distribution of demand and capacity for classes 3 and 5 for the case where the wind velocity is 200 mph. The area of the section where the demand overlaps the capacity is the probability of failure.

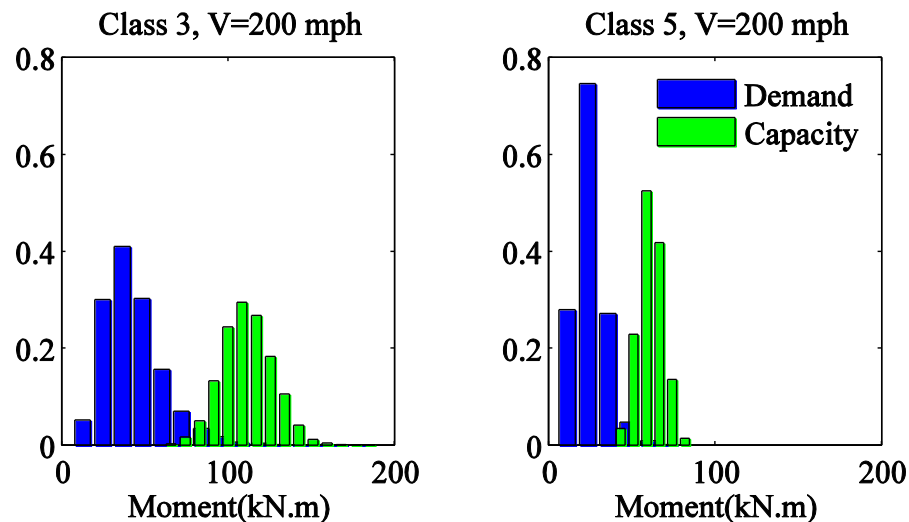


Figure 40 A plot of probability versus moment: The distribution of moment demands of wind loads and moment capacities of wood poles of classes three and five for V=200 mph in power distribution networks.

Repeating this analysis for a range of wind velocities will provide an estimate of the probability of failure of wood poles conditioned on the wind velocity which is essentially the fragility curve of the utility poles. The fragility curves of the poles in classes three and five are shown in Figure 41. Considering the distribution of the parameters involved in capacity and demand estimation phases, the standard deviation around the mean fragility curves is determined and the curves one standard deviation above and below the mean fragility curves are shown in Figure 41.

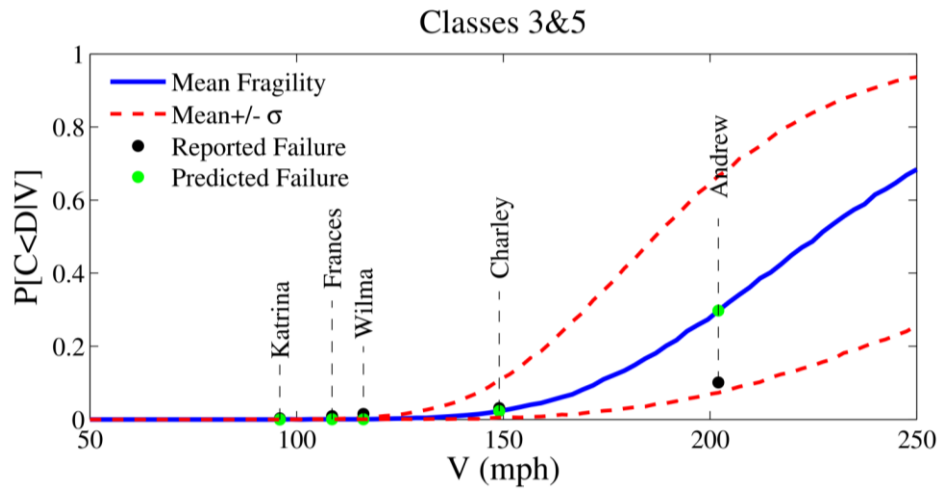


Figure 41 Combined mean fragility curve of Class 3 and 5 wood poles (solid line) ± 1 standard deviation (dashed lines), excluding wind loads on overhead lines.

The reported wood pole failures are from Florida Power and Light (FPL) as was published following the huge devastation caused by Hurricane Wilma of 2005. Table 6 shows the proportion of failures of wood poles replaced by Florida Power and Light as a result of five historical hurricanes [6]. The maximum wind speeds experienced within the time window of Wilma were used for validation. The reported failure rates fall within one standard deviation of the fragility curve so obtained. It must be noted that the fragility assessment was done ignoring the effect of load sharing and the ages of the poles. The assessment models uncertainty of the capacity of *new* wood poles of different dimensions and classes.

Table 6 Proportion of Florida Power and Light wood pole restorations after five real hurricanes [6].

Year	Name	Poles exposed to 74+ mph wind speeds	% of exposed poles that failed	Max Wind Speed (in mph)
2005	Katrina	343,200	0.3	96.6
2004	Frances	397,134	0.9	108.1
2005	Wilma	773,700	1.5	117.3
2004	Charley	222,666	3.1	149.5
1992	Andrew	203,500	10.1	201.25

3.2.2 A Second Predictor of Stress-Based Failure Risk: Age

The vulnerability of components is expected to increase as they age, especially under extreme weather conditions. Incorporating aging into fragility assessment implies two dominant predictors of failures of components. Age is an important variable to model since the physical assets owned by electric power utilities are becoming very old. The average age of transformers, substations, underground cables, overhead lines and wood poles in the utilities continue to increase. Several around the country are already in excess of 40 years old, way past their average design lifetimes of approximately thirty-five years [7]. An age-dependent fragility model of the components would require that the capacity of the components be written as a function of time.

Let the capacity of a component be represented by some random variable R , the age of the component by random variable T , and the *aged state* of the component at time t by some binary random variable D . The *aged state* could be severe decay in the component. Let D be equal to one if the component is found to be in the aged state that would make it highly vulnerable to failure and zero if it is not. Then the conditional expected capacity of the poles can be written as

$$E[R|T] = E[E[R|D, T]] = \sum_d E[R|D = d, T]P(D = d|T).$$

Since random variable D is binary, the summation can be written as a sum of two expressions as given in

$$E[E[R|D, T]] = E[R|D = 0, T]P(D = 0|T) + E[R|D = 1, T]P(D = 1|T). \quad (18)$$

The first expectation to the right of the equality sign can be read as the expectation of the initial capacity of the component, assuming the aged state is the only time-dependent factor that affects its capacity. The last probability is the probability that the component is in an aged state. Finding the relationship between the initial capacity and the capacity after time t is useful in estimating the time-dependent expected residual capacity of the components at its aged state. By so doing, samples of demand and capacity can be generated using the Latin hypercube sampling technique. These are compared as in the previous subsection to develop the fragility model for a range of wind speeds.

Sample Results: Wood Poles

Deterioration of wood poles may occur over time (as they increase in age) from such environmental conditions as moisture content, temperature and oxygen. This is expected to reduce their capacity or strength. Environmental conditions differ all over the country. However, due to the scarcity of a variety of field test data, analysis of wood pole decay carried out in [8] is used in developing a model for age-dependent fragility assessment of wood poles.

The age-decay relationship found in [8] was derived using field test data consisting of age, initial strength, and one effective ground-line circumference per pole for 13,940 wood poles. The ages of the poles ranged from 1 to 79 years with a mean age of 30 years. From these data, it was found that the percentage of decayed poles increases linearly and takes the form

$$Per(t) = b_1 t - b_2,$$

where $Per(.)$ is a time-dependent function that describes the percentage of decayed poles. From the regression analysis of field test data, the values of b_1 and b_2 are found as 0.004 and 0.04 respectively. By extracting and analyzing values of the data-points in their publication, an exponential model was found to provide a good fit to the data.

Clearly, $Per(.)$ is the conditional probability that a wood pole at age t is decayed, that is,

$$P(D = 1|T = t) = \min(\max(Per(t), 0), 1).$$

Therefore,

$$P(D = 0|T = t) = \min(\max(1 - Per(t), 0), 1).$$

The use of min-max functions ensure that the postulates of probabilities are obeyed.

Finally, $E[R|D=1, T]$ in Equation (18) is the expected residual capacity of the decayed poles at age t . The loss in the capacity R can be written as the multiplication of the initial capacity by the percentage loss as a function of age.

$$E[R|D = 1, T] = E[R_0(1 - L)|D = 1, T],$$

where L is a time-dependent function describing the percentage loss in the initial capacity of the poles. Assuming that the decay process is statistically independent from the initial capacity of the poles as is realistic,

$$E[R|D = 1, T] = E[R_0|D = 1, T]E[(1 - L)|D = 1, T].$$

Considering the fact that the current state of decay as well as the age of the pole does not impact the initial capacity of the poles, the equation can be further expanded as

$$E[R|D = 1, T] = E[R_0](1 - E[L|D = 1, T]).$$

Using field test data, [8] showed that the strength loss percentage of wood poles follows also follows a linear model,

$$Lspm(t) = a_1 t - a_2,$$

where $Lspm(.)$ is the lost strength percentage mean, and a_1 and a_2 are the regression parameters found as 0.014418 and 0.10683 respectively. Consequently, one can write

$$E[L|D = 1, T = t] = \min(\max(Lspm(t), 0), 1).$$

Substituting the equations into Equation (18),

$$E[R|T = t] = E[R_0][1 - \min(\max(a_1 t - a_2, 0), 1) \min(\max(b_1 t - b_2, 0), 1)].$$

Figure 42 shows the expected residual capacity of the poles as a function of poles age for the case of linear and exponential fits to $Per(t)$.

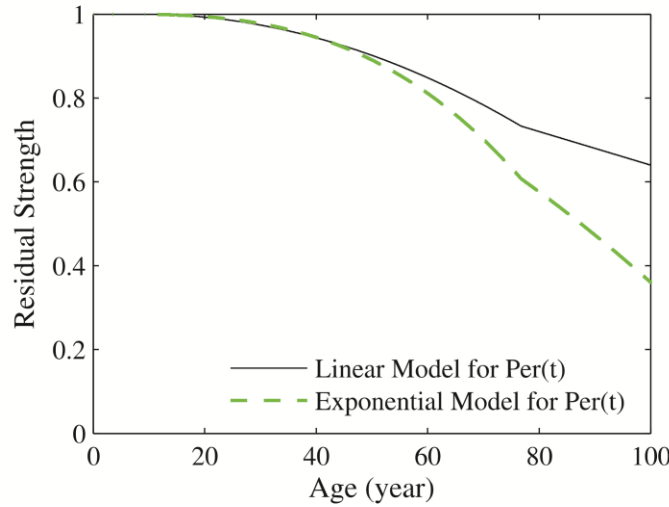


Figure 42 Expected residual strength of the poles as a function of age.

Incorporating the time-dependent capacity model of Equation (18) in the formulation of fragility models, a set of fragility curves are developed for wood poles of Classes 3 and 5 for various ages and for the two models of $Per(t)$. These relationships are derived using Monte Carlo simulations with 20,000 samples generated through Latin hypercube sampling method. These fragility curves are shown in Figure 43.

As the age of the poles increases, their probability of failure increases. This increase is not noticeable from age 0 to age 25 but the increase in the probability of failure becomes significant for older wood poles especially for the exponential models (Figure 43 b, d). The curves also show that newer distribution wood poles are not

susceptible to damage from low category hurricanes (under 100 mph) as are aged poles. In the coming years, electric utilities may start to experience annual failures of old wood poles rather than failures in years of severe (category 3+) hurricanes as the average ages of their older poles increase.

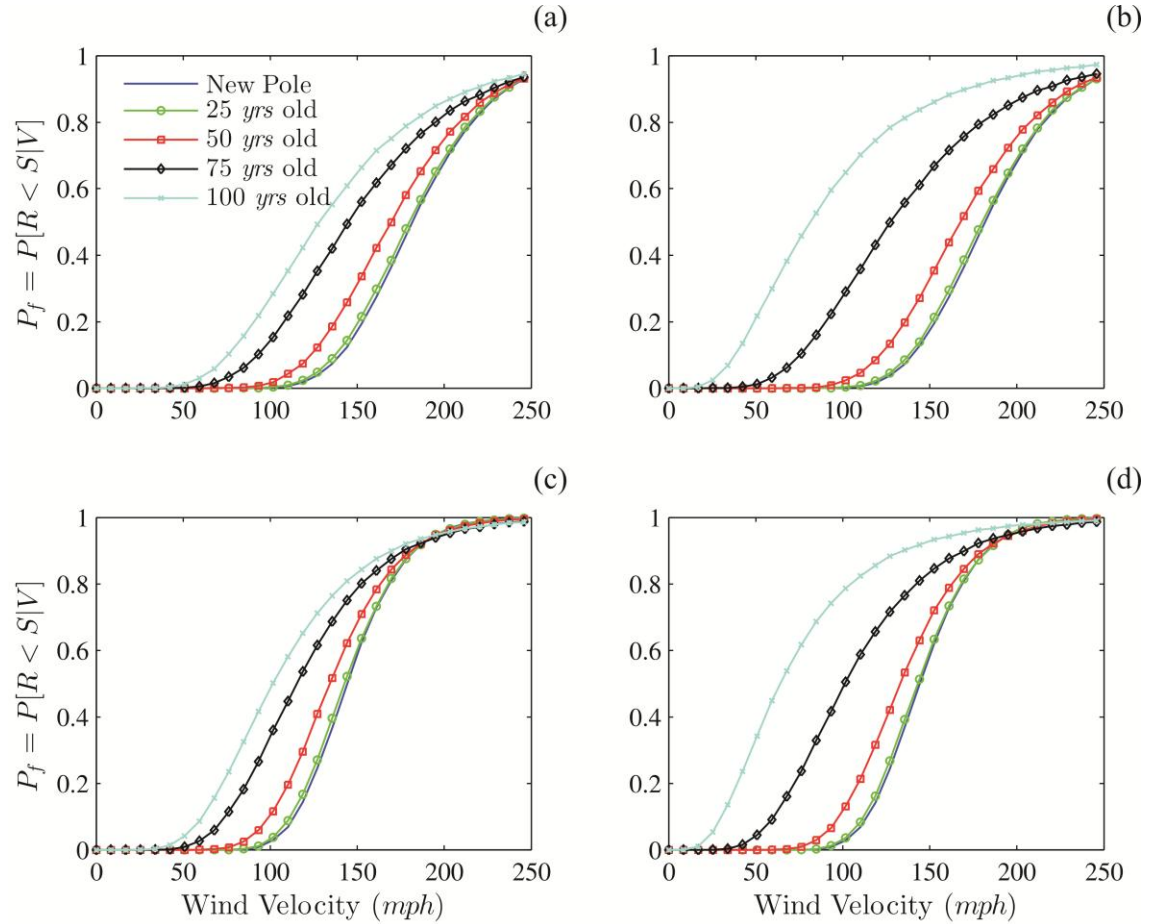


Figure 43 Fragility curves of utility poles of (a) class 3 with linear model for $Per(t)$, (b) class 3 with power model for $Per(t)$, (c) class 5 with linear model for $Per(t)$, (d) class 5 with power model for $Per(t)$. “C” is the capacity of the wood pole and “D” is the demand of wind on the poles.

3.2.2.1 Onset of Increasing Fragility by Age

The developed framework for age-dependent fragility assessment of components is useful for understanding the impact of preventive maintenance on components as they age. For each possible simulated wind speed, the difference between the probability of

damage or failure of a newly installed component and that of a component that has been in service for a longer period reflects the benefit of preventive replacement of older components. It represents failures that have the potential to be averted.

In the age-based scenario of failure analysis, which assumes mild or regular environmental stress, the suggested start-age of component inspections is logically the threshold or onset of increasing failure risk. Similarly, the start-age in a combined stress and age-based failure scenario is recommended to be the age at which the difference in fragility or the probability of wind-failures becomes evident. The failures of components that have been in service for short periods of time cannot be regarded as negligence. They constitute random failures (with or without preventive measures). Thus, inspecting them would be of no benefit to the electric utilities.

Significance may be measured in terms of a simple average of the difference between pairs of fragility curves (for an older component relative to a new component) or statistically using a two-sample t-test for means. Modifying the notations of Equation (17) from $P(C < D | IM)$, where C is the capacity of the component, D is the demand of the wind load on the component and IM is the intensity measure (wind speed) to $P_f(w, t)$, where w is the wind speed and t is the age of the component, the measure of difference in fragility curves is defined as

$$MD(t) = \frac{\sum_w [P_f(w, t) - P_f(w, 0)]}{n_w}, \quad t = 1, 2, \dots, w > 0,$$

where $MD(t)$ is the average of the differences between fragility curves of age t and the fragility curve of a new component, $P_f(w, t)$ is the failure probability at wind speed w and age t , and n_w is the number of simulated wind speeds used in generating the fragility curves.

For age t at which the difference is small, the measure $MD(t)$ is expected to be negligible. However, at some age of significance, the measure should become noticeably

greater than zero. The t-test may be used to provide some information on the statistical significance of their differences.

Sample Results: Wood Poles

The age-dependent fragility curves generated for wood poles are analyzed to estimate the age at which difference in fragility becomes evident. The measure $MD(t)$ is found for $t = 25, 50, 75$ and 100 years using the linear and power models of degradation. Results are shown in Table 7 using 100 simulated wind speeds from 0 mph to 300 mph, with both extremes being very unlikely in high-wind events in the United States. The results show a relatively low value of 0.005 obtained for $t = 25$ years when all the wind speeds are used.

Table 7 The average difference of the fragility curves are shown as $MD(t)$ for $t = 25, 50, 75, 100$ years using 100 simulated wind speeds . The recommended minimum age for inspections is highlighted in bold.

Ages of poles	25 years	50 years	75 years	100 years
Linear model $MD(t)$	0.006	0.044	0.126	0.183
Power model $MD(t)$	0.005	0.049	0.186	0.349

The ratios between consecutive measures $MD(t)$ may be regarded as the increase in failure risk for the 0 – 250 mph range of hurricane wind speeds between the pole ages. For the linear model, where $MD(100)$ is about 30 times $MD(25)$, it could be said that the failure risk for the simulated speeds increases by a factor of about 30 between the ages 25 and 100 years. The age 25 years is recommended as the minimum age for pole inspections given the estimates in the table.

Running a two-sample t-test on the fragilities per age, where one sample is that of 0-year old poles and the other is one of 25 - 100 years old, the p-values of the t-statistics can be used to determine statistical significance of the differences between the fragilities.

Table 8 shows the results. For both the linear and power models, the p-values are below the 0.05 level for ages ≥ 50 years. However, the p-value reduces by about half between ages 25 and 50 years. The table provides some statistical corroboration to decision-making using the metric $MD(t)$.

Table 8 Mean of fragilities per age and p-values of the t-statistics of the difference between fragilities with the null hypothesis being that the fragilities are equal and $\alpha = 0.05$.

t (years)	0	25	50	75	100
Mean $P_f(t,w)$ for linear model	0.267	0.273	0.311	0.395	0.453
p-value for linear model	-	0.918	0.486	0.049	0.005
Mean $P_f(t,w)$ for power model	0.267	0.272	0.317	0.456	0.622
p-value for power model	-	0.932	0.439	0.004	0.000

3.2.3 A Strategy for Prioritizing the Geographic Scheduling of

Component Inspections using Environmental and Fragility Risk:

Development of the Inspect Index

Exposed power system components that are geographically spread around a large area served by a power utility, like a state or several counties in the state, tend to be inspected geographically. This works best logistically for inspection companies. Usually, the region with the oldest population of components is inspected first. Then, in consecutive years of inspection, the surrounding regions are inspected progressively based on proximity and inspector allocation till the number budgeted to be inspected in a year is met. This method makes it easy to track the components in the schedule that have already been inspected. The technique, though sensible, can be improved by the inclusion of fragility and environmental stress risk assessment of the components. This involves both age, as has partly been discussed in the former subsection, and historical hurricane information.

Component age data in different regions of a power utility can be extracted from geographical information systems (GIS), while hurricane data on each region may be obtained from different hurricane databases. One such database is the Historical Hurricane Track archive from the National Hurricane Center of the National Oceanic and Atmospheric Association (NOAA). Unisys and HurricaneCity are other sources of hurricane information. Using the age-dependent model on fragility assessment of a power system component, an index by which to identify regions of a power utility to inspect annually is developed.

The decision index developed is called an “inspect index” for identical environmentally stressed aging components per district or county served by a power utility. The intention is that inspections of the components can be scheduled geographically by sorting respective districts in descending order of the inspect index. The inspect index per district or region is defined as the evaluation of the fragility curve at a point s for age a .

$$Inspectindex = P_f(w_a, t_a), \quad (19)$$

where w_a is the average of all historical hurricane wind speeds in the region greater than the maximum no-effect wind speed derived from the fragility curve of the component and t_a is one standard deviation more than the average component age (or the average component age past the minimum age of inspections) in the same region. The maximum no-effect wind speed is the minimum wind speed at which the loading or moment demand of the hurricane can exceed the moment capacity of the component. In other words, for greater wind speeds than the maximum no-effect wind speed, the fragility of the component is greater than zero. The standard deviation is used as a worse care scenario age since its failure risk would be higher than the average age.

See Figure 44 for illustration.

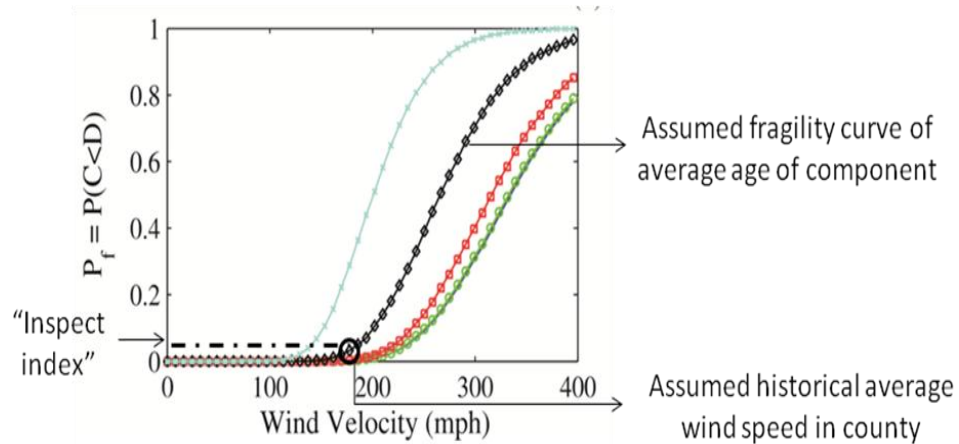


Figure 44 Illustration of derivation of the “inspect index” for a component in a county or district, signifying the fragility-risk of components in that region given historical hurricane information.

A flowchart of the scheduling of components for inspections is shown in Figure 45. It begins with a selection of cities served by a power utility. Data on component locations and hurricane data are extracted from GIS and hurricane databases per city. The respective ages and wind speeds are used in determining the inspect index as illustrated in Equation (19) and Figure 44. The indices are sorted in *descending order* for selection of the highest priority cities. In each city or district, only n_{ci} components that are over the recommended start-age of inspections as found in subsection 3.2.2.1 are selected for inspection. Assuming the inspection interval or schedule is set at c years, the excess of $\sum n_{c[i]}$ (the sum of all components over the start-age in the ordered cities) from the number of components planned to be inspected per year (N divided by c) is scheduled for the following year(s). If there are no excesses, the cumulative cost of components not inspected represents savings in adopting this strategy relative to the traditional strategy of the utility.

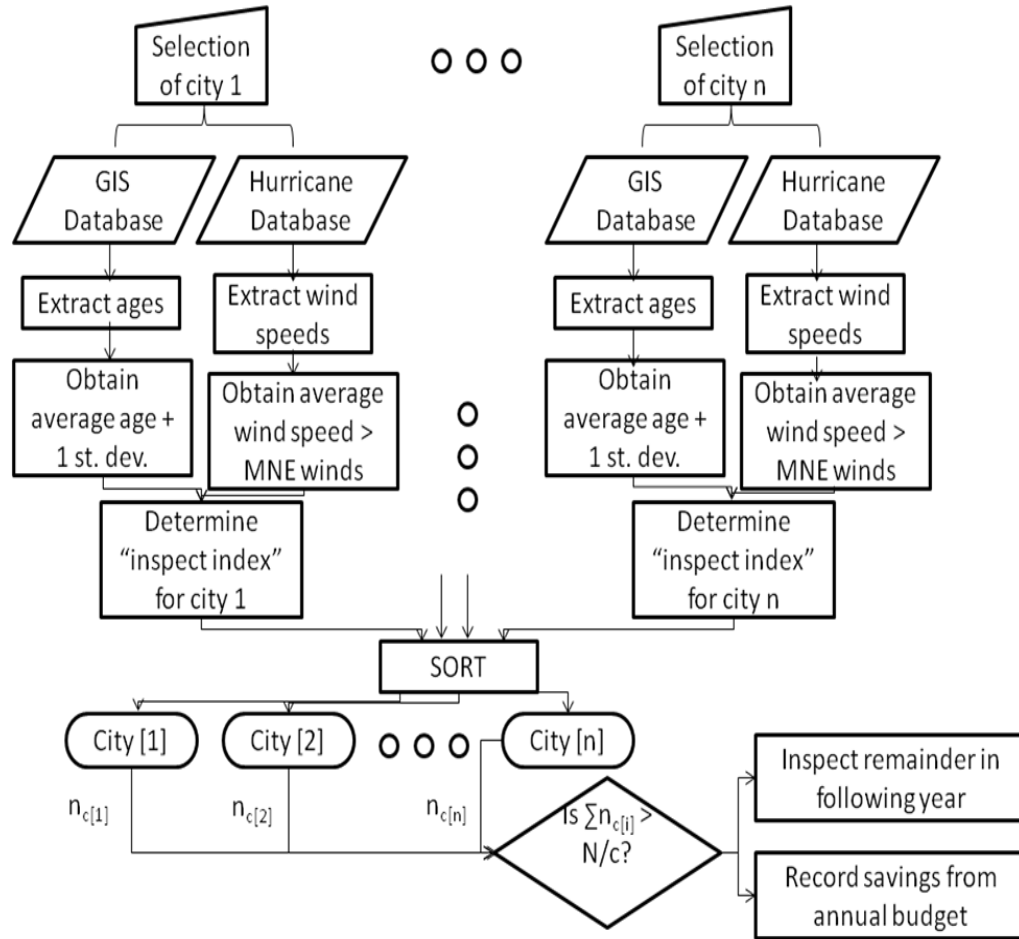


Figure 45 Flowchart of strategy for scheduling inspections of environmentally stressed aging components in electric power utilities, where MNE stands for maximum no-effect winds, City [i] is an ordered list of cities, $n_{c[i]}$ is the number of components over the inspection start-age of components for the ordered list of cities, N is the total number of components owned by the electric utility and c is the inspection interval or schedule.

The inspect index should be determined each year since the age distributions (average ages) of the components per county or region may change annually with replacements of failed components or components replaced preventively. The proportion of components in each ranked region inspected in a year should be a function of their relative inspect-indices. For instance, if two regions have the same index, an equal number of either should be inspected in that year. The number of components inspected per year should coincide with the utility budget and should not include those recently inspected. The inspection intervals for most components are often set by inspection

companies and communicated to the power utilities. Scheduling represents a consulting need for power utilities. Thus, it is defined to be a choice variable in Chapter 5 for the optimization of preventive replacement programs for environmentally stressed aging components.

3.2.4 Economic Effects of Component Replacements

The replacement of any component comes with a charge. It includes the cost of purchasing new or spare components, the cost of locating the component for “change-out,” the accessibility of the component at the time of change-out, the labor involved in replacing the component, the time of replacement, among other charges. Corrective replacement is best when the consequences and costs arising from failures of a component are low. On the other hand, preventive replacement is best if component failures lead to severe costs and consequences.

3.2.4.1 Replacing Components Correctively versus Preventively

By running components to failure, the utility incurs costs of corrective replacement. During inspections, components that are diagnosed to be below some safety, electrical or reliability standard may be scheduled for replacement (preventive replacement). The cost of replacing components after a failure is sometimes greater than the cost of replacement during inspections. This depends on a number of factors, some of which are listed below.

- Time of replacement of the component: When critical components fail, technicians may be sent to the site of the component at off-hours in the day. This could result in overtime hours billable to the utility. The United States Department of Labor (DOL) has rules covering certain employees under the Fair Labor Standards Act (FLSA). Using specified guidelines, such employees may receive overtime of at least one and one-half times

their regular rates of pay. When the replacement occurs during normal working hours (as for preventive replacements or corrective replacements that occur during normal working hours), overtime pay does not count. In addition, when the number of components that fail exceeds the number of spare components, the cost of rush deliveries of the components may be significantly higher than the cost of normal deliveries of the components.

- Union or non-union workers: Say employees performing corrective or preventive replacements in a certain state are union workers. The difference in labor costs for either type of replacement remains unchanged during normal work hours. However, the difference arises outside normal hours where overtime rates may apply. Unions have been known to make significant efforts in securing good employee benefits. Thus, overtime rates for change-out technicians in a union may be higher than those of the technicians not in a union.
- Accessibility of components: The grid as a whole is not yet completely “smart”. As such, it is not always easy to locate failed components. When they are located, gaining access to the component may be difficult, especially when failures occur during bad weather conditions like floods and storms. This may cause the cost of replacement in such circumstances to be higher than normal. On the other hand, when the replacement occurs during times of inspection, the “change-out” personnel typically knows the location of the component and often comes prepared (with the right tools) to do the job.
- External utility workers employed: When replacements occur following harsh storms, utilities are sometimes forced to solicit help from neighboring utilities and contractors. Press releases and studies show that the number from past hurricanes sometimes exceeds one thousand.

- Location of the utility and components in the United States: The cost of replacing bulky power system components in the West coast is different from the cost in the East cost. For instance, preventive replacements of wood poles in the Southeast may cost less than \$4,000 per pole versus over \$12,000 in California. The cost of replacing a wood pole after a storm may be over five times its regular cost in the southeastern region of the country, while the factor may be smaller in the West.

In general, the cost of replacing components correctively especially in times of harsh weather could be significantly higher than the cost of replacing components during inspections. However, because of the variety of components replaced in such conditions, it is often difficult to distinguish between the costs specific to one particular group of components. If component failures were self-announcing, the location and timing of replacements would be straightforward, but since most failures are not self-announcing, the challenges in replacements continue.

In prior research, the cost ratio between corrective and preventive replacements has been used to show potential benefits from choosing preventive measures of component management over corrective measures. As the cost ratio increases, the benefit of preventive measures increases also.

Sample Results: Wood Poles

The locations and ages of wood poles inspected in the years 2000 and 2001 in a real power utility were extracted from the utility inspection database. Historical hurricane information on these locations was found and the average of the high intensity hurricanes (those over “Category 3”) obtained. Note that Category 3 hurricanes are over 96 knots or 111 mph. This speed is used as a threshold because fragility assessment showed that the probability that a newly installed wood pole that meets NESC and ANSI specifications hit by a hurricane under 96 knots is negligible.

The average wind speeds and one standard deviation of the average ages of the poles in the locations of the wood poles over the two years of inspection are shown in Figure 46. Data extracted from both years of inspections show that the same general areas were inspected twice, though the same wood poles were not inspected. The different regions (dashed boxes) shown in the figure were allocated to different inspectors in each year.

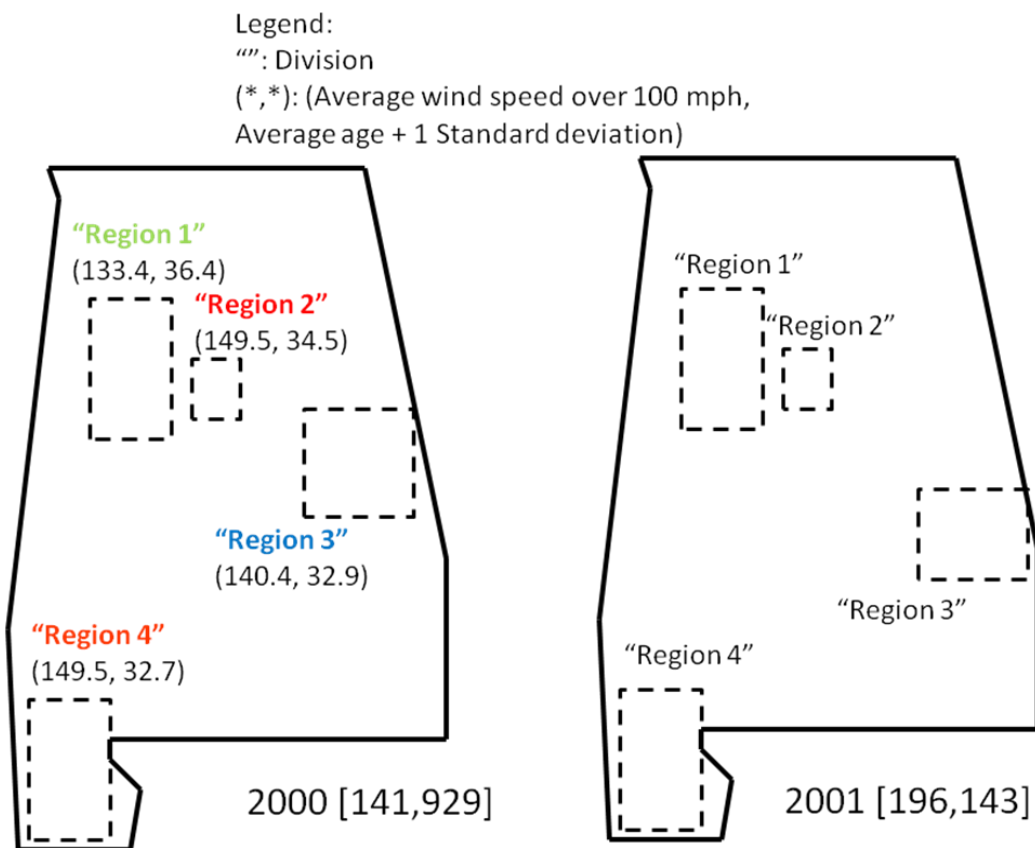


Figure 46 Regions of a power utility, where wood poles were inspected in 2000 and 2001, showing the numbers in parentheses (#, #) as the historical averages of wind speeds from 1900 to present in mph and one standard deviation of the average ages (in years) of the poles inspected in the regions respectively. While Region 2 has the highest inspect index, Region 1 has the least.

The color-coding of the regions in the figure: red for Region 2, orange for Region 4, blue for Region 3 and green for Region 1, is used to show the suggested priority order of inspections, where red represents high priority and green, low priority. The wind

speeds and average ages were compared to the appropriate fragility curves and ordered. The risk of poles in Region 2 through Region 1 in order is evaluated at inspect-indices equal to 2.8×10^{-4} , 2.7×10^{-4} , 1×10^{-5} and 0, using the linear model of degradation for the age-dependent fragility curves. The rankings are the same even when the exponential model is used though the specific risks differ. Since the indices for Region 2 and Region 4 are very close and much greater than that of Region 3, an equal number of either should be inspected in one year.

Using a minimum age of 25 years as the minimum age for inspections of wood poles as from Chapter-section 3.2.2.1, the power utility saves some cost and reduces its pool of searching for poles most at risk. The latter increases the likelihood of inspecting and potentially replacing at-risk poles before they fail. The numbers and costs of excluding either under-20 or under-25 year old poles are shown in Table 9. The search volume is reduced to 40% of the original inspection volume by inspecting poles over the age of 25.

The table shows that for a unit cost of inspections C_i of \$20, the utility stood to save over \$1.5 million in 2000 by excluding under 20 or 25 year old poles; in 2001, over \$2 million. In a year of high wind speeds (105 mph), assuming the number of poles under 20 or 25 years are uniformly distributed, the numbers failing for either differ by less than 100, translating into \$200,000 savings for inspecting poles between 20 and 25 years old. While the total cost is less for 25 year old poles as highlighted in the table, either is recommended as the minimum age for inspections.

Table 9 Number of wood poles inspected in years 2000 and 2001 from an inspection database and costs of excluding poles less than 20 or 25 years using an inspection cost C_i of \$20 per pole. The cost of failures less than either minimum age is used with the inspection cost to determine the benefits of recommending either 20 or 25 years.

INSPECTION YEAR	2000	2001
N poles	141,929	196,143
C_i	\$20	\$20
Total for inspecting all N	\$2,838,580	\$3,922,860
Cost of inspecting $N > 25$ years	55,809 => \$1,116,180	77,426 => \$1,548,520
Cost of inspecting $N > 20$ years	68,736 => \$1,374,720	94,567 => \$1,891,340
Cost of one failure	\$2,000	\$2,000
At 105 mph, cost of failures for fragility ≈ 0.005 on average for $N \leq 25$	431 => \$862,000	594 => \$1,188,000
At 105 mph, cost of failures for fragility ≈ 0.005 on average for $N \leq 20$	366 => \$732,000	508 => \$1,016,000
Cost of inspecting $N > 25$ only + Failure cost for $N \leq 25$	\$1,978,180	\$2,736,520
Cost of inspecting $N > 20$ only + Failure cost for $N \leq 20$	\$2,106,720	\$2,907,340

Since the inspect-indices of the regions in Figure 46 were found to be 2.8×10^{-4} , 2.7×10^{-4} , 1×10^{-5} and 0, the proportion of poles to be inspected in each region (Region 2, Region 4, Region 3 and Region 1) in order is approximately 50%, 50%, 0% and 0% of the number of poles budgeted for inspections in that year.

3.3 Summary

Useful decisions that may affect utility expenditure on component replacement are made by acquiring and analyzing complete failure information on the components. For failures that are approximately age-based (possibly failures in mild environmental conditions), an exponential assumption on time segments of times-to-failure of the components can be very helpful in predicting the distribution function of future failures and for detecting an onset of increase in the failure risk.

In the earlier part of the chapter, an exponential failure dataset is analyzed as a precursor to failure datasets that have an increasing failure risk with age. Using a Gamma distribution, which is a conjugate prior to an exponential likelihood function, an algorithm was developed to obtain a closed-form solution to the predictive distribution of times-to-failure of components surviving by some present time instant using Bayesian statistics. The benefit of the methodology is that it does not rely on the rate parameter of the failure dataset, which is unknown at the beginning of observation. Instead, it uses the observed times-to-failure, the original number of components in the population, the time instant of analysis, and the parameters of the Gamma prior.

One limitation to the algorithm for application to real components is the unavailability of complete failure records. When the times-to-failure of proportions of a population of components are missing and say, missing completely at random, the algorithm tends to overestimate the expected time-to-failure of the surviving components. There are no known methods for compensating for the loss of information since it is not possible mathematically to produce or invent information that does not exist, though assumptions could be made from known data. An evaluation of the extent of wrong information can be deduced from missing failure data is conducted in this chapter. For a population where only 0.5% of the population had failed by the time of analysis, assuming that 50% of the failure was missing completely at random, an overestimation of

the rate parameter by about 27% was observed. The lesson in this evaluation is the importance of storing failure data: a mechanism that automatically records the time-to-failure of power utility components in a system would be beneficial.

Another limitation to the developed algorithm is a mix of the restriction of the prior density of the rate parameter to the Gamma density function and the assumption that the hazard risk of a component is exponential; when in general, most components show some type of bathtub shape. However, the ease of use of the exponential distribution may be exploited by modeling times-to-failure of components within an appropriately sized window as though they were from an exponential distribution of unknown rate parameter.

An algorithm was developed to estimate this rate parameter by assuming a doubly censored exponential dataset within a sliding time window. The failure risk of the component was therefore built gradually over time. Given that power utilities can take timely actions in performing preventive measures on their components, an algorithm that detects the onset of failure risk was integrated into the hazard estimation algorithm. For real replacement data and a synthesized dataset consisting of times-to-failure of different population-sizes generated using a Weibull distribution of scale parameter 50 and shape parameter 7 with a simulated onset at 30 years, the developed algorithm detected major increases in failure risk. A Kalman filter was used to determine the dynamic state of the failure process (smoothing the time-varying exponential rates). Extreme value theory was applied to detect the onset of increasing failure risk.

The high failure-volumes of sets of components under environmental stress prompted research into reliability of those components under stress. In structural engineering, this field deals with the fragility assessment of the components. The reliability analysis here is no longer with respect to time or age but the intensity of stress. Fragility curves were developed for new wood poles and the results compared to actual failure rates from known historical hurricanes. The fragility curves were found to estimate the actual events well within one standard deviation of the mean curves.

A key issue facing the power industry of late is that of aging infrastructure. So, aging of the environmentally stressed components was incorporated into the fragility model to evaluate failure risk. By applying an age-deterioration function from literature to the fragility function of the components, a framework was developed for generating age-dependent fragility curves of the components. For some intensities of stress, the performance of all ages of components was unchanged. However, for increasing intensities and ages larger than a certain threshold, the difference in performance became evident.

The developed frameworks and fragility curves were used in generating cost-efficient strategies for component inspections and creating inspection schedules. The mean differences between fragility curves of different ages compared to the fragility curve of a new component were used in identifying the age at which probability of failure across several levels of environmental stress became significant. This became a measure to define the recommended minimum age at which the environmentally stressed components should be inspected. The recommendation could potentially lead to millions of dollars in annual savings for the utility. It also improved the chances of inspecting, detecting and preventively replacing high-risk components. The fragility with respect to average age of components in a given region and the average maximum wind speed in that region were used as an index for ranking regions for inspection according to failure risk.

3.4 REFERENCES

- [1] R. E. Brown, (2006, June 21). IEEE GM 2006 Working Group on System Design [Online]. Available: <http://grouper.ieee.org/groups/td/dist/sd/doc/2006-06-Hurricane-Wilma.pdf>

- [2] *American Society of Civil Engineers (ASCE): Minimum Design Loads for Buildings and Other Structures*, ASCE 7-10, 2010.

- [3] *American National Standards Institute (ANSI), Specifications and dimensions for wood poles*, ANSI O5.1, 2008.

- [4] R. W. Wolfe, J. Bodig, and P. K. Lebow, “Derivation of nominal strength for wood utility poles,” U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, Madison, WI, Rep. FPL-GTR-128, 2001.

- [5] B. R. Ellingwood and P. B. Tekie, “Wind load statistics for probability-based structural design,” *ASCE Journal of Structural Engineering*, vol. 125, no. 4, pp. 453-463, April 1999.

- [6] KEMA Inc., “Post Hurricane Wilma Engineering Analysis,” KEMA Inc., Arnhem, Netherlands, KEMA Proj. 05-349, Jan. 2006.

- [7] R. E. Brown, “Aging infrastructure,” in *Electric Power Distribution Reliability*, 2nd ed. Boca Raton, FL: CRC Press, 2009, ch. 8, sec. 6, pp. 449.

- [8] Y. Li, S. Yeddanapudi, J. McCalley, A. Chowdhury, and M. Moorehead,
“Degradation-path model for wood pole asset management,” in Proc. 37th North
Amer. Power Symp., Ames, IA, Oct. 2005, pp. 275–280.

4 PREDICTIVE MAINTENANCE: DIAGNOSTICS AND PREVENTIVE REPLACEMENT OF HIGH-RISK COMPONENTS

Introduction

The terms “inspections,” “tests,” and “diagnostics,” are used to describe processes of qualification, verification and quantification of the extent to which a component is functioning and meeting design intent. Inspection focuses on condition parameters that can be visually, audibly or otherwise sensed by human beings and that have relation to deterioration (e.g. corrosion, rot, discoloration, signs of leakages, etc). A test on the other hand could be defined as a verification of a certain aspect of component condition or health. Tests normally involve measurement equipment, focus on a single condition parameter, and have clear procedures and standards against which the results can be screened. Yet, both inspection data and test data can be used to assess the overall component condition. Lastly, diagnostics involve measurement equipment and typically focus on combining several condition parameters and have less defined pass/fail standards as is the case with single parameter tests.

Each one of these techniques, or all combined (as data collection input) allow for the determination of the condition or health of a component with the conclusion of deciding whether a maintenance operation should be performed on the component. Recommended actions or predictive maintenance operations are logical follow-throughs of diagnostics. Yet one technique that provides more accurate data may be more invasive and expensive than the other. Hence, this chapter suggests a framework for the optimal combination of these techniques. The three terms will be used interchangeably in this chapter.

This chapter deals with developing a risk-informed process for making decisions on performing maintenance operations on components. The risk here is one of the

inaccuracies of diagnostic tests, and the maintenance operation covered is the preventive replacement of a component diagnosed to have fallen below performance requirements. Diagnostic tests are supposed to be non-destructive evaluations of the health of components. They should produce the same results of health as destructively testing the components. Here, destructive tests are used as the “gold standard” of the health of a component; comparison with diagnostic tests will give information on the accuracy of the diagnostic test.

Maintenance operations vary by components, costs and effect on failure-risk of components. Preventive replacements are by far the most expensive maintenance operation on a component. Some researchers do not view it as a maintenance operation while others do. It is classified as a preventive maintenance operation here since it is one possible follow-through from running diagnostics. It could be the single recommendation when other maintenance operations do not improve performance of the component or are unknown.

Navigating this Chapter

The focus of this chapter is the increase of diagnostic validity and accuracy in identifying components that fail to meet reliability standards by combining condition-classifications of multiple diagnostic tests. The complexity of the algorithms that compute the combined accuracies increases with the number of testing methods carried out serially on one sample of identical components.

Components sampled from the group already scheduled for preventive replacement are suggested for long-term assessment of diagnostic accuracy so that the loss of capital to the electric utility is small. This increases the sample size of components for destructive testing though narrowing assessment to only one type of accuracy

measure. The effect of length of the accuracy assessment period on utility spending is evaluated.

Statistical and subjective approaches found in literature are applied to accuracies of multiple diagnostic tests for ranking purposes, but are less practical from a business point of view in selecting the optimum diagnostic testing method. An economic analysis, on the other hand, is performed as a more persuasive and practical decision tool.

Statistical and combinatorial logic approaches are applied to classifications of individual diagnostic testing methods relative to classifications from destructive tests to obtain the best combination of tests that minimize misclassification costs in replacements or failures of components.

4.1 Analyzing the Accuracy of Individual Diagnostic Tests

The techniques applied in monitoring the conditions or health of power system components are numerous and varied. Components with several mechanical parts may receive multiple diagnostic testing on each part. The outcome of which may involve preventive replacement of parts at different times during the lifetime of the component. The decision to preventively replace the entire component may then be based on the series or parallel connection of the parts or the health of critical parts of the component.

Selection of a diagnostic testing company for each group of components depends on the cost of the test, sensitivity of the test, location of diagnostic vendors (for import/export reasons), among others. Testing methods differ with components. Megger, TTR and Doble tests are applied on transformers for predictive maintenance purposes. Visual inspections, sound and bore tests, and transmission of signals are a few tests performed on wood poles, which support overhead lines, to determine deterioration, reliability and strength of the poles.

The fields of epidemiology and biostatistics are replete with statistical and mathematical methodologies for the assessment of diagnostic test accuracies and

classifications of test subjects as healthy or not healthy. A control group is used as a baseline for comparing symptoms or progression of a disease in a test group. Similarly, for non-living subjects: power system components in this case, two groups used for assessing the accuracy of diagnostic tests could be a healthy group that passes destructive tests and a non-healthy group that does not. For ease in referring to the two groups, components that meet reliability standards or pass a destructive test are called *good* components. Those that do not are called *bad* components. Note that destructive tests are used here because the focus is on components that fail mechanically (from environmental loading). Destructive tests yield accelerated results compared to waiting for natural hazards to occur before analysis of accuracy.

4.1.1 Measures of Diagnostic Accuracy

Common measures used to characterize the accuracy of a diagnostic test are

- Sensitivity: The probability that a bad component is detected by a diagnostic test, $P(\text{positive test} \mid \text{bad component})$.
- Specificity: The probability that a good component is accurately classified by a diagnostic test, $P(\text{negative test} \mid \text{good component})$.
- Positive predictive value (PPV): The probability that a component that tested positive is bad, $P(\text{bad component} \mid \text{positive test})$.
- Negative predictive value (NPV): The probability that a component that tested negative is good, $P(\text{good component} \mid \text{negative test})$.

Since the measures are probabilities, they must obey the postulates of probability and fall within the $[0, 1]$ interval. An illustration of the classifications between destructive and diagnostic tests is given in Figure 47.

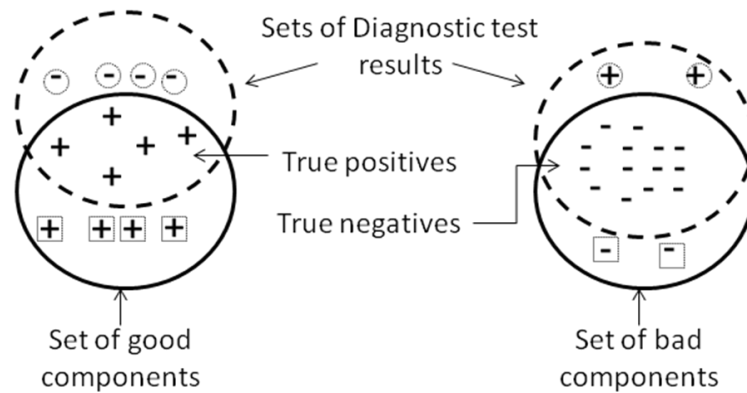


Figure 47 Illustration of the classification of components as good and bad based on destructive test results (lower sets), and predictions of conditions of the components using diagnostic testing (dashed sets). The circled markers are inaccurate diagnoses of the conditions marked by squares.

Notice that the areas of intersection in the Venn diagrams differ. This reflects higher predictions of good components (diagnostic negatives) than bad components (diagnostic positives). That is, the diagnostic tests are less likely to classify a good component as bad, than a bad component as good.

Mathematically, given N components in a field trial classified as follows: x components that test positive, y components that test negative, m bad components and n good components, a 2×2 contingency table for classification is in Table 10.

Table 10 2×2 contingency table for classification of conditions and diagnostic test results of components.

Destructive test (Health of the component) →			
Diagnostic test result ↓	Bad component	Good component	Total
Positive	a	b	x
Negative	c	d	y
Total	m	n	N

Estimates of the measures of accuracy are given below.

$$Sensitivity = \frac{a}{m}, \quad (20)$$

$$Specificity = \frac{d}{n}, \quad (21)$$

$$PPV = \frac{a}{x}, \quad (22)$$

$$NPV = \frac{d}{y}. \quad (23)$$

Bootstrapping may be applied to a small set of sample results to determine the statistics of the measure of accuracy by assuming observations from the component population are independent and identically distributed. The re-sampling method usually produces a larger sample size per run than the original size. A 95% confidence interval of the accuracy measure would be [0.025*K-th element, 0.975*K-th element] of the sorted elements from the bootstrapping experiment, where K is the number of bootstrapped samples.

A binomial proportion confidence interval (CI) may also be generated for the measures by assuming that *success* is the accurate diagnosis of the condition of a component and *failure* is the inaccurate diagnosis for each trial on the N components. This method relies on the assumption of a binomial distribution. Using a normal approximation interval for large enough sample size N, the CI found by approximating the binomial distribution with a normal distribution is

$$\hat{w} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{w}(1-\hat{w})}{N}},$$

where \hat{w} is an estimate of any of the previously mentioned measures of diagnostic accuracy, $z_{1-\alpha/2}$ is the 1- $\alpha/2$ percentile of a standard normal distribution, and N is the number of components on test.

4.1.2 Sampling Power System Components for Destructive Tests

An arbitrary, usually small, number of components are often used in running destructive tests with the goal of either learning about components: how they react in different environments, or evaluating the accuracies of diagnostic testing procedures on the components. Some examples are five components in [1], fifty components in [2], and eighty-eight components in [3]. The advantage of using small samples is that testing costs are low; the disadvantage, that bias in results may occur. In general, the following are important issues to consider.

- Identification of samples to test: Since destructive tests used in evaluating diagnostic tests incur costs for the purpose of gaining information, the components selected for testing should not affect power delivery unjustifiably. They should not cause shutdown of power.
- Determining a sufficient number of test samples to avoid bias in estimation: The dimensions, material and other properties of the same type of some components vary greatly. Their performance in destructive tests may therefore differ. Without appropriate representation of components, especially when sample size is very small, estimated diagnostic accuracies will be biased.
- Justification for carrying out destructive tests: of the idea of destructive testing of utility components for evaluating diagnostic tests is a hard idea to sell to an industry that has been conducting service in the same way for decades. The value or worth of a diagnostic test should be clearly stated. Using dollar values is expected to be more persuasive than presenting the statistics of accuracy of a contracted diagnostic test provider.

These points are addressed in this subsection.

4.1.2.1 Smart Selection of Destructive Test Samples

The cost of performing destructive tests may seem looming in an industry that tends to be rigid in its management operations. This is especially so when the benefit of the tests is unquantifiable or stated incomprehensibly. Thus, to keep costs low, external researchers sometimes request only a few components from the utilities to run tests. The utilities are not always cooperative.

Rather than persuading utilities to relinquish their in-service or operational components for destructive tests, a natural selection of potential destructive test samples occurs annually for aging utility components. For instance, the public service commission in the state of Florida mandates that wood poles be inspected every eight years. During these inspections, wood poles considered below standard are scheduled for replacement. Since these wood poles or other such components are not recommended for further use in the electric utility industry, they can be transported to external researchers for testing. Either the utility or the testing agency, as the utility sees fit, may then bear the responsibility of the costs of transportation, testing and eventual disposal of the component. Figure 48 is an illustration of the components that are a part of an annual utility inspection database and also candidates for destructive testing. This selection of components should therefore not influence system reliability when they are given to researchers.

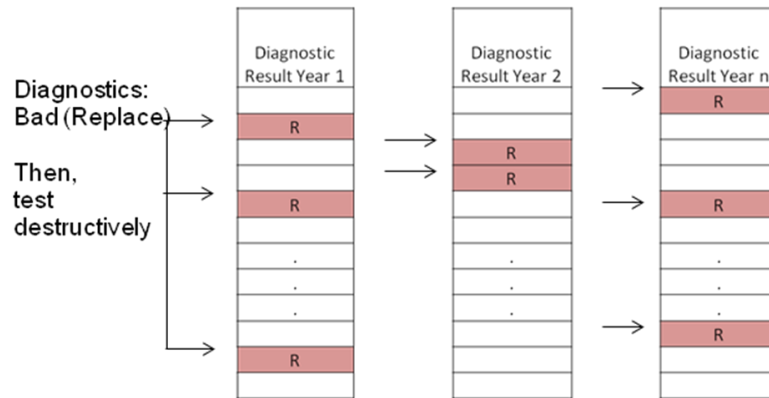


Figure 48 Illustration of annual destructive tests performed on some proportion of unique components (with label “R”) that test positive over n years of inspections for estimation of the positive predictive value (PPV).

The setback to this set of components scheduled for replacement is that only the positive predictive value (PPV) of a diagnostic test can be estimated. The selection of components for evaluation of the negative predictive value (NPV) may be accomplished using the former challenging strategy of writing proposals to the electric utility. That is, components that passed diagnostic test evaluations or that were not scheduled for inspection in some present year may be destructively tested.

4.1.2.2 Determining a sufficient sample size of destructive tests for assessment of diagnostic positive predictive value

A small sample size may lead to a biased estimate of positive predictive values (PPV), but will cost relatively little. Though increasing the sample size may improve the estimate, it will increase the cumulative cost of the destructive tests. Assume that the utility is interested in sponsoring the research on the evaluation of diagnostic test accuracy. Since management budgets are typically revised and allocated yearly, the assessment of diagnostic test accuracy may be a long-term process. In other words, a portion of each *annual* budget can be assigned to destructively testing a proportion of replaced components. In each consecutive year, more tests are carried out and integrated

with test results from prior years. The number of years that experimentation should be carried out depends on these factors.

- Steady-state (stability) in the estimate of the positive predictive value (PPV)
- Economic value to the power utility of conducting the destructive tests.

Technically, conducting the tests does not directly reduce or increase replacement costs or safety. Performance results of diagnostic test procedures are instead provided to support better decision making, leading to improved diagnostic process performance and better allocation of utility expenses.

4.1.2.2.1 Long-term assessment of diagnostic accuracy: positive predictive value

The estimate of the predictive value of a diagnostic test using a small sample size is expected to be biased. This is especially the case when the types of components tested have large variability in material properties or effects of age on their performance, though the components are otherwise identical. Interpretation of the statistical results in such cases may be misleading. For instance, during the NEETRAC evaluation of wood-pole diagnostic testing procedures, where 50 wood poles were broken to measure pole strengths, a diagnostic test provider considered leading in the market was found to be largely insensitive in detecting poles below ANSI standard.

In the provider's defense, the material properties of poles and the differentiation among pole species cause decay to occur in sometimes invisible areas. The leading provider, under the assumption that the majority of poles fail in the ground-line, restricted identification of decay, which is considered related to strength, to the ground-line. Meanwhile, the majority of the tested poles had properties that caused them to fail above the ground-line. Thus, the provider was unable to detect the majority of these poles. Publication of such performance results on diagnostic providers requires qualification of

the types of components likely to be undetected by the provider, and what types are likely to be detected. This entails large sample sizes for experiments.

An increased sample size can be accomplished in either one year or annually by combining new destructive/diagnostic test results with results of prior years, depending on the utility budget. Large sample sizes would go a long way to reduce bias in PPV estimation early and provide more confident information. In the absence of real-life results, long-term assessment of a diagnostic test is simulated in this section under the following assumptions.

1. The accuracy of a diagnostic test is fixed from year to year.
2. The components tested are identical and do not have much variability.
3. The accuracy of a diagnostic test for all ages and material properties of components is the same. That is, the test does not perform better for certain ages for instance, compared to others.

The estimate of the PPV for each consecutive year is a combination of the results from the previous year and the destructive test results in the present year. Thus, with each year, new information is acquired and the confidence interval of the mean is reduced since the sample size increases with every year.

Assuming initial bias from the first-year estimate of the positive predictive value of the diagnostic test, the optimum time to terminate destructive tests or have confidence in the estimated PPV is when steady-state characteristics are observed. In [4], Pawlikowski provides a listing of methods in literature used to determine the initial transient period in queuing theory: time before steady-state occurs. Since effectiveness of the decision rules varied by application or were unavailable, the coefficient of variation, which is sometimes found to be superior to variance in some applications, like in [5], is used here.

At steady-state,

$$|\overline{PPV}_{est}(n) - PPV_x| < \varepsilon_x,$$

where the first term in the inequality is the average PPV found for the n-th year of simulation, PPV_x is the steady-state estimate of the PPV: represents the true value of the PPV and ε_x is some threshold of maximum allowable difference between the annual estimates and the steady-state response: implies approximately no bias in estimation of the PPV. The coefficient of variation (CoV) per year of running the destructive tests is

$$CoV(n) = \frac{\sigma_{est}(n)}{\overline{PPV}_{est}(n)},$$

where the numerator is the standard deviation of the mean of the PPV in the n-th year and the denominator is the mean or average of the estimate in the n-th year. The objective in this section is defined as estimation of the number of years within which the coefficient of variance is small enough, for a given proportion of replaced components that are candidates for destructive tests.

For presentation to management, the coefficient of variation in this context is analogous to a measure of risk per unit return on evaluating the PPV with a proportion of the preventive replacements. A PPV of 0.6 implies that 60% of components preventively replaced were truly below specifications, standards or requirements. So, if the utility spent \$1,000,000 in preventive replacements of 500 components at a rate of \$2,000 per preventive replacement, the positive impact is in 60% of that cost or \$600,000. Similarly, the wasted investment is in \$400,000 of components that were replaced prematurely. The utility would work to reduce the latter cost by either finding another diagnostic test provider or encouraging the current one to improve.

Using a wide margin of uncertainty in PPV, say 0.5 to 0.7, the utility would estimate a loss between \$300,000 and \$500,000, which gives a \$200,000 difference. A much smaller variance implies a better estimate of the positive predictive value, and

therefore a better estimate of the cost of retaining the utility's current inaccurate diagnostic test.

4.1.2.2.2 A Framework for Defining the Economic Value of Conducting Destructive Tests

It was previously mentioned that a justification for conducting destructive tests can be challenging since the power industry tends to be inflexible in management practices. Unfortunately, researchers are often more interested in novel intellectual contributions using sometimes complex tools than in providing complete decision analysis tools that may be directly relevant to a manager. Thus, research findings that do not clearly translate into costs are usually not of interest to the utilities.

Literature in engineering can be found on experiments that use destructive testing to evaluate performance of components under different conditions, assess diagnostic testing procedures, and statistically analyze and compare diagnostic testing procedures. The interests are similar in the medical field, where diagnostic tools are applied in clinical trials to determine how well the tools identify disorders or illnesses in patients. Prior research focuses more strongly on accuracy values than on economic value. Analysis performed in this chapter presents the economic value of conducting destructive tests and serves to encourage power utilities to promote destructive testing program. This way, they will make better decisions on selecting diagnostic testing procedures for their components.

Given annual utility budgets on component management, the cumulative cost of running destructive tests for a number of years on a portion of components scheduled for preventive replacement is equal to

$$C_{dest} = \sum_{i=1}^{n_d} (C_{dt} \times num_replaced_i \times prop_i), \quad (24)$$

where n_d is the number of years of the destructive test program before steady-state occurs as discussed in subsection 4.1.2.2.1, C_{dt} is the cost of one destructive test, d is the interest rate of money, $num_replaced_i$ is the number of components replaced in year i , and $prop$ is the proportion of replaced components that are destructively tested in the year. It is realistic for the choice of proportions to be dependent on utility budget.

This cumulative cost represents a cost to the utility for gaining information on the accuracy of a diagnostic test. What does the utility gain by learning the PPV of a diagnostic test? A test with a PPV less than one replaces good components prematurely and leads utilities to waste their annual budgets. These components are actually above safety or reliability standards but were misdiagnosed.

Therefore, assuming it takes a destructive test program n_d year(s) as in Equation (24) for the mean of the PPV to settle, the benefit of obtaining a good estimate of the PPV is in providing the power utility with *confident information* about the economic impact of the utility's current diagnostic test provider. A PPV estimate tells the power utility how much of its annual investments on preventive replacements are lost to inaccuracy of the diagnostic test. The cost of inaccuracy from first year of destructive testing to n_d th year, when steady-state is reached is

$$C_{pmr} = (1 - PPV_x) \times C_{pr} \times num_replaced \times n_d, \quad (25)$$

where C_{pmr} is the annual cost of premature replacements: inaccurate classifications of components that are above specifications as being below specifications, PPV_x is the steady-state estimate of PPV: mean of the PPV found by replicating a destructive test

experiment (using Monte Carlo sampling if strictly theoretical), C_{pr} is the cost of one preventive replacement of a component, and *num_replaced* is the number of components replaced in one year.

Recall that when a large proportion of replaced components are selected for destructive testing per year, it takes a shorter amount of time to reach steady state compared to when a small proportion of components is destructively tested annually. By making early decisions on the positive predictive value (PPV) of a diagnostic test, a power utility can correspondingly take early actions in either changing a diagnostic test provider for a more accurate diagnostic provider or being convinced of a current accurate one.

In considering the economics of sizing the destructive test samples, enumerating the savings or lack thereof that result from potentially switching diagnostic providers, is an important decision-making tool for a power utility. An illustration of these concepts when 5%, 10%, 30% or 50% of annual preventive replacements are destructively tested follows. In the absence of real experiments, a numerical example is presented. Monte Carlo sampling is applied in repeating assessments of diagnostic test accuracies. A unique set of components are selected for destructive testing during each trial. Since the experiment is theoretical, the true conditions of all replaced components are predetermined given a PPV. The assumptions for this experiment are same as those listed at the beginning of the subsection. For the algorithm given, it is assumed that logic “0” represents components that are above specifications but were scheduled for replacement. In the same light, logic “1” represents components that are below specifications and were accurately scheduled for replacement.

Sample Results: Hypothetical Dataset

One hundred thousand of a certain type of components are simulated to be owned by a power utility. Of these, the number of components inspected or scheduled for

replacement per year by a diagnostic test is obtained as a proportion of the variables listed in Table 11. The destructive test results of all the replaced components are assumed to be known and equal to logic “1” for the number of components equal to PPV_true times the number of replaced components. They are logic “0” for the remainder of the components.

Table 11 Parameters of simulated test for estimating the positive predictive value (PPV) of a diagnostic test over a 10-year inspection cycle. MC size is the Monte Carlo size, PPV_true is the simulated PPV, C_{dt} is the cost of a destructive test, C_{pr} is the cost of a preventive replacement, U represents a uniform distribution.

N	100,000
MC size	10,000
Number of years	10
PPV_true	0.6
C _{dt} (\$)	200
C _{pr} (\$)	2000
% inspected	U ~ (1/12, 1/8)
% replaced	U ~ (0.02, 0.07)

Out of all the simulated replaced components, 5% to 50% are randomly selected for destructive testing up to 10,000 times which represents the Monte Carlo sample size for each year. The means and standard deviations of the estimates of the PPV for each year of simulation are found.

Descriptions of the variables used in the algorithm are

- num_of_insp_years: this is the number of inspection years to simulate.
- num_insp: this is the number of wood poles inspected during the inspection year analyzed. It is designed to be a random variable, distributed as the product of a fixed variable N components and a uniformly distributed percentage inspected. Here, num_insp ~ U(0.08*N, 0.125*N). This is based on limits from real utility inspection databases.

- num_replaced: this is the number scheduled for diagnostic replacement. It is designed to be a random variable, distributed as the product of num_insp and a uniformly distributed percentage replaced. Here, num_replaced $\sim U(0.02, 0.07) * \text{num_insp}$. This is based on limits of percentages replaced from real utility inspection databases.
- num_d: number of components destructively tested.
- diag_positive: this is a vector of all ones representing poles scheduled for replacement and left in service respectively during the inspection year analyzed.
- true_condition: this is a vector of ones and zeros for poles simulated to have failed and those simulated not to fail respectively using a given PPV.
- PPV_true: this is a fixed variable representing the true PPV of a diagnostic test.
- PPV_EST: this is output of computed positive predictive values (PPVs) for either one year of inspection or a combined number of years of inspections.
- Cdt: cost of a destructive test.
- Cpr: cost of a preventive replacement.
- Info_cost: cost of gaining information on prediction accuracy using destructive tests.

The algorithm to calculate the PPVs for single or combined years of inspections and simulated destructive tests is as follows:

PPV_INDIVIDUAL_COMBINED (N)

1. **for** i = 1 ... num_of_insp_years
2. N = 100,000
3. num_insp $\leftarrow \{0.08 + 0.045 * \text{rand}(1)\} * N$

```

4.    num_replaced (i) ← round ({0.02 + 0.05*rand(1)}* num_insp)
5.    diag_positive ← [ones(1, num_replaced)]
6.    PPV_true = 0.6
7.    true_condition ← diag_positive
8.    for j = 1: 10000
9.        locate ← random sample (1:num_replaced,
                                (1 – PPV_true)*num_replaced)
10.       true_condition (locate) ← 0
11.       Cdt ← 200; Cpr ← 2000
12.       num_d (i) ← proportion*num_replaced
13.       locate1← random sample (1:num_replaced, num_d)
14.       num_true_positives (i, j) ← sum {true_condition(locate1)}
15.       PPV_EST(i,j) ← sum {num_ true_positives(:, j)}/sum(num_d)
16.       Info_cost ← Cdt * num_replaced
17.    end
18.       E_PPV_EST (i) ← mean (PPV_EST, 2) % row-averages (per year)
19.       SD_PPV_EST (i) ← stdev (PPV_EST, 2) % row-deviations (per year)
20. end

```

The mean of the 10,000 positive predictive values (PPV) found per year of testing for the four proportions of selected components simulated is shown in Figure 49, with only four standard deviations of each proportion indicated.

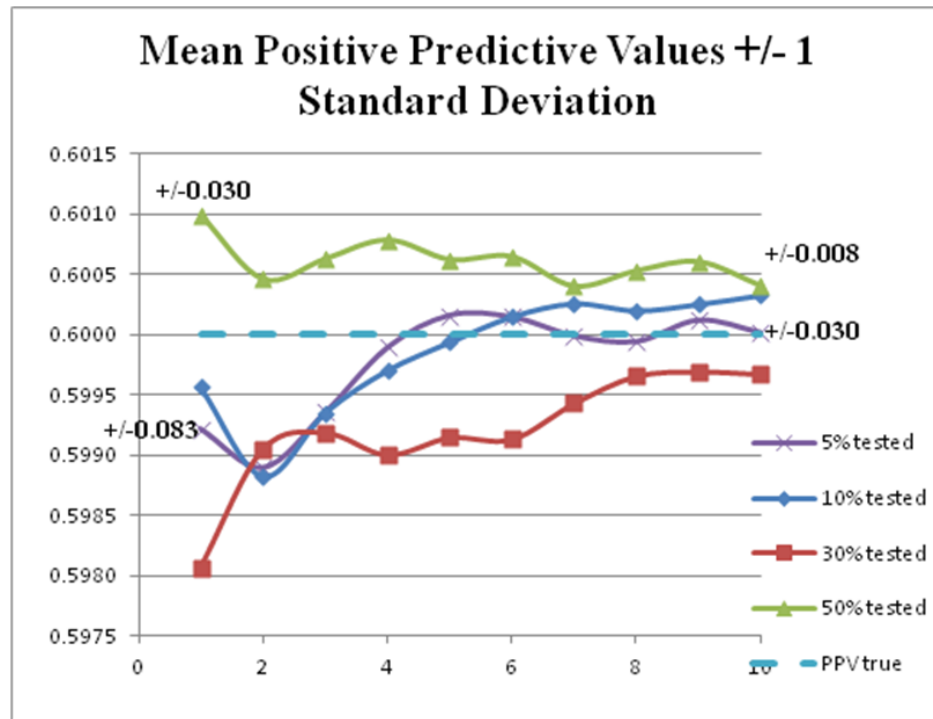


Figure 49 A plot of mean positive predictive value (PPV) versus number of years. Expected PPV of a diagnostic test per year of analysis for different sizes of destructive tests. The standard deviations of 5% and 50% PPV estimates for years 1 and 10 are also shown.

The table of means and standard deviations of the PPV estimates per proportion of preventive replacements is shown in Table 12.

Table 12 Table of means (E_PPV) and standard deviations of PPV estimates for a Monte Carlo sampling size of 10, 000 showing 4 proportions of preventive replacements that were destructively tested.

Years	1	2	3	4	5	6	7	8	9	10
5%										
E_PPV	0.5992	0.5989	0.5994	0.5999	0.6002	0.6001	0.6000	0.5999	0.6001	0.6000
Standard Deviation	0.0835	0.0664	0.0508	0.0436	0.0415	0.0384	0.0362	0.0343	0.0313	0.0302
10%										
E_PPV	0.5996	0.5988	0.5994	0.5997	0.5999	0.6002	0.6003	0.6002	0.6003	0.6003
Standard Deviation	0.0745	0.0464	0.0387	0.0330	0.0288	0.0255	0.0247	0.0226	0.0209	0.0197
30%										
E_PPV	0.5981	0.5990	0.5992	0.5990	0.5991	0.5991	0.5994	0.5997	0.5997	0.5997
Standard Deviation	0.0386	0.0244	0.0207	0.0190	0.0169	0.0155	0.0141	0.0129	0.0117	0.0110
50%										
E_PPV	0.6010	0.6005	0.6006	0.6008	0.6006	0.6006	0.6004	0.6005	0.6006	0.6004
Standard Deviation	0.0298	0.0188	0.0154	0.0126	0.0110	0.0102	0.0094	0.0088	0.0085	0.0081

The standard deviations of the means are observed to reduce with each year of the destructive test program and are commensurate with the increasing number of samples. The coefficients of variation (CoV) for each proportion of preventive replacements are shown in Figure 50.

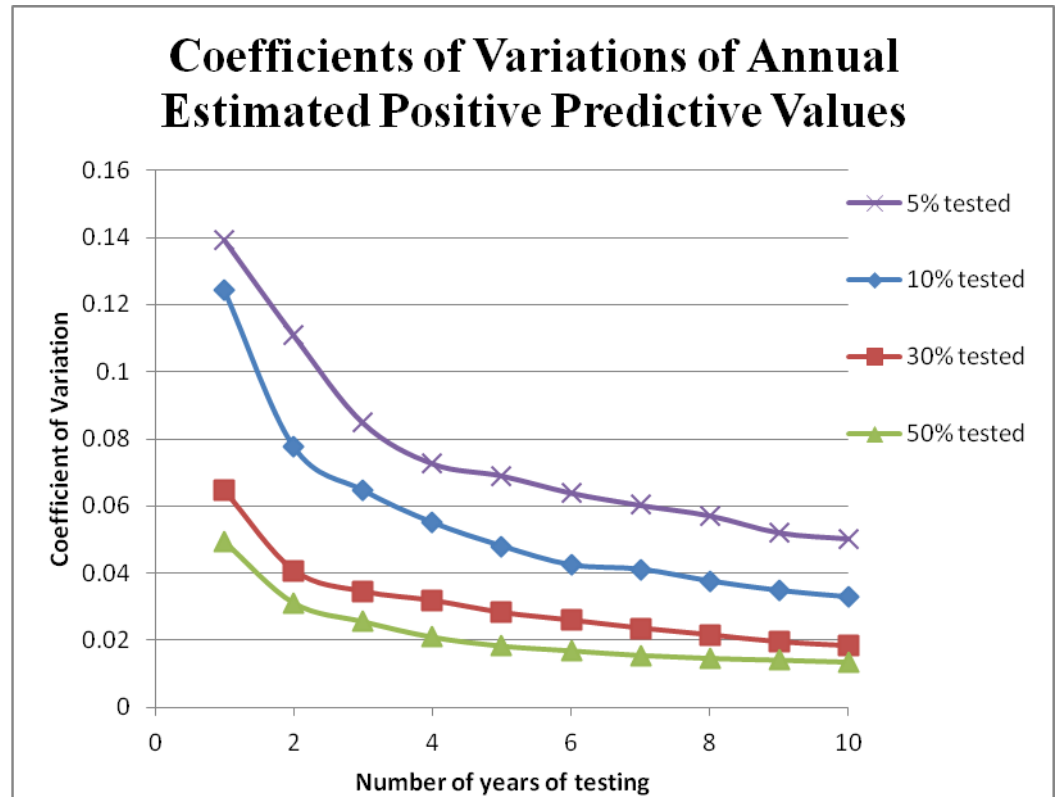


Figure 50 Coefficients of variation for PPV estimates for 1 through 10 years of destructive testing simulations, showing that the least coefficients were found when 50% of preventive replacements were destructively tested.

Using a minimum CoV of 5% as a threshold for confidence on the PPV estimate, the number of years n_d it takes to accept the expectation of the PPV estimate PPV_x is shown in Table 13. Note that the experiment makes a number of assumptions that caused the variances of the PPV estimates to be small among the proportions. In a real experiment, the variances are expected to be very large.

Table 13 Table of the length of period in years it takes to reach minimum coefficient of variation of 5%.

Proportion tested	5%	10%	30%	50%
Number of years to reach steady state (n_d)	10	5	2	1

Simplifying the problem for illustration, it is assumed that the same number of components (500) was replaced preventively for the n_d years for each proportion of destructive tests. Also, that the cost of destructive tests stays the same. The total costs of destructive testing for n_d and premature replacements estimated up to n_d years is shown in Table 14. Then, if the utility switches the diagnostic test provider to one that is higher, say 0.8, the reduced cost of premature replacements is shown.

Table 14 Table of costs of a destructive testing program when 5% - 50% of annual preventive replacements are destructively tested in each year. The number replaced is assumed to be 500 per year, the cost of one destructive test is \$200, one preventive replacement is \$2,000, the initial PPV is 0.6 and later 0.8 for proportions 10%- 50% after n_d years.

Proportions	5%	10%	30%	50%
PPV _x	0.600			
n_d (years)	10	5	2	1
Cost of destructive tests for n_d years (C_{dest})	\$ 50,000.00	\$ 50,000.00	\$ 60,000.00	\$ 50,000.00
Cost of premature replacements for n_d years (C_{pmr})	\$ 4,000,000.00	\$ 2,000,000.00	\$ 800,000.00	\$ 400,000.00
PPV _{new} after n_d years up to 10 years	-	0.8 for latter 5 years	0.8 for latter 8 years	0.8 for latter 9 years
New cost of premature replacements for remaining years (10 - n_d)	-	\$1,000,000	\$1,600,000	\$1,800,000
Sum over 10 years of destructive test program and preventive replacements	\$4,050,000.00	\$ 3,050,000	\$2,460,000.00	\$2,250,000.00

The analysis shows there is a \$1.8 million dollar difference in destructively testing 50% of the 500 components preventively replaced per year relative to a 5% replacement. Fifty-percent sample size testing implies early decisions on whether to retain a diagnostic testing provider based on the PPV, while five percent testing implies a longer decision time. Assumptions in the analysis include that the number of components replaced preventively per year stays constant and that the electric utility has knowledge of a diagnostic testing method that has a higher PPV. The latter may be the result of another electric utility assessing the accuracy of a competing diagnostic method concurrently. The analysis excludes the present value theory of money. While the latter may change the dollar amounts, the decision of 50% versus 5% will remain unchanged.

Inclusion of the NPV is important for complete decision-making on retaining or switching a diagnostic testing provider. The NPV gives information on the number of inspected components left in service that may fail in a short amount of time and incur large corrective replacement costs. The NPV is difficult to measure accurately since it depends on either the occurrence of sufficiently strong environmental hazards or very expensive destructive testing. It would take a much longer time to measure NPV confidently than the PPV.

The conclusion of the results is that it may be more beneficial to destructively test a large number of components per year rather than a few components. This encourages the power utility to take timely actions in switching or seeking new diagnostic testing companies for their components and averting future losses in premature replacements. The utility actions would create more demand for accurate diagnostic techniques, reduce monopoly in the field, and encourage more research and development of improved techniques, enhancing the chances of increased safety of in-service components in general.

4.1.3 Ranking Accuracies of Diagnostic Tests using Statistical Tools and Economics of Diagnostic Decisions

The rules ranking accuracies of diagnostic tests differ in literature. The statistical tools that can be applied in decision-making include logistic regression analysis, use of the receiver operating characteristics (ROC) curve and discriminant analysis. These are discussed in this section with applications to real power system components.

4.1.3.1 Ranking Using Deviance in a Logistic Regression Analysis

Logistic regression is used to predict the response of a categorical variable from one or more explanatory variables. In the case of diagnostics, the destructive test condition of a component (“good” or “bad”) is the response variable to be predicted from diagnostic test results. The analysis generates the coefficient estimates β , and their standard errors and p-values for the function in Equation (26).

$$\text{logit}(P[\text{bad}]) = \ln\left(\frac{P[\text{bad}]}{1 - P[\text{bad}]}\right) = \beta_0 + \sum_{i=1}^p \beta_i X_i, \quad (26)$$

where $P[\text{bad}]$ is the probability that a component is bad (and will fail a destructive test), X_i is the condition diagnosed by a diagnostic test, and p is the number of diagnostic tests. For one diagnostic test, as studied in this subsection, the value of p is one.

The p-value of the regression coefficients is used to determine whether a diagnostic test contributes significantly to the prediction of the condition of the component. If the p-value is less than a specified characteristic value, the diagnostic test is considered to be significant in prediction of the response variable.

Inference for logistic regression analysis is often based on deviance (twice the log-likelihood ratio statistic). A large deviance indicates that the logistic model does not fit the data well, and that another model may be appropriate. Asymptotically, the deviance has a χ^2 distribution. Hypothesis tests involving deviance are therefore

compared to percentiles of a χ^2 distribution. The number of degrees of freedom is the number of observations minus the number of parameters estimated.

Sample Results: Wood Poles

Six wood-pole diagnostic companies (called providers here) were evaluated by National Electric Energy Testing Research and Applications Center (NEETRAC). Initially, 52 poles were chosen for evaluation. Due to mechanical or transportation issues, which affected confidence of some results, a few of the poles were later eliminated from reports. The evaluation methodology involved prediction of strengths of the wood poles by the non-destructive (diagnostic) test providers, and then the breaking or bending of the wood poles to measure their moduli of rupture (true strengths of the poles). The breaking tests are destructive tests.

Based on American National Standards Institute (ANSI) and National Electric Safety Code (NESC) standards for allowable wood pole strength, the diagnostic tests were used to classify components as either bad (schedule for replacement) or good (leave in service). The difference between the means of the diagnostic test scores for bad components and good components is obtained per test. A two-sample t-test is carried out to find out whether the differences between the means are statistically significant; that is, if the classification or cut-off score was obvious. The null-hypothesis is stated to be equality between the means of either class. The results of the test are given in Table 15.

Table 15 Means and standard deviations of strengths of wood poles predicted by diagnostic tests one through six. They show the separations between predicted strengths for poles that failed the destructive test and those that passed, for significance level $\alpha = 5\%$.

Mean psi (Standard Deviation)	Diagnostic Test 1	Diagnostic Test 2	Diagnostic Test 3	Diagnostic Test 4	Diagnostic Test 5	Diagnostic Test 6
Bad components	5920 (3232)	3718 (2534)	4149 (2316)	4558 (2612)	4372 (1495)	6254 (2653)
Good components	7897 (565)	5836 (1170)	6826 (1139)	6872 (1543)	5750 (957)	7684 (484)

Any intersections between the classes are small. In general, the mean scores differed substantially between bad and good components. For a significance level of 5%, the null hypothesis was rejected. In other words, classification is clear among the diagnostic results.

The correlation coefficients between the predicted strengths and moduli of rupture were found for each diagnostic test. The correlation coefficients and corresponding accuracies of the tests are shown in Table 16 in order of NEETRAC's ranking for this experiment.

Table 16 Sensitivity and specificity of six diagnostic testing techniques applied to at most fifty distribution wood poles. The tests are arranged in the order that NEETRAC considered best performance.

Diagnostic Test	Correlation Coefficient for Strength Prediction	% of Reject Poles Correctly Diagnosed	% of Non-Reject Poles Correctly Diagnosed (34 poles were Non-Reject)
5	0.49	81 (13 of 16)	100
3	0.69	53 (8 of 15)	100
1	0.32	38 (6 of 16)	100
6	0.27	19 (3 of 16)	100
4	0.53	53 (8 of 15)	97
2	0.32	69 (11 of 16)	91

NEETRAC placed high priority on not diagnosing any good poles as bad (premature replacements). Thus, though diagnostic test 2 has a general accuracy (ratio of sum of true positives and negatives to the total number of wood poles tested) only second to diagnostic test 5, it has the worst ranking. The priority order for this decision rule is specificity first and then sensitivity. On one hand, the rule appears subjective; on the other hand, the rule is strict on specificity because of the difficulty in measuring the negative predictive value or the sensitivity from a small sample of poles.

Deviances of each test provider are obtained using logistic regression analysis. The coefficients of the analysis: β_0 for intercept and β_1 , the standard errors of the estimates, respective p-values and the deviance per model per diagnostic test, are shown in Table 17. The values in each column are obtained by using each of the diagnostic tests as sole predictors in the regression model.

Table 17 Statistics of coefficient estimates of logistic fits of distinct diagnostic scores. The response variable is the classification of poles as weak or strong by destructive tests. The diagnostic scores are the predictor variables. Lower deviances are better than higher deviances.

Statistics of Logistic Regression Model per Diagnostic Test	Diagnostic Test 1	Diagnostic Test 2	Diagnostic Test 3	Diagnostic Test 4	Diagnostic Test 5	Diagnostic Test 6
β_0	3.5766	2.1167	4.21	2.7091	6.4937	6.3599
Standard error of estimate	2.1218	0.9884	1.6906	1.2583	2.6897	4.5123
p-value for estimate	0.0919	0.0322	0.0128	0.0313	0.0158	0.1587
β_1	-0.0006	-0.0006	-0.0009	-0.0006	-0.0014	-0.001
Standard error of estimate	0.0003	0.0002	0.0003	0.0002	0.0005	0.0006
p-value for estimate	0.0919	0.0013	0.0021	0.0045	0.0058	0.096
Deviance	55.6161	53.2695	41.707	46.0632	50.6889	54.8678

Ranking predictions of the binary response (a destructive test result of weak or strong for the poles) per diagnostic test using deviance shows the following order.

- Best to worst diagnostic test: 3, 4, 5, 2, 6, 1.

These rankings differ greatly from NEETRAC's. The logistic regression analysis looks at overall prediction of the response variable (whether a wood pole is weak or strong if destructively tested) from the diagnostic tests.

4.1.3.2 Ranking Using Receiver Operating Characteristic Curves

The receiver operating characteristics (ROC) plot is a tool for visualizing the accuracy of a diagnostic test. It is a plot of the true positive rate TPR of a diagnostic test against its false positive rate FPR given a threshold that is used in classifying a diagnostic test result as either right or wrong. The TPR is mathematically equivalent to sensitivity while the FPR is the complement of specificity. The ROC plot area is a 1-by-1 space in the x-y plane; that is, the maximum value on either the vertical or horizontal axis is one. Using the ROC curve, a given diagnostic test can be easily compared to a perfect test or the result of a random guess. A perfect test should have a vertically-inverted L-shape, with its vertex on the (0, 1) point of the ROC plot area. Meanwhile, a random guess would be a diagonal line connecting points (0, 0) and (1, 1) on the ROC plot area.

The area under the curve (AUC) is a quantitative measure of the ROC that can be used in ranking diagnostic tests. The AUC typically has a value ranging from 0.5 (area of a right-angled triangle for a random guess) to 1 (area of a square of length 1). ROC analysis can be applied to logistic regression results by using the response variable in Equation (26) to denote the classification of a component. The closer an AUC is to the value one, the better it is.

Sample Results: Wood Poles

The ROC curves of the diagnostic tests on the wood poles tested by NEETRAC are shown in Figure 51. The classification variable used for the ROC analysis is the response variable of the logistic regression analysis. The figure shows that all diagnostic tests fall above the random-guess line, which runs between the points (0, 0) and (1, 1). The curves are seen to intersect making it difficult to identify what tests are superior to others. From the figure, however, the curves of diagnostic tests three and five appear to be the most consistent in terms of proximity to the (0, 1) point on the plane. The visual

qualitative method of ranking is inefficient and inconclusive; so the area under each curve is estimated. Results are in Table 18.

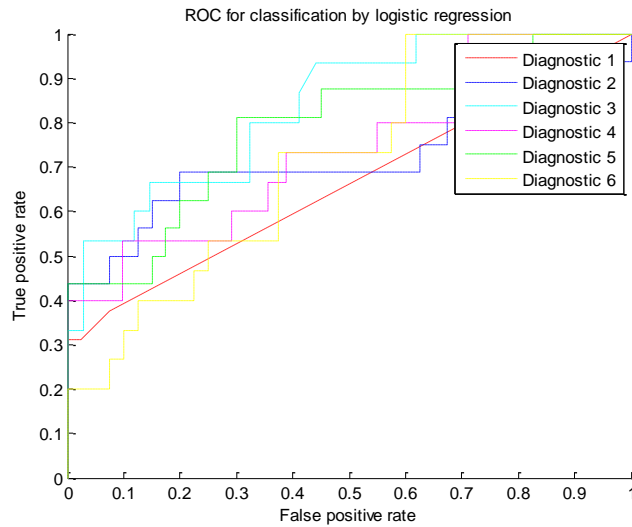


Figure 51 Receiver operating characteristics curves of six diagnostic tests using the results of logistic regression analysis for classification of components.

Table 18 Area under the curve (AUC) for logistic regression model of each diagnostic test and combined prediction of tests one and three.

Diagnostic	1	2	3	4	5	6
AUC	0.6609	0.7281	0.8363	0.7398	0.7828	0.7150

The table shows the following descending order of superiority:

- Best to worst diagnostic test: 3, 5, 4, 2, 6, 1.

The order is similar to the results of the logistic regression model, since the model was used in generating the ROC curves.

4.1.3.3 Ranking using the Economics of Decisions

The costs of corrective replacement of a component, especially following extreme weather conditions, and the costs of preventive replacement of the same type of component are sometimes substantially different.

Ranking diagnostic tests using any of the statistical methods applied in the previous sections does not account for the effects of misclassification costs. The cost of replacing a good component prematurely (false positive) is a sum of costs of preventive replacement and diagnostics. The cost of not identifying a component at high risk of failure is the sum of costs of corrective replacement and diagnostics. Even if the costs of preventive and corrective replacements are assumed equal, the ranking of the superiority of diagnostic tests may still differ from the results of logistic regression and ROC analysis.

The costs used in ranking the diagnostic tests include the costs of individual diagnostic tests and the costs and probabilities of misclassifications. A tree diagram showing the progression of decisions in diagnostics of one component is shown in Figure 52. The total cost shown per leg of the tree is one scenario of either a false positive or a false negative (misclassification of components).

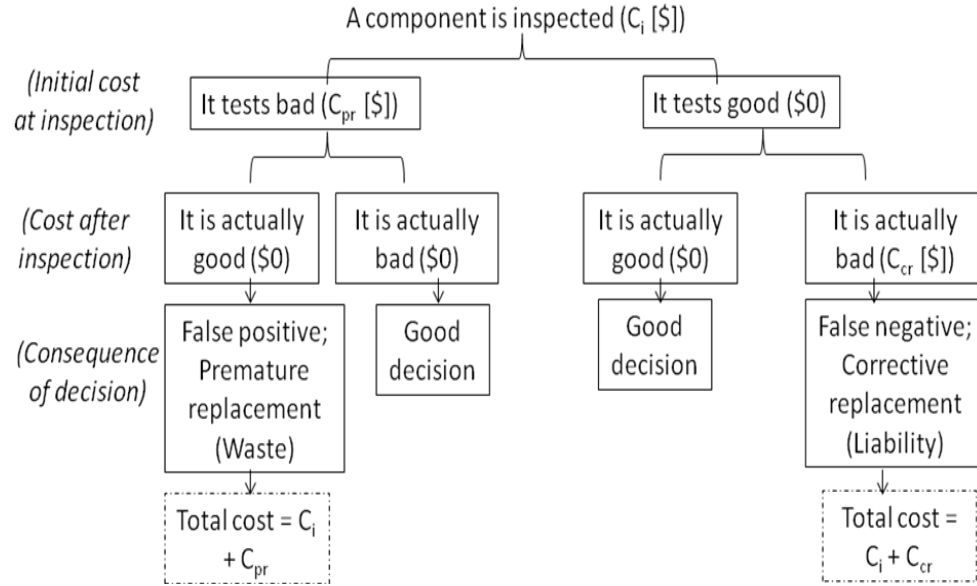


Figure 52 A tree diagram showing costs and decisions of diagnostics and replacements of a component with costs in parentheses.

The total cost for accuracy-informed decisions on selection of diagnostic tests is defined as follows.

$$C_i + P[\text{tests bad}] * P[\text{good}|\text{tests bad}] * C_{pr} + P[\text{tests good}] * P[\text{bad}|\text{tests good}] * C_{cr}, \quad (27)$$

where C_i is the cost of a diagnostic test, C_{pr} is the cost of a preventive replacement, and C_{cr} is the cost of a corrective replacement. Note that the costs of replacement are weighted by the probabilities of making erroneous decisions. The results of destructive testing experiments can be used in obtaining estimates of each prior probability. The diagnostic test with the least cumulative cost would be the most superior diagnostic test.

Sample Results: Wood Poles

The results of diagnostic and destructive testing on the wood poles listed in Table 16 are analyzed for ranking based on economics. It is assumed here that the costs of all the diagnostic tests and the costs of preventive and corrective replacements are equal. The probabilities and costs of each diagnostic test are shown in Table 19.

Table 19 Error probabilities of diagnostic decisions taken by six testing procedures. The tests were compared to destructive test results for evaluation of accuracy. The performance ranking in ascending order of costs is 5, 3, 4, 2, 1, 6.

	P[good tests bad]	P[bad tests good]	P [tests good]	P[tests bad]	Cost (\$)
Diagnostic 1	0	0.2273	0.880	0.120	410.00
Diagnostic 2	0.2143	0.1389	0.720	0.280	330.00
Diagnostic 3	0	0.1622	0.804	0.196	270.87
Diagnostic 4	0.1000	0.1714	0.778	0.222	321.11
Diagnostic 5	0	0.0811	0.740	0.260	130.00
Diagnostic 6	0	0.2609	0.939	0.061	499.80

Diagnostic tests five and three were found to have top rankings when economic effects are considered. This coincides with NEETRAC rankings. The margin of difference between the cost of diagnostic five and three is large (almost double). This implies that diagnostic test five is significantly better than diagnostic test three (and other tests for that matter) based on the assumptions used in the calculations. In contrast, the costs of diagnostic tests two and four are relatively close with a difference of less than nine dollars between them. Diagnostic tests six and one have consistently been ranked lowest by the statistical methods, and now, the economic method.

The economic method is the most practical method of all those mentioned in this chapter since it accounts for misclassification costs. It is suggested that the technique be used for making decisions on selection of diagnostic tests.

4.2 Improving Diagnostic Accuracy and Validity by Combining Diagnostic Test Results

Accurate diagnosis eliminates the need for destructive testing of components. As was seen in the previous subsection, the usefulness of a diagnostic test can be quantified

by the likelihood and costs of condition-misclassification for that test. When the cumulative cost in Equation (27) is unacceptably high, there are a number of ways to improve the accuracy of the diagnostic test. One way is choosing a more accurate test. Another is combining the decisions of that test with another diagnostic test that adds to the information derived from the first test.

A consecutive increase of the number of diagnostic tests used in making a decision of preventive replacement may increase accuracy. However, it will increase the expense of diagnostic testing, which is still just an information-gathering system and does not directly improve reliability of the component. The number and selection of diagnostic tests that yield the minimum cumulative cost of diagnostic testing and misclassification cost are optimum design variables for making decisions on preventive replacement of a component. The objective function C_{opt} is given in Equation (28).

$$C_{opt} = \min_n \left[\sum_{i=1}^n C_i + P_n[\text{tests bad}] * P_n[\text{good}|\text{tests bad}] * C_{pr} + P_n[\text{tests good}] * P_n[\text{bad}|\text{tests good}] * C_{cr} \right], \quad (28)$$

subject to

$$n \geq 1, n \in \mathbb{Z},$$

where n is the number of diagnostic tests combined and therefore an element of a set of integers \mathbb{Z} , $P_n[.]$ is the probability of some event when n tests are combined. The optimization problem can be solved by dividing it into sub-problems and solving sequentially. The probabilities change for each combination of diagnostic tests. Therefore, different estimates of the total cost will be found per combination. The costs can then be compared after calculations to output the objective cost and corresponding objective design variables n and the selection of diagnostic tests that produce C_{opt} .

With two classes of the condition of a component: “bad” or “good,” let the vector of classifications predicted by n tests for one single component be $\mathbf{D}_i = [D_{i1}, D_{i2}, \dots, D_{in}]^N$ for the i -th component. Similarly, each element is either “bad” or “good,” represented using binary variables: “1” or “0” respectively. If the class of the components as

determined from a destructive test is X , then the contribution of the i th component to the likelihood is

$$P(D_i = d_i) = P(X_i = 0)P(D_i = d_i | X_i = 0) + P(X_i = 1)P(D_i = d_i | X_i = 1),$$

noting that the dependencies between the n tests can take the form of

$$P(D_1 \cap D_2 \cap D_3 \dots \cap D_n) = P(D_1)P(D_2 | D_1)P(D_3 | D_1, D_2) \dots P(D_n | D_1, D_2, \dots, D_{n-1}).$$

It is obvious how the complexity increases with the number of tests.

Correlations between tests are expected to affect the relevance of combining tests for the purpose of improving classification of components by their conditions. Perfectly correlated tests do not provide additional information in accuracy. However, the tests may increase the confidence of any decisions to preventively replace (or not replace) a component depending on the accuracy of the tests. Combining diagnostic tests with correlation coefficients less than one will provide new information. Based on the decision rule for combination of the diagnostic tests, the accuracy of tests may improve significantly, improve by a small degree, or actually decrease.

The *degree* of correlation between events D_i (classifications of distinct individual diagnostic tests) can be represented as the ratio of the conditional probability of D_2 given D_1 (for 2 tests) to the unconditional probability of D_2 alone. This indicates the degree to which D_1 influences D_2 .

$$g = \frac{P(D_2 | D_1)}{P(D_2)} = \frac{P(D_1, D_2)}{p(D_1)p(D_2)}.$$

If g is equal to one, the events are independent and not correlated. If it is greater than one, they are positively correlated. If it is less than one, they are negatively correlated, and if it is zero, if D_1 occurs, then D_2 will not. When data is present, combinatorics may be used as the likelihood estimates of each probability.

Various methods exist for obtaining the best classification of the condition of components from some mathematical combination of diagnostic test results. The response variable would of course be the true class of the component, which we take to

be the destructive test result of the component. The explanatory or predictor variables are the diagnostic test results. Statistical techniques or models include logistic regression, linear discriminant analysis (LDA), naïve Bayes classifier, and support vector machines. The steps taken in applying any of these techniques involve the following.

1. Training the classifier (fitting/training the model using observed values of the predictor variables and the response variable),
2. Measuring the classifier accuracy (finding error estimates between predicted response and observed response using non-training dataset),
3. If the accuracy of the trained classifier is not high enough, any of the other statistical techniques may be applied,
4. Else, the classification model can be used for new datasets in predicting the response.

Some of the listed methodologies are tedious to compute and yield results that are hard to interpret or apply in practical. For instance, linear discriminant analysis in conjunction with analysis of area under a receiver-operating curve (AUC) was applied by Pepe and Thompson in [6] to find the optimum linear combination between two diagnostic tests for cancer screening research. The equation analyzed was similar to $\alpha Y_1 + \beta Y_2$, where α and β are weights, and Y_1 and Y_2 are two diagnostic test results for the same class of conditions. The weights that optimized the AUC were obtained but did not present a solution implementable by a decision-maker seeking to either replace or not replace a component.

The results of regression, LDA and similar methodologies are new quantities on completely different scales from the original scales of the individual diagnostic tests. This is one reason outcomes of the techniques are difficult to interpret. A decision-maker requires an intuitive implementable rule or model.

Logic combinations of diagnostic test classifications are easier rules to implement; in which case OR, AND, and NOR combinations can be applied to identify

conjunctive or compensatory relationships between the tests. The response is still a binary classification and is therefore on the same decision scale as the individual diagnostic tests.

While statistical and logic combination rules have been applied to fields of biostatistics research, it has proved difficult finding applications to preventive replacements of components in literature. In addition, the algorithms found in peer-reviewed publications (even in biostatistics research) have been focused on combination of only two diagnostic tests. The algorithms developed here are designed specifically for application to evaluations of diagnostic tests using destructive tests and for multiple combinations of diagnostic tests.

Sample Results: Wood Poles

The Pearson correlations between predicted strengths and predicted classifications of wood poles from the six diagnostic test providers and procedures introduced in Table 16 are given in Table 20 and Table 21. The strengths are continuous numeric values while the classifications are binary variables with logic values zero for poles that tested “good” and one for poles that tested “bad.” There are discrepancies between the coefficients. The lowest correlation coefficient in Table 20 is 0.549 while that in Table 21 is 0.425. The diagnostic tests with the highest correlations are diagnostic tests three and four in both tables. These tests are actually provided by the same company but using different procedures. Other than these two, diagnostic tests three and five follow closely in both tables.

Table 20 Correlation coefficients between predicted strengths of wood poles from six different diagnostic procedures.

	Diagnostic Test 1	Diagnostic Test 2	Diagnostic Test 3	Diagnostic Test 4	Diagnostic Test 5	Diagnostic Test 6
Diagnostic Test 1	1	0.685	0.651	0.604	0.715	0.744
Diagnostic Test 2	0.685	1	0.652	0.656	0.690	0.653
Diagnostic Test 3	0.651	0.652	1	0.903	0.749	0.558
Diagnostic Test 4	0.604	0.656	0.903	1	0.713	0.549
Diagnostic Test 5	0.715	0.690	0.749	0.713	1	0.687
Diagnostic Test 6	0.744	0.653	0.558	0.549	0.687	1

Table 21 Correlation coefficients between predicted classifications of wood poles from six different diagnostic procedures using logic values zero for poles that tested good and one for poles that tested bad.

	Diagnostic Test 1	Diagnostic Test 2	Diagnostic Test 3	Diagnostic Test 4	Diagnostic Test 5	Diagnostic Test 6
Diagnostic Test 1	1	0.592	0.623	0.734	0.623	0.684
Diagnostic Test 2	0.592	1	0.542	0.603	0.646	0.425
Diagnostic Test 3	0.623	0.542	1	0.802	0.786	0.575
Diagnostic Test 4	0.734	0.603	0.802	1	0.721	0.533
Diagnostic Test 5	0.623	0.646	0.786	0.721	1	0.448
Diagnostic Test 6	0.684	0.425	0.575	0.533	0.448	1

The effect of the correlations on tests combined using logic rules will be investigated in the coming sections.

4.3 Analyzing the Accuracy of Combined Diagnostic Tests

4.3.1 Applying Logistic Regression Analysis to Multiple Diagnostic Test Predictions

A stepwise or sequential logistic regression model can be used to test whether the regression coefficients of diagnostic tests from Equation (26) add significantly to prediction of the response. It does this by adding different predictors to a null model, which consists of an intercept only, and checking if there is a statistically significant improvement in prediction. Inference for logistic regression analysis is often based on

deviance (twice the log-likelihood ratio statistic). A large deviance indicates that a logistic model does not fit the data well, and that another model may be appropriate. Asymptotically, the deviance has a χ^2 distribution. Hypothesis tests involving deviance are therefore compared to percentiles of a χ^2 distribution. The number of degrees of freedom is the number of observations minus the number of parameters estimated.

Stepwise regression can be conducted by starting with a null model, where all the β -coefficients of the diagnostic scores are assumed to be zero, and then adding diagnostic scores. A significance level or tolerance is used to test whether the diagnostic test added at each step statistically improves significance or accuracy of prediction. Sequences of F-tests, t-tests and other techniques have been used for hypothesis testing. This method of adding to the null model is called forward selection. The reverse is called backward elimination. Here, all the diagnostic test scores are included in the initial model. Then, using a specified significance level, diagnostic tests that are not statistically significant are removed from the model. Another method may include a combination of forward selection and backward elimination. The results from each method are expected to be different.

Sample Results: Wood Poles

The predicted strengths of the fifty wood poles tested by six diagnostic tests evaluated by NEETRAC were plotted on a normal probability plot to test whether linear discriminant analysis (LDA) could be applied on the tests. (Normally distributed test scores are a requirement for LDA.) The plot of the scores of each diagnostic test is shown in Figure 53. The Anderson-Darling statistics and the alignment of the scores show the scores may not be efficiently represented by a normal distribution. Thus, LDA was not applied to the sets.

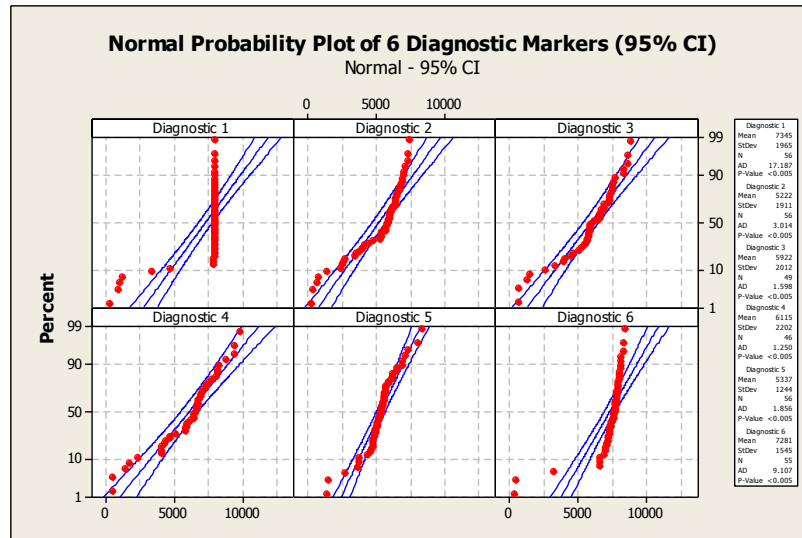


Figure 53 Normal probability plots of six diagnostic testing scores of strengths of utility wood poles. The plots show that a normal distribution may not be the best fit for the scores.

Assuming the scores had shown a Gaussian distribution, LDA may have been used to find a threshold that best separates observations of bad components from good components using diagnostic test scores. The following steps would have been applied.

- Separation of components (wood poles) into two classes using destructive test results: those that are bad and those that are good.
- Obtaining the mean (μ_1, μ_2) and covariance matrices (Σ_1, Σ_2) of the diagnostic test scores for each class of components separately.
- The coefficients of the linear combinations of the diagnostic test scores can be optimized as the product of the inverse of the sum of the covariance matrices of the two classes ($[\Sigma_1 + \Sigma_2]^{-1}$) and the difference between the vectors of the means of the scores in the two classes ($\mu_1 - \mu_2$).

Instead of LDA, a forward selection algorithm is applied to the diagnostic prediction strengths as predictor variables. The response variable is the classification from destructive test results. Starting with a null model, the deviance of the model is calculated. By running F-tests to find the likelihood of a non-zero slope for any of the

diagnostic tests, the third test is found to add the most to prediction (reduces the deviance). It is therefore the first to be added to the null model. This is seen in Table 22.

The current model (null + diagnostic test three) is compared to the other predictors, with the p-values of the F-statistics used to check whether addition of any other predictor will improve the deviance. Adding diagnostic predictor four improves the deviance significantly. So, it is included in the model. Comparison of the two-predictor model (three and four) to other predictor variables shows there will be no more significant improvements given a five percent entrance tolerance. Therefore, the forward selection algorithm ends here.

Table 22 Progression of deviance of a logistic regression model using a forward selection algorithm. The final intercepts and slopes of the included predictions are in the last rows of the table.

Initial deviance	67.006						
Step 1	Add diagnostic test 3 as predictor						
New deviance	41.707						
Step 2	Add diagnostic test 4 as predictor						
New deviance	36.955						
Reduced model	Intercept	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Coefficients	5.193	0	0	-0.002	0.000	0	0
Standard errors	2.174	0	0	0.001	0.001	0	0

The change in deviance with each step is shown in Figure 54.

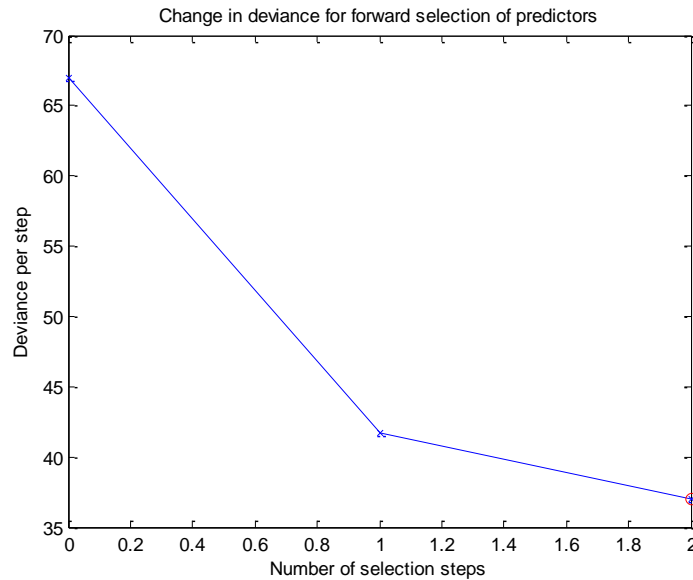


Figure 54 Change in model deviance from an initial null model to inclusion of the first predictor (diagnostic test three) and then, inclusion of the second predictor (diagnostic test four). The other predictors were not found to significantly improve the model deviance, as per hypothesis tests. A forward-selection algorithm is used.

It turns out that diagnostic test one is a traditional inspection technique applied by a good number of utilities, even though it is relatively poor in strength prediction of wood poles. The use of this technique is another support that power utilities are prone to carrying out operations in the old way versus actively seeking better and more effective methods of managing their components.

A second experiment of logistic regression analysis is carried out with diagnostic test one forced into the initial (null) model. Then, coefficients of the other predictors are compared to the new model to find which, if any, will significantly improve the model deviance. Results are shown in Table 23.

Table 23 Statistics and deviance of a forward selection logistic regression model that starts by including diagnostic test one in the initial model. The deviance increases to about 38 with inclusion of predictor three.

Initial deviance with diagnostic 1	55.616						
Step 1	Add diagnostic test 3 as predictor						
New deviance	38.502						
Reduced model	Intercept	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Coefficients	95.575	-0.012	0	-0.001	0	0	0
Standard errors	138.092	0.017	0	0.000	0	0	0

The table shows that including diagnostic test one to the null model reduced the deviance of the null model from 67.00 in Table 22 to 55.62: an improvement of only 17%. This is compared to improvement of about 38% when using diagnostic test three as the first inclusion in Table 22. However, by using the third test as the complementary predictor to diagnostic one, the final deviance is found to be 38.5016. It is only slightly worse (4%) than the final deviance of 36.95 in Table 22, but still a lot better than the deviance of the null model (67.00).

The stepwise regression analysis shows that classification of components by their conditions can be improved by using multiple diagnostic tests rather than one test. While the best prediction comes by using both diagnostic tests three and four, the “comfort-zone” behavior of the utilities might encourage the use of diagnostic tests one and three instead. Interestingly, diagnostic techniques three and four are highly correlated (correlation coefficient is over 0.8). It is likely the reason that diagnostic test four did not significantly improve prediction in the second analysis. Though the analysis encourages combination of diagnostic test scores, it does not provide a means of implementation.

4.3.2 Applying Receiver Operating Characteristics (ROC) Analysis to Multiple Diagnostic Test Predictions

The area under receiver-operating characteristic (ROC) curves can be used for analyzing predicted and true classifications of strengths of a component. This was discussed in Section 4.1.3.2. The results of a sequential logistic regression model can be evaluated using AUC analysis. Comparison with single test predictions may show improvement in predictions. That is, if the AUC of a model with combined predictors is higher than the AUC of a model with only one predictor.

Sample Results: Wood Poles

The intercept and slope coefficients of the combined diagnostic test one and three logistic regression model from the previous subsection are used for classification in ROC analysis. The curve is shown in Figure 55.

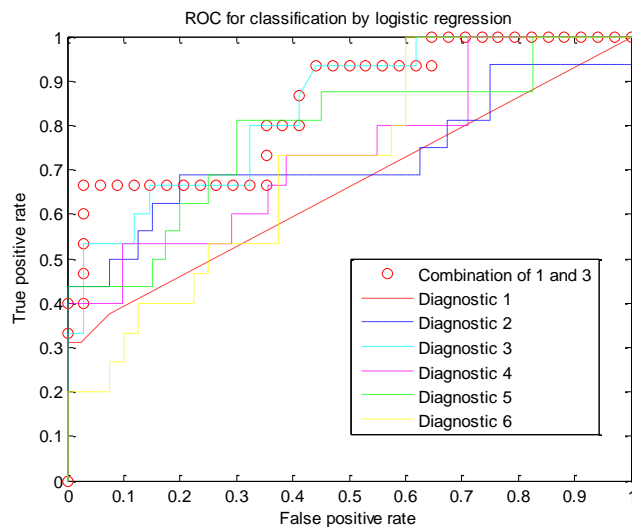


Figure 55 Receiving operating characteristics (ROC) curves for classification of condition of distribution wood poles by logistic regression modeling of diagnostic test scores one through six. The logistic model that combines predictions from diagnostic test one and three is shown to have a generally higher ROC curve than the model of the individual tests.

Except in very small regions, the ROC of the combined model is above or as good as the ROC of the individual tests. The AUC of individual diagnostic tests and the combination (diagnostic one and three) is in Table 24.

Table 24 Area under the curve (AUC) for logistic regression model of each diagnostic test and combined prediction of tests one and three.

Diagnostic Test	1	2	3	4	5	6	Combination (1 and 3)
AUC	0.6609	0.7281	0.8363	0.7398	0.7828	0.7150	0.8461

The area under the curve of the combined model is higher than all the individual tests. However, it is only 1.2% higher than diagnostic test 3. This is not a large improvement. Again, the results of this kind of analysis do not present any practical contribution since implementation in the field is not clear. Logic combinations and a corresponding analysis of economics of the decisions are more practical.

4.3.3 Applying Combinatorial Logic Rules

The combinatorial logic analysis relies on “OR” and “AND” operations among predicted classes of tested components. While prior analysis (logistic regression) has been performed by using scores from the diagnostic tests as continuous variables, this analysis uses classification of the components (a binary variable). Logic one is used to denote components that test positive (test “bad”) while logic zero is used for components that test negative (test “good”).

The “OR” rule is a compensatory rule. Diagnostic tests compensate for each other: a component that tests bad based on one diagnostic test will be preventively replaced whether or not the component tests bad using other tests. So, if the sensitivities of two tests are relatively high and one test misses a bad component detected by the other, the “OR” rule would lead to a good decision. It would be inappropriate to use on tests with low sensitivities, and does poorly in reducing false positive errors.

The “AND” rule is conjunctive. So, components are preventively replaced only if they test positive for all combined tests. When a test with high sensitivity is combined

with a test that has a low sensitivity, a good number of bad components may be left in service. The “AND” rule does poorly in reducing false negative errors.

Truth tables are useful in analyzing how decisions are made using logic combinations. See Table 25 for “AND” and “OR” combinations of two diagnostic (predicted) classes of components. The table shows when decisions to replace a component occur (logic one outputs on the combinatorial rules). Truth tables for combinations of more than two logic inputs (diagnostic tests) can be generated similarly.

Table 25 “OR” and “AND” combinatorial rules between predictions of two diagnostic tests. Logic one stands for components that test positive and logic zero stands for components that test negative.

Prediction of diagnostic test 1	Prediction of diagnostic test 2	OR	AND
0	0	0	0
0	1	1	0
1	0	1	0
1	1	1	1

The probabilities P_n in Equation (28) are found by comparing results of destructive tests to the decisions found by applying combinatorial rules to multiple diagnoses. The cost of diagnostics increases with the number of diagnostic tests. The decision to combine tests as well as the selection of diagnostic tests will depend on the probabilities and costs that result from combination of the diagnostic tests. Algorithms are designed in this chapter to compute and display results of combinations; they become more complex as the number of diagnostic tests increases.

For any combination of n_1 out of n available diagnostic tests combined, the number of resulting cost-decisions to compare is equal to $n - \text{combination} - n_1$. This is the same as

$$\frac{n!}{(n - n_1)! n_1!}$$

where n is the total number of available diagnostic test procedures and n_1 is the number of combinations to be analyzed. This is a combination problem and not a permutation problem since all permutations of the same set of diagnostic tests, excluding any relevance of order in which the tests are carried out, will yield the same cost-results.

4.3.3.1 Algorithm for Combining Multiple Tests

The tests were carried out serially: one after the other so that the most invasive test (which was still only minimally invasive) was performed last. So we may assume some independence between the tests. Let the number of combinations to analyze be n_1 equal to two through n , which is the total number of available diagnostic tests. Then, the optimum number (and selection) of diagnostic tests that yields the minimum cost C_{opt} in Equation (28) can be found using the following algorithm.

For each n_1 ,

1. Store all combinations $n - C - n_1$ with n_1 elements as rows in a matrix `comb_n`. For instance, given $n = 6$ available tests and n_1 equal to 2 (logic inputs), the matrix `comb_n` will have size 15-by-2. In MATLAB, “`nchoosek`” will produce all combinations of a given sequence $\{1, 2, \dots, n\}$.
2. Obtain the probabilities P_{n1} which are $P_n[\text{tests bad}]$, $P_n[\text{good}|\text{tests bad}]$, $P_n[\text{tests good}]$, $P_n[\text{bad}|\text{tests good}]$ in Equation (28) for each row of the matrix in step 1. This means errors rates of different selections of diagnostic tests are being found. The error rates are estimated by comparing the logic output using the “OR,” “AND,” or some combination of either rule on pairs of the diagnostic predictions and the destructive test results.

3. Assuming equality of all diagnostic costs, solve for the cost in Equation (28) per row of combinations. Exclude the minimization operation. This yields a cost vector C_n for the combination of n_1 diagnostic tests. The elements of this vector are all possible costs from combining any n_1 distinct diagnostic tests.
4. Sort the elements of C_n in ascending order keeping track of the indices. Output the minimum cost and selection of tests that yield the minimum cost. The latter (the selection) is a set of the row-elements of the matrix $comb_n$ that has the same row-index as the first element of the ordered cost vector C_n .

After every combination is found for n_1 equal to two through the total number of available diagnostic tests n , plot the minimum cost per combination against n_1 combinations. The minimum cost of all the minimum costs per combination is C_{opt} . Likewise, the optimum selection of diagnostic tests to combine will be clear.

Sample Results: Wood Poles

The predicted strengths of the wood poles tested in the NEETRAC experiment were compared against a set threshold (the ANSI standard for acceptable remaining strength of in-service wood poles). The comparison resulted in the classification of the wood poles as bad (to be replaced) or good (to be left in service) using the six different diagnostic procedures. This results in two binary classifications of the poles; one classification is a prediction from a diagnostic test and the other is the true condition from destructive testing.

The diagnostic tests were all non-destructive and were supposed to have been conducted in such a way as not to influence testing from other diagnostic procedures. So, we assume independence between the tests and their results. In reality, they might have

had some influence on each other. However, we assume that any such effect is so negligible that independence is acceptable.

A table that shows the resulting costs from combining two or three diagnostic tests using the “OR” rule is shown in Table 26. Using values from surveys filled out by power utility engineers, the cost of one replacement (whether preventive or corrective) is assumed to be \$2,000 while the cost of inspection one pole is assumed to be \$10.

Table 26 Selections of two or three diagnostic tests combined using the “OR” logic rule for making decisions on preventive replacements of wood poles. The decision costs of each combination are shown in the column labeled “Total.” “CNPV” is $P[\text{bad}|\text{tests good}]$ and “CPPV” is $P[\text{good}|\text{tests bad}]$.

OR	Two tests		CNPV	CPPV	Total	Three tests			CNPV	CPPV	Total
	1	2	0.139	0.214	\$ 340.00	1	2	3	0.097	0.200	\$ 290.87
	1	3	0.139	0	\$ 237.39	1	2	4	0.133	0.267	\$ 385.56
	1	4	0.171	0.100	\$ 331.11	1	2	5	0.061	0.177	\$ 230.00
	1	5	0.081	0	\$ 140.00	1	2	6	0.139	0.231	\$ 356.53
	1	6	0.209	0	\$ 387.35	1	3	4	0.147	0.091	\$ 296.67
	2	3	0.097	0.200	\$ 280.87	1	3	5	0.061	0	\$ 116.96
	2	4	0.133	0.267	\$ 375.56	1	3	6	0.139	0	\$ 252.22
	2	5	0.061	0.177	\$ 220.00	1	4	5	0.065	0.071	\$ 163.33
	2	6	0.139	0.231	\$ 346.53	1	4	6	0.171	0.111	\$ 348.18
	3	4	0.147	0.091	\$ 286.67	1	5	6	0.081	0	\$ 152.45
	3	5	0.061	0	\$ 106.96	2	3	4	0.103	0.250	\$ 341.11
	3	6	0.162	0	\$ 286.67	2	3	5	0.067	0.188	\$ 247.39
	4	5	0.065	0.071	\$ 153.33	2	3	6	0.097	0.214	\$ 296.67
	4	6	0.171	0.111	\$ 338.18	2	4	5	0.071	0.235	\$ 296.67
	5	6	0.081	0	\$ 142.45	2	4	6	0.133	0.286	\$ 393.64
						2	5	6	0.061	0.188	\$ 234.08
						3	4	5	0.065	0.071	\$ 163.33
						3	4	6	0.147	0.100	\$ 302.73
						3	5	6	0.061	0	\$ 118.89
						4	5	6	0.065	0.077	\$ 166.36

Similar results are available for the larger number of “OR” combinations (four through six). These can be found in Table 27 and Table 28.

Table 27 Selections and decision-costs of four diagnostic tests combined using the “OR” logic rule for making decisions on preventive replacements of wood poles.

OR	Four tests				Total
	1	2	3	4	\$ 351.11
	1	2	3	5	\$ 257.39
	1	2	3	6	\$ 306.67
	1	2	4	5	\$ 306.67
	1	2	4	6	\$ 403.64
	1	2	5	6	\$ 244.08
	1	3	4	5	\$ 173.33
	1	3	4	6	\$ 312.73
	1	3	5	6	\$ 128.89
	1	4	5	6	\$ 176.36
	2	3	4	5	\$ 306.67
	2	3	4	6	\$ 358.18
	2	3	5	6	\$ 262.22
	2	4	5	6	\$ 312.73
	3	4	5	6	\$ 176.36

Table 28 Selections of five or six diagnostic tests combined using the “OR” logic rule for making decisions on preventive replacements of wood poles.

OR	Five tests						Total
	1	2	3	4	5		\$ 316.67
	1	2	3	4	6		\$ 368.18
	1	2	3	5	6		\$ 272.22
	1	2	4	5	6		\$ 322.73
	1	3	4	5	6		\$ 186.36
	2	3	4	5	6		\$ 322.73
	Six tests						Total
	1	2	3	4	5	6	\$ 332.73

The minimum costs per combination of two through six diagnostic tests are plotted in Figure 56. The first data-point in each subplot is the decision-cost when using either diagnostic one or diagnostic five alone. The costs are from Section 4.1.3.3. The figure shows that a combination of two tests (diagnostic tests three and five) will produce the minimum decision costs for preventive replacements of wood poles. The difference between using this combination and diagnostic five alone is less than \$25 per pole. However, when thousands of poles are put into consideration, the value becomes clearer. In contrast, there is a large difference of about \$300 per pole when the combination of

tests three and five is compared to using diagnostic one alone (most popular diagnostic procedure on wood poles).

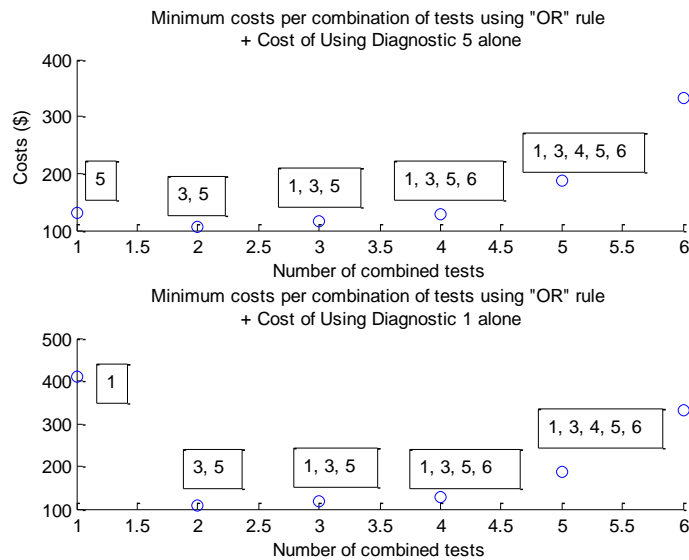


Figure 56 Plot of the minimum costs that result from preventive replacements of wood poles using single (“Diagnostic 1 alone” or “Diagnostic 5 alone”) or logic OR-combined diagnostic tests. The x-axis shows the number of tests combined (2, 3, 4, 5 or 6). The selections yielding each minimum cost are shown in text boxes in the figure.

The effect of correlations between the diagnostic tests on the probability of misdiagnosing a wood pole is shown in Figure 57 and Figure 58. In Figure 57, the probability of missing a bad wood pole, $P[\text{bad} | \text{tests good}]$ is plotted against correlations. The plots have been separated for each diagnostic test. For diagnostic test one, which had an initial error probability of 0.2273, the probability appears to decrease with decreasing correlation coefficients (upward slope). The relationship appears reversed for the other tests (downward slope). Notice that the new rates are less than or equal to the initial rate.

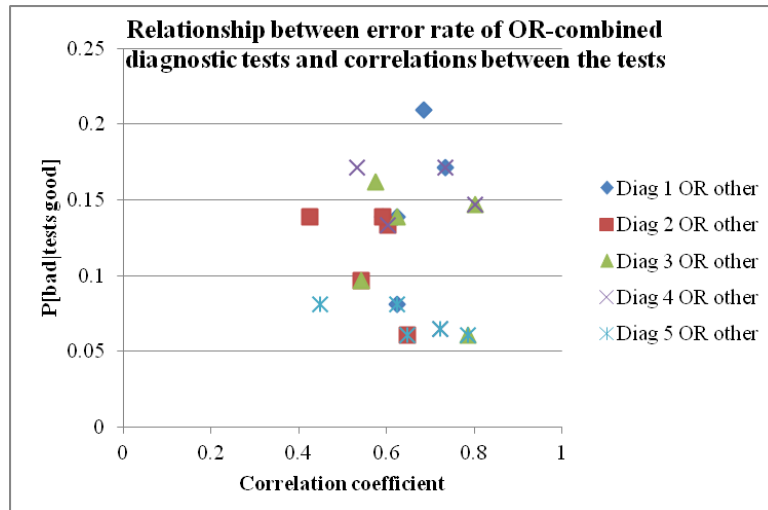


Figure 57 Plot of error rate of missing a bad wood pole after a compensatory combination of diagnostic tests against the correlations between the combined tests. Initial error rate when diagnostic test one is used alone is 0.2273; two, 0.1389; three, 0.2143; four, 0.1622.

In Figure 58, other than the rates on the x-axis, the error rates appear to decrease with increasing correlation coefficients (downward slope). Comparing the initial false positive rates, that is when the tests are used alone, to the new rates when the tests are combined, it can be seen that the “OR” rule sometimes worsens false positive rates.

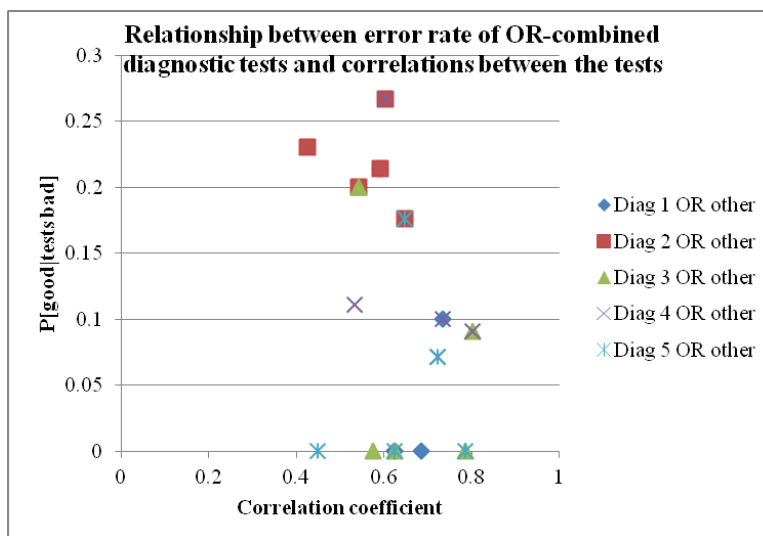


Figure 58 Plot of error rate of prematurely replacing a good wood pole after a compensatory combination of diagnostic tests against the correlations between the combined tests.

The decision-costs when the tests are combined using an “AND” rule are shown in Figure 59. The first data-point on each subplot is the cost when diagnostic five or diagnostic one is used alone.

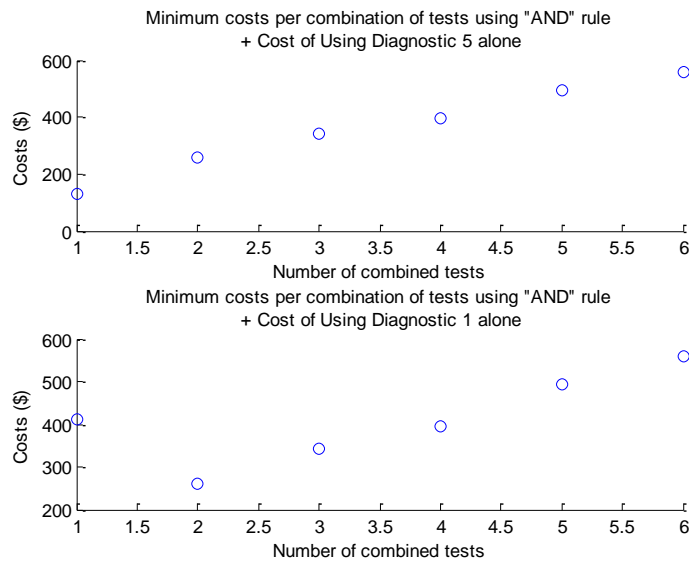


Figure 59 Plot of the minimum costs that result from preventive replacements of wood poles using single (“Diagnostic 1 alone” or “Diagnostic 5 alone”) or logic AND-combined diagnostic tests. The x-axis shows the number of tests combined (2, 3, 4, 5 or 6). The selections yielding each minimum cost are shown in text boxes in the figure.

It is obvious from the plots that using the “AND” rule improves on using diagnostic test one alone (up to three combinations of tests). It however, produces costs that are higher than the cost of using diagnostic five alone. The decision-costs increased by combining the diagnostic tests using the “AND” rule. The range of minimum costs using the “AND” rule is now [\$260.00, \$560.00] compared to [\$106.96, \$332.73] from the “OR” rule. The error rate of missing a bad wood pole increased as was expected while the error rate of replacing a wood pole prematurely was zero for all combinations. This includes those that already had a zero premature-replacement rate.

Combinations of “OR” and “AND” rules were applied to the diagnostic tests to investigate if such combinations would lead to a smaller decision cost than that found by

“diagnostic test three OR five.” The combinations did not improve the costs. The minimum costs found for a few such combinations are shown in Figure 59.

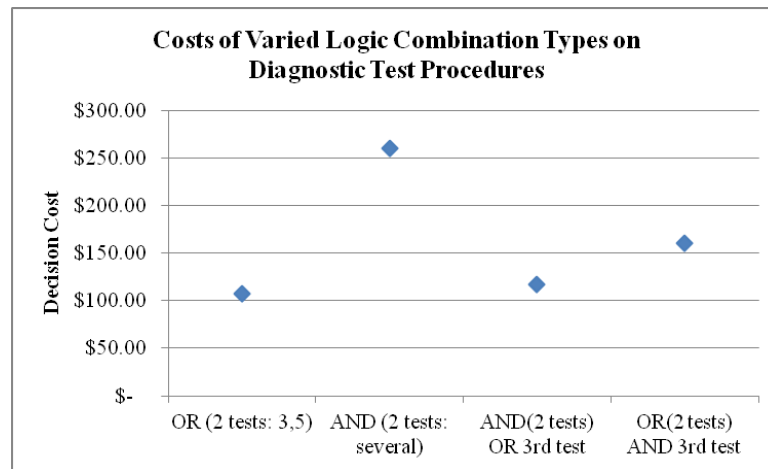


Figure 60 Decision-costs generated by combining diagnostic tests using logic “OR” rules, “AND” rules or some combination of the rules. The third and fourth data-points are derived by combining the first two tests using the “AND” rule or the “OR” rule. Then, the results are combined with the last third test using the “OR” or “AND” rule correspondingly.

The minimum decision costs have been observed when diagnosis from two test procedures (test three and five) are combined using the “OR” rule. The “AND” rule is not a competitive rule for improving diagnostic accuracy. Mixing the “AND” and “OR” rules produce better results than using single diagnostic tests. However, they have not been found to improve decision costs as much as the “OR” combination of diagnostic tests three and five.

4.4 Effects of Cost-Relationships on Decisions of Preventive

Replacements from Multiple Diagnostic Testing

The optimum number and selection of combined diagnostic tests may be sensitive to the actual cost of diagnostics, the relationship between diagnostic costs of distinct diagnostic procedures, costs of preventive replacements, or the relationship between costs of corrective replacements and preventive replacements. The influence of these costs on

decisions should be evaluated before making any conclusions on minimization of diagnostic errors using multiple diagnoses.

Assuming the cost of diagnostics is equal for all diagnostic procedures and the cost goes up, a decision-solution that is sensitive would change. Likewise, when the cost of corrective replacement is known to be higher than the cost of preventive replacements by some factor a_1 , an evaluation of the influence on the decision to use multiple diagnoses should be conducted. These features will be investigated in this section.

Sample Results: Wood Poles

For the same wood pole experiment, which uses diagnoses from six different diagnostic procedures, the cost of corrective replacement is assumed to be a multiplicative factor a_1 equal to two greater than the cost of preventive replacements. The effect of this change on finding the minimum decision cost (and selection of tests) for preventive replacements of wood poles using multiple diagnostic test procedures is shown in Figure 61.

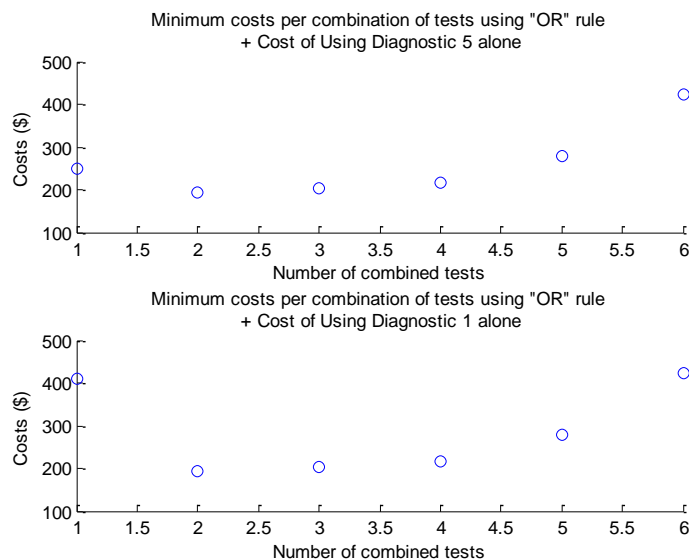


Figure 61 A plot of minimum costs of different combinations of diagnostic tests used in making decisions on preventive replacement of wood poles. Here, the ratio of corrective to preventive replacement costs is two.

The figure shows that though the range of decision-costs increased from [\$106.96, \$332.73] for a unit value of a_1 to [\$193.90, \$423.60] here, the optimum number of combinations is still two. The selected tests are still tests three and five. Other values of the factor a_1 set to be greater than one showed similar results.

With a_1 back to value one and cost of inspections changed from \$10 to \$20, the optimum number of diagnostic tests to combine remains unchanged at two. See Figure 62. The difference in costs between using three tests and two tests increases. The range of the minimum decision-costs is [\$127.00, \$392.70].

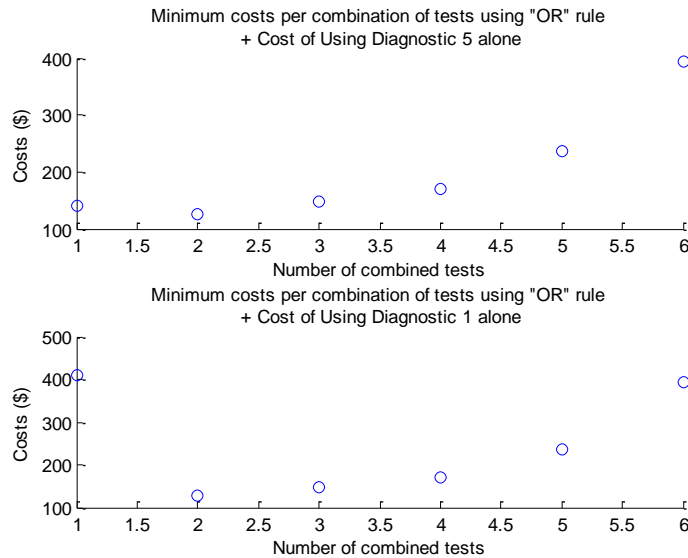


Figure 62 A plot of minimum costs of logic “OR” combinations of diagnostic tests used in making decisions on preventive replacement of wood poles. Here, the cost of individual inspections is twice the initial cost.

Assume the costs of each of the diagnostics are related to their individual accuracies using the following formula.

$$\frac{C_{i+1}}{C_{i_b}} = \frac{a_{i+1}}{a_b},$$

where C_{i_b} is the base cost of an inspection, say \$10; C_{i+1} is the cost of some other diagnostic procedure; a_b is the accuracy of the base test and a_{i+1} is the accuracy of some other diagnostic procedure. Let the base test be the one with the least accuracy. Accuracy

is defined as the ratio of true diagnosis (sum of true positives and true negatives) to the total number of components on test.

The ratio of accuracies and therefore costs of diagnostics is shown in Table 29.

Table 29 Accuracies and resulting inspection-cost ratios for six diagnostic testing procedures. The base accuracy (for test six) is 0.755.

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Accuracy	0.800	0.840	0.870	0.844	0.940	0.755
Inspection cost or accuracy ratio	1.060	1.112	1.152	1.118	1.245	1.000

Using these ratios for calculating individual diagnostic costs of the tests, the resulting plot of minimum decision costs per combination of tests is shown in Figure 63. The minimum cost has now increased to \$110.90 from \$106.96. The decision however remains unchanged. That is, the optimum number of tests to combine is two.

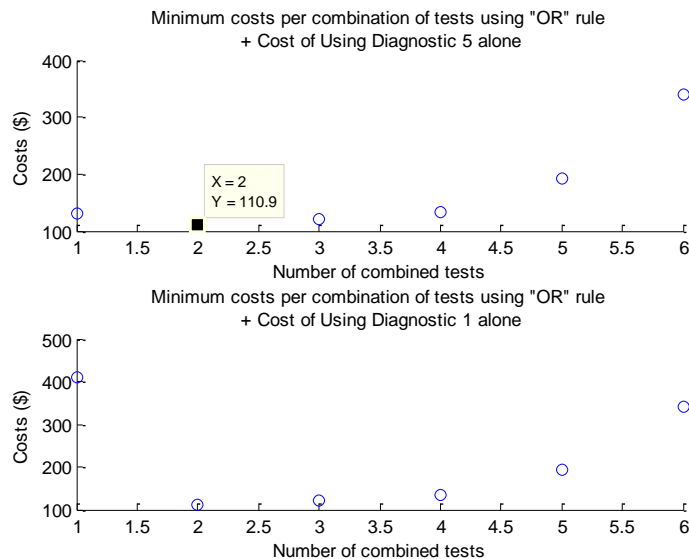


Figure 63 A plot of minimum costs of logic “OR” combinations of diagnostic tests used in making decisions on preventive replacement of wood poles. Here, the costs of diagnostic procedures are directly proportional to the accuracies of the procedures.

The choice of the number of combinations of the tests and the selection of the tests that yield the minimum costs are insensitive to the costs of diagnostics and preventive replacements. The optimum decisions in preventive replacements of wood

poles would be made by using the prediction of the logic output of diagnostic test three OR five.

4.5 Summary

Power system components are expected to be maintained to within National Electrical Safety Code (NESC) standards. Electric utilities therefore employ expertise of diagnostic contractors to investigate what components meet requirements and which do not. The best compliance test is a gold standard test: that precisely and accurately evaluates the condition of components. However, these tests are often invasive and expensive. One such type of test is destructive test applied to components that fail mechanically from say, environmental stress loading. These tests render components unusable. To dispense with destructive tests, less expensive and invasive tests are needed: diagnostic tests. However, accuracy is an issue in gaining information on the true conditions of populations of the components.

To learn the accuracy of the diagnostic tests, samples of components are selected from a population and classifications of the samples using destructive and diagnostic tests, compared. A small sample usually results in large bias in assessment of accuracy. This work suggests the use of proportions of components already selected for preventive replacements since diagnostics is a regular process in the lifetime of populations in keeping with NESC standards. Experimental studies on such samples will yield estimates of only the positive predictive value of the diagnostic test. The positive predictive value (PPV) gives the ratio of components scheduled for replacement that actually do not meet NESC standards, where the latter is known based on destructive test results.

Since destructive tests are costly, though not as costly as preventive replacements themselves, only a proportion of annual preventive replacements may be able to receive destructive testing based on annual utility budgets. The coefficient of variation, which

can be seen as a measure of the risk of returns on destructive testing investments, is used to determine when the estimate of PPV has reached a steady-state. The cumulative cost of destructive testing, the cost of premature replacements until steady-state is reached and the potential reduced cost of premature replacements after actions are taken to mitigate low PPVS is used to show the benefit of shorter time assessments over long time assessments. The analysis performed excludes the negative predictive value (NPV), which may take even longer to determine. Both measures of accuracy would give a more complete picture of the benefit, if any, of switching contracted diagnostic test companies.

Having determined the accuracy of real competing diagnostic tests from a sample of components assumed to be unbiased, rules for ranking superiority of the distinct tests are compared based on subjective judgment, statistical tools and economics of costs of misclassification. Logistic regression analysis and receiver operating characteristics (ROC) curves are some of the tools applied. The economical tool provides intuitive rankings likely to be preferred by management.

Statistical and economic tools are also applied in determining the relationship between combined diagnostic test scores that gives the most accurate predictions in classification of components by whether they meet or fail NESC requirements. Deviance of forward and backward propagation step logistic regression analysis was used for inferring the combinations of the test scores that were statistically significant in predicting classifications. The tool found a combination of two of a total of six tests as appropriate using real data. The ROC method did not provide new results since it was not independent of the regression methodology. The economical method was defined by combining enumerations of all six tests using “OR” and “AND” rules to find the combination that yielded the minimum cumulative costs of combined diagnostics and costs of misclassification of the components.

The optimum number of combined testing methods was two based on an “OR” rule, as similarly found by the logistic regression method. The use of the optimum

combination of two specific diagnostic tests out of the six available tests yielded savings up to \$300 per component tested than the use of the one traditional testing method. Increasing the number of tests combined led to an increase in the cumulative cost of testing and misclassifications compared to the optimum of two tests. The economic method, which applied logic rules to tests combined serially, was again considered superior to the statistical methods since the latter are difficult to implement in the field.

4.6 REFERENCES

- [1] M. F. Miller and G. S. Hosford, "Fiberglass distribution poles: A Case Study," IEEE Transactions on Power Delivery, Vol. 10, No. 1, pp. 497-503, Jan. 1995.
- [2] P. Springer, "Wood Pole In-Service Assessment Methods," National Electric Energy Testing, Research and Applications Center Baseline Project, 2008.
- [3] S. Jeng et al., "Accelerated destructive degradation tests robust to distribution misspecification," IEEE Transactions on Reliability, Vol. 60, No. 4, pp. 701-711, Dec. 2011.
- [4] K. Pawlikowski, "Steady-state simulation of queuing processes: a survey of problems and solutions," ACM Computing Surveys, Vol. 22, No. 2, pp.123-170, June 1990.
- [5] E. U. Weber *et al.* "Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation," Psychological Review, Vol. 111, No. 2, pp. 430 – 445, Apr. 2004.
- [6] Pepe, M.S.; Thompson, M.L., "Combining diagnostic test results to increase accuracy," Biostatistics, Vol. 1, No. 2, pp. 123-140, 2000.

5 STOCHASTIC OPTIMIZATION OF PREVENTIVE MAINTENANCE PROGRAMS FOR POWER SYSTEM COMPONENTS

Introduction

Stochastic optimization tools have played a significant role over the past few years in decision-making where financial markets are concerned. Decision-makers in equity markets and in the field of Asset and Liability Management require specialized algorithms and models to process a large amount of financial data for deciding on how to move assets for investors in order to produce gains for the investors. An example of the magnitude of financial data and assets is seen in the United States (US) Pension Asset Reserves as analyzed in [1]. The decision-making tools applied in the financial markets have more recently been applied in energy planning and lifecycle cost planning. Decisions include the choosing between make and models of structures or equipment to purchase, build or install; choosing between maintenance procedures to apply on equipment, among others. The objectives may be the minimization of lifecycle costs of equipment or the maximization of some score function like profits and gain.

The algorithms used for analysis may be based on numerical methods like Monte Carlo methods or analytical if the functions are well defined continuous functions. Monte Carlo methods are very popular across financial management because of the possibility to directly explore effects of specific scenarios of decision-parameters. Accounting for uncertainties in parameters while exploring scenarios is what differentiates stochastic optimization from deterministic optimization.

While deterministic optimization techniques are very popular in engineering, stochastic optimization is not as common because of computational complexities associated with the process. However, stochastic optimization is important because making decisions under uncertainty is more realistic than following a deterministic

approach. For instance, the exact number of components that will fail in a year is a random variable, which may be modeled using an expected value, variance and other statistics of a distribution. Each possible number of annual failures could be set up as one scenario to analyze and may yield differing optimum decisions in a larger optimization problem.

Power utilities are faced with complex annual decisions on the management of their physical assets or components. They are interested in the choice of a cost-effective schedule of component inspections and replacements. There are uncertainties in costs, probabilities of failures of the components, among other things. In this chapter, a Monte-Carlo approach is applied in the stochastic optimization of preventive maintenance programs. Because the run-to-failure program of management for certain kinds of components is considered highly competitive, the objective of the optimization will be the minimization of the annual cost of a preventive maintenance program relative to a run-to-failure (RTF) program. The kinds of components considered are those that are non-critical to power delivery when they fail individually but could result in millions of dollars in replacements and penalties when they fail in large numbers. The safety risks posed to the public from unexpected failures of these components are also high.

The RTF program is not sustainable and could carry a high safety risk for the public, especially as a significant number of power system components are now considered aging infrastructure. In fact, preventive maintenance programs for such assets as distribution wood poles are now mandatory in states like Florida. In states with flexible management programs on their components, finding parameters for which maintenance decisions are optimum would be very beneficial. The decision parameters in this chapter are the frequency of inspections of components (inspection-schedules) and the proportion of inspected components to replace preventively. The decisions to replace components are assumed to be made based on recommendations from diagnostic techniques.

5.1 Optimizing the Component-Management Process

The run-to-failure (RTF) program is arguably a competitive management program. It requires no expertise in decision-making, no data or data analysis, and no additional annual costs of component inspections or maintenance. In regions of low risks of failures of components resulting from some environmental stress, the number of components failing per year might be very low excluding the effect of age could be quite low. As an illustration, see the wind zone map in Figure 64. While failures and replacements of wind-exposed and wind-vulnerable components might be high in states like Alabama, the number for similar components in Utah might be very low. Decisions of inspections and replacements among the zones will therefore, vary.

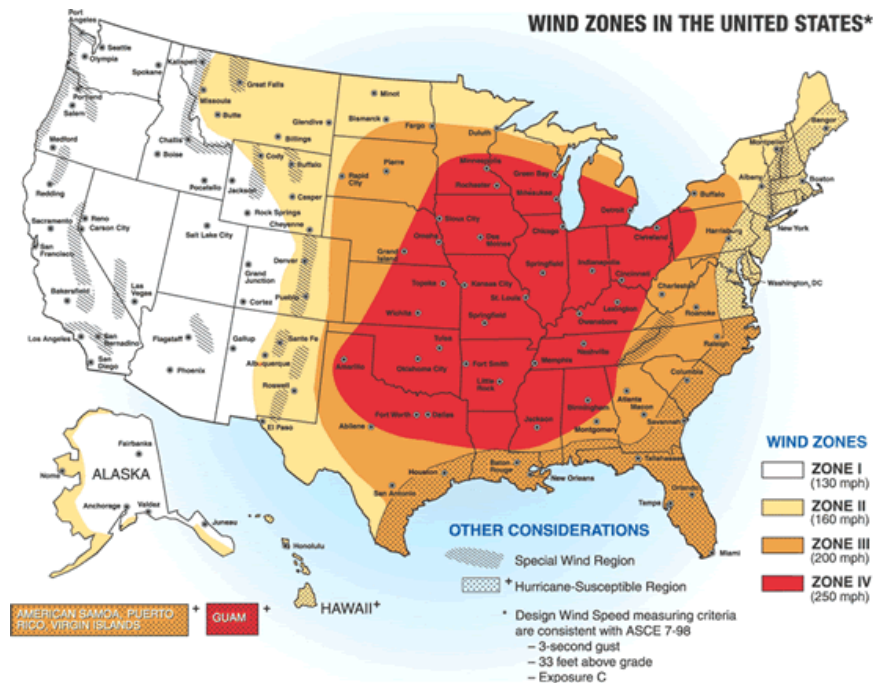


Figure 64 Wind zone map of the United States by Federal Emergency Management Agency (FEMA), showing relative risks of high winds among states [2].

The preventive maintenance program, in comparison, for environmentally stressed components could involve a significant amount of annual investment in inspection and maintenance costs. Sometimes, the benefit of maintenance programs is not

immediately evident or obvious. The maintenance program is somewhat of a long-term investment. It takes a long time before it pays off in terms of how much money is invested into the program. However, when the impact on public perception and power utility reputation is factored, the benefit of the program becomes clearer.

While the RTF program might be competitive in cost, it is deficient in safety planning and is unattractive to the public. Power utilities may incur very high costs of penalties when they (the utilities) are considered to have been negligent in their duties. For instance, civil penalties of up to \$1 million per day may be charged to utilities if they are found to deliberately violate reliability standards or environmental protection acts.

The choice of decision parameters that will minimize the annual cost of a preventive maintenance program relative to a run-to-failure (RTF) program depends greatly on the choice and set-up of constraint functions, fixed parameters, and the bounds of the random variables defined in the problem. In the sections that follow, the decision parameters (design variables) and their bounds are defined, the random variables and their distribution functions are defined, and any constraint functions needed for computation will be defined. Finally, the objective function will be defined.

The use of Monte Carlo methods requires a large selection of possibilities (scenarios) picked from continuous probability density functions. To reduce the complexity of computation and because finding the global optimum is not critical (the alternative decisions are not expected to have as much impact as stock trading or hydrothermal generation decisions, for instance), some optimization parameters will be defined as a small countable number of values. That is, they will be defined as discrete variables with a non-small number of steps between the lower and upper bounds. Local optima will be found.

5.1.1 Costs that affect Decision-Making in Management of Environmentally Stressed Aging Components

The lifecycle of a typical component that experiences inspections at least once before its end-of-life is shown in Figure 65. This represents a component from a preventive maintenance program. The component is assumed to be in operational state from installation to its end-of-life. Each milestone and replacement has associated with it a cost. In lifecycle cost analysis, all costs that occur before a time of analysis, for instance, the present time instant, are neglected since they do not affect future costs or management decisions. In economics, these costs are called *sunk costs*. In Figure 65, the costs of acquisition and installation as well as the cost of the inspection before the present time instant are neglected in decision-making. It is easy to see that the lifecycle of a component undergoing a run-to-failure program does not incur inspection costs.

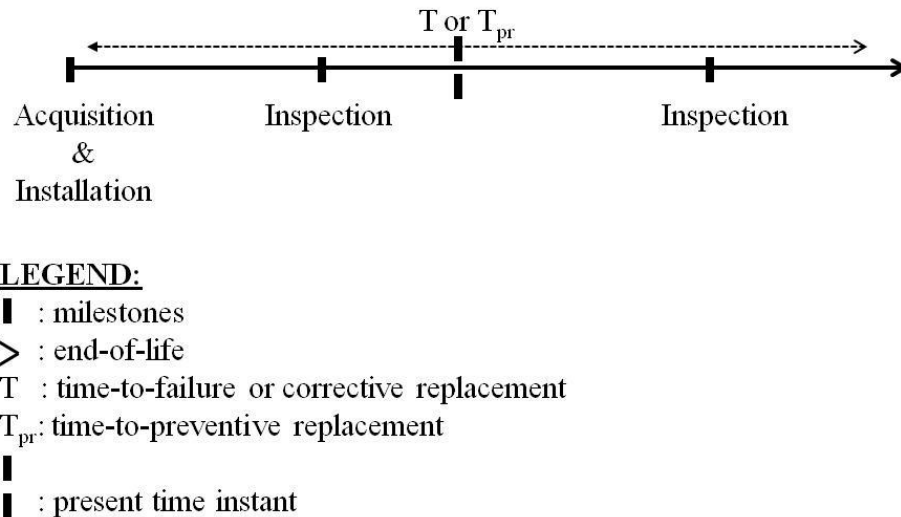


Figure 65 Lifecycle of a typical inspected component up to its end-of-life either from failure or diagnostic decision for preventive replacement.

The end-of-life of a power utility component may be directly related to a number of resulting costs. This includes

- Cost of replacement (labor and capital),
- Cost of energy not served (lost revenue), and

- Cost of switching between lines when damage to a component causes disruption in power flow in some initial line.

When the end-of-life of a component is forced by some environmental stress that causes the failure of other non-identical components and an outage results, it becomes very difficult to estimate lost revenue pertaining to any class of components. Also, in the event of severe storms, switching may be impossible. That is, all possible lines in the contingency design may have failed. Thus, outages often last very long during severe storms: restoration might include replacement of a significant number of non-identical components. For this reason, the replacement costs, which are the most quantifiable costs for component end-of-life, are the only costs considered here in management decision-making.

Going from a microscopic view on the lifecycle of a component to the broader view of the consequences of power utility preventive actions on a population of environmentally vulnerable identical components, inspection costs and replacement costs will be accrued annually. The sum of all the costs represents the costs relevant in the decision-making process tackled in this chapter for the run-to-failure program or the preventive maintenance program.

5.1.2 Defining the Decision Parameters

Decision parameters may also be called design variables. The optimum decision in an optimization problem is made by evaluating and comparing solutions of an objective function from a sufficient number of possibilities of the decision parameters. An optimal solution varies for different sets of decision parameters. The design parameters must reflect decisions within the power, control and authority of the power utility. For instance, while a power utility can define the number of components to inspect in a year, it cannot control the environment: weather conditions are unpredictable.

Given the definition of the power utility management-problem described in the introduction of this chapter, the decision parameters are the following.

1. The component inspection-schedule
2. The proportion of inspected components that are preventively replaced

5.1.2.1 The Component Inspection-Schedule

The schedule affects the number of components inspected annually. Assuming the inspection-schedule is regular and static, that is the same schedule is used each year, the number of components inspected N_I per year n is defined as

$$N_I[n] = \frac{N[n]}{c},$$

where $N[n]$ is the total number of identical components to be managed in year n , and c is the inspection-schedule. In this work, it is assumed that the total population of components in each year remains unchanged. In other words, growth or decline in the population size of the components is ignored. Therefore, for all years n ,

$$N_I[n] = N_I \text{ for all } n.$$

In this work, the inspection-schedule c is defined to be a bounded integer. Because giving c the value of zero implies no preventive maintenance and makes $N_I[n]$ undefined, c is defined not to include the value zero. Also, the upper bound of c is defined so that there are not a large number of years between inspections of components. The latter would show a low public safety-interest from the power utilities. Thus, c is defined as

$$a < c < b,$$

$$a > 0; a, b \in \mathbb{N},$$

where a is the lower bound of c , b is the upper bound of c , and \mathbb{N} is the set of natural numbers (positive integers greater than zero).

It should be noted that the results of age-dependent fragility assessment of some components may show a recommended start-age for inspections. This would reduce the total number of components at the “benefit” of inspections from N to $\sum_{t \geq a} N(t)$, where t is time or age and a is the recommended minimum age of inspected components. So, the formulation of the population-size depends on the component analyzed.

5.1.2.2 The Choice of Preventive Replacements

The proportion of inspected components that are preventively replaced in each year will affect power utility spending on replacements in that year. Preventive replacement of all inspected components will result in maximum replacement-expenses and is unjustifiable especially given utility budgets. By prioritizing scheduling of components as discussed in Chapter-section 3.2.3, it is possible that a large number of high-failure-risk components may be inspected in the first few years. However, with the low *percentages* of historical failures of components exposed to severe environmental stress, it is not sensible to replace all inspected components. Note that though the percentages of failures may be small, the total number failing may reach thousands, translating into millions of dollars in expenditures for certain types of power utility components.

The minimum possible proportion of preventive replacements σ_{\min} is zero-percent: no components replaced preventively. The justification for a drastic decision to avoid replacements may depend on the accuracy of diagnostics. For instance, if accuracy is high, components identified to be replaced could save the utility in corrective replacement costs and system reliability. However, since the power utility faces the problem of aging infrastructure and should start progressively replacing very old components to avoid much larger failure risks in the future, σ should be chosen to be greater than zero. The choice of preventive replacements here will be defined as some

proportion of the expected number of failures of inspected components. This will be applied to the annual cost of preventive replacements in Section 5.1.4.

It is suggested that at least the proportion of inspected components expected to fail in one year should be replaced preventively. Assuming the number of components expected to fail in one year is evenly distributed around the regions of inspection per year, the lower bound of σ is

$$\sigma_l = \frac{E\left[\sum_{t \geq a} \{\mathbf{f}(t) \times N(t)\}\right]}{\sum_{t \geq a} N(t)},$$

where $E[.]$ is expected value, a is the recommended minimum age of inspected components, and $\mathbf{f}(t)$ is in boldface because it is a random variable.

Because the expected number of failures is likely very small relative to the number of components inspected, the upper bound of σ is chosen to be much larger than its lower bound, say by a factor of 100. Thus, the bounds of σ are

$$\sigma_l \leq \sigma \leq 100\sigma_l.$$

(29)

5.1.3 Defining the Random Variables

The comparative annual costs of either a run-to-failure (RTF) program or a preventive maintenance program depend on certain variables that are uncertain and not fixed. Defining the parameters as fixed makes the problem deterministic, while as uncertain makes the problem stochastic. The variables that are more realistically defined as stochastic are listed below.

- Costs of inspections and replacements: preventive and corrective,
- Probability of failure,

- Sensitivity of a diagnostic test (an expression of the accuracy of diagnostic techniques applied on components),
- Discount rate of money.

The variables are defined in more detail in the subsections that follow.

5.1.3.1 Discount Rate

The time value of money is a concept of finance that accounts for the premise that societies and governments place a lower value on future dollars than on present dollars. The discount rate is commonly used to model the present value of costs that may be incurred in future years. This rate reflects preference for present rather than future consumption. It is essential for the multi-stage or multi-year optimization of the decision-problem tackled in this chapter.

If the year for which a cost is to be estimated is n years from the present time, then the present value of the cost C may be defined as

$$C_n = \frac{C}{(1+d)^n}, \quad (30)$$

where C_n is the present value, n is the year being analyzed and d is the discount rate.

Interest rate forecasts are not high precision numbers. Long-term and short-term rates depend on the global economy as well as the United States (US) economy. During bad economies, where incentives for investments and borrowing are needed, interest rates are not high. To tackle issues with the economy in the present year, 2012, the Federal Reserve announced that it was likely to leave short-term interest rates at rock-bottom levels [3].

Inflation is a variable that shows how the price of commodities and/or services rises over time. Putting the economy and unemployment in perspective over the past few years, inflation is not expected to rise by much in the next few years. Companies do not

and will not be under any pressure to increase the wages and salaries of their employees. Though inflation can be defined as a random variable, it will be used as a fixed parameter of value 1.6%. The value is an average of the 2012 and 2013 forecasts of inflation from the Congressional Budget Office (CBO) publication. Using inflation, Equation (30) becomes

$$C_n = \frac{C \times (1 + i)^n}{(1 + d)^n},$$

where i is the inflation rate.

5.1.3.2 Costs of Inspections and Replacements

The costs of inspections have been described in chapter-section 4.1.3.3 and the costs of replacements in 3.2.4. Estimates of the costs can be obtained from accounting reports within power utilities, accounting reports from inspection companies, or survey responses from power utilities. Defining the cost of inspections and replacements as a random variable requires some knowledge of how the costs vary or change over time. In the absence of this information, the factors that affect the cost of labor (wages, salary, income and the like) may be used to estimate variation in the costs.

Though unemployment rate, the cost of fuels and therefore transportation, and competition in market, may affect the cost of labor or services, it is assumed that the Employment Cost Index (ECI) causes the only significant variation. The ECI measures change over time in labor costs, the inflation of wages and employer-paid benefits, in a sense. Quarterly values of past ECIs can be obtained from CBO to learn the statistics of distributions of ECIs. The most recent value of the ECI, which will be measured in terms of percentage increase in compensation cost, will be used for projecting future labor costs, affecting both costs of inspections and replacements. Then, any cost of interest (inspections or replacements) can be calculated as

$$C_n = C \times \frac{ECI_n}{ECI_{n-1}}, \quad (31)$$

where ECI is a random variable described by a discrete probability distribution function, n is the current year, $n - 1$ is the prior year, and C is defined as a fixed parameter with a known or projected average value. Let ECI be equal to 0.5% and defined as either a uniform distribution or triangular distribution.

The ratio of cost of corrective replacements to the cost of preventive replacements is some variable a_1 defined as

$$a_1 = \frac{C_{cr}}{C_{pr}},$$

where C_{cr} is the cost of corrective replacements and C_{pr} is the cost of preventive replacements. The variable a_1 is defined as a uniformly distributed random variable with lower and upper bounds at two and five respectively. This reflects the minimum overtime pay required for service outside of a normal 40-hour work week or peak hour work. More detail can be found in Chapter-section 3.2.4. Note that these costs are defined as fixed parameters, only made random by a_1 (if corrective) and the ECI.

5.1.3.3 Probability of Failure of Components from Fast-Moving Hurricanes

The age-dependent fragility assessment of components developed during the course of this dissertation provides expected probabilities of failure of different ages of components across a wide range of wind speeds. The hazard curve of wind speeds (probability density function of high winds) varies by location. The state of Georgia, for instance, has received a lower number of high intensity land-falling hurricane strikes than the state of Alabama or Florida, according to archival records kept by the National Oceanic and Atmospheric Association (NOAA).

The probability of failure of some age t of a component can be estimated using both the age-dependent fragility curve of that component and the hurricane hazard curve of the component. The procedure is fully explained in [5]. The probability of failure of a component of age t out of a population of identical components under a spectrum of possible hurricane winds is

$$F(t) = \int P(C < D|w, t) f(w) dw, \quad (32)$$

where $P(\cdot)$ is the mean probability that the demand (wind load) on the component exceeds its capacity (strength) given the age t and a wind speed w , and $f(w)$ is the probability that a hurricane-induced wind speed of value w will be experienced in the region: it is the hurricane hazard curve. The former probability has already been discussed in Chapter 3. The latter is discussed briefly in this subsection. Note that because the fragility curve is a cumulative distribution function and it has been observed that it translates (or shifts) to the left for increasing ages, at some age of the component, the fragility curve will have a value of one for all hurricane wind speeds. That is,

$$P(C < D|w, t) \rightarrow 1 \quad \forall w \text{ as } t \rightarrow \infty.$$

Thus,

$$\lim_{t \rightarrow \infty} F(t) = \int f(w) dw = 1.$$

The probability $F(t)$ can be interpreted as the proportion of exposed components of age t that will fail under wind load.

The formulation for a wind hazard model differs significantly in literature. The uncertainty in the models is discussed in [5], where three distinct pairs of Weibull parameters are presented for the estimation of the wind hazard model for Southern Florida. In a report by Lai and Kiremidjian [6], severe wind hazard was analyzed for

Hong Kong. The *Lai* model was split into two sub-models: a hurricane recurrence model and an extreme wind model. The mean rate of hurricane occurrences was estimated using Poisson processes, while the extreme wind model was described by Gumbel and Weibull distributions. The Gumbel distribution was found to fit observed data better than the Weibull distribution. The hazard curve generated was an exceedance probability: that is, the probability that the annual hurricane speed for a specified return period would exceed some minimum speed of damage. The application of the probability to a fragility curve was however not stated.

In this work, the risk assessment method in [5] for the Vickery model in [7] will be used for estimating the hurricane hazard curve. The hazard curve is region-specific; so, parameters of the density function that define the curve will also vary per region. For instance, in the southern end of Florida State, a two-parameter Weibull density function with shape parameter equal to 1.769 and scale parameter of 61.07 is given for the Vickery's model. By convolving the fragility curve with the wind speed probability density function to obtain the probability of failure, the expected number of components that will fail in one year Y is

$$Y = \sum_t N(t) \times F(t), \quad (33)$$

where $N(t)$ is the age distribution of the population of components and t is age.

The uncertainty in the probability of failure could be modeled in different ways. For one, a uniform band of uncertainty, say 10%, can be applied around the hurricane hazard curve, the fragility curve or the probability of failure. For another, confidence intervals of a regression fit of the generated failure probability for distinct ages can be used to describe the uncertainty in the failures. Alternatively, enumeration may be used for combining mean and corresponding one-standard-deviations of generated distinct age-specific failure proportions per age. This gives a distribution of failure numbers that

can be fitted to a probability density function to account for stochastics in the parameter. The equation of age-specific failure proportions may be estimated using simple regression analysis.

Appropriate implementation of regression analysis requires a normality assumption on the residuals. A log-transformation model may be used on the response variable (failure probability here) to improve the regression fit. With a simple linear regression fit with the estimate of the intercept \hat{b} and estimate of the slope \hat{m} , the confidence band has a hyperbolic form given by the equation

$$\hat{b} + \hat{m}\tau \pm t_{n-2}^* \sqrt{\frac{1}{n-2} \sum_i \hat{\varepsilon}_i^2 \left(\frac{1}{n} + \frac{(\tau - \bar{t})^2}{\sum (t_i - \bar{t})^2} \right)},$$

where τ is the independent variable for age, n is the number of data points used for the regression fit, ε_i is each one of the residuals for each data point, t_{n-2} is the inverse of the t -distribution for $(n-2)$ degrees of freedom, and t represents the observed data points for age.

Sample results: Application to wood poles

The provided Weibull parameters in [7] are for the hurricane wind speed probability density function in the southern end of Florida. Using the parameters and the age-dependent fragility curves of wood poles, the probabilities of failure $F(t)$ for ages 0, 25, 50, 75 and 100 are shown in Table 30. A simple linear regression model was attempted on the points but the residuals did not show normality.

Table 30 Derived failure probabilities of wood poles for distinct ages 0 – 100 years.

Ages (Years)	0	25	50	75	100
Probability of failure	$1.091 \cdot 10^{-3}$	$1.224 \cdot 10^{-3}$	$2.808 \cdot 10^{-3}$	$1.810 \cdot 10^{-2}$	$6.200 \cdot 10^{-2}$

The failure probability for these mean failure proportions was log-transformed, fit using a simple regression model, and the residuals were tested using a normal probability plot. The equation that defines the regression model is

$$\ln F(t) = b + mt,$$

where b is the intercept, m is the slope and t is the age of the wood pole in years. The probability plot of the residuals is shown in Figure 66. The low Anderson-Darling statistic and the high p -value imply the normality assumption holds.

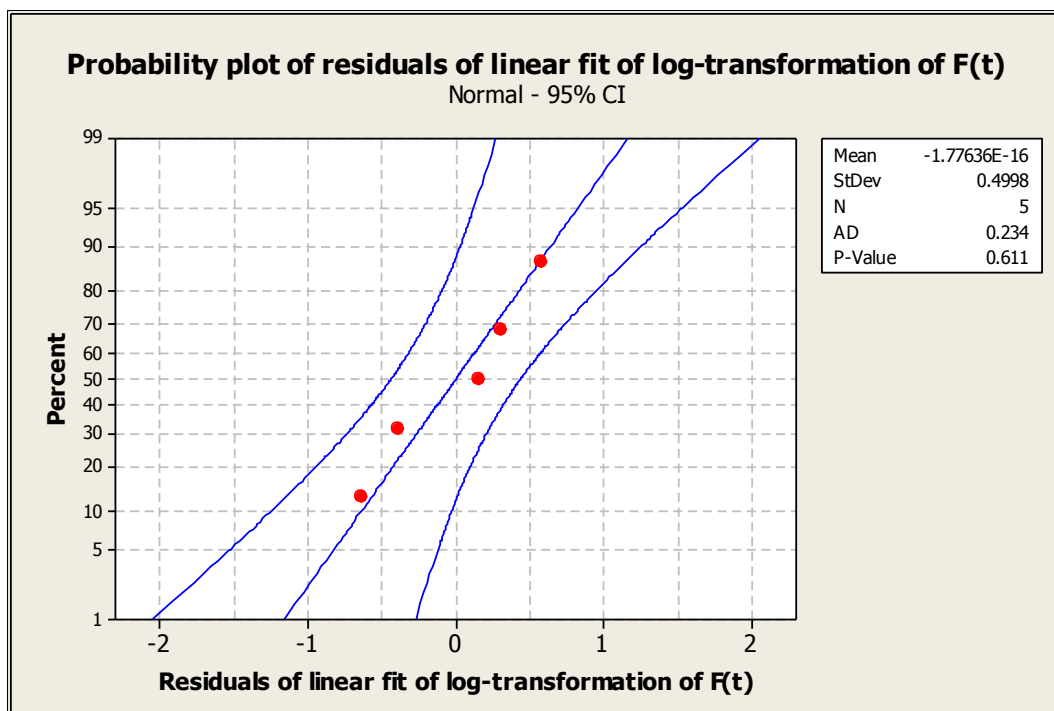


Figure 66 Normal probability plot of residuals of a simple regression fit of a log-transformed $F(t)$.

The plot of the fits is in Figure 67, with a high correlation coefficient of 0.92. All data points fall within the confidence bounds.

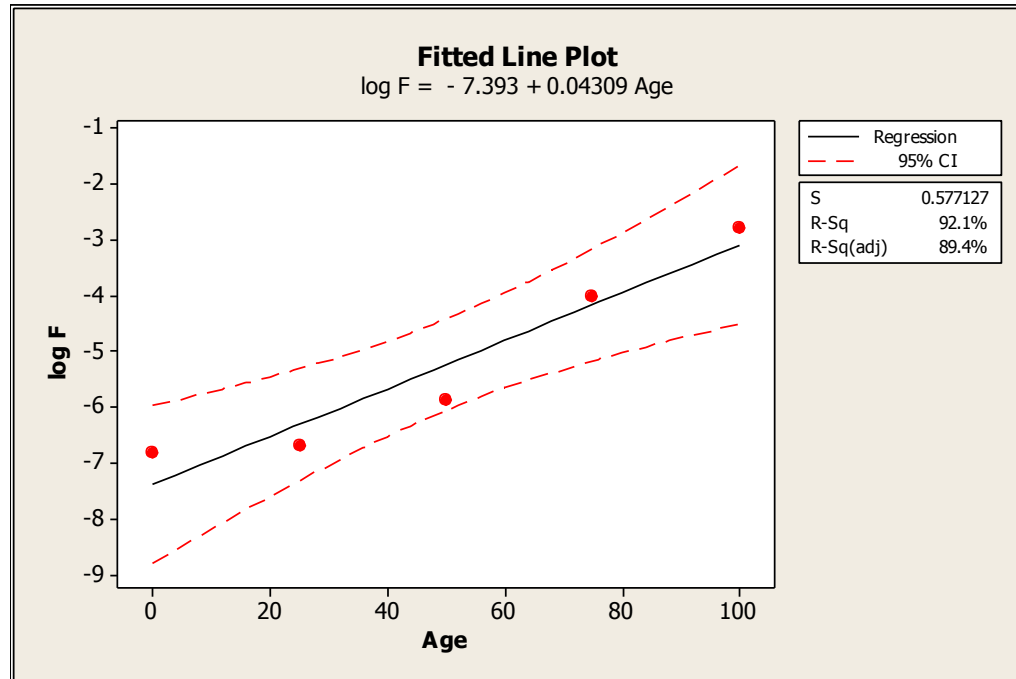


Figure 67 Simple regression fitted plot of the log-transformed failure probability and its confidence bounds.

To obtain the values of $F(t)$ from the fits, the exponential of the right-hand side, that is $(-7.393 + 0.043t)$, must be found. To ensure that the fit never exceeds the value of one, the regression fit is modified so that for the estimates of the coefficients, $\log\text{-}F$ is

$$\min[0, \ln F(t)],$$

where $\min[.]$ represents the minimum between the natural log transformation and zero.

Note that a zero in the exponent gives a value of one. A prediction for $F(t)$ is

$$\hat{F}(t) = e^{\hat{b} + \hat{m}t},$$

which is the exponent of the sum of the intercept estimate and the product of the slope estimate and a known age t . A similar procedure is used to find the prediction confidence bounds of $F(t)$.

The age-specific proportion of failure $F(t)$ and its 95% bounds are shown in Figure 68. From the figure, the interval increases very rapidly after about 50 years. A closer view of the lower ages is shown in Figure 69.

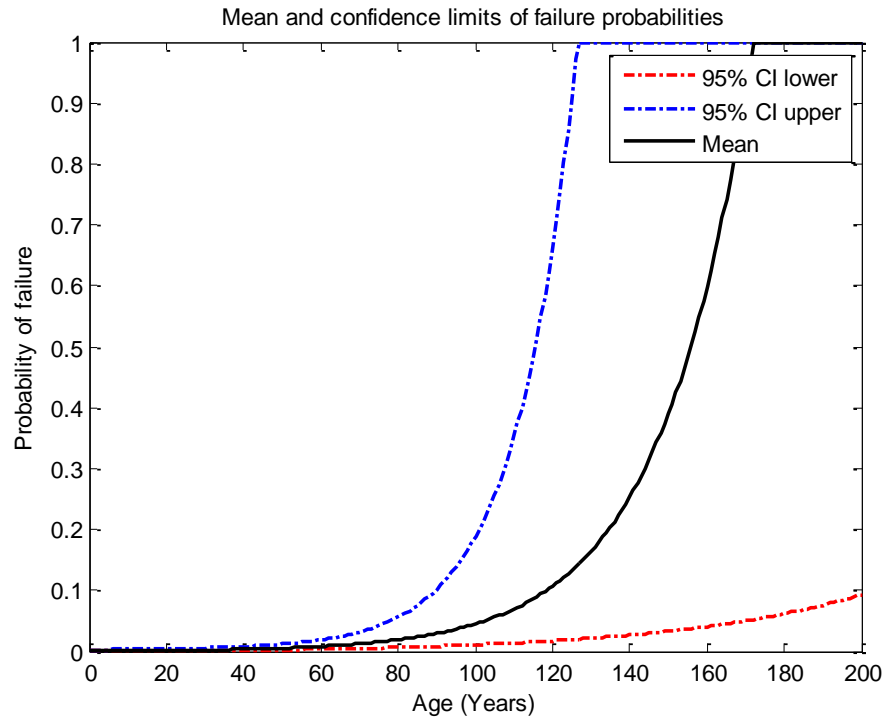


Figure 68 Mean and 95% confidence interval of age-specific proportions of exposed wood poles that may fail under possible hurricane wind speeds.

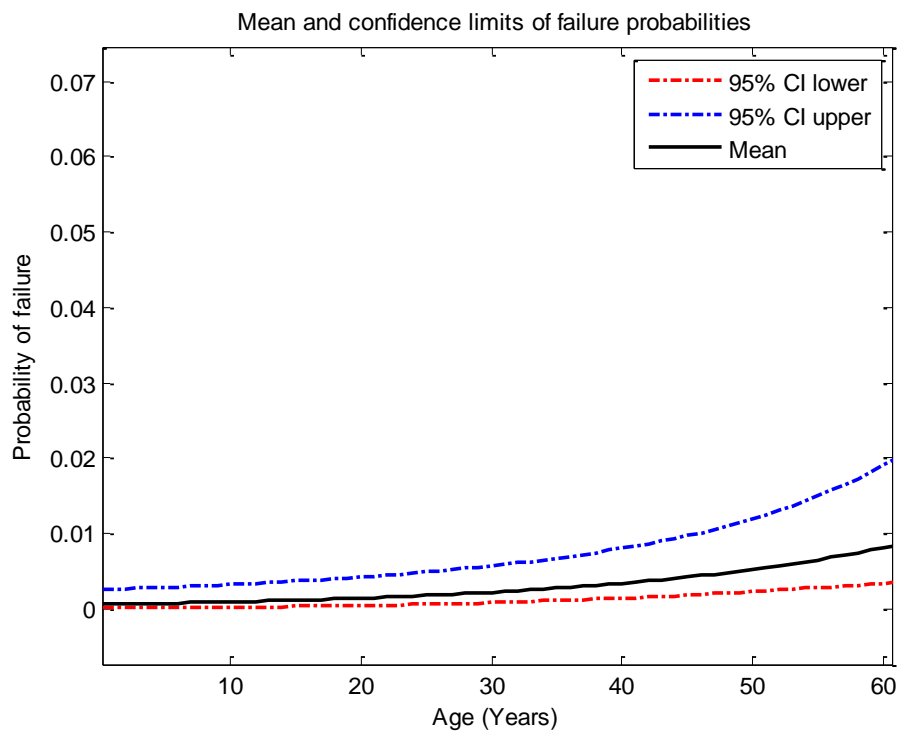


Figure 69 A closer view of $F(t)$ estimate for lower ages.

To illustrate the effect of the failure proportions to a realistic population of wood poles, the mean proportions are multiplied by a scaled age distribution of wood poles from a real power utility. The distribution is shown in Figure 70. The use of the probability density function of wind speeds for convolution with the fragility curves for a location like Florida State includes a hurricane exposure factor. Thus, no scaling is required for component-exposure to some hazard analyzed. A plot of the annual expected failures by age is in Figure 71. Because the products of the number of poles per age and the failure densities may not be integers, the products are rounded up to represent realistic quantities. The total number of wood poles in this power utility is 974,383; the number of them that are at least 20 years old is 701,079 (72% of the entire population).

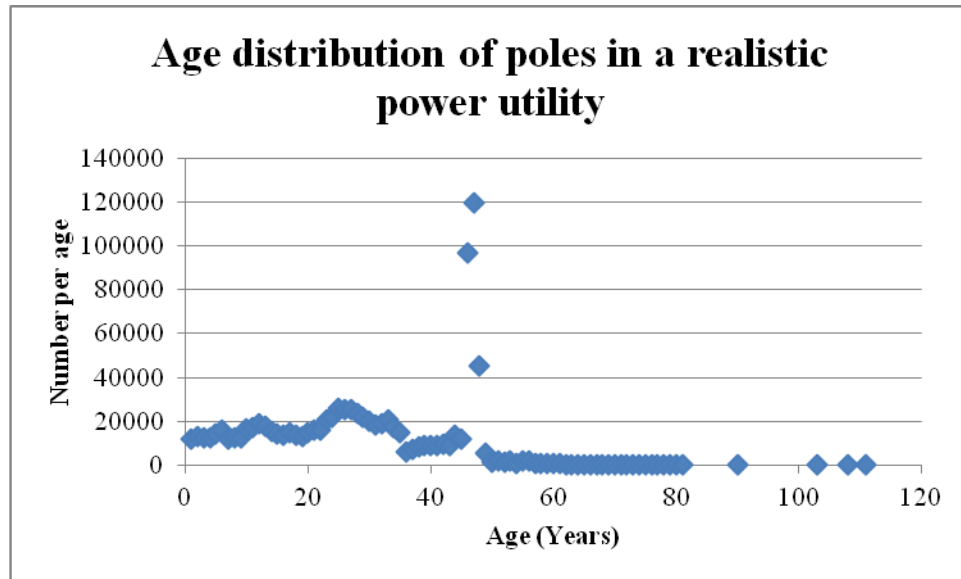


Figure 70 Age distribution of wood poles in a real power utility used in comparing three stochastic sampling approaches.

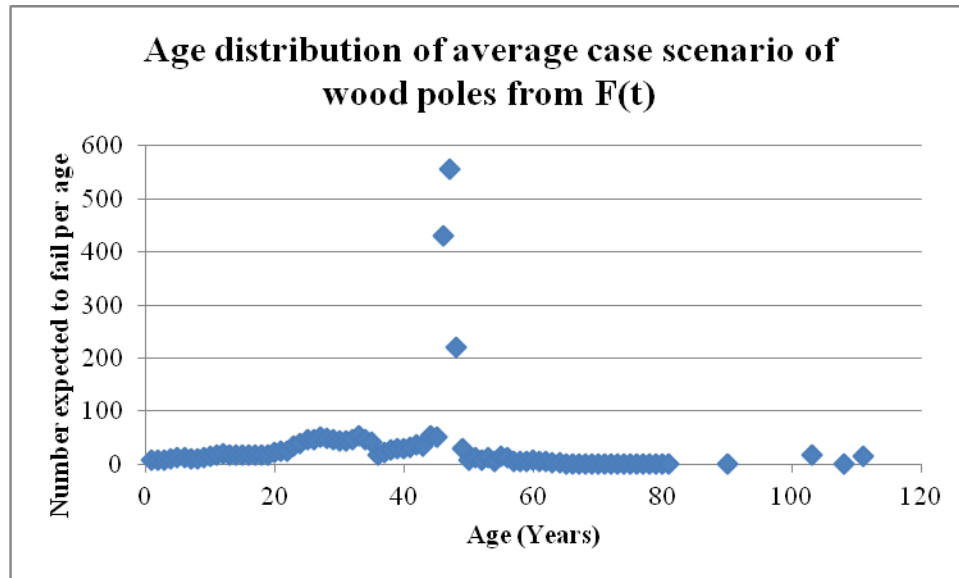


Figure 71 Age distribution of expected failures during a hurricane event for a southern Florida hurricane probability distribution.

To quantify uncertainties in failure proportions, other regression fits are found using enumeration: combinations of the means and one standard deviations of fragilities per age (0, 25, ..., 125 years). The values are shown in Table 31. With six distinct ages and three possible proportions per age, there are 3^6 or 729 such combinations of the proportions. One instance of combination is 0L, 25L, 50L, 75L, 100M, 125U, where each number is age, “L” is “lower,” “U” is “upper,” and “M” is “mean.”

Table 31 Mean and one standard-deviation limits for age-specific failure proportions $F(t)$ estimated for distinct ages 0 – 125 years.

Age	0			25			50		
	Lower	Mean	Upper	Lower	Mean	Upper	Lower	Mean	Upper
F(t)	0.000	0.001	0.003	0.000	0.001	0.003	0.001	0.003	0.006
ln F(t)	-7.920	-6.820	-5.914	-7.786	-6.706	-5.823	-6.863	-5.875	-5.132
Age	75			100			125		
	Lower	Mean	Upper	Lower	Mean	Upper	Lower	Mean	Upper
F(t)	0.009	0.018	0.028	0.042	0.062	0.078	0.192	0.205	0.210
ln F(t)	-4.679	-4.012	-3.578	-3.179	-2.781	-2.552	-1.650	-1.587	-1.559

The combination of (0L, 25M, 50L, 75L, 100L, 125L) yields Figure 72.

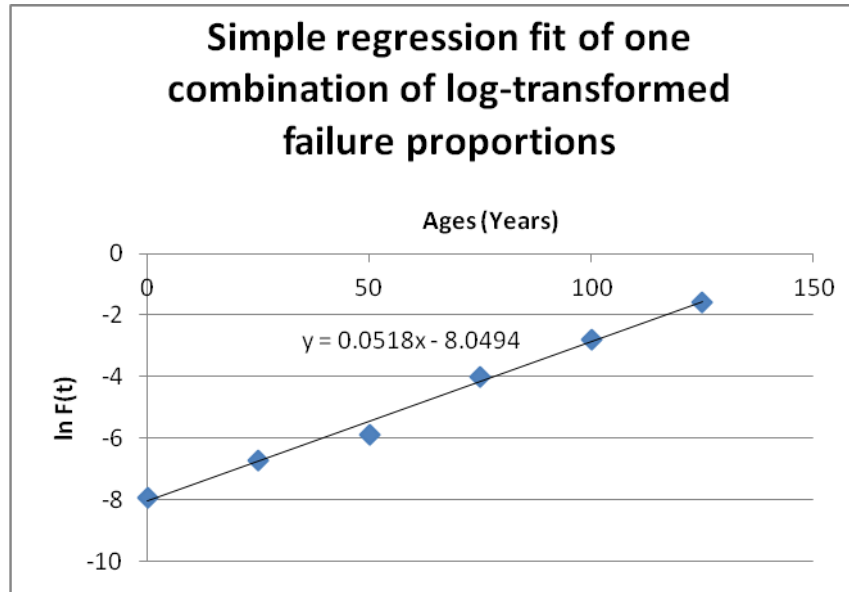


Figure 72 Simple regression fit of one of 729 possible combinations of age-specific failure proportions.

The density of the expected annual failures and those at least 20 years old is shown in Figure 73. Note that the recommended minimum age of inspected poles is 20 years, obtained from inspection of the fragility curve and age-specific proportions of failures. The annual expected failures are defined one random variable in simulation. The number expected to fail past age “a” or 20 years in this case is

$$Y_a = \sum_{t \geq a} [N(t) \times F(t)],$$

where “a” is the recommended minimum age of inspected components based on the results of fragility assessment.

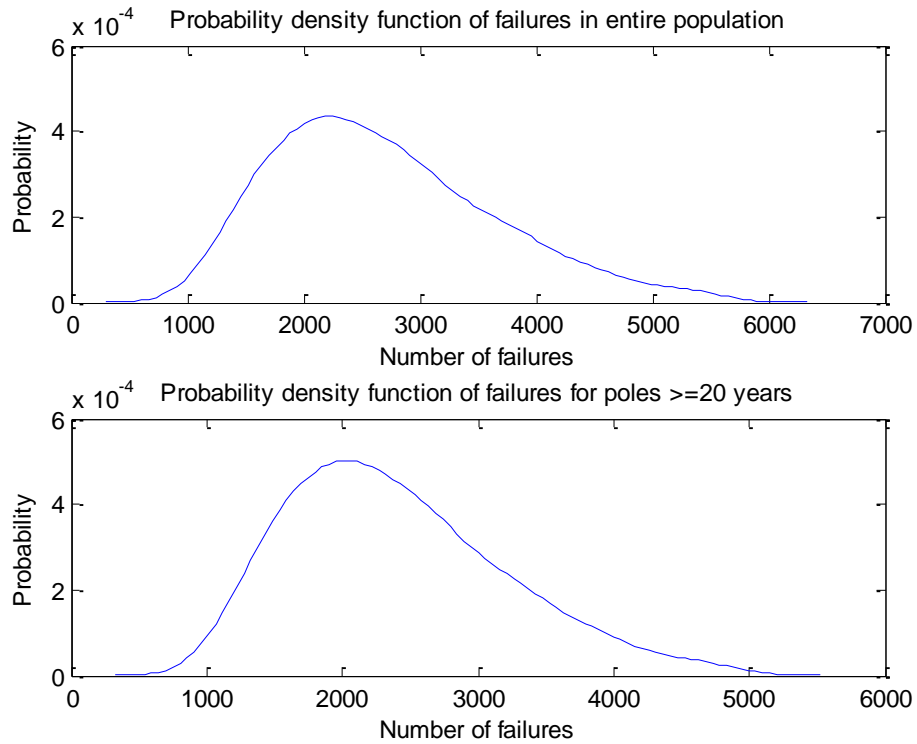


Figure 73 Densities of annual expected failures using simple regression fits of enumeration of means and one standard deviation limits of the age-specific failure proportions.

Goodness-of-fit tests are applied to determine parametric fits to the densities.

While lognormal and Weibull fits were the best distributions, the former exceeded the latter in both Anderson-Darling statistics and p-values. Fits are shown in Figure 74 and Figure 75.

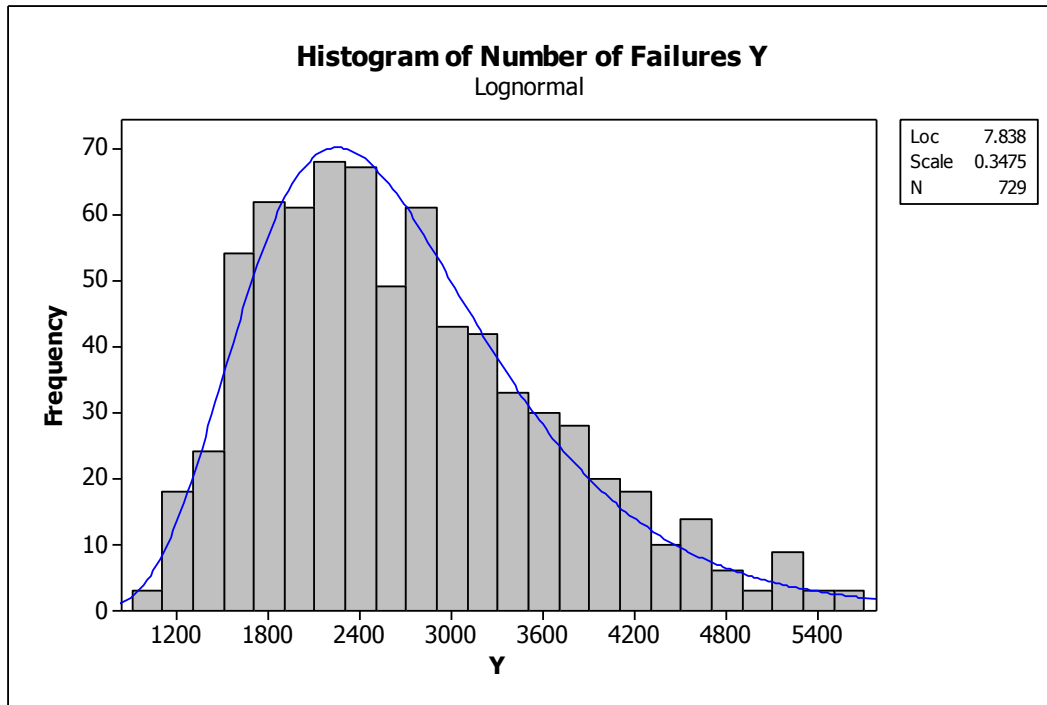


Figure 74 Histogram and lognormal regression fit of the expected failures.

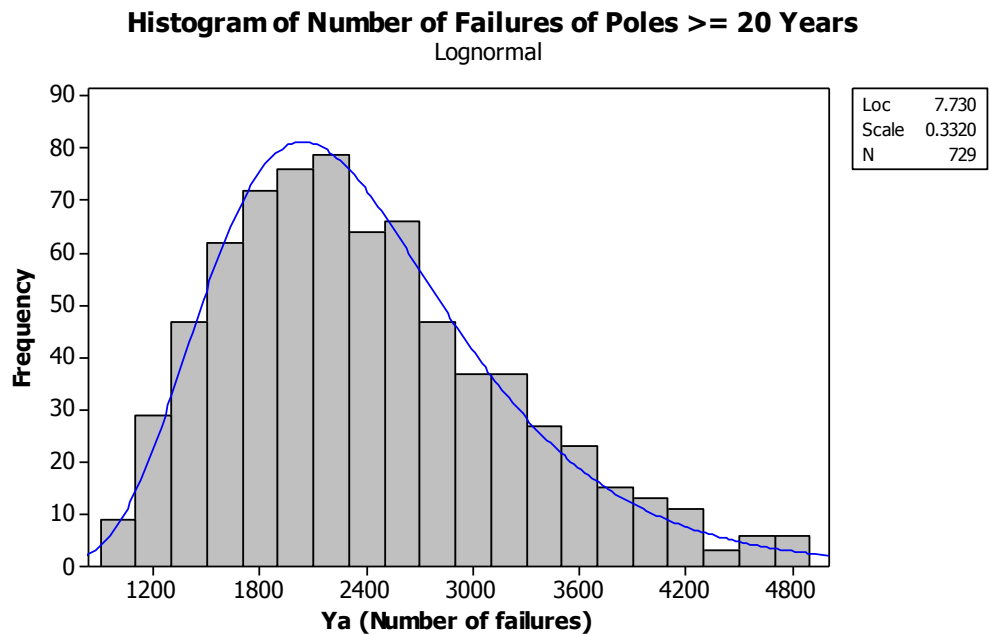


Figure 75 Histogram and lognormal regression fit of the expected failures for ages ≥ 20 years.

5.1.3.4 Sensitivity of a Diagnostic Procedure

The sensitivity of a diagnostic procedure is a measure of the accuracy of the procedure. It measures the fraction of components known to be below reliability and/or safety standards that are accurately detected by a diagnostic procedure (test). The possible bounds of sensitivity are zero (lower) and one (upper). A diagnostic test with a high value is one that detects a large number of inspected below-standard components. In Chapter 4, the possibility of improvement in the accuracy of a diagnostic test was sought through the combination of diagnostic testing decisions using combinatorial logic. Sensitivity will be defined as a uniformly distributed random variable in this work with the lower and upper bounds as some experimental estimate of the sensitivity of a test and the simulation estimate of the maximum sensitivity of combined diagnostic decisions respectively.

5.1.4 The Objective Function

The cumulative cost of corrective replacements for a run-to-failure management approach on wind-stressed components is negligible in a year where there are no high intensity hurricanes, relative to the corresponding cost for a preventive maintenance program. The maintenance program is a proactive approach to component failures and encourages safety and good public perception. During very strong storms, even a preventive maintenance program may not prevent newly replaced components from failing. In addition to the strength of a storm and the corresponding probability of failure of components, the sensitivity of diagnostic procedures, which are the decision-making for the preventive maintenance approach, affects the benefit a maintenance program offers over a run-to-failure (RTF) program.

As discussed in prior sections of this chapter, it is valuable to obtain the optimum inspection schedule and preventive replacement that minimize the differential cost

between a traditional RTF program and a preventive maintenance program under uncertainty. An enumeration of the competing costs is shown in Figure 76.

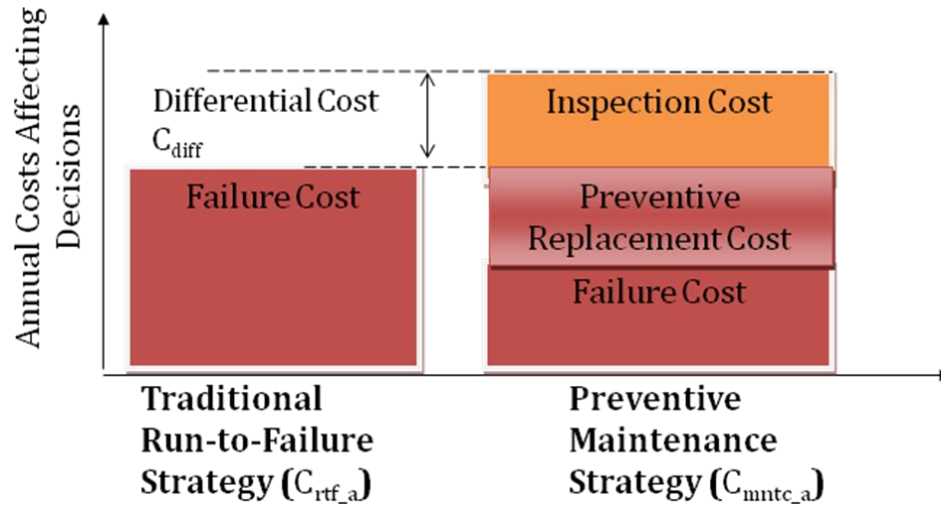


Figure 76 Annual costs of competing strategies (run-to-failure RTF and preventive maintenance) that play a role in component management decision-making, ignoring the cost of penalties.

The objective in this chapter is the minimization of the difference between both strategies. It is expected that the cost of the preventive maintenance approach will be more expensive than the traditional run-to-failure approach. The differential cost C_{diff} is

$$C_{diff} = C_{mntc_a} - C_{rtf_a}, \quad (34)$$

where C_{rtf_a} is the cost of corrective replacements of identical components that fail as a result of high wind loads in a year for the RTF strategy and C_{mntc_a} is the cumulative cost of diagnostics, preventive and corrective replacements of components in a year for the preventive maintenance strategy. The maintenance strategy, for cost-efficiency, should be made as small as possible in comparison to the RTF strategy, making the problem the minimization of the differential cost.

The expected RTF cost is a product of the expected cost of one corrective replacement and the number that fail in a year from Equation (33). Thus,

$$C_{rtf_a} = Y \times C_{cr} = C_{cr} \times \sum_t N(t) \times F(t), \quad (35)$$

where $N(t)$ is the age distribution of the identical environmentally stressed components owned by a utility, $F(t)$ is the age-specific expected proportion of failure of components exposed to environmental stress in the power utility, and C_{cr} is the cost of one corrective replacement. This cost concludes the relevant costs for the RTF program. The costs listed onward are costs in the preventive maintenance program.

The expected annual cost of inspections for a preventive maintenance program, assuming a uniform number of components is inspected per year using a regular inspection cycle c is

$$C_{insp} = \frac{\sum_t (N(t) \times I_{t \geq a}) \times C_i}{c},$$

where

$$I_{t \geq a} = \begin{cases} 1, & \text{for } t \geq a \\ 0, & \text{otherwise} \end{cases}$$

where $N(t)$ is the age distribution of components, $I_{t \geq a}$ is an indicator function, a is the recommended start-age of inspections, c is the inspection cycle and C_i is the unit cost of inspections. Note that the summation in the equation can also be written as a dot product of a vector of the age distribution \mathbf{N} and a vector of ones and zeros of the indicator function \mathbf{I} .

The expected annual cost of preventive replacements C_{pr_a} is flexible since it consists of a decision variable. The proportion of components σ preventively replaced will be defined as some proportion of inspected components expected to fail. Thus,

$$C_{pr_a} = \frac{\sum_t (N(t) \times I_{t \geq a}) \times C_{pr} \times \sigma}{c},$$

where σ is a decision variable (proportion to replace preventively) and C_{pr} is the cost of a single preventive replacement.

Finally, the cost of corrective replacements in the preventive maintenance program is of the number of components that would likely fail in the year excepting those of them that were accurately replaced preventively. The sensitivity of the diagnostic tests will play a role in the “avoided” failures. The following assumptions are important in defining the failure cost for the program.

- The number of components over the start-age of inspections expected to fail in each year is evenly distributed across each year and region of inspection within a cycle. Thus, the number that will fail in each inspected region per year is a proportion of Y divided by c .
- The likelihood of the sensitivity of the diagnostic procedure applied on the components is same across each year of inspection.

Using these assumptions, the expected cost of corrective replacements per year C_{r_a} is

$$C_{r_a} = \left[Y - s_1 \times \frac{\sum_t (N(t) \times F(t) \times I_{t \geq a})}{c} \right] \times C_{cr},$$

where s_1 is the sensitivity of the diagnostic program and C_{cr} is the cost of one corrective replacement.

Since the annual cost of the preventive maintenance program is the sum of the annual costs of inspections, preventive replacements and corrective replacements for the program, Equation (34) becomes for the differential cost,

$$C_{diff} = \left[\frac{\sum_t (N(t) \times I_{t \geq a}) \times C_i}{c} \right] + \left[\frac{\sum_t (N(t) \times I_{t \geq a}) \times C_{pr} \times \sigma}{c} \right] - \left[s_1 \times \frac{\sum_t (N(t) \times F(t) \times I_{t \geq a}) \times C_{cr}}{c} \right].$$

(36)

This equation shows that the differential cost between the RTF strategy and the maintenance strategy is in the number of “good” or beneficial replacements made. These replacements are the averted or avoided failures of the high-risk components.

The aggregate of the optimization problem including uncertainties in the variables is

$$\min_{c, \sigma} E(C_{diff}),$$

which reads “minimize the expected value of the differential cost with respect to the inspection cycle and preventive replacement proportion”. The solution to this problem also gives the minimum annual cost of a preventive maintenance program for a single-stage formulation. While the inspection cycle is an integer variable, the proportion of preventive replacements is not and can have any value within defined realistic bounds as described in Section 5.1.2.2.

5.2 Comparisons between Stochastic Sampling Approaches

In [8], three methods of stochastic sampling are compared in selecting values of input variables for modeling the depressurization of a straight pipe filled with water after the pipe ruptures. The methods, which will also be compared for selecting random variables in this work, are

- Simple random sampling,
- Stratified sampling, and
- Latin hypercube sampling.

Let the differential cost C_{diff} to be modeled be represented by S , for notational simplicity, and the random variables listed in the former sections be represented by vector \mathbf{w} . If one sample of each random variable within a defined range is seen as a *scenario* in generating the result S_1 of the differential cost, then, several other samples could be seen

as other possible scenarios under which the differential cost S may be obtained.

Generating enough scenarios can be used in numerically solving problems such as integrals that may not be easily solved analytically. The selection of appropriate samples in order to keep the variance in estimating the output S sufficiently low is the focus of this section. The procedures for sampling are explained in the following subsections.

5.2.1 Simple Random Sampling

Under simple random sampling, assuming the random variables in \mathbf{W} are independent and identically distributed, an estimate S' of the expected value of the differential cost is

$$S' = \frac{1}{N} \sum_{i=1}^N S(W_i), \quad (37)$$

where N is the sample size and for k distinct random variables in \mathbf{W} , the samples are drawn from the associate probability density functions defining the variables. As $N \rightarrow \infty$, $S' \rightarrow E[S]$. Represent for simplicity,

- The cost of one inspection C_i as x_1 ,
- The cost of one preventive replacement C_{pr} as x_2 ,
- A vector of the probability of failure of distinct ages of a component $f(t)$ as \mathbf{x}_3 ,
- The sensitivity of a diagnostic test s_1 as x_4 , and
- The cost of one corrective replacement C_{cr} as x_5 .

Then, one random sample used in generating one $S(w_i)$ is $w_i = \{x_{1i}, x_{2i}, \mathbf{x}_{3i}, x_{4i}, x_{5i}\}$ for the simple random sampling approach. The estimate in Equation (37) is a point estimate and is not necessarily accurate. The interval for the estimate can be found from the variance of S' , which depends on the sample size chosen. The variance is

$$Var(S') = \frac{1}{N} \sum_{i=1}^N (S_i - S')^2,$$

(38)

It is well known that the ratio $N/(N - 1)$ times the variance in Equation (38) is an unbiased estimator of the variance of S . Say the probability density function of S (\mathbf{W}) is found using numerous samples of the random variables, and the probability of some interesting value (maybe the mean) of S is p' . The 95% confidence interval of the mean S' for large N may be estimated using

$$S' \pm t^* \sqrt{\frac{Var(S')}{N}},$$

where t^* is from the t-distribution. When N is sufficiently large, t^* at confidence of 95% or $\alpha = 0.05$ is $1.96 \approx 2.0$. Since the population mean and standard deviation are often unknown, the sample means and standard deviations may be estimated using some smaller sample size, like $N = 10,000$ versus millions to generate an initial value of the mean and standard deviation.

From the estimate of the confidence interval, it is easy to see that in order to increase the precision of the simple random sampling or crude Monte Carlo approach, N must also be increased. The “curse” of \sqrt{N} is that for an improvement in the precision by one-tenth, 100 times more Monte Carlo replications would be required.

5.2.2 Stratified Sampling

The stratified sampling approach can be an improvement to the simple random sampling approach when strata are chosen appropriately. Recall that it is assumed that $f(W)$ is known and that the transformation of W to S is known based on Equation (36). Let the range space R of \mathbf{W} be partitioned into I non-overlapping subsets of size $p_i = P$

($\mathbf{w} \in R_i$), with the sum of p_i over i equal to one. Let W_{ij} , $j = 1, \dots, n_i$ be a random sample from stratum R_i . The strata means μ_i and variances σ_i^2 are denoted by

$$\mu_i = E[S(W_{ij})] = \int_{R_i} S(\mathbf{w}) \left(\frac{1}{p_i} \right) f(\mathbf{w}) d\mathbf{w},$$

$$\sigma_i^2 = Var[S(W_{ij})] = \int_{R_i} [S(\mathbf{w}) - \mu_i]^2 \left(\frac{1}{p_i} \right) f(\mathbf{w}) d\mathbf{w}.$$

Thus, the estimate of the mean for the stratified sampling approach S_S , which is an unbiased estimator of the overall mean is

$$S_S = \sum_{i=1}^I \left(\frac{p_i}{n_i} \right) \sum_{j=1}^{n_i} S(W_{ij}),$$

with variance

$$Var(S_S) = \sum_{i=1}^I \left(\frac{p_i^2}{n_i} \right) \sigma_i^2.$$

One sample can be chosen per stratum so that $n_i = 1$. Note that when $I = 1$, this approach becomes the simple random sampling approach. For $n_i = 1$,

$$Var(S_S) = Var(S') - \frac{1}{N^2} \sum_{i=1}^N (\mu_i - S_S)^2.$$

(39)

This shows that the stratified sampling (SS) approach offers an improvement to the former approach and that the variance reduction is a function of the differences between the strata means μ_i and the overall mean S_S . The variance of the SS approach is less than or equal to that of the SRS approach when the stratified plan uses equal probability strata and one sample per strata.

If the range of k random variables were divided into two equal probabilities and only one sample is used per stratum, an algorithm for the stratified sampling approach

would be as follows. Assume that the k -th root of N is an integer. D is a matrix consisting of integers from 1 to the k -th root of N in each of K columns. The combination of each sample for each random variable sums up to N , for comparison with the simple random sampling approach.

1. $D = \{1, 2, \dots, \sqrt[k]{N} ; 1, 2, \dots, \sqrt[k]{N} ; \dots ; 1, 2, \dots, \sqrt[k]{N} \}$
2. for $k = 1, \dots, K$
3. for $i = 1, \dots, \sqrt[k]{N}$
 - a. $U_{ik} \sim U(0, 1)$
 - b. $M_{ik} = \frac{D_{ik} - U_{ik}}{\sqrt[k]{N}}$
 - c. $W_{ik} = F_{W_k}^{-1}(M_{ik})$
4. end
5. end

In the algorithm, F^{-1} is the inverse of the distribution function of the random variable being sampled. All combinations of the variates for the K random variables are used to compute the output S_{ik} , yielding N samples of S .

5.2.3 Latin Hypercube Sampling

The Latin hypercube sampling (LHS) method is an extension of quota sampling as seen in [9]. Like with stratified sampling, it ensures that all portions of the range space R of the random variables are covered. Even more, each of the input variables W_k has all portions of its distribution represented in the sampling. The range of each W_k (the costs, sensitivity and failure density) can be divided into N strata of equal marginal probability $1/N$, and sample once from each stratum. Let this sample be W_{kj} , where $j = 1, \dots, N$. So, the W_k component, for $k = 1, \dots, K$ is formed in \mathbf{W}_i , $i = 1, \dots, N$. The components of the W_k 's are matched at random. The advantage of the approach is that each of the

components is fully represented in a stratified manner irrespective of what components are the most important.

The process for Latin hypercube sampling can be automated in the following way as seen in [10]. Remember that N is the sample size, K is the number of random variables and W is a matrix of samples of the random variables. Let Π be a set of size N by K of permutations of N integers from 1 to N in each of k columns.

1. $\Pi_1 = \text{PERMUTE } \{1, \dots, N\}$
2. Repeat independently up to Π_K .
3. $\Pi = \{ \Pi_{kj} \}^{N \times K}$
4. $U_{kj} \sim U(0, 1)$, for $k = 1, \dots, K, j = 1, \dots, N$ independently
5. $U_w = \{U_{kj}\}^{N \times K}$
6. $W = \{W_{kj}\}^{N \times K} = \frac{1}{N} [\Pi - U_w]$

The estimate of the mean for the LHS method S_L , which is also an unbiased estimator, is

$$S_L = \frac{1}{N} \sum_{j=1}^N S(\mathbf{w}_j),$$

where each W for one j is a row-vector consisting of k random variates for each random variables. These are inputs to solving for one sample of S . According to [8], the variance of the LHS method is not easily comparable to the SS method. However, it is known to provide an improvement to the crude Monte Carlo when the function of the random variables is monotonic in nature. In comparison to the variance of the crude Monte Carlo, the variance of the LHS method, adapted from [8], is found as follows.

$$\text{Var}(S_L) = \text{Var}(S') + \frac{N-1}{N} \left[N^{-K} (N-1)^{-K} \sum_R (\mu_i - S') (\mu_j - S') \right], \quad (40)$$

where $\text{Var}(S')$ is the variance of the simple random sampling method, K is the number of random variables, R is a restricted space of $N^K(N-1)^K$ pairs (μ_i, μ_j) corresponding to samples that are not common, and μ_i is the output of the LHS selection of a set of the K random variables. Obviously, the variance of the LHS method is less than or equal to that of the simple random sampling method if the expression in [.] is less than or equal to zero.

5.2.4 Simulation Results: Application on wood pole data

5.2.4.1 The effect of sample size on simple random sampling approach

As previously mentioned, an increase in the sample size of a simple random sampling (SRS) approach or a crude Monte Carlo method will decrease the confidence interval of the estimate of the expected value of some transformation or function of the random variables. This is because of the \sqrt{N} in the denominator used in estimating the interval. An experiment is performed in this section to observe the influence of sample size on the SRS method for the differential cost problem applied to a population of wood poles.

The inspection cycle in this simulation is assumed to be 10 years and the proportion of inspected wood poles replaced preventively is 3%. It is assumed that the same proportion is used in each year during the cycle. The estimates of the means and variances of the differential cost C_{diff} or S are shown for different sample sizes in Table 32.

Table 32 Computation time, expected value and variance of differential cost S between a run-to-failure program and a preventive maintenance program for increasing sample sizes.

Sample size	1000	5000	10000	50000	100000	250000	500000	1000000
E [S], M\$	8.277	8.280	8.276	8.281	8.284	8.282	8.282	8.283
V [S], T\$	0.241	0.232	0.237	0.235	0.237	0.237	0.236	0.236
Comp. time	0.013s	0.014s	0.016s	0.029s	0.050s	0.098s	0.333s	0.186s

A plot of the means and confidence intervals is shown in Figure 77. The results show that given the selection of the inspection cycle (10 years) and the proportion of inspected components replaced preventively (5%), there appears to be no financial benefit to the preventive maintenance program. That is, the differential cost is positive, even on the extremes of a 95% confidence interval. The analysis excludes the revenue lost, unquantifiable public perception (reputation), and the reduction of safety risk and anxiety, that would greatly improve the benefit of the program over a run-to-failure program.

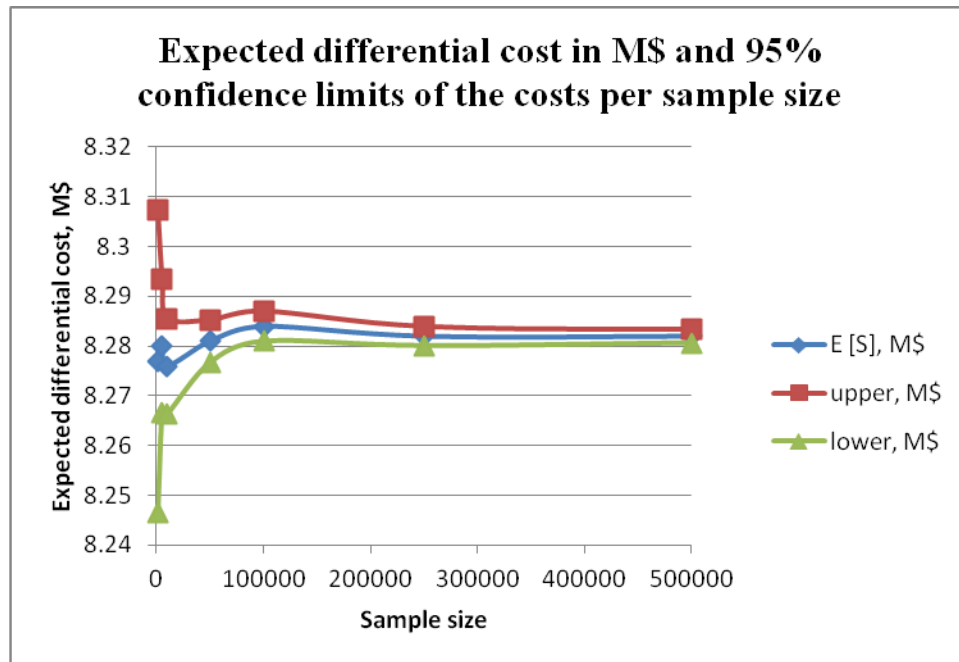


Figure 77 A plot of the estimate of the mean and 95% confidence intervals of the mean for increasing sample sizes of the random variables for the simple random sampling scheme.

5.2.4.2A comparison of three stochastic sampling approaches

With the number of random variables now reduced to five, comparisons of the stochastic sampling approaches are again performed. For ease of understanding, the differential cost function between the competing management approaches C_{diff} or S in this section is shown here using the newly defined Y_a .

$$S = C_{diff} = \frac{1}{c} \times \left\{ \left[\sum_{t \geq a} N(t) \times c_i \right] + \left[\sum_{t \geq a} N(t) \times c_{pr} \times \sigma \right] - [s_1 \times Y_a \times c_{cr}] \right\}.$$

The random variables are emboldened in the equation. It is obvious from the formulation that the decision of the inspection interval does not affect whether there will be a net benefit to the preventive replacement program; that is, for S to be less than zero. It however acts as a scaling to the net benefit, whether positive or negative. The choice of proportion of preventive replacement may on the other hand be able to affect the possibility of savings from the replacement program.

The performance of the three stochastic sampling approaches for the following sample sizes is evaluated.

- 5^5 : Each of the 5 random variables is divided into 5 strata of equal probability. Thus, the sample size for each method is 5^5 : 3,125.
- 6^5 : Each of the 5 random variables is divided into 6 strata of equal probability. Thus, the sample size for each method is 6^5 : 7776.
- 7^5 : Each of the 5 random variables is divided into 7 strata of equal probability. Thus, the sample size for each method is 7^5 : 3,200,000.
- 10^5 : 10 strata of equal probability, a total of 100,000.
- 15^5 : 15 strata of equal probability, a total of 759,375.

The estimates of the expectation, variance, computation time, and memory space for each of the methods are shown in Table 33.

Table 33 Statistics and properties of three sampling plans: simple random sampling (SRS), stratified sampling (SS) and Latin hypercube sampling (LHS) applied to the management cost problem and implemented in MATLAB, where “^”: power-exponent.

Sample size	5^5			6^5			7^5			10^5			15^5		
	3125	3125	3125	7776	7776	7776	16807	16807	16807	100000	100000	100000	759375	759375	759375
	SRS	SS	LHS	SRS	SS	LHS				SRS	SS	LHS	SRS	SS	LHS
E, M\$	8.292	8.284	8.282	8.288	8.283	8.282	8.280	8.281	8.283	8.281	8.282	8.283	8.282	8.282	8.283
V, T\$	0.230	0.226	0.231	0.238	0.230	0.236	0.236	0.233	0.236	0.238	0.234	0.234	0.236	0.235	0.236
95% Lower, M\$	8.264	8.267	8.265	8.277	8.272	8.271	8.272	8.273	8.275	8.278	8.279	8.280	8.281	8.281	8.282
95% Upper, M\$	8.298	8.300	8.299	8.299	8.293	8.293	8.287	8.288	8.290	8.284	8.285	8.286	8.283	8.283	8.284
Range, \$	33612	33357	33696	21690	21310	21618	14684	14582	14676	6050	5994	6002	2186	2180	2184
Difference, \$		-255.521	83.284		-380.069	-71.639		-101.533	-7.162		-56.012	-47.686		-6.767	-2.130
Time, s	0.014	0.234	0.015	0.016	0.564	0.021	0.019	1.201	0.030	0.049	7.091	0.132	0.263	53.415	0.982
Memory, MB	125	150	625	311	373	1306	672	807	2824	4000	4800	16800	30375	36450	127575

The expected values found using the three approaches show that the stratified sampling and Latin hypercube sampling (LHS) techniques provide for faster convergence than the simple random sampling technique. While it takes only about 3,000 samples to get close to a steady-state value of \$8.283 million dollars using the non-simple random sampling (non-SRS) methods, it takes about 750,000 samples for SRS. The factor of improvement in needed sample sizes is about 250.

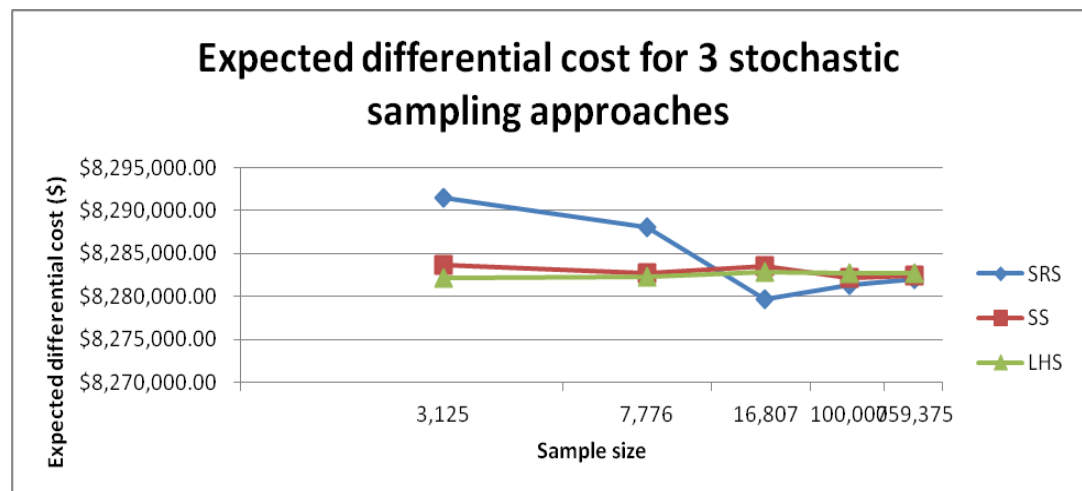


Figure 78 Expected annual differential cost of competing pole-management approaches by sampling using three approaches: simple random sampling (SRS), stratified sampling (SS) and Latin hypercube sampling (LHS).

The upper and lower 95% confidence limits for each sample size and sampling approach were found. However, the ranges (upper limit minus lower limit) were so small in comparison that the differences for the approaches were not obvious. Instead, the differences between the ranges for the non-SRS methods and the SRS method were obtained. These are shown in Figure 79. The figure shows that the range of the limits for the LHS method was initially (at 3,125 sample sizes) slightly greater than the LHS method. They were less afterwards. However, the confidence-limit ranges for the SS method were consistently smaller (and better) than the SRS method.

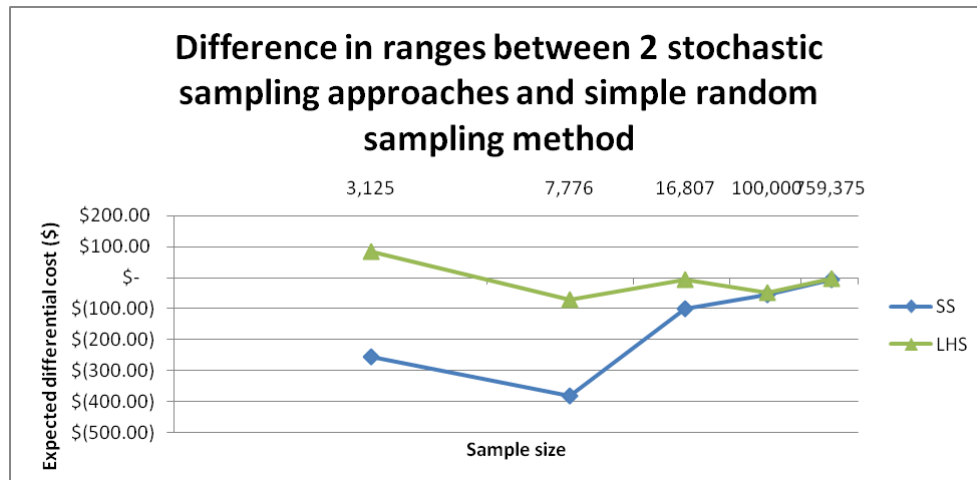


Figure 79 Difference in ranges of 95% confidence limits between the stratified sampling (SS) and Latin hypercube sampling (LHS) methods, and the simple random sampling (SRS) method. The SS provide provides consistently smaller and relatively significant confidence ranges while the LHS range improves after 3,125 samples.

While the LHS and SS methods are superior in variance and convergence to the SRS method, they are inferior to the latter method in memory, labor and computation time requirements. A plot of memory requirements is shown in Figure 80 while a plot of computation time requirements is shown in Figure 81.

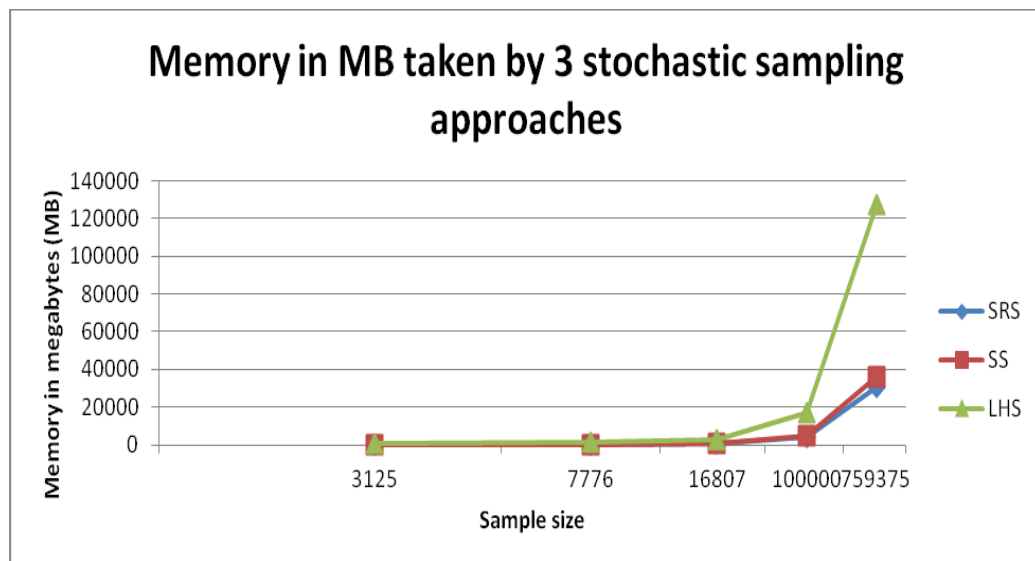


Figure 80 Comparisons of memory requirements for three stochastic sampling methods.

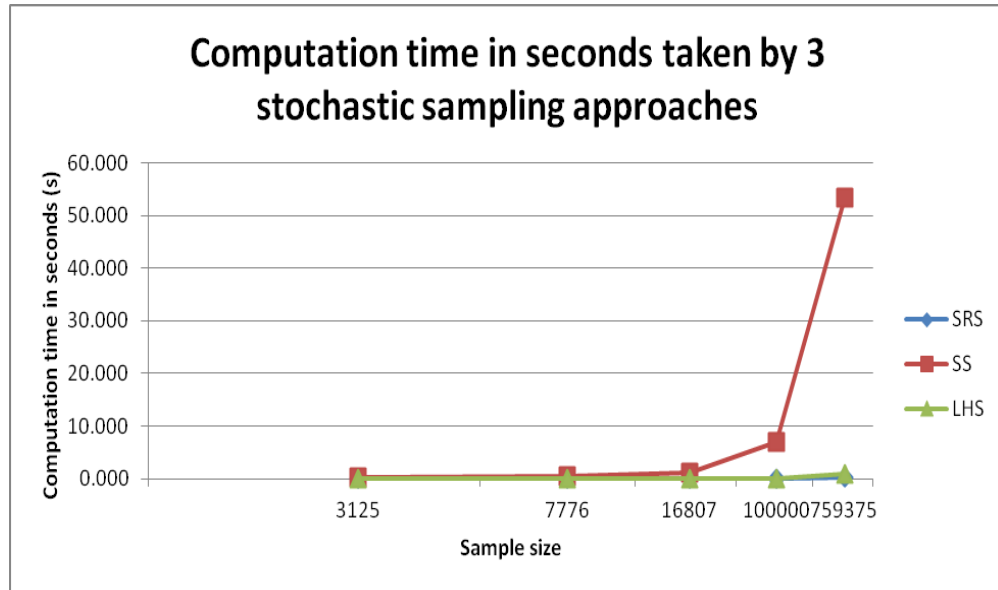


Figure 81 Comparisons of computation time requirements for three stochastic sampling methods.

The performance of stratified sampling exceeds that Latin hypercube sampling in memory requirements, while the latter is an improvement over the former in computation time. Taking a product of the ratios of each of the two methods and factoring in convergence, the Latin hypercube method may in general be chosen over stratified sampling. However, at the time when memory was very expensive, the latter method may have been chosen.

5.3 Single-Year Optimization

The simple random sampling method has been found to be reasonable for analysis of the stochastics of the differential cost function, relative to two other sampling approaches (stratified sampling and Latin hypercube sampling). The method will be used in this section for the optimization of the differential cost between the run-to-failure program and the preventive maintenance program. The random variables, constraints and decision parameters have been discussed in more detail at the beginning of the chapter. The objective function is included below showing the five random variables in boldface.

$$S = C_{diff} = \frac{1}{c} \times \left\{ \left[\sum_{t \geq a} N(t) \times c_i \right] + \left[\sum_{t \geq a} N(t) \times c_{pr} \times \sigma \right] - [s_1 \times Y_a \times c_{cr}] \right\}.$$

The problem here is to find the optimum proportion of components to inspect per year and the optimum proportion inspected to replace preventively so that the objective function is optimized on average. Here, this is the maximization of the expected net benefit or minimization of the expected net financial loss from a preventive maintenance program. Thus,

$$\min_{c, \sigma} s(c, \sigma) \equiv \min_{c, \sigma} E[S(c, \sigma, \mathbf{W})], \quad (41)$$

where $E[.]$ is the expected value, c is an inspection cycle in years, σ is the proportion of inspected components to replace preventively, and \mathbf{W} is a vector of random variables.

The problem is neither convex nor linear. Given the bounds, a stochastic sampling approach may be easily used to solve the problem.

To hedge against variability, one may consider the following optimization problem.

$$\min_{c, \sigma \geq 0} \{s_\beta(c, \sigma) \equiv E[S(c, \sigma, \mathbf{W})] + \beta \text{Var}[S(c, \sigma, \mathbf{W})]\},$$

where β represents the weight given to the conservative (possibly secondary) part of the decision. If $\beta = 0$, the minimization problem becomes the same as in Equation (41), while a large β implies that the solution with minimal variance is sought. More details on similar formulations can be found in [11]. In [11], a maximization of profits was described. So, while an addition operator is used in this formulation, a subtraction operator and “max” are used for profit maximization. Because variance can itself be defined as an expected value, from a mathematical point of view, the latter formulation is similar to the former one. Other properties of random variables like quantiles may be

defined similarly in the formulation of an optimization problem without increasing the complexity of the problem.

This formulation is applied to the wood pole problem.

5.3.1 Simulation Results

The optimization formulations are applied to a realistic database of power utility wooden poles. There are numerous (hundreds of millions) of wood poles used by power utilities around the country. Several power utilities in the United States have recorded estimates of a varied number of pole failures over the years following hurricane strikes on their facilities. Some failure numbers include

- 8,800 from Hurricane Hugo in September 1989 [17],
- 5,500 from Hurricane Fran in September 1996 [17], and
- Over 20,000 from Hurricane Andrew in 1992 [18].

The aged population is taken to be poles over the age of 60 years in the simulations.

In the absence of a time series of information on historical costs, failures and evaluations of diagnostic accuracy, subjective judgment must be relied on for estimation of the range of each item as discussed in [12]. Unfortunately, a uniform density function produces uninteresting results that show no clear direction on decision-making. However, a uniform density function may be used as a prior density function for the random variables and then updated to find posterior density functions for the variables as new data is acquired over time. Some information on the random variables may be obtained from utility surveys (for costs) or analytical modeling (for failure predictions). In the future, a Bayesian approach may be used for updating parameters of posterior density functions of the random variables as data are acquired over time.

5.3.1.1 Assumptions for the single year optimization formulation

The following assumptions are made in formulating the cost function for the optimum net benefit of the preventive maintenance program on wood poles.

- The population of components remains unchanged during the optimization period: there are no new installations.
- Environmental hazards are the primary and possibly the sole causes of failures of the components: no replacements for purposes of construction or vehicular or other man-made accidents cause failures of the components.
- The number of components expected to fail per year are evenly distributed by the regions inspected in each year: the number of high-risk components can be scaled by the inspection cycle.
- Preventive replacements are done prior to hurricane season (or season of the environmental hazard of interest) so that some failures and therefore costs of some otherwise corrective replacements are avoided by the maintenance program.
- The unit costs of inspections and replacements are primarily affected by the increase of labor salaries and wages per year.
- Independence of all non-cost random variables is assumed. The costs are simulated to be perfectly correlated: since they are dependent on trends in labor rates.

5.3.1.2 Approach for parameter estimation of costs of single inspections, preventive replacements or corrective replacements

The United States Department of Labor's Bureau of Labor Statistics (BLS) provides a procedure for determining the escalation of wages and salaries for bargaining agreements for businesses or business contracts from the employment cost index (ECI).

This can be found in [13]. This procedure will be used in this work to obtain possible unit costs of management from a time series of compensation increases provided by ECI information. The general guidelines in [13] are

1. Specifying the costs to be escalated: In this case, costs of single inspections C_i , preventive replacements C_{pr} , and corrective replacements C_{cr} .
2. Identifying the appropriate index and series: Since the workers often employed by the power utility for maintenance of equipment, specifically for wood poles, are often contracted externally, the total compensation ECI for private industry workers will be used here.
3. State the frequency of adjustments: The frequency here is annually as the stochastic optimization is for annual net benefits of the preventive maintenance program.
4. Compute the percentage increase: The index number for the most recent year is divided by the index number for the prior year to determine the percentage increase. Multiply the percentage increase by the base cost to determine the “escalated” cost. This is as in Equation (31). The base costs in this work will be the costs extracted from utility surveys or contractor information.

The base costs from utility surveys, which will be used in this work are

- Cost of single inspections C_i : \$20,
- Cost of single preventive replacements C_{pr} : \$2000, and
- Cost of single corrective replacements C_{cr} : \$6000,

Sensitivity and failure information are in this work assumed to be uniformly distributed until further information is obtained and more appropriate density functions are found for the variables. Parameters used from a combination of subjective judgment and analytical models of failure and evaluations of diagnostic accuracy are

- Sensitivity (an element of diagnostic accuracy) $s_1: \sim U(0.25, 0.40)$, and
- Number failing on average per year as a result of Southern Florida-type hurricanes $Y_a: \sim \text{lognormal}(\mu=7.73, \sigma=0.3329)$.

The time series of ECI are provided in [14]. The December indices for 1995 to 2012 will be used to determine percentage increases. While a decade of data might be more reasonable to simulate recent occurrences, going back another half decade will provide more data points for analysis. Outliers may be neglected in the process of analysis. A summary of the statistics of escalated costs is shown in Table 34. Triangular and Erlang distributions are fairly widespread and familiar as mentioned in [12]. They were also recommended in [15] and [16]. The minima, maxima and mean will be used as the parameters of a triangular distribution to describe the costs. The minima of each cost are re-scaled to account for net present value. That is, future money having less value than present money. Equation (30) is used for scaling.

Table 34 Statistics of escalated costs of inspection, preventive replacements and corrective replacements as a product of base costs and ECI percentage increases using ECIs for total compensation for private industry workers. “*”: These costs are re-scaled by discount rates.

C_i		C_{pr}		C_{cr}	
Mean	15.4704	Mean	2062.721	Mean	6188.162
Minimum	15.17906*	Minimum	2023.875*	Minimum	6071.625*
Maximum	15.63591	Maximum	2084.788	Maximum	6254.364

Optimization will be carried out for two scenarios of diagnostic accuracy: one of low sensitivity (between 0.25 and 0.40) and one of high sensitivity (between 0.85 and 1.00). Recall that the bounds of the decision parameters are 5 to 20 years for the inspection cycle c or as in Equation (29) for the proportion of preventive replacements.

5.3.1.3 Optimization results for a low sensitivity diagnostic test

Careful observation of the objective function shows that depending on the number of failures expected per year, the accuracy (sensitivity) of the diagnostic test applied to the components may affect the net benefit of a preventive maintenance approach. A highly sensitive test will lead to a reduction in the differential cost. Because the inspection cycle c mathematically is a divisor and the proportion of preventive replacements is in a numerator in the function, increasing the cycle and/or decreasing the proportion will decrease the expectation of the objective function. Thus, the bounds affect the optimum decisions. With the bounds for the inspection cycle set deterministically from subjective judgment (5 to 20 years based on creating a band around present inspection cycles in real utilities) or constraints set based on performance restrictions, a stochastic sampling approach is used for optimization.

Discrete steps for the decision variables were selected.

- Inspection cycle c : 5 – 20 years with a discrete step of 1,
- Proportion of preventive replacements: 0 to 0.1 with a discrete step of 0.003.

Thus, in all, there are about 530 combinations of the decision parameters.

For a single-year simulation, where candidates (components) for inspection are selected in the first year within an inspection cycle and not allowed to grow over time, Table 35 shows the constraints for the proportion of preventive replacements and other fixed parameters used for the simulation.

Table 35 Parameters of population properties for single-year optimization.

Description	Value
Number of poles	974,383
Number ≥ 20 years: candidates for inspection	701,079
Ratio of failures expected for poles 20 years or greater to the number of candidates	0.0035
Number ≥ 60 years: aged population	3,968
Ratio of aged components to candidates for inspection	0.0057

Given the table and the constraint functions of the decision parameters, the proportion of preventive replacements must be at least 0.57% times the chosen proportion of the aged population planned to be eliminated per year. Crystal Ball software was used in running the optimization. The software allows for the enumeration of numerous possible samples of each random variable simultaneously. The random variables are assumed to be independent and identically distributed, except for the costs. Perfect correlation among the three costs is factored into the analysis. This is because any increase in salaries of the employees in one year will affect the costs in a similar way. For 100,000 trials of a Latin hypercube sampling approach, the optimal decisions are found to be

- Optimum proportion of inspections per year c^* : 5% or a 20-year inspection cycle, and
- Optimum proportion of inspected components to replace preventively σ^* : 0.6%.

The expected value of the differential cost per year is \$641, 828.28 with a variance of \$10.856 million. The cumulative distribution function of the costs is shown in Figure 82. The 95% confidence bound is found to be (\$423 298.29, \$820 126.30). Solely looking at these numbers, one may presume then that the power utility loses financially

by adopting a preventive maintenance approach. However, when lost revenue from energy not served, negative public perception from serving customers with unreliable aged infrastructure, penalties from ignoring a state mandate, and safety are factored into the results, it becomes obvious that the preventive maintenance approach will by far exceed its run-to-failure counterpart.

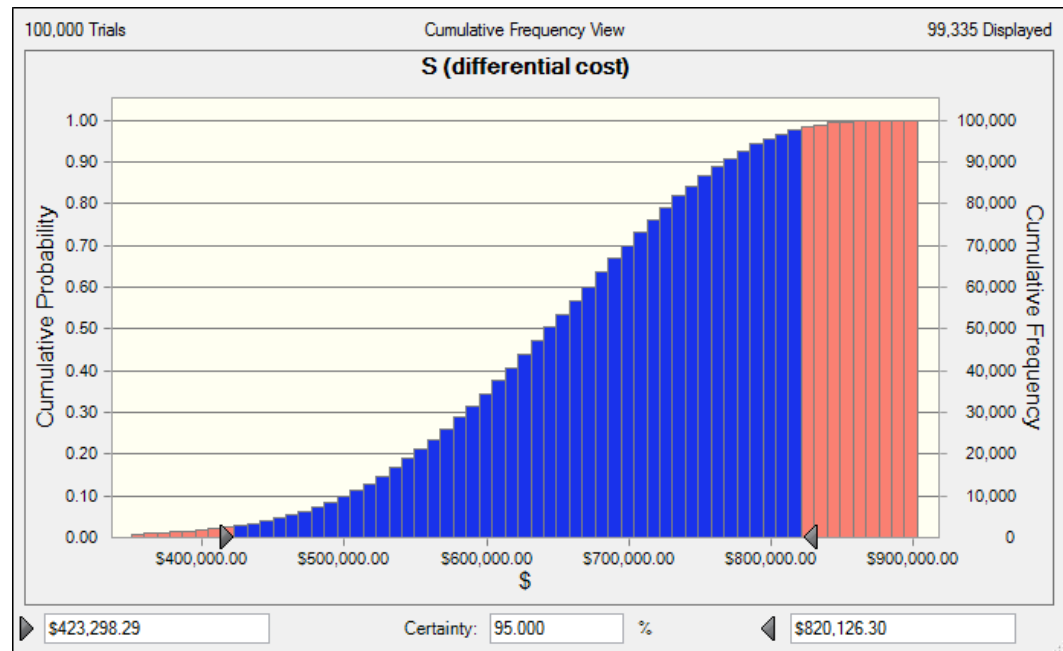


Figure 82 Cumulative distribution function of optimum expected differential cost between two competing maintenance approaches, showing that excluding lost revenue, negative public perception from a run-to-failure (RTF) program, penalties for failing to run a preventive maintenance program, the RTF appears to be more financially viable than a preventive program.

Using a correlation approach, the sensitivity of the expected differential cost to the five random variables is shown in Figure 83. The change in the objective with the random variables is observed using their respective correlation coefficients. Thus, negative sensitivity signifies an inverse relationship (increase in random variable leading to decrease in objective); whereas, positive sensitivity signifies a direct relationship. Note the equality in sensitivities of the costs, resulting from perfect correlations among the three. The costs and number of failures appear to have the most influence on the annual differential costs between the competing management approaches.

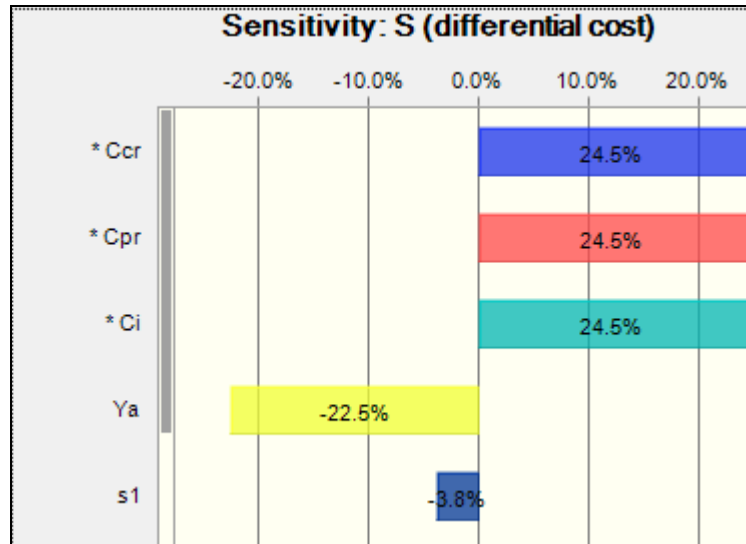


Figure 83 Sensitivity of expected differential cost of five random variables: cost of unit corrective replacement Ccr, cost of unit preventive replacement Cpr, cost of unit inspection Ci, expected annual number of failures out of inspected components Ya, and sensitivity s1, using a correlation approach.

Using the same probability distribution functions of the random variables used for optimization, the distribution of differential costs for one instance of present power utility practices on preventive maintenance is shown in Figure 84. The realistic inspection cycle used is 10 years and a proportion of preventive replacements of 3%, a mean from historical records of a real power utility. The mean of the distribution is \$6,509,583.99 with a relatively large variance of \$196 billion or a standard deviation of about \$440,000. The difference in means (or expected values) is about \$6 million, representing the savings achievable by implementing the optimum decision. The present utility strategies are flawed by the inaccuracy or inability of current diagnostic tests to appropriately detect or distinguish between components at high risks of failure. This becomes obvious in the sensitivity ranks of the objective to the individual costs. The differential cost was positively correlated to the costs by 31.8% each and negatively to the number of expected failures Y_a by a small amount: -3.9%. The annual costs of inspection for the current practice are significantly more than that of the expected corrective replacement costs. Thus the number of failures has a smaller impact on the costs.

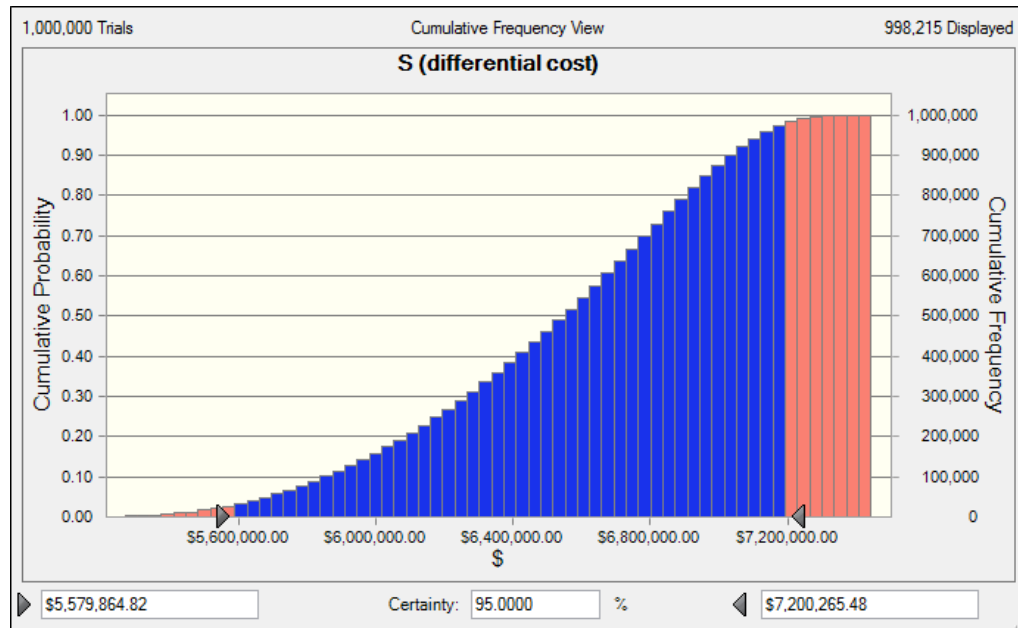


Figure 84 Cumulative distribution function of differential costs between the present utility preventive maintenance approach and a run-to-failure (RTF) approach, showing both a large difference favoring RTF and a large variance.

5.3.1.4 Optimization results for a high sensitivity diagnostic test

The power utility's decision or choice of diagnostic test providers may affect their annual expenditures especially where corrective replacements are concerned. In this subsection, a highly sensitive diagnostic test is assumed in optimization. The sensitivity is chosen to have a uniform distribution with lower and upper limits at 0.85 and 1.0 respectively. The range is the same as that used for the low sensitivity case.

The results show that the same optimum decision parameters are chosen: 20 years and 0.6% for inspection cycle and preventive replacement proportion respectively. The choice is both a factor of the constraint functions for preventive replacements and the structure of the objective which has the inspection cycle as a denominator. The expected differential cost using a highly sensitive (accurate) diagnostic test is \$251,567.19. However, the variance is relatively high at over \$44 billion. A plot of the cumulative distribution function is shown in Figure 85.

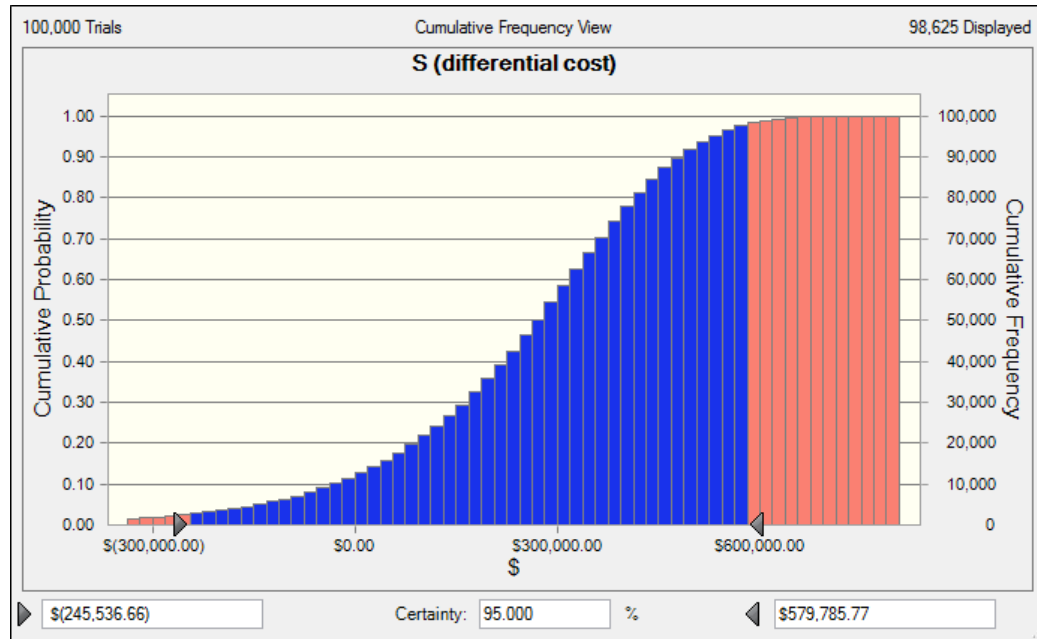


Figure 85 Cumulative distribution function of expected management differential cost for a highly sensitive diagnostic test.

The upper 95th confidence bound here is only slightly greater than the corresponding 25th percentile in the low sensitivity case from Figure 82. With a highly sensitive diagnostic test, the power utility has a better chance of maximizing the net benefit of its preventive maintenance program to the run-to-failure program. Again, this excludes potential penalties to the utility for failing to adopt a maintenance program, lost revenue from power outages during storms, and negative public perception. These factors may cause the net benefits to be in the millions of dollars.

5.3.1.5 Solving the Mean-Value Problem: Single Year Optimization

The optimization problem may be solved deterministically using only one scenario of each of the random variables. However, this yields only one number and gives no information as to the distribution of the net benefit of (or loss from) the preventive maintenance program. The mean values of each random variable, excluding net present value, are shown in Table 36.


Table 36 Values of the random variables for the mean-value problem, where “a,” “b” and “c” are modes, or upper or lower bounds of the distributions.

Random variable	Distribution and parameters	Formula for mean	Value
Cost of inspection	Triangular (a,c,b)	$(a + b + c)/3$	\$15.43
Cost of preventive replacement	Triangular (a,c,b)	$(a + b + c)/3$	\$2,057.13
Cost of corrective replacement	Triangular (a,c,b)	$(a + b + c)/3$	\$6,171.38
Number of failures	Lognormal (μ , σ)	$e^{\mu+0.5\sigma^2}$	2,406
Diagnostic test sensitivity	Uniform (a, b)	$0.5 (a + b)$	0.325

The optimal solution comes out to be the same as for the stochastic optimization problem. However, the differential cost is found to be \$732,334.34, which is about the 70th percentile of the stochastic sampling approach for a low-sensitive diagnostic test.

The ranks of the decisions are shown in Table 37.

Table 37 Ranks of 15 feasible solutions of the deterministic optimization problem. The last column (σ) is the 3-digit form of the proportion shown more fully under the “Constraints” column.

Rank 	Solution #	Objective	+ Constraints				= Decision Variables	
		Minimize Final Value S (differential cost)	Sheet1!D4	>=	Sheet1!D4	<=	c	σ
1	1	\$732,334.34	0.006	>=	0...	0...	20.00	0.01
2	118	\$770,878.25	0.006	>=	0...	0...	19.00	0.01
3	86	\$813,704.82	0.006	>=	0...	0...	18.00	0.01
4	44	\$861,569.81	0.006	>=	0...	0...	17.00	0.01
5	41	\$915,417.92	0.006	>=	0...	0...	16.00	0.01
6	5	\$948,665.72	0.009	>=	0...	0...	20.00	0.01
7	40	\$976,445.78	0.006	>=	0...	0...	15.00	0.01
8	51	\$1,046,191.91	0.006	>=	0...	0...	14.00	0.01
9	249	\$1,116,077.32	0.009	>=	0...	0...	17.00	0.01
10	52	\$1,126,668.21	0.006	>=	0...	0...	13.00	0.01
11	15	\$1,164,997.11	0.012	>=	0...	0...	20.00	0.01
12	50	\$1,220,557.23	0.006	>=	0...	0...	12.00	0.01
13	19	\$1,264,887.63	0.009	>=	0...	0...	15.00	0.01
14	91	\$1,294,441.23	0.012	>=	0...	0...	18.00	0.01
15	54	\$1,331,516.98	0.006	>=	0...	0...	11.00	0.01

The single-year optimization approach excludes a growing population of inspection-candidates over time. In other words, only the population of components that is at least “a” years (recommended minimum age of inspections) at the beginning of the optimization period is considered. The components that shift to the inspection range in subsequent years are not included. Example: A component that is 19 years in the present year will be 20 years in the future. While not a candidate for inspection in the previous year, it may become one in the following year. The management decisions would be more useful if the growing number of inspection-candidates is integrated into analysis. This is termed a “multi-year optimization approach” in this work.

5.4 Multi-Year Optimization

An asset manager would be interested in optimizing the net benefit of a preventive maintenance program on sets of components he manages for at least his term in service. Let the length of the planning period within which he wishes to optimize the benefit be p . By incorporating the growing population of components that is at least the recommended minimum age of inspection, the age distribution in year p from the present year is

$$N(t, p) = \begin{cases} - , & \text{for } t < a, \\ N(t-1, p-1) \times \{1 - F(t)\}, & \text{for } t = a, \\ \left[\frac{c-p}{c} \right] \times [1 - F(t)] \times N(t-1, p-1), & \text{for } t > a, \end{cases}$$

where $N(t-1, p-1)$ is the population size of the prior age group $(t-1)$ in the prior year, $F(t)$ is the age-specific failure proportion, c is the inspection cycle. The first case in the equation is a “don’t care” case since components of age less than “a” are not inspected. The second case represents components that were less than “a” in the prior year and did not fail. The third case represents those that were at least “a” years in the prior year but

were not inspected, and did not fail from environmental hazard events in the prior year.

Then, the total number of components to inspect in each subsequent year is

$$\sum_{t \geq a} N(t, p).$$

The differential cost in each year p from the present year is found using

$$S_p = \left[\frac{\sum_{t \geq a} N(t, p) \times C_i}{c} \right] + \left[\frac{\sum_{t \geq a} N(t, p) \times C_{pr} \times \sigma}{c} \right] - \left[s_1 \times \frac{\sum_{t \geq a} [N(t, p) \times F(t)] \times C_{cr}}{c} \right].$$

It is realistic to assume that at the beginning of each year, the number of failures in the prior year will be known. This would remove randomness in the failure proportions in the prior years and reduce complexity in the stochastic optimization problem. Since the failures per age group differ, a function is still required to determine the number of components per age group that will fail per year. The mean age-specific failure proportion is used in estimating the age distribution in the beginning of each subsequent year. The optimization problem is then defined as

$$\min_{c, \sigma} \sum_{i=0}^p \mathbb{E}[S_i],$$

where S_i is the differential cost at the present (0) or p -th year from the present year. Since the differential cost in each year is affected by decisions of inspections in the former year, the optimization problem could also be written as

$$\min_{c, \sigma} \{ \mathbb{E}[S_0] + \mathbb{E}[S_1 | c_0, \sigma_0] + \dots + \mathbb{E}[S_p | c_{p-1}, \sigma_{p-1}] \},$$

where $c_0 \dots c_{p-1}$ is the inspection cycle at the beginning of each prior year. Note that the conditional expectations are independent of σ since preventive replacements in prior years do not affect the age distributions of uninspected components in future years.

The multi-year formulation is done in such a way as to create ease of implementation for an asset manager: the decisions are defined to be static or remain unchanged during one planning period or horizon. The period could be the expected length of the present manager's administration, for instance. So,

$$c = c_0 = c_1 = \dots = c_p,$$

$$\sigma = \sigma_0 = \sigma_1 = \dots = \sigma_p.$$

This definition will cause the distributions of differential costs in prior years to differ from similar distributions in future years.

5.4.1 Simulation Results: Multi-Year Optimization

The optimization formulations are applied to a realistic database of power utility wooden poles. There are numerous (hundreds of millions) of wood poles used by power utilities around the country. Several power utilities in the United States have recorded estimates of a varied number of pole failures over the years following hurricane strikes on their facilities. Some failure numbers include

- 8,800 from Hurricane Hugo in September 1989 [17],
- 5,500 from Hurricane Fran in September 1996 [17], and
- Over 20,000 from Hurricane Andrew in 1992 [18].

The aged population is taken to be poles over the age of 60 years in the simulations and is found to exceed the number of expected failures per year. The optimization results for a low sensitivity diagnostic test and one of high sensitivity are presented in the following subsections. Recall that the goal is to find the single decision of the combination of inspection cycle and proportion of preventive replacements to be made in the first year of a planning period to ensure that the net benefit of a maintenance program is optimized with respect to its competing corrective program.

5.4.1.1 Assumptions for the single year optimization formulation

The following assumptions are made in formulating the cost function for the optimum net benefit of the preventive maintenance program on wood poles.

- The population of components remains unchanged during the optimization period: there are no new installations.
- Environmental hazards are the primary and possibly the sole causes of failures of the components: no replacements for purposes of construction or vehicular or other man-made accidents cause failures of the components.
- The number of components expected to fail per year are evenly distributed by the regions inspected in each year: the number of high-risk components can be scaled by the inspection cycle.
- The age distribution in each region inspected per year is identical: each age group can be scaled in subsequent years by the inspection cycle.
- Preventive replacements are done prior to hurricane season (or season of the environmental hazard of interest) so that some failures and therefore costs of some otherwise corrective replacements are avoided by the maintenance program.
- The unit costs of inspections and replacements are primarily affected by the increase of labor salaries and wages per year.
- Independence of all non-cost random variables across the planning period is assumed. The costs are simulated to be perfectly correlated since they are dependent on trends in labor rates.

5.4.1.2 Optimization results for a low sensitivity diagnostic test: Static or known function of age-specific failure proportions in prior years of simulation

For $p = 0$ year and using the parameters of the random variables from the previous subsection, also shown in Table 38 for convenience, a multi-year optimization is carried out. In the first year, net present value does not apply. Thus, the range of the unit costs of inspections and replacements is small as seen in the table. Consequently, the variance of the distribution of differential costs is higher than in the single-year optimization.

Table 38 Parameters of the random variables, excluding net present value scaling.

Random variable	Distribution and parameters	Value
Cost of inspection(\$)	Triangular (a,c,b)	(15.18, 15.47, 15.64)
Cost of preventive replacement (\$)	Triangular (a,c,b)	(1562.84, 2062.72, 2084.79)
Cost of corrective replacement (\$)	Triangular (a,c,b)	(4688.51, 6188.16, 6254.36)
Number of failures	Lognormal (μ , σ)	(7.73, 0.333)
Diagnostic test sensitivity	Uniform (a, b)	(0.25, 0.40)

The Crystal Ball software, which uses advanced search algorithms to find the optimum objective, converged to the solution within 40 simulations. For all 512 enumerations, it took 4.5 minutes to run. The solution was found to again be a 20-year inspection cycle or 5% of inspections in this first year, and 0.6% of preventive replacements in that year. The distribution is shown in Figure 86. The mean of the

differential cost is \$732,291.35: very close to the mean value solution. The variance is however about \$8.0 billion.

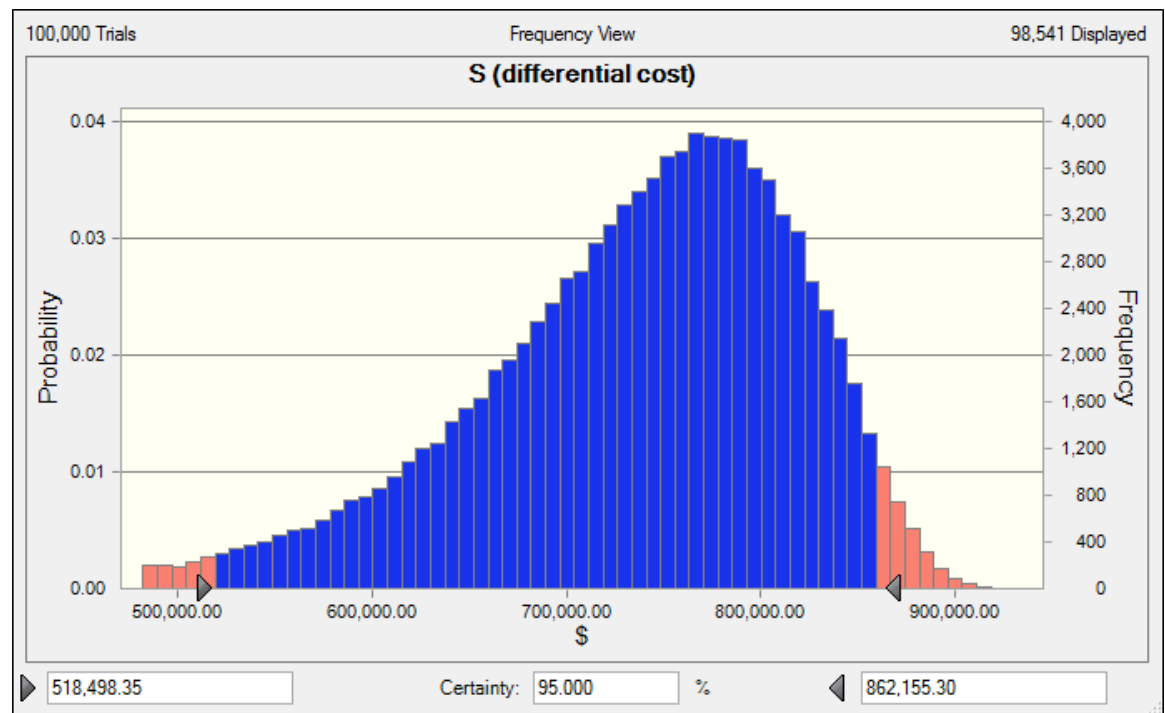


Figure 86 Distribution of differential costs for the first year of optimization excluding net present value scaling.

The sensitivity of the differential cost to the random variables is shown in Figure 87. The cost is seen to change most significantly with the number of failures, less with the test accuracy and barely, with the costs. This may be a result of the small cost ranges.

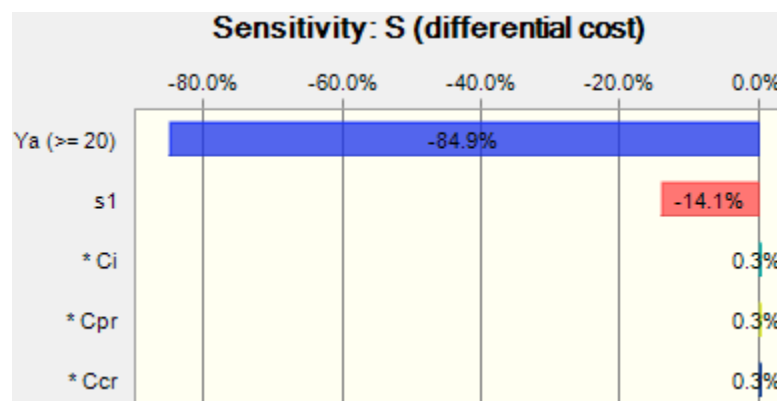


Figure 87 Sensitivity of differential cost to five random variables: number of failures for poles at least 20 years old (Ya), accuracy of diagnostic test (s1), cost of inspection (Ci), cost of preventive replacements (Cpr) and cost of corrective replacement (Ccr), using a correlation approach.

The variance gives some information as to risk of financial loss to the preventive maintenance program. Thus, it may be used to define requirements to the solution of optimization problem. For instance, if the distribution of differential costs is very flat or with very wide tails, the one-standard-deviation range of possible differential costs would be so large it may become unbeneficial. On the other hand, with a small variance, the confidence margin would be small. The range of one standard deviation from the expected value of the differential cost estimated for five feasible solutions of combinations of the decision parameters is shown in Table 39.

Table 39 Expected values, variances and standard deviations of five feasible solutions of combinations of the decision parameters: inspection cycle c and proportion of inspections to replace preventively σ .

c (years)	σ (%)	$E[.]$	$V[.]$	$E[.] - 1$ standard deviation	$E[.] + 1$ standard deviation
5	0.60	\$2,929,165.41	\$127,589,977,431.57	\$2,571,968.02	\$3,286,362.80
10	0.60	\$1,464,582.70	\$31,897,494,357.89	\$1,285,984.00	\$1,643,181.40
15	0.60	\$976,388.47	\$14,176,664,159.06	\$857,322.67	\$1,095,454.27
20	0.60	\$732,291.35	\$7,974,373,589.47	\$642,992.00	\$821,590.70
20	1.50	\$1,381,285.50	\$8,026,112,759.37	\$1,291,696.93	\$1,470,874.07

The expectations and variances of the solutions are shown in Figure 88. The variances for the 20-year inspection cycles are minimal compared to those of other cycles. This implies that their standard deviations are closer to their means, having narrower tails than for 5 – 15 year inspection cycles.

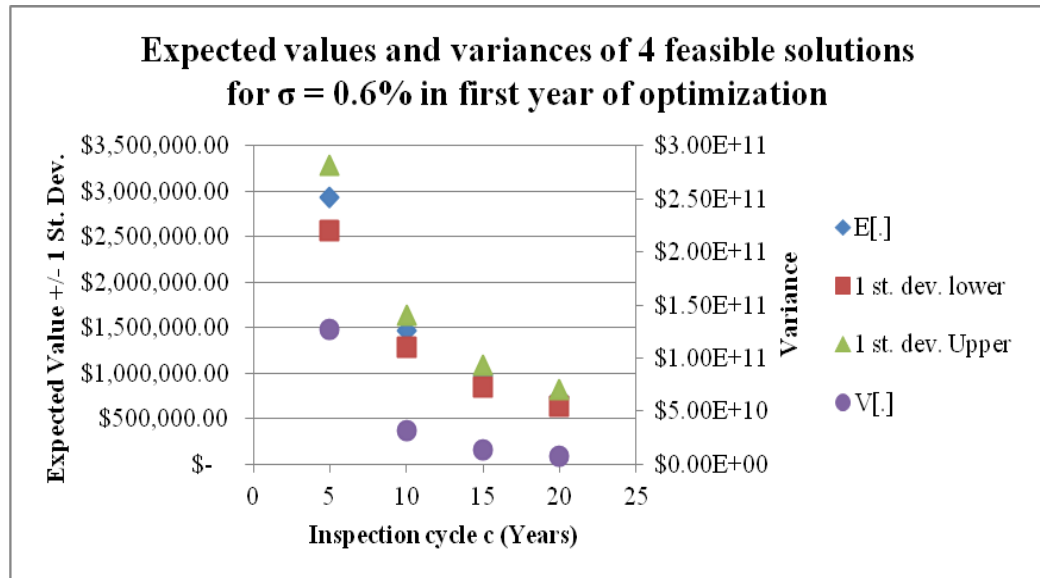


Figure 88 Expectation and variances of feasible solutions of the first year for the optimization problem, showing an inspection cycle of 20 years as the solution.

Managers are sometimes more interested in minimizing the risk of loss; that is, the probability that loss will exceed a certain percentage, than the expectation. The concept is called minimizing *value-at-risk* (VAR). Rephrased, it is also minimizing a certain percentile of a profit-loss function. In the case of the cost function defined in this chapter, the 95% percentile will be minimized rather than minimizing the 5% percentile often used because the profit is the right tail of most profit-loss distributions in finance.

When the planning period is increased to two years, that is, $p = 2$, and the optimization goal is to minimize the 95th percentile of the distribution of the sum of the differential costs in the first and second year, the optimal solution changes only slightly to

- A 20-year inspection cycle and a 0.9% proportion of preventive replacements. The optimum objective value is \$2,023,868.12, which is still smaller than the amount of money the power utility currently spends in one year on inspections and preventive replacements alone.

As can be seen in Figure 89, the distribution of differential costs in the first year ($p = 0$) has a narrower tail than either the sum or the distribution in the second year. This is because by the end of the first year, it is assumed that the function of age-specific failure proportions is known, rather than a random variable.

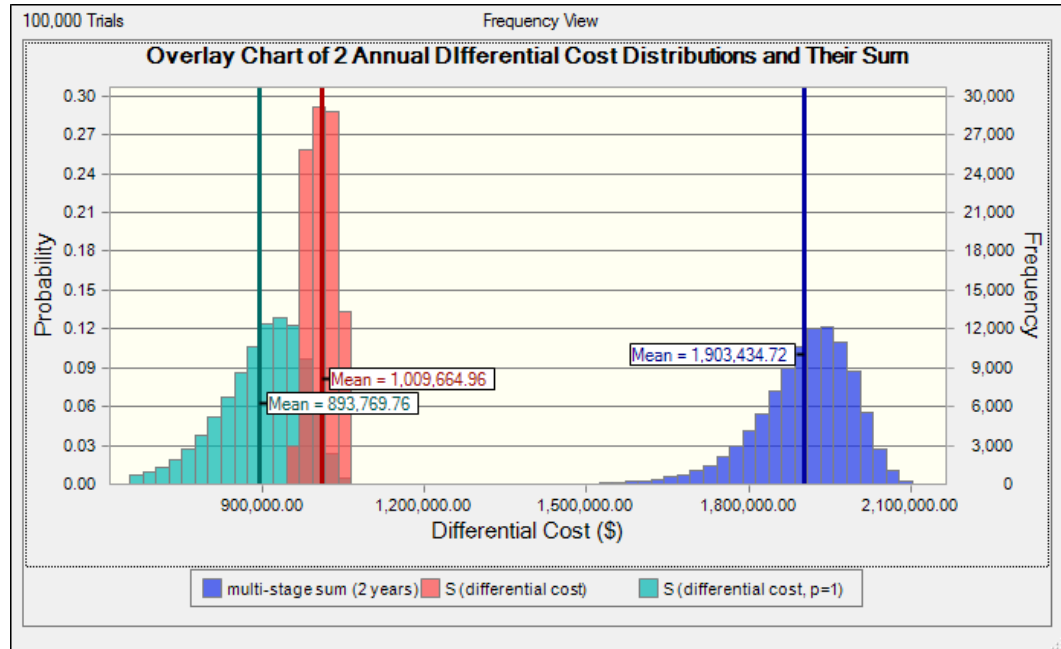


Figure 89 Overlay chart of annual differential costs (left distributions) between competing management approaches for two years of optimization including the distribution of their sums (right).

Also note the decrease in the expectations of the differential costs across the 2 years. This is the result of a few changes in the distribution of components inspected in the second year. The total number of candidates of inspection (those over the age of 20 years) increased from 701,080 in the first year to 677,500 in the second year. However,

- The number inspected per year for the 20-year optimal cycle reduced from 35,054 to 33,875 in the second year.
- The number replaced preventively decreased from 316 to 305 in the second year.

- The number of inspected components simulated to have failed within the inspected region barely changed: 120 to 119 in the second year. The distribution of diagnostic test sensitivity was same for both periods.

The two-stage sum was found to change most closely with the number of failures in the second year. (Those in the first year are assumed to be known by the second year.) The accuracies of the diagnostic tests also appear to affect the two-stage sum more than the unit costs, as seen in Figure 90.

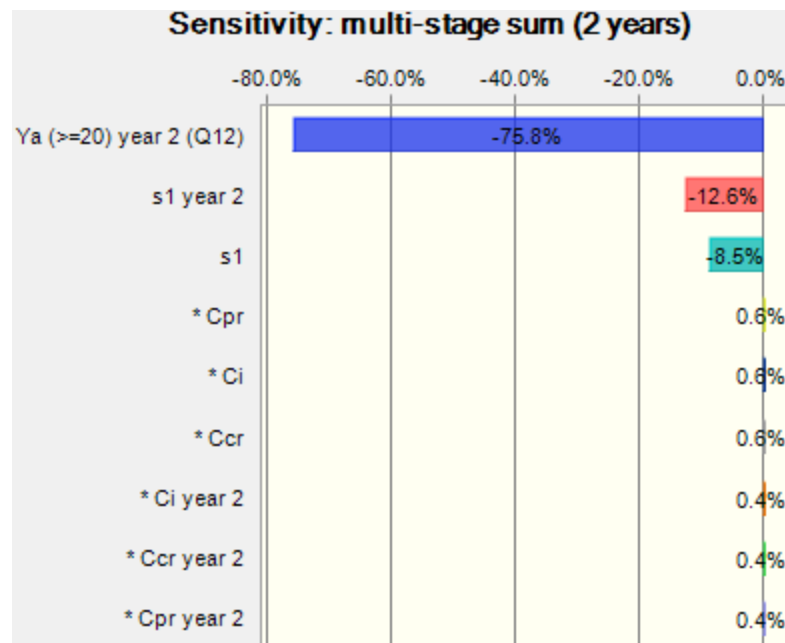


Figure 90 Sensitivity of the sum of two annual differential costs to nine random variables using a correlation approach.

The sensitivity of the multi-stage sum and the differential cost in the third year of optimization is shown in Figure 91. Both costs change significantly with the number of failures, but the sum changes with the diagnostic test sensitivities (accuracies) in the three years more than the differential cost in the third year alone. The latter cost changes more with diagnostic test sensitivity in the third year than the sensitivities in the former years.

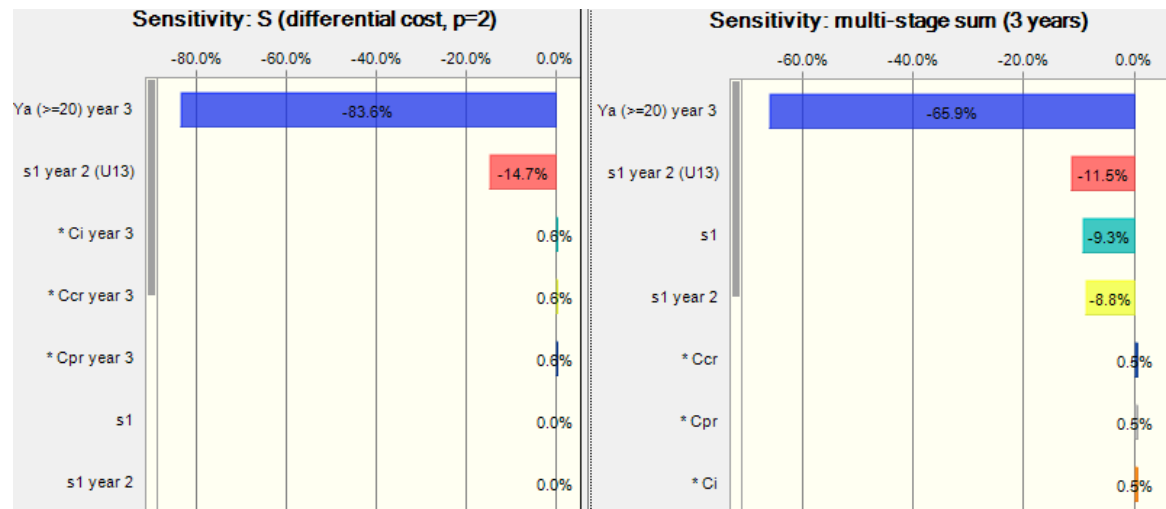


Figure 91 Sensitivity of the sum of three annual differential costs to seven shown random variables using a correlation approach. “U13” is for the sensitivity s1 in the 3rd year.

As with Figure 89, in Figure 92, the distributions in the first years ($p = 0$ and 1) have relatively narrow tails than either the sum or the distribution in the third year. This is because randomness is removed in the failure information for the prior years. It is assumed that the function of age-specific failure proportions is known.

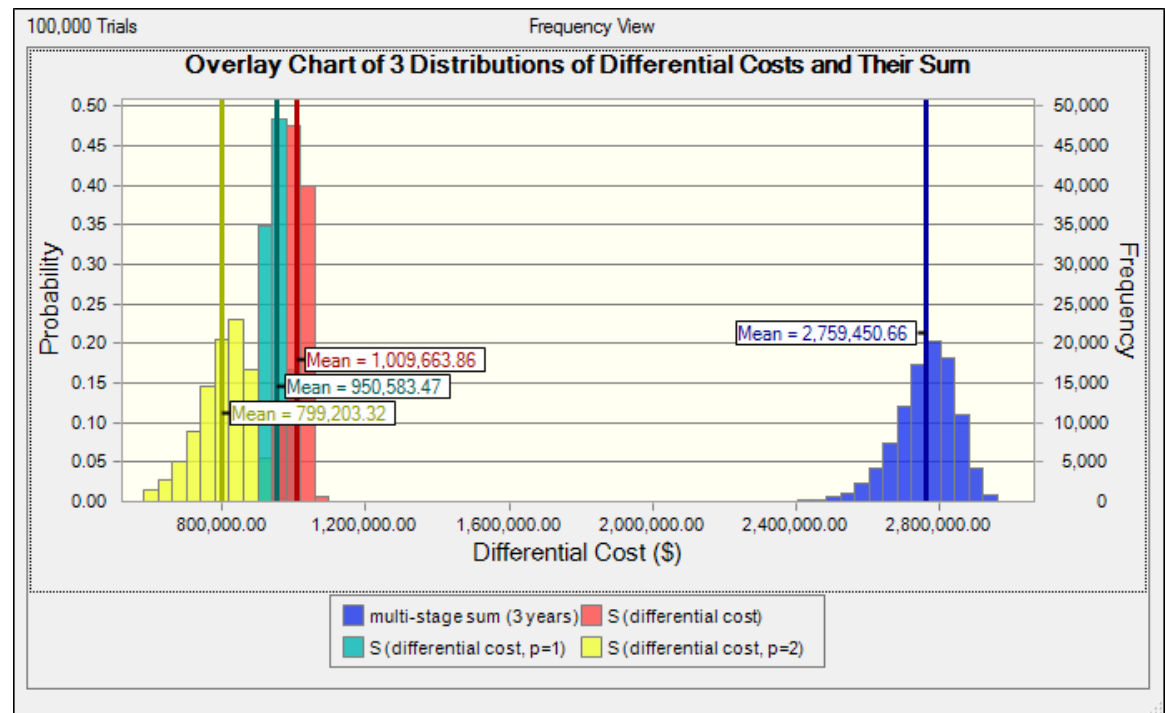


Figure 92 Overlay chart of annual differential costs (left distributions) between competing management approaches for three years of optimization including the distribution of their sums (right).

The optimal solution is the same at (20 years, 0.9%). Note that the proportion of preventive replacements is constrained on the aged population and the expected failures. Because the aged population is in general greater than the expected failures, the former increases more than the latter per year as younger poles consistently get older. The minimum 95th percentile of the third year sum of the differential costs is \$2,879,247.66.

The optimal solution even after four years of optimization is (20 years, 0.9%). The minimum 95th percentile of the third year sum of the differential costs is \$3,605,988.61. The overlay of all four distributions is shown in Figure 93. The sensitivities as before are in Figure 94.

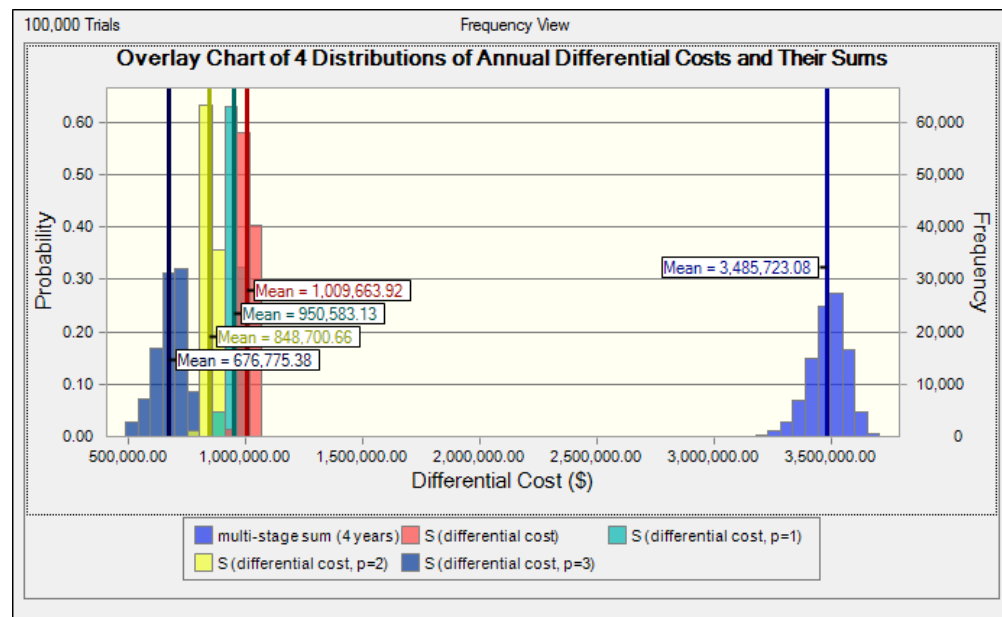


Figure 93 Overlay chart of annual differential costs (left distributions) between competing management approaches for three years of optimization including the distribution of their sums (right).

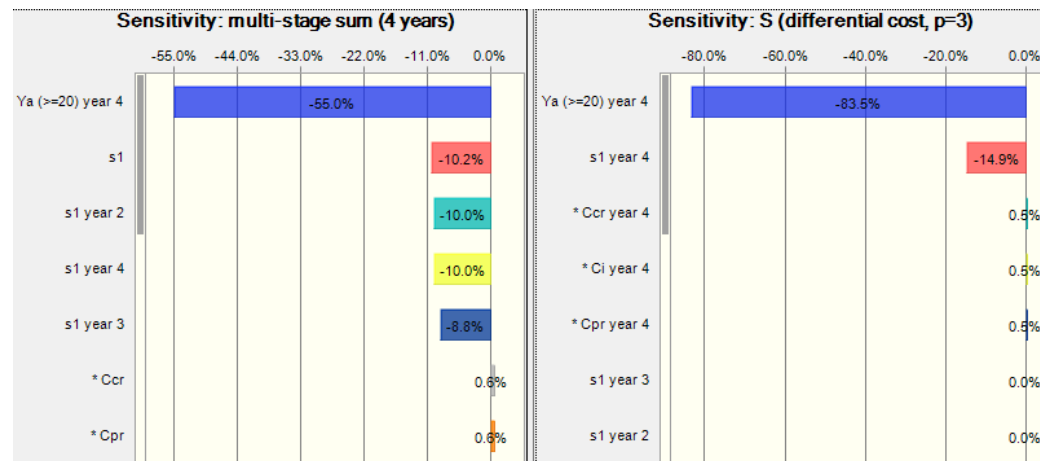


Figure 94 Sensitivity of the sum of three annual differential costs to seven shown random variables using a correlation approach.

5.4.1.3 Optimization results for a low sensitivity diagnostic test: Defining the age-specific failure proportions in prior years as random variables

In the former simulations, the number and function of age-specific failure proportions is assumed to be known prior to each year of optimization. However, at the beginning of the first year of optimization projecting forward to the next year, the failures would be random. The population of the poles has been truncated in the first year to have 80 years as the age of the oldest poles. Thus, between 20 years and 80 years, there are 61 random variables! This, in addition to the other 4 random variables makes a total of 65 random variables. The goal of this new experiment is to find out whether defining the failure proportions per age as individual random variables will affect the optimal decision.

The 729 regression fits of the combinations of log-transformed failure proportions using the mean and one standard deviations of each of 0, 25, 50, 75, 100, and 125 aged poles in subsection 5.1.3.3 are used here. However, instead of finding the number of failures per year from the age distributions, the proportions itself per age are obtained. A probability plot of F(80): the proportion of 80-year old poles expected to fail is shown in

Figure 95. It shows the best fit being for the lognormal or 3-parameter lognormal distributions.

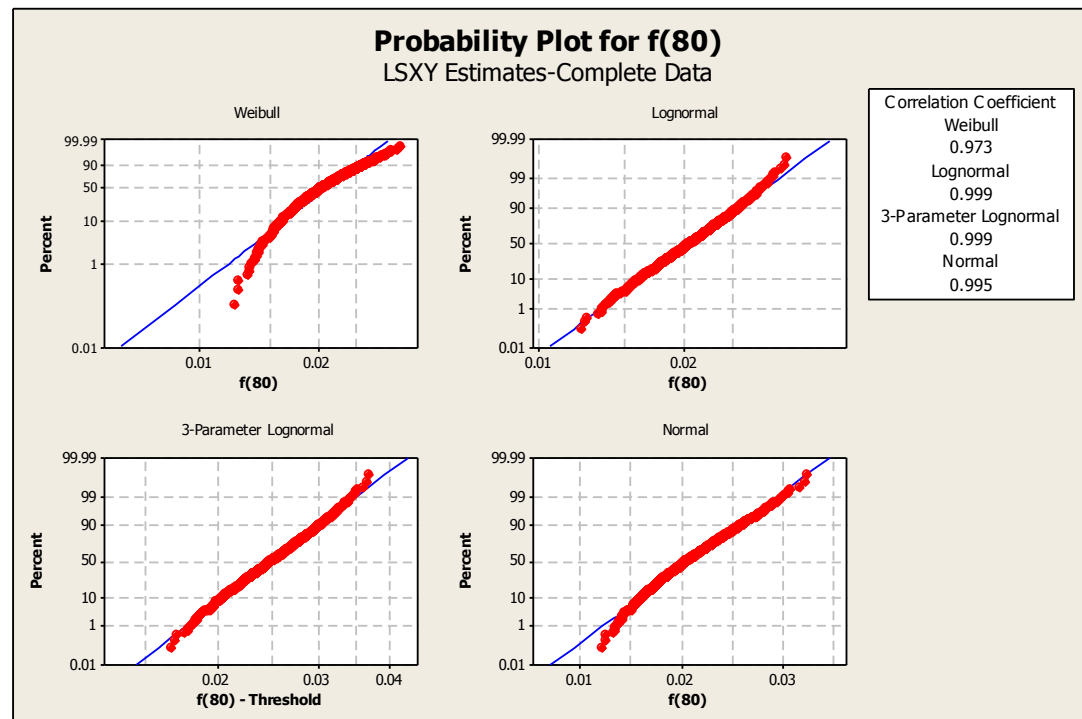


Figure 95 Probability plot of distribution of failure proportions for 80-year old poles.

The goodness-of-fit statistics also show that the lognormal distribution is the best fit for the sample: the Anderson-Darling statistic is very low with a high p-value as shown in Table 40.

Table 40 Goodness-of-fit statistics for probability plots of the distribution of F(80).

Goodness-of-Fit; Distribution	Anderson-Darling (adj)	p-value	Correlation Coefficient
Weibull	13.345	<0.005	0.973
Lognormal	0.336	0.489	0.999
Exponential	512.606	-	-
Normal	1.684	-	0.995

Four instances of the densities of F(t) for ages 59 – 62 are shown in Figure 96.

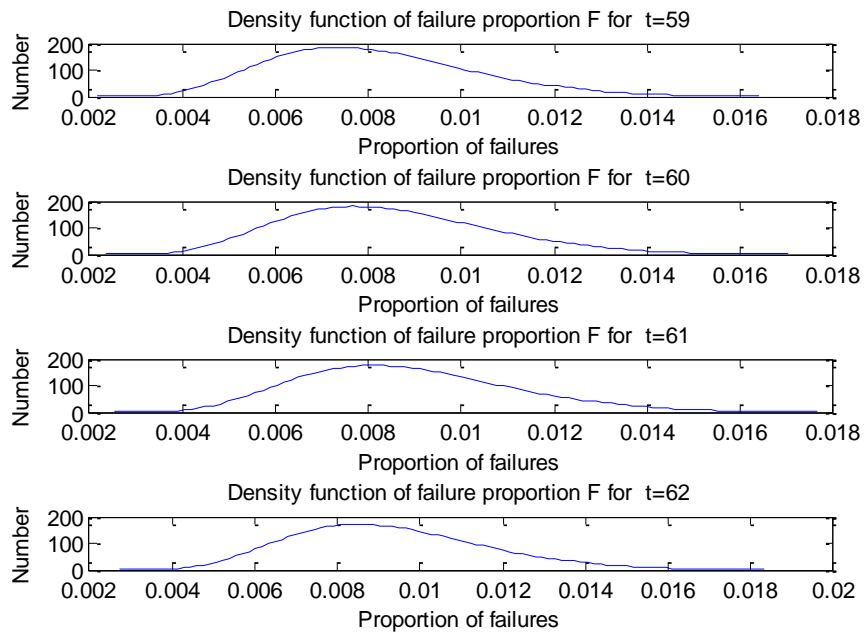


Figure 96 Four instances of distributions of failure proportions for ages 59 – 62.

With randomness in the failure proportions for two years of optimization, an overlay of the resulting distributions is shown in Figure 97.

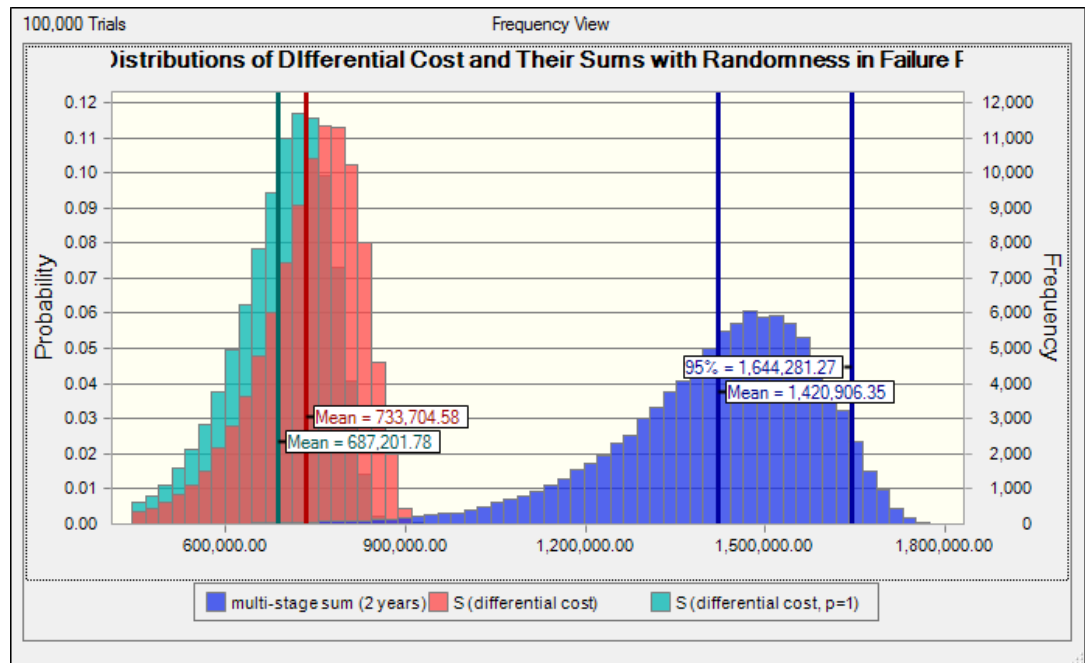


Figure 97 Overlay chart of annual differential costs (left distributions) between competing management approaches for two years of optimization and the distribution of their sums (right).

The sensitivity chart in Figure 98 shows that the annual differential costs and their sum primarily change with the failure proportions.

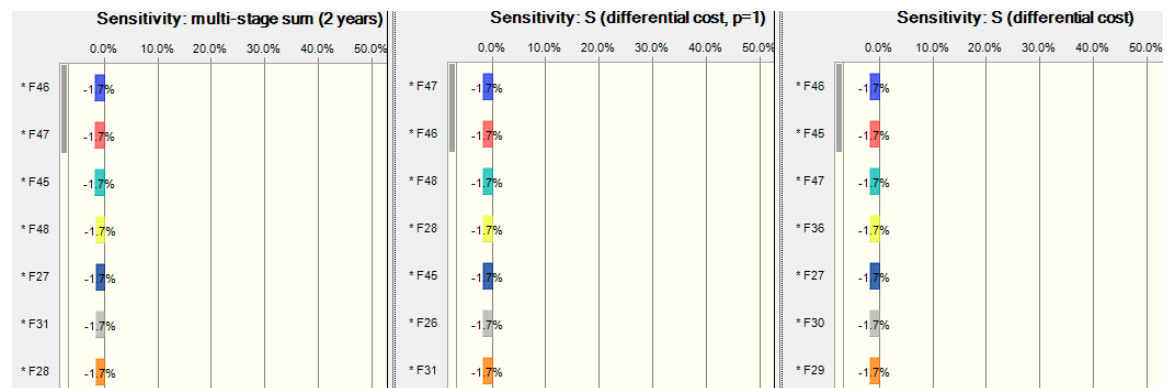


Figure 98 Sensitivity charts of the two-year sum of differential costs, the differential cost in the first year ($p=0$) and the cost in the second year ($p=1$).

The solution is the same at (20 years, 0.9%) with the minimum 95th percentile at \$1,644,281.27. This is almost \$400, 000 less than that of the distribution without the failure randomness in the prior years of simulation. However, with the inspection cycle and proportion of replacements staying the same, running a simulation for additional years may be unnecessary.

5.4.1.4 Summary of results for a low-sensitivity diagnostic test using the multi-year optimization formulation

In general, the optimal solution did not change significantly with additional planning (or budgeting) years. The constraints on the aged population and the expected failure rates relative to the number of components that were candidates for inspection affected the choice of σ : the proportion to replace preventively, but not the choice of the inspection cycle. For a four-year planning period or horizon, the optimal decision and solution for the age distribution of poles implemented is

- 20-year inspection cycle and 0.9% of inspected components replaced preventively per year.

The expected value of the sum of the differential costs between the preventive maintenance program and a run-to-failure program is less than \$2.8 million dollars for all four years. This value excludes any penalties imposed on the power utility for each period of failing to perform a maintenance program. It also excludes lost revenue from energy not served during power outages from failures of the components, and the negative public perception the power utility would receive for not managing its assets adequately. The result provided was for excluding randomness in failure proportions in $(p-1)$ years prior to the length of the optimization planning period. The inclusion of the randomness in the previous experiment showed that the mean may be less when the 60-point randomness is added to the simulation.

The effect of sensitivity or accurate detection of high-risk components is evaluated in the next subsection, where a highly sensitive test is used in optimization.

5.4.1.5 Optimization results for a high-sensitivity diagnostic test

With the distribution parameters of the sensitivity of the diagnostic test applied to the component increased to $(0.85, 1.00)$, the mean costs reduce by about \$1.4 million as shown in Figure 99. The savings are significant.

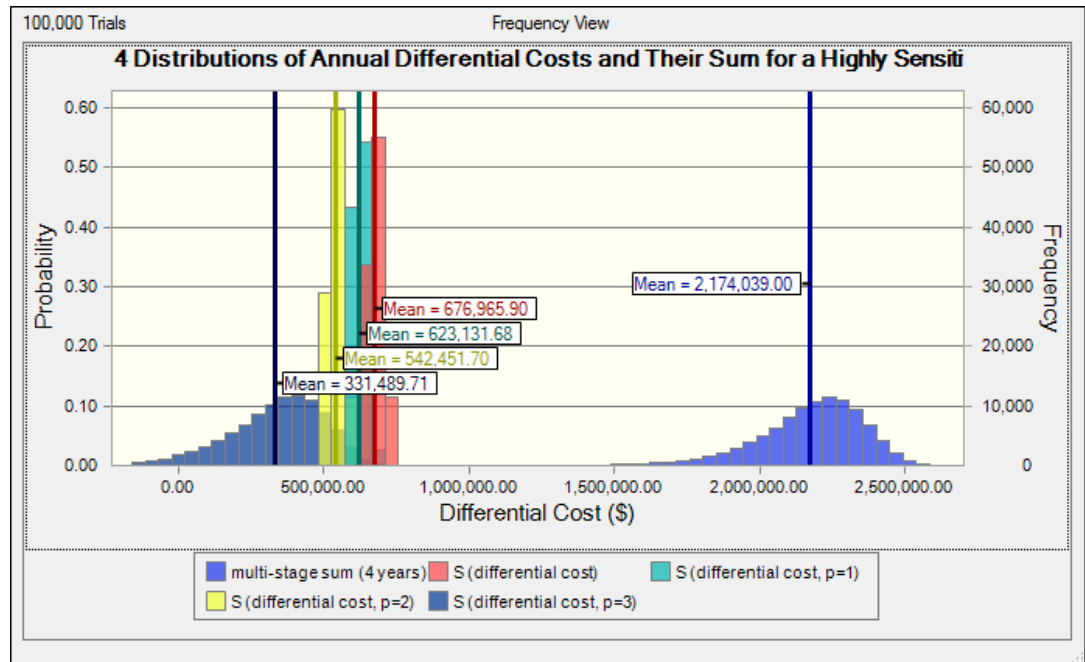


Figure 99 Overlay of distributions of differential costs for a highly sensitive diagnostic test.

The corresponding sensitivity chart for the sum and the differential cost in the fourth year is shown in Figure 100. As is expected, the sum and annual costs change significantly with the number of failures in the fourth year.

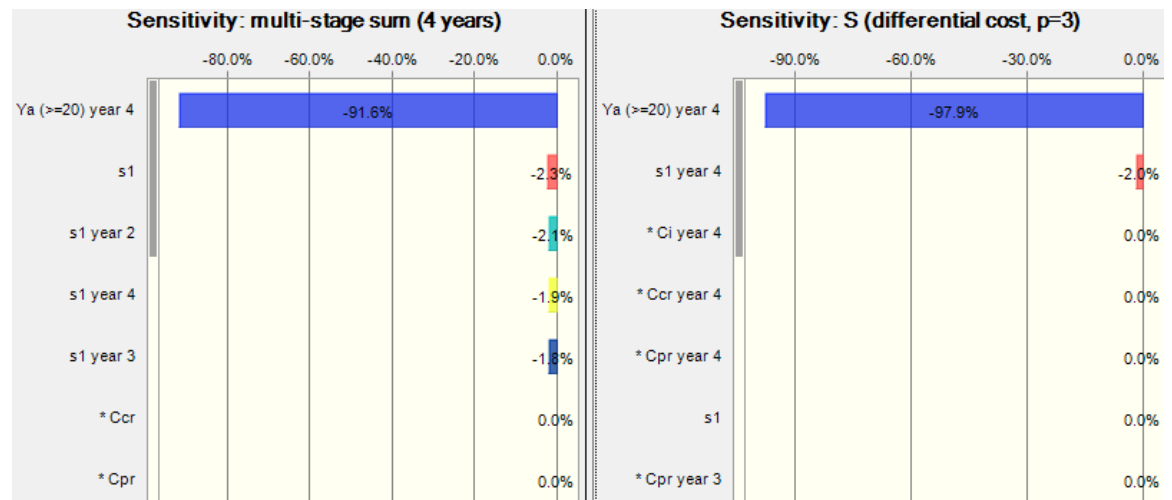


Figure 100 Sensitivity chart of the four-year sum of differential costs.

5.5 Summary

Decisions the electric utility makes in each year on preventive maintenance of their environmentally stressed aging assets can be grouped mainly into the inspection cycle of identical assets and loosely in the proportion of the assets that receive some kind of maintenance. When preventive replacements are the sole or dominant maintenance operation for bringing an asset to as-new condition, the proportion of inspected components that will be replaced preventively in each year becomes an important decision. The importance is both in estimating the number of annual failures averted and consequently annual corrective replacement and failure consequence costs, and also in improving safety conditions of a population of utility assets. The inspection cycle calls for integer programming since in this case, it is given in years. The proportion of preventive replacements on the other hand is not an integer.

The difference between the quantifiable costs in a preventive maintenance process and those of a corrective, more vulnerable process (run-to-failure) depends on unit costs of inspections, corrective replacements and preventive replacements, diagnostic accuracy, and failure estimates. If these parameters were known, optimizing the net benefits of the maintenance process to the run-to-failure (RTF) process would have been deterministic. The problem could have been solved as a number of linear programs since the possible decisions on the inspection cycle are countable. However, uncertainty in the parameters implies the use of stochastic optimization algorithms. These algorithms require that the probability density functions of the parameters are known or at least can be estimated.

With the availability of data, empirical methods may be used to determine the probability density functions. Other methods, like Bayesian approaches, may also be applied to update the functions over time as new information is acquired, given some prior knowledge on the parameters of the density functions. In other cases where

historical data cannot be compiled, subjective judgment like in the estimating distribution functions of costs may be relied upon. Triangular distributions are, for instance, popular for describing costs as shown in prior research. While single data points on unit costs of services on components studied are available from utility surveys, the distributions were estimated by assuming that labor costs in each year are the dominant influences on the annual unit costs. Based on this assumption, historical data on employment cost indices from the United States Congressional Budget Office were used to approximate parameters of the triangular distribution for each service cost.

Mean age-dependent fragility curves of the components under environmental hazards and the probability density function of hurricanes in some state or region of the electric utility may be integrated to estimate the annual age-specific proportions of failures of the components. The cumulative of the product of the proportion and a known age distribution of the components gives the expected number of failures per year. However, using the standard deviations of the fragility curves, other scenarios of failures per age of the components can be explored. Normal probability plots of residuals of regression analyses of log-transformed failure proportions for distinct ages simulated showed that residuals followed a normal distribution. Thus, regression was used for each enumeration of the standard deviations and means of the fragility curves to estimate the distribution of the number of failures per year. A lognormal distribution was found to be the best fit compared to other distributions including Weibull and normal.

The novelty of evaluations of diagnostic accuracy led to the use of a uniform distribution for the uncertainty in accuracy. The bounds of the distribution were obtained from research analysis of accuracy and subjective judgment.

In financial applications, Monte Carlo or stochastic sampling approaches are often used when decisions must be made under uncertainty. The approach is both intuitive and allows for specific and unambiguous scenario analysis. However, as the number of random variables for analysis increases, computational speed and memory sometimes

suffers. Two stochastic sampling approaches were compared to the conventional Monte Carlo technique for selecting samples from the distributions of the random variables. They are Latin hypercube sampling (LHS) and stratified sampling. Analysis showed that both approaches converge very quickly to steady-state, within 3000 samples compared to about 700,000 samples for the conventional method in the experiment conducted in this chapter. The variance and memory requirements of the stratified sampling technique were, to an extent, superior to those of the Latin hypercube sampling technique. The LHS method on the other hand was superior in speed and convergence. The latter reasons led to the selection of the LHS method rather than the conventional stochastic sampling method for stochastic optimization.

5.6 REFERENCES

- [1] A. Collomb, “Dynamic asset allocation by stochastic programming methods,” Ph.D Dissertation, Dept. Mngt. Sci. and Eng., Stanford Univ., Stanford, CA, 2004.
- [2] Federal Emergency Management Agency (FEMA). (2012, June 14). *Wind zones in the United States* [Online]. Available: <http://www.fema.gov/safe-rooms/wind-zones-united-states>
- [3] D. Lee. (2012, Jan.) *Fed expects to keep interest rates near zero* [Online]. Available: <http://articles.latimes.com/2012/jan/26/business/la-fi-fed-forecast-20120126>
- [4] Congressional Budget Office (CBO). (2012, Jan.) *The budget and economic outlook: Fiscal years 2012 to 2022* [Online]. Available: http://www.cbo.gov/sites/default/files/cbofiles/attachments/01-31-2012_Outlook.pdf
- [5] Y. Li and B. R. Ellingwood, “Hurricane damage to residential construction in the US: Importance of uncertainty modeling in risk assessment,” *Engineering Structures*, vol. 28, pp. 1009-1018, Jan. 2006.
- [6] H.C.J. Lai and A.S. Kiremidjian, “Wind hazard analysis in hurricane-prone regions,” Stanford Univ., Stanford, CA, Rep. No. 110, Dec. 1993.
- [7] P. J. Vickery *et al.*, “Simulation of hurricane risk in the United States using empirical track modeling technique,” *Journal of Structural Engineering*, vol. 126, no. 10, pp. 1222-1237, 2000.

- [8] M. D. McKay *et al.*, “A comparison of three methods for selecting values of input variables in the analysis of output from a computer code,” *Technometrics*, vol. 21, no. 2, pp. 239 – 245, May 1979.

- [9] H. A. Steinberg, “Generalized quota sampling,” *Nuclear Science and Engineering*, vol. 15, pp. 142 – 145, Feb. 1963.

- [10] K. Fang *et al.*, “Latin hypercube sampling and its modifications,” in *Design and Modeling for Computer Experiments*, 1st edition, London, UK: CRC, 2005, ch. 2, pp. 47 – 66.

- [11] A. J. Kleywegt and A. Shapiro. (2003, Aug. 24). *Stochastic optimization* [Online]. Available: <http://www2.isye.gatech.edu/so/resources/AntonStochOpt.pdf>

- [12] M. Elkjaer, “Stochastic budget simulation,” *International Journal of Project Management*, vol. 18, pp. 139 – 147, 2000.

- [13] Bureau of Labor Statistics, BLS. (2008, June 17). *How to use the employment cost index for escalation* [Online]. Available: <http://bls.gov/ncs/ect/escalator.htm>

- [14] Bureau of Labor Statistics, BLS. (2012, July 31). *Employment cost index historical listing: Continuous occupational and industry series September 1975 – June 2012* [Online]. Available:

- [15] R. Flanagan *et al.*, “Lifecycle costing and risk management,” *Construction Management and Economics*, vol. 5, pp. 53 – 71, 1987.

- [16] S. Lichtenberg, "New project management principles for the conception stage,"
Journal of Project Management, vol. 7, no. 1, pp. 46 – 51, 1989.
- [17] B. Johnson, "Utility storm restoration response," Edison Electric Institute, U.S.A.,
Jan. 2004.
- [18] KEMA Inc., "Post-hurricane Wilma engineering analysis," KEMA Inc., NC, Final
Rep., Proj. 05-349, Jan. 2006.

6 FUTURE WORK AND CONTRIBUTIONS

6.1 Future Work

This research may be further developed in the following ways.

- The effect of non-replacement (non-terminal) preventive maintenance measures on the fragility or failure risk of *newly installed* components under environmental hazards should be studied.
- The effect of non-replacement preventive maintenance measures on the fragility of *aging* components under environmental hazards would also be an interesting study. Both studies may yield a less costly alternative to preventive replacements.
- Using wood poles as a case study, alternatives to preventive replacement include remedial chemical treatments and steel reinforcements. Historical decay levels (and conditions) of aging wood poles may be derived from inspection databases. Poles that received remedial treatment may show less deterioration than those that had not received treatment. Incorporating such findings into fragility assessment will show whether treatments affect the performance of wood poles under high category hurricanes.
- Depending on the results of the above research, alternative measures of preventive maintenance may be included into a stochastic optimization model. Decision parameters for this potential model should include the proportion and conditions or properties of components to select for each preventive maintenance operation. The optimization may yield benefits to the preventive maintenance program that are superior to the benefits from using preventive replacements.
- In this research, a logical method of selecting components to determine the positive predictive value (PPV) of diagnostic tests was discussed. This

method involves running tests on components scheduled for replacement per year. Over time, the confidence in the PPV estimate improves. A more logical and convincing method for selecting components used in estimating and improving the confidence of the negative predictive value (NPV) of tests should also be found. This may involve testing components that fail subsequent to high category hurricanes. Estimates of NPV could then be generated using prior diagnostic recommendations (replace or not) on the failed components. Time between last inspection and failure from environmental hazards would play a role in the estimate. An appropriate function between the time interval and the failure indicator of the components at the time of hazard impact should be found.

- The effect of incomplete data on the bias of the NPV estimate should also be studied. Incomplete data here refers to information from the components that were not exposed to environmental hazard in each year of severe hazards.

6.2 Contributions

The contributions of this research include the following.

- Statistical analysis of electric utility inspection databases on components.
 - Evaluated discrepancies between reliability function estimated for components using inaccurate and incomplete diagnostic replacement databases and reliability function using possible instances of failure information obtained by applying a stochastic sampling algorithm on diagnostic databases with diagnostic inaccuracy information.
- Modeling failure rates.

- Developed predictive distribution function of lifetimes of surviving components sampled from an exponential dataset by applying Bayesian theory on observed lifetimes assuming a Gamma prior density function on the unknown exponential rate parameter.
- Assessed the effect of missing historical failure information on estimation of the predictive distribution function of lifetimes of surviving components sampled from an exponential dataset using Bayesian statistics.
- Developed a methodology for real-time detection of increasing failure risk of a population of components that follow a bathtub hazard curve by estimating time-varying rate parameters using the maximum likelihood estimation method for a doubly censored exponential dataset.
- Developed age-dependent fragility functions of environmentally stressed power utility components.
- Developed a framework for prediction of age-specific failure proportions of environmentally stressed power utility components in their geographic locations by applying conventional hazard loss method of integrating hazard probability function and fragility functions to age-dependent scenario.
- Predictive maintenance: Diagnostics and preventive replacements.
 - Evaluated costs and benefits of long-term assessment of positive predictive values of individual diagnostic testing methods on utility components. Components used for assessment are logically selected as those scheduled for replacement in each year.
 - Evaluated benefits of ranking diagnostic testing methods using misclassification costs and probabilities rather than the use of strict

statistical approaches that include logistic regression analysis and receiver operating characteristics curves.

- Developed algorithm for combining multiple (up to six) diagnostic testing methods using a combinatorial “OR” rule while simultaneously evaluating the cumulative costs of inaccuracies of the combined methods for ranking purposes.
- Stochastic optimization of preventive maintenance programs on environmentally stressed aging power utility components.
 - Conducted comparative analysis of three stochastic sampling approaches: simple random sampling, stratified sampling and Latin hypercube sampling, in the selection of inputs of five random variables for estimating the distribution of net benefits (or costs) of performing a preventive maintenance program over a run-to-failure program.
 - Developed framework for optimization of expected net benefits (or costs) of performing a preventive maintenance program over a run-to-failure program by selecting best realistic management decisions on inspection cycles and proportion of annual preventive replacements of inspected components under uncertainty for a one-year planning horizon.
 - Developed framework for optimization of expected net benefits (or costs) of performing a preventive maintenance program over a run-to-failure program by selecting best realistic management decisions on inspection cycles and proportion of annual preventive replacements of inspected components under uncertainty for a multi-year planning horizon.

6.3 Publications

The following journal papers are planned for submission prior to or around time of defense.

- A. Shafieezadeh, U. Onyewuchi, M. Begovic and R. DesRoches, “Age-dependent fragility models of utility wood poles in power distribution networks against extreme wind hazards.”
- U. Onyewuchi and M. Begovic, “Detecting the onset of growing failure risk.”

The following conference papers have either been accepted or been submitted awaiting response.

- A. Shafieezadeh, U. Onyewuchi, M. Begovic and R. DesRoches, “Fragility assessment of wood poles in power distribution networks against extreme wind hazards,” ATC-SEI Advances in Hurricane Engineering Conference, Accepted for October 2012 presentation in Florida.
- U. Onyewuchi, A. Shafieezadeh, M. Begovic and R. DesRoches, “A stochastic framework to optimizing net benefits of a wood pole preventive maintenance program,” 11th International Conference on Structural Safety & Reliability (ICOSSAR) 2013, Awaiting acceptance decision.

Other potential journal papers will address

- The influence of missing data on predicting distributions of future lifetimes of surviving components sampled from an exponential dataset,
- The improvement of diagnostic validity and possibly classification accuracy by combining multiple diagnostic tests,
- Stochastic optimization of a preventive replacement program,

- Stochastic optimization of a preventive maintenance program that includes a combination of maintenance operations.

VITA

URENNA ONYEWUCHI

ONYEWUCHI was born in Houston, Texas but spent her infant and adolescent years in the eastern part of Nigeria. She enjoyed Mathematics, Calculus and Physics in high school and looked forward to studying Electrical Engineering in the university. She relocated to Fairfax, Virginia where she received a B.S. in Electrical Engineering from George Mason University with high honors in 2005. Urenna worked at Orbital Sciences Corporation as an assistant ancillary engineer for six months before moving to Georgia to pursue a Master's degree still in Electrical Engineering from Georgia Institute of Technology. While at GA Tech, she was persuaded to stay on for a doctorate degree. She earned her Master's degree in 2008 and interned at a number of companies in the summers, including General Electric and Schlumberger. She will work for Corning after graduating from GA Tech. When not working on research, she enjoys traveling, reading novels, and mentoring young students. She plans on starting a non-profit organization that encourages Nigerians to provide solutions for Nigerian economic, educational and technology issues sometime in the future.