Dynamic Loading of Substation Distribution Transformers:

Detecting Unreliable Thermal Models

and Improving the Accuracy of Predictions

by

Shruti Dwarkanath Rao

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved July 2014 by the
Graduate Supervisory Committee:

Daniel Tylavsky, Chair
Keith Holbert
George Karady

ARIZONA STATE UNIVERSITY

August 2014

ABSTRACT

The existing deregulated market structure for electricity necessitates that utilities make the generation, transmission and distribution of electricity cost-effective. This encourages investment in technological upgrades to utilize the equipment optimally, thus reducing the operation and maintenance costs, while ensuring an extended operational life. This goal can be achieved for a transformer with the help of dynamic loading. Dynamic loading of a transformer implies optimally loading it given available load, cooling and ambient conditions. This can be of significance in maintaining the reliability of the electric supply.

Dynamic loading allows the utility to load a transformer above its nameplate rating for a specified duration of time, such that its service life is not unduly reduced. Over-heating is more often than not the cause behind premature insulation breakdowns and insulation breakdowns often lead to overhaul or replacement of transformers. Hottest-spot temperature (HST) and top-oil temperature (TOT) are reliable indicators of the insulation temperature. The objective of this project is to use thermal models to estimate the transformer's maximum dynamic loading capacity without violating the HST and TOT thermal limits set by the operator. In order to ensure the optimal loading, the temperature predictions of the thermal models need to be accurate. A number of transformer thermal models are available in the literature. In present practice, the IEEE Clause 7 model is used by the industry to make these predictions. However, a linear regression based thermal model has been observed to be more accurate than the IEEE model. These two models have been studied in this work.

This document presents the research conducted to discriminate between reliable and unreliable models with the help of certain metrics. This was done by first eyeballing the prediction performance and then evaluating a number of mathematical metrics. Efforts were made to recognize the cause behind an unreliable model. Also research was conducted to improve the accuracy of the performance of the existing models.

A new application, described in this document, has been developed to automate the process of building thermal models for multiple transformers. These thermal models can then be used for transformer dynamic loading.

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my advisor, Dr. Daniel J. Tylavsky, for giving me the opportunity to work on this project. I sincerely thank him for his invaluable guidance on this research work. His dedication and honesty towards his profession is very inspiring. I consider myself lucky to have had him as a mentor.

I would like to thank my committee members, for taking the time to give their valuable feedback.

In addition, I would like to express my appreciation to Salt River Project, for providing financial support and supplying the data and industrial guidance. I am very grateful to the following engineers at SRP for their important and insightful contributions to this work: Ken Alteneder, Kenneth E. Brown, Jason Gunawardena, Thomas W. LaRose, Gary A. McCulla and Sam Ortega.

Finally, I would like to thank my family and friends, for supporting me emotionally and financially through my educational process. I am deeply indebted for their encouragement and motivation. I will always remain grateful to my parents for the unfaltering love and support they gave me.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

ε	Error term of the linear regression model
θ_{amb}	Ambient temperature (°C)
θ_{fl}	Top-oil rise over the ambient temperature at rated load (°C)
θ_h	Hottest-spot temperature rise over top-oil temperature (°C)
θ_{hu}	Ultimate hottest-spot temperature rise over top-oil temperature for load I (°C)
θ_{hi}	Initial hottest-spot temperature rise over top-oil temperature (°C)
θ_{hr}	Hottest-spot temperature rise over top-oil temperature at rated load (°C)
θ_{hst}	Hottest spot temperature
θ_i	Initial top-oil rise for t=0 (°C)
θ_o	Top-oil rise over ambient temperature (°C)
θ_{top}	Top oil temperature (°C)
θ_u	Ultimate top-oil rise for load I (°C)
σ^2	Variance of the error term in the linear regression model
β	Linear regression coefficient or model parameter
β_j	The j^{th} linear model parameter
Δt	sampling period
τ_{oil}	Time constant for the top oil temperature
C	Thermal capacity (Wh/°C)
C_{jj}	The j^{th} diagonal element of the matrix $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$
FA, ONAF	Oil-natural-air-forced cooling mode

FAFA	Forced-air Forced-air cooling mode
I_{pu}	p.u. load current
I_{rated}	Rated RMS load current (A)
k	Time step index
K_1, K_2, K_3	Coefficients of the linear TOT model
L_{1m}, L_{2m}	Coefficients of the linear HST model
m	Empirically derived exponent for each cooling mode, to approximately account for changes in resistance and viscosity with temperature.
n	Oil exponent
OA,ONAN	Oil-natural-air-natural cooling mode
P_{fl}	Total loss at rated load (MVA)
R	Ratio of load loss to no-load loss at rated load
R^2	Coefficient of Determination
S	Load under consideration (MVA)
S_{rated}	Rated load (MVA)
SS_{Res}	Sum of squares of residuals
SS_{Tot}	Sample variance
SSL_{Max}	Maximum steady-state load
SRP	Salt River Project
τ_{oil}	Thermal time constant at rated KVA (hours)
T_{amb}	Ambient temperature
T	Time duration of load (min)

T_h	Winding time constant at hottest-spot location (min)
V	RMS load voltage under consideration (V)
V_{rated}	The rated RMS load voltage (V)
VIF	Variance inflation factor
x	Predictor or regressor variable in a regression model
$\boldsymbol{X}$	An n by p matrix containing the k regressor variables
x_{ij}	Value of regressor variable x_j at the i^{th} observation
x_j	The j^{th} regressor variable
y	Response variable in a regression model
$\boldsymbol{y}$	An n by 1 vector of the response variable y_i
y_i	Measured value of the response variable at the i^{th} observation

CHAPTER 1

INTRODUCTION

1.1 Introduction

The deregulated market structure requires the utilities to submit their bids in the energy market and the lowest bids are selected to meet the demand. This makes it important for a utility to reduce the total cost of supplying electricity in order to ensure that its bid is low enough to be selected while still supplying electricity to the consumers profitably. Thus utilities desire to use their equipment optimally without unduly reducing their shelf lives. In most utilities, the transformers are rarely loaded to their optimum capabilities. For example, at SRP, transformers are typically loaded only up to 80% of their rated capacities for 90% of the time. Hence the utilities desire to optimally load the transformers, where optimal loading takes into account dynamic loading calculations. Millions of dollars can be saved by even a moderate size utility if their transformers can be loaded even 2% to 3% higher than the limits established using traditional methods. However a tradeoff exists between loading a transformer more heavily to defer capital cost versus prolonging its service life through lighter loads. Dynamic loading of transformers is the term used when the optimum loading capacity is calculated with the help of an appropriate thermal model and taking into account the load magnitude, load shape, thermal limits and external cooling conditions. In order to optimally utilize their substation distribution transformers and, consequently, minimize cost, Salt River Project (SRP) has provided financial support for the development of a software application which can be used by system operators and load specialists to perform dynamic loading of substation distribution transformers for load planning and scheduling. This dynamic loading application is called the operator's tool or the (Dynamic Loading of Transformers Application) DLTA. It ensures that the transformers are optimally loaded without degrading the insulation beyond acceptable limits since insulation breakdown in transformers is often the reason behind transformer failures. In order to delay the overhaul or replacement of a transformer, one needs to ensure that the thermal limits of the insulation are not violated when loading a transformer.

The deterioration of insulation with time and the hottest spot temperature are related by the Arrhenius reaction rate theory and the equation for per unit life is given by [1].

$$Per\,unit\,life = Ae^{\left[\frac{B}{\theta_H + 273}\right]} \tag{1.1}$$

where A, B are constants and θ_H is the hottest-spot temperature.

The primary reason behind the heating of the insulation is no-load and load losses. Thermal models can predict the top-oil temperature (TOT) which is a proxy measure of the insulation temperature and the hottest spot temperature (HST). Both TOT and HST are used as the limiting criteria at SRP to decide the maximum loading capability of their transformers.

Hence the dynamic loading application allows the transformer to be over-loaded until either of the TOT or HST predictions reach their limits. The TOT and HST are dependent on many factors such as the load shape, load magnitude, ambient temperature, thermal limits and external cooling conditions. The external cooling conditions typically refer to the cooling mode in which the transformer is operating, however wind and rain also affect transformer cooling. The transformers considered in this project typically have three cooling modes which are as follows:

- FAFA cooling mode: All fans are on.
- FA cooling mode: Half of the fans are on.
- OA cooling mode: All fans are off.

The transformers considered in this project do not have pumps.

Dynamic loading involves loading a transformer until the TOT or HST limits are reached. The dynamic loading application aids the operator in optimum dynamic loading by predicting the maximum loading that the transformer can sustain without exceeding the TOT and HST limits determined to be acceptable by the operator. Thus it predicts the TOT and HST and calculates the maximum loading that the transformer can sustain before the TOT or HST reach their limits. When a transformer is operating close to its thermal limits, it is expected to be operating in FAFA cooling mode. Thus it is important to have accurate temperature prediction in the FAFA cooling mode. Although typically HST is the limiting criterion, TOT reaches its limit earlier if the load shape is flat. Also it has been observed that sometimes in the OA and FA cooling modes, the TOT is greater than the HST. The cause behind this is the reduced viscosity of oil at lower temperatures which prevents the oil from circulated well and hence the TOT may be higher than the HST at times.

A number of transformer thermal models have been developed and tested in the past. The traditional IEEE Clause 7 model [1] is usually preferred in the industry to predict the TOT and HST, since it only requires parameters which can be easily obtained from the available transformer heat-run test reports. However, it does not accurately account for the dynamic behavior of ambient temperature [2], [3]. The linear regression based thermal model developed at ASU uses measured TOT and HST data to build the models.

TOT measurements are easily available and the HST can also be measured with the help of fiber-optic sensors and fluoro-optic thermometers available today. Thus the linear-regression model can also be implemented easily. In order to have accurate dynamic loading results, it is important to have accurate and reliable thermal model predictions. It has been observed that the linear regression models are more accurate than the IEEE models [3].

1.2 Summary

The dynamic loading application called DLTA requires the detailed thermal models for each transformer to which it is to be applied. These models which include the IEEE Clause 7 model and the linear model constructed using linear regression, are constructed using a separate (as yet unnamed) model building application. The model building application developed as a planning tool in an earlier project, is capable of reading in measured data for a desired transformer, building a thermal model with historical data and giving comments about the quality of the model to the user. The dynamic loading application, DLTA, also known as the operator tool, uses these models to estimate the maximum dynamic load the transformer can sustain without violating its thermal limits. However the model building application previously developed was not automatic and required the models to be built one by one by the user by selecting the suitable options. Consequently, a major task of the project reported upon here is the development of an automatic model building tool.

This research is primarily focused on improving the accuracy of model predictions, identifying the unreliable models and possible causes behind their poor predictions.

1.3 Thesis Outline

The organization of the chapters in this document is as follows. Chapter 2 gives a literature review on the subject. Chapter 3 describes the difficulties involved in determining the cooling mode in which the transformer is operating and methods to detect a V-shaped residual plot. Chapter 4 contains a detailed analysis of the metrics obtained to screen unreliable thermal models. Chapter 5 describes the efforts to improve the performance of the existing models and to make the predictions more accurate. An experiment to test the performance of the linear model in OA and FA cooling modes is described. Possible improvements in accuracy of predictions by using least absolute value based regression instead of least squares method is discussed. Also slight improvements in accuracy of predictions by using incremental and decremental models are discussed. In Chapter 6, the conclusions and scope for future work is provided. The appendix contains a detailed description of the model building application.

CHAPTER 2

LITERATURE REVIEW

The substation distribution transformer dynamic loading application being developed in this project requires the TOT and HST predictions to be accurate in order to ensure that the amount of permissible load that the transformer is subjected to does not cause the insulation temperature to exceed its prescribed thermal limit. A number of thermal models have been studied in the past. The two models analyzed in detail in this document are the IEEE model and the linear model.

2.1 IEEE Models

The thermal models widely used in the industry are the top-oil rise and hottest-spot temperature models given in the IEEE Guide for Loading Mineral-Oil-Immersed Transformers [1], which are commonly referred to as IEEE Clause 7 models in the literature. They are referred to as the "IEEE" models in this document. The IEEE models require certain parameters which are easily obtained from the transformer heat-run test reports. In this section, both the top-oil-rise and hottest-spot temperature models are discussed.

2.1.1 IEEE TOT Model

The IEEE Top-Oil-Rise (TOR) model is defined by the equation

$$\tau_{oil}\frac{d\theta_o}{dt} = -\theta_o + \theta_u \tag{2.1}$$

The equation (2.1), can be solved to yield,

$$\theta_o = (\theta_u - \theta_i)(1 - e^{-(t/\tau_{oil})}) + \theta_i \tag{2.2}$$

where τ_{oil} is the top-oil-temperature time constant given by

$$\tau_{oil} = \frac{C\theta_{fl}}{P_{fl}} \tag{2.3}$$

Thus the transformer TOT is given by,

$$\theta_{top} = \theta_o + \theta_{amb} = (\theta_u - \theta_i)(1 - e^{-(t/\tau_{oil})}) + \theta_i + \theta_{amb} \tag{2.4}$$

where,

$$\theta_u = \theta_{fl}\left(\frac{K^2 * R + 1}{R + 1}\right)^n \tag{2.5}$$

$$K = \frac{S}{S_{rated}} = \frac{VI}{V_{rated} I_{rated}} \tag{2.6}$$

C is the thermal capacity given by

C = 0.0272 (weight of core and coil assembly in kilograms)

+ 0.01814 (weight of tank and fittings in kilograms) (2.7)

+ 5.034 (liters of oil)

If the voltage V is assumed to be constant at rated value, (2.6) can be written as,

$$K = \frac{I}{I_{rated}} = I_{pu} \tag{2.8}$$

Discretization can then be used to convert the differential equations to difference equations which are then used to obtain the model parameters through linear regression. Equation (2.1) can be discretized using the Backward Euler approximation given by,

$$\frac{d\theta_o}{dt} \approx \frac{\theta_o[k] - \theta_o[k-1]}{\Delta t} \tag{2.9}$$

On substituting (2.5), (2.8) and (2.9) into (2.1), we can obtain

$$\theta_o[k] = \frac{\tau_{oil}}{\tau_{oil} + \Delta t}\theta_o[k-1] + \frac{\Delta t \cdot \theta_{fl}}{\tau_{oil} + \Delta t}\left(\frac{\left(\frac{I[k]}{I_{rated}}\right)^2 \cdot R + 1}{R+1}\right)^n \tag{2.10}$$

It was observed that for the 28 MVA substation distribution transformers considered in this project, the top-oil-temperature time constant is on the order of 2.5 hours.

If $R > 1$ and $I_{pu}^2 R >> 1$, (2.10) can be approximated as:

$$\theta_o[k] = \frac{\tau_{oil}}{\tau_{oil} + \Delta t}\theta_o[k-1] + \frac{\Delta t \cdot \theta_{fl}}{\tau_{oil} + \Delta t}\left(\frac{I[k]}{I_{rated}}\right)^{2n} \tag{2.11}$$

Equation (2.11) can be rewritten as follows

$$\theta_o[k] = (1 - K_2)\theta_o[k-1] + K_1\left(I[k]\right)^{2n} \tag{2.12}$$

where

$$K_1 = \frac{\Delta t \theta_{fl}}{\left(\tau_{oil} + \Delta t\right)\left(I_{rated}\right)^{2n}} \tag{2.13}$$

$$K_2 = \frac{\Delta t}{\tau_{oil} + \Delta t} \tag{2.14}$$

However, the IEEE top-oil-rise model considers only the load as a varying factor in determining the top-oil rise over ambient temperature [2], [3]. Consequently, the ambient temperature is simply added to the top-oil rise to get the top-oil temperature. Thus it fails to accurately model the time-domain response of the oil temperature to the time-domain variations in ambient temperature. In addition to this failing, the oil exponent in the model is inappropriately placed [5], [6], [9], [10] and [11]. Thus a number of more accurate thermal models have recently been reported in the literature.

Lesieutre et al. developed an improved version of the IEEE model that better accounted for the dynamics of the ambient temperature, oil viscosity and various types of thermal losses [3]. Swift, Molinski and Lehn developed a TOT model based on heat-transfer theory using an analogy between an electric circuit and a thermal system [5], [6]. It used a current source analogy to represent the heat generated due to load losses and nonlinear resistor analogy to represent the cooling mechanism. Ambient temperature is modeled as an ideal voltage source in this model.

The authors of the papers [10] and [11] used various metrics to compare the thermal models given by [1], [2], [5] and [9]. The metrics used were the eigenvalues, parameter sensitivities, R^2 values, the Variance Inflation Factor (VIF), the maximum steady state load predicted by the model known as SSL_{Max} and residual plots. Some of these quantities are standard metrics used in regression analysis to test the reliability of a linear regression model which will be explained in Chapter 4. The authors came to the conclusion that the linearized top-oil model is the most accurate thermal model of all the models which were compared.

2.1.2 IEEE HST Model

The IEEE HST model is defined by,

$$T_h \frac{d\theta_h}{dt} + \theta_h = \theta_{hu} \tag{2.15}$$

The above equation can be solved to yield,

$$\theta_h = (\theta_{hu} - \theta_{hi})\left(1 - e^{-t/T_h}\right) + \theta_{hi} \tag{2.16}$$

where,

$$\theta_{hu} = \theta_{hr} K^{2m} \tag{2.17}$$

The suggested values of the exponents used in the IEEE model equations are given in Table 1.

Table 1 Suggested Values of Exponents for IEEE Model

Type of cooling	m	n
OA	0.8	0.8
FA	0.8	0.9
Non-directed FOA	0.8	0.9
Directed FOA	1.0	1.0

where FOA stands for forced oil and forced air cooling mode.

However the transformers considered in this project do not have pumps. They are only air-cooled transformers.

Using the Backward Euler approximation given by (2.9), (2.15) can be modified to obtain,

$$T_h\left(\frac{\theta_h[k] - \theta_h[k-1]}{\Delta t}\right) = \theta_{hu}[k] - \theta_h[k] \tag{2.18}$$

Substituting (2.8) and (2.17) into (2.18) and rearranging gives,

$$\theta_h[k] = \left(\frac{T_h}{\Delta t + T_h}\right)\theta_h[k-1] + \left(\frac{\Delta t}{\Delta t + T_h}\right)\theta_{hr}\left(I_{pu}[k]\right)^{2m} \tag{2.19}$$

The winding time constant T_h , can be projected with the help of the resistance cooling curve which is obtained in heat run test reports [1]. The 28 MVA substation distribution transformers studied in this project have winding time constants on the order of five minutes.

The IEEE HST model assumes that, under overloaded conditions, the temperature of oil in the cooling ducts is the same as the temperature of the oil at the top of the tank. However Pierce showed that, during overloads, the temperature of the oil in the winding cooling ducts rises rapidly and exceeds the top oil temperature in the tank [7]. This temperature difference causes the IEEE HST predictions to be lower than the actual winding HST. Pierce then developed an HST model using bottom-oil temperature measurements, a model which accounts for the type of fluid, cooling mode, winding-duct-oil-temperature rise, resistance, and viscosity changes [8]. The challenge of using Pierce's model is that it requires parameters and measurements that are usually not available to most utilities.

The Susa et al. thermal models [9], accounted for the nonlinear thermal resistance of the transformer oil. They used an empirically derived exponent for each cooling mode to account for the variation in oil viscosity and winding resistance with changes in temperature and load. The changes in the time constants due to changes in oil viscosity and variation of loss with temperature are also accounted for.

It has been found that of all these models, the linear models (introduced below) trained on measured field data are the most acceptable thermal models in terms of accuracy and reliability [10], [11]. The advantage of using linear regression based models for both TOT and HST is that they are based on actual field data. Due to this, a lot of factors which vary and cannot be captured in other models, such as inoperative fans/pumps/heat exchangers are captured in these models.

2.2 The Linear Regression Model

A thermal model constructed using linear regression and measured data can be found which accounts for the ambient temperature variations along with the dynamics of the transformer load [10], [11]. This model may be derived using linear regression analysis described in section 2.3. The linear-regression model-building procedure uses measured data to obtain the model coefficients. Provided the measured data quality is good, the linear models account for the thermodynamics of a transformer more accurately than the IEEE model. When the data is not of high quality, the accuracy of model predictions can be improved by using data quality control as shown in [12], thus increasing the model reliability.

2.2.1 Linear TOT Model

The linear TOT model is more accurate than the IEEE TOT model since it models the variations in load and ambient temperature more accurately than the IEEE TOT model. The linear TOT model is governed by the differential equation,

$$\tau_{oil} \frac{d\theta_{top}}{dt} = -\theta_{top} + \theta_{amb} + \theta_u \tag{2.20}$$

Equation (2.20) can be solved to obtain

$$\theta_{top} = \left(\theta_u + \theta_{amb} - \theta_{topi}\right)\left(1 - e^{-t/\tau_{oil}}\right) + \theta_{topi} \tag{2.21}$$

where θ_{topi} is the initial value of TOT.

Using the Backward Euler approximation given by (2.9), and substituting (2.5) into (2.20), we can obtain,

$$\theta_{top}[k] = \left(1 - \frac{\Delta t}{\tau_{oil} + \Delta t}\right)\theta_{top}[k-1] + \frac{\Delta t}{\tau_{oil} + \Delta t}\theta_{amb}[k] + \frac{\Delta t}{\tau_{oil} + \Delta t}\theta_{fl}\left(\frac{I_{pu}^2[k] \cdot R + 1}{R+1}\right)^n \qquad (2.22)$$

Equation (2.22) can be simplified to obtain a linear equation as follows:

$$\theta_{top}[k] = K_1 I_{pu}^2[k] + (1 - K_2)\theta_{top}[k-1] + K_2\theta_{amb}[k] + K_3 \qquad (2.23)$$

The model coefficients K_1, K_2 and K_3 are obtained using linear regression analysis. Since measured data is used to obtain these coefficients, this model is more accurate than the IEEE TOT model if the measure data is accurate.

2.2.2 Linear HST Model

The linear HST model is governed by the differential equation,

$$T_h \frac{d\theta_{hst}}{dt} + \theta_h = \theta_{hu} \qquad (2.24)$$

where

$$\theta_h(t) = \theta_{hst}(t) - \theta_{top}(t) \qquad (2.25)$$

Substituting (2.25) in (2.24), we obtain

$$T_h \frac{d\theta_{hst}}{dt} + \theta_{hst} = \theta_{hu} + \theta_{top}(t) \qquad (2.26)$$

The solution to (2.26) is

$$\theta_{hst} = \left(\theta_{hu} + \theta_{top}(t) - \theta_{hi}\right)\left(1 - e^{-t/T_h}\right) + \theta_{hsti} \qquad (2.27)$$

Equation (2.26) can be discretized using the Backward Euler approximation given by (2.9) to obtain

$$T_h\left(\frac{\theta_{hst}[k]-\theta_{hst}[k-1]}{\Delta t}\right)=\theta_{hu}[k]+\theta_{top}[k]-\theta_{hst}[k] \tag{2.28}$$

This can be further simplified to

$$\theta_{hst}[k]-\theta_{hst}[k-1]=L_{1m}\left(\theta_{top}[k]-\theta_{hst}[k-1]\right)+L_2\left(I_{pu}[k]\right)^{2m} \tag{2.29}$$

where,

$$L_{1m}=\left(\frac{\Delta t}{\Delta t+T_h}\right), L_2=\left(\frac{\Delta t\cdot\theta_{hr}}{\Delta t+T_h}\right)$$

2.3 Linear Regression

Regression analysis is a statistical technique which is used to establish a mathematical relation between a dependent and one or more independent variables. If this relation is linear, the regression thus performed is called linear regression.

A multiple linear regression model that best fits the measured data is given by

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots\beta_k x_k+\varepsilon \tag{2.30}$$

where β_j, $j=0,1,..n$ are the regression coefficients and x_j are the model variables. The variable ε represents the normally distributed error term which has a mean 0 and a constant variance of σ^2.

The regression coefficients are obtained by fitting the model to the measured data. Normally the number of measurements is greater than the number of variables i.e. $n>k$. Let y_i denote the i^{th} observed value of the dependent variable y and x_{ij} denote the i^{th} observation of the independent variable x_j. The errors are assumed to be mutually uncorrelated. The least-squares method can be used to estimate the regression coefficients of equation (2.30) which is rewritten as,

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2.31}$$

where $\boldsymbol{y}$ is an $n\times1$ vector of the observed values of the response variable y, $\boldsymbol{X}$ is an $n\times k$ matrix of the independent variables, $\boldsymbol{\beta}$ is a $k\times1$ vector of the regression coefficients to be calculated and $\boldsymbol{\varepsilon}$ is an $n\times1$ vector of uncorrelated errors.

Thus, the least-squares estimator of $\boldsymbol{\beta}$ can be found by minimizing the function:

$$\boldsymbol{S}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = (\boldsymbol{y} - \boldsymbol{X\beta})^T(\boldsymbol{y} - \boldsymbol{X\beta}) = \|\boldsymbol{y} - \boldsymbol{X\beta}\|_2^2 \tag{2.32}$$

which can be further simplified to

$$\boldsymbol{S}(\boldsymbol{\beta}) = \boldsymbol{y^T y} - \boldsymbol{2\beta^T X^T y} + \boldsymbol{\beta^T X^T X\beta} \tag{2.33}$$

To find the least square, the derivative of (2.33) is equated to 0.

$$\frac{\partial \boldsymbol{S}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\boldsymbol{X}^T\boldsymbol{y} + 2\boldsymbol{X}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}} = 0 \tag{2.34}$$

Thus $\boldsymbol{\beta}$, is obtained as

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} \tag{2.35}$$

The least-squares method is used to obtain the model coefficients. To use this method, (2.23) is rewritten as,

$$Y[k] = K_1 X_1[k] + K_2 X_2[k] + K_3 \tag{2.36}$$

where the k (= 1,2,…N) index in the above equation represents an independent measured value associated with time step k,

$$Y[k] = \theta_{top}[k] - \theta_{top}[k-1] \tag{2.37}$$

$$X_1[k] = I[k]_{pu}^2 \tag{2.38}$$

$$X_2[k] = \theta_{amb}[k] - \theta_{top}[k-1] \tag{2.39}$$

The constant K_1 is representative of the heat generated by the load in time Δt, K_2 is representative of the heat lost to the atmosphere in time Δt, and K_3 is representative of heat generated by no-load losses.

Thus the objective function needed to find the coefficients that minimize $S(\beta)$ is:

$$\hat{K} = \min_k \left\| \left(\theta_{top}[k] - \theta_{top}[k-1]\right) - \begin{bmatrix} I_{pu}^2[k] & \left(\theta_{amb}[k] - \theta_{top}[k-1]\right) & 1 \end{bmatrix} \cdot \begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix} \right\|_2^2 \tag{2.40}$$

The transformer coefficients K_1, K_2, K_3 can be estimated using equation (2.40).

Similarly the model coefficients for the ASU HST model, L_{1m} and L_2 can be obtained by using the equation given below.

$$\hat{L} = \min_k \left\| \left(\theta_{hst}[k] - \theta_{hst}[k-1]\right) - \begin{bmatrix} \left(\theta_{top}[k] - \theta_{hst}[k-1]\right) & \left(I_{pu}[k]\right)^{2m} \end{bmatrix} \cdot \begin{bmatrix} L_{1m} \\ L_2 \end{bmatrix} \right\|_2^2 \tag{2.41}$$

2.4 Model Screening Metrics Used In the Existing Application

An application designed for dispatchers and load specialists has been developed over the years at ASU under the guidance provided by Dr. Tylavsky and the engineers at SRP, called TTeMP. It is also referred to as the “planning tool.” This application performed several functions. It built the thermal models for the desired transformer. It also screened the models based on a number of metrics to determine if the linear regression model produced was reliable and it presented the results to the user. Some of these metrics used by this application are standard metrics used in regression analysis to determine the reliability of a linear regression based model. The metrics used in the application are described below.

2.4.1 Maximum Steady-State Load (SSL_{Max})

Maximum steady-state load is the maximum load to which a transformer can be subjected, without violating the defined TOT or HST limits under steady state conditions i.e., constant load and temperature conditions. Under steady-state conditions, the top-oil temperature remains constant, implying that $\theta_{top}[k+1] = \theta_{top}[k]$. Assuming that TOT is the load-limiting criterion, this constraint can be used in equation (2.23) and, after some simplification, to obtain the maximum steady-state load limited by TOT:

$$SSL_{Max} = \sqrt{\frac{K_2\left(TOT_{Max} - T_{amb}\right) - K_3}{K_1}} \tag{2.42}$$

Similarly if HST is the limiting criterion, the maximum steady-state load is obtained as:

$$SSL_{Max} = \sqrt[2m]{\frac{L_{1m}\left(HST_{Max} - TOT_{Max}\right)}{L_2}} \tag{2.43}$$

As per recommendations by SRP, TOT_{Max} is assumed to be 95°C, HST_{Max} is assumed to be 110°C. The ambient temperature, T_{amb}, is assumed to be the worst case condition of 117°F, to get a conservative value of maximum steady state load, SSL_{Max}.

At present, if the SSL_{Max} obtained is greater than 1.3 p.u., the model is rejected since values of steady state load greater than this are unrealistic. It was also observed that if the

SSL_{Max} is below 1 p.u., the model is unreliable. Thus only those models whose SSL_{Max} are within this range are considered as potentially acceptable.

2.4.2 Variance Inflation Factors (VIF)

It is important for linear regression models that the predictor variables be completely independent of each other. If the predictor variables are correlated with each other, this condition is known as multicollinearity. If the multicollinearity is severe, the model coefficients change erratically in response to minor changes in the data. Thus the model built will be highly sensitive to noisy data. A high degree of multicollinearity will also cause the $\boldsymbol{X}^T\boldsymbol{X}$ matrix to have a large condition number, leading to possible inaccuracy in numerical evaluation of the pseudo-inverse, which is necessary in the calculation of model coefficients. The Variance Inflation Factor (VIF) quantifies the severity of multicollinearity.

If $\mathbf{X}$ is the matrix of the predictor variables, the VIF for the jth parameter is defined as

$$VIF_j = C_{jj} \tag{2.44}$$

where C_{jj} is the diagonal element of the matrix $(\mathbf{X}^\mathbf{T}\mathbf{X})^{-1}$.

The application developed at ASU rejects a model by considering it unreliable if the VIF of a variable is greater than 10. This is because it implies that the corresponding model coefficient is poorly estimated since the predictor variable is highly dependent on one or more other predictor variables.

2.4.3 Coefficient of Determination R^2

The coefficient of determination R^2 determines how well the data points will fit a linear regression model. The coefficient of determination R^2 is defined as,

$$R^2 = 1 - \frac{SS_{\mathrm{Res}}}{SS_{Tot}} \tag{2.45}$$

where SS_{Tot} is the sample variance given by,

$$SS_{Tot} = \sum_i \left(y_i - \bar{y} \right)^2 \tag{2.46}$$

and SS_{Res} is the sum of squares of residuals given by

$$SS_{\mathrm{Res}} = \sum_i \left(y_i - \hat{y}_i \right)^2 \tag{2.47}$$

The closer the value of R^2 to 1.0, the more reliable the model is. In this application, models with $R^2 < 0.95$ are discarded as being unreliable.

Whether using the IEEE model or a linear model, the accuracy of the data used to build the model is critical. Although data quality control algorithms can be used to flag certain types of bad data, it is not possible for the well-known traditional procedures to eliminate all bad data [4]. Thus metrics are proposed in Chapter 4 which can be used to identify unreliable models built from measured data.

2.5 Conclusions

The IEEE models are used predominantly in the industry today since the parameters required to build the IEEE models are easily obtained in the heat run test reports. However, the IEEE model predictions are not very accurate. The ASU linear models account for the ambient temperature variations and other undetected phenomena more accurately since the linear model parameters are obtained from measured data using linear regression. However, sometimes bad data leads to unreliable models. Hence certain metrics such as the SSL_{Max}, *VIF and* R^2 are used to distinguish between reliable and an unreliable thermal models. In addition to the metrics described above, some metrics have been developed that are described in Chapter 4 that can refine the model screening process further.

CHAPTER 3

DIFFICULTIES IN COOLING MODE ESTIMATION AND DETECTING “V-shaped” RESIDUAL PLOTS

The transformers studied in this project typically have three cooling modes:

- FAFA cooling mode (Forced Air Forced Air): All fans are on.
- FA cooling mode (Forced Air): Half the fans are on.
- OA cooling mode (also known as ONAN - Oil Natural Air Natural): All fans are off.

The transformers considered in this project do not have pumps and thus there is no oil circulation. The FAFA cooling mode has all fans on which means that the forced air cooled by radiators increases the rate of oil circulation within the transformers, which prevents the insulation temperature from increasing rapidly.

The parameters used to build the IEEE thermal models are different for different cooling modes. Similarly, the linear models for each cooling mode use different parameters. These parameters are distilled using linear regression techniques operating on measured data corresponding to that cooling mode. If the data used to build a model for any particular cooling mode is mixed with data for other cooling modes, the model built will not accurately represent any of the cooling modes. Thus accurately predicting the cooling mode in which a transformer is operating is very important for building reliable linear and IEEE thermal models.

3.1 Method Used to Determine Cooling Mode

In the transformers studied, the measured HST is compared to a threshold setting and fans are turned on/off accordingly by hardware in the field. Thus, in separating measured data according to cooling mode regime, the HST measured values must be used since fan contactor information is not monitored. Since some transformers do not have fiber optic sensors to measure the hottest-spot temperature, a simulated HST is used to turn on/off fans and we use this simulated HST likewise to determine the cooling mode in which the transformer is operating. The simulated HST at an observation *k* is given by equation (3.1) for the transformers in our sample.

$$\mathrm{SHST}[k] = TOT[k] + Load_{pu}[k] * TGR \tag{3.1}$$

where $Load_{pu}[k]$ is the per unit load at the *k*th observation, $TOT[k]$ is the top-oil temperature at that observation and *TGR* is the rated full scale rise of the HST over TOT of the transformer at rated load.

The transformer fans are set to switch on or off as per certain fixed set points. The terminology used for these set points is as follows:

- AllTurnOnTemp: Temperature at which all fans turn on (Typically 75°C).
- HalfTurnOffTemp: Temperature at which half the fans turn off (Typically 70°C).
- HalfTurnOnTemp: Temperature at which half the fans turn on (Typically 65°C).
- AllTurnOffTemp: Temperature at which all fans turn off (Typically 60°C).

The algorithm used to decide the cooling modes is described in the flow chart given in Figure 3.1.

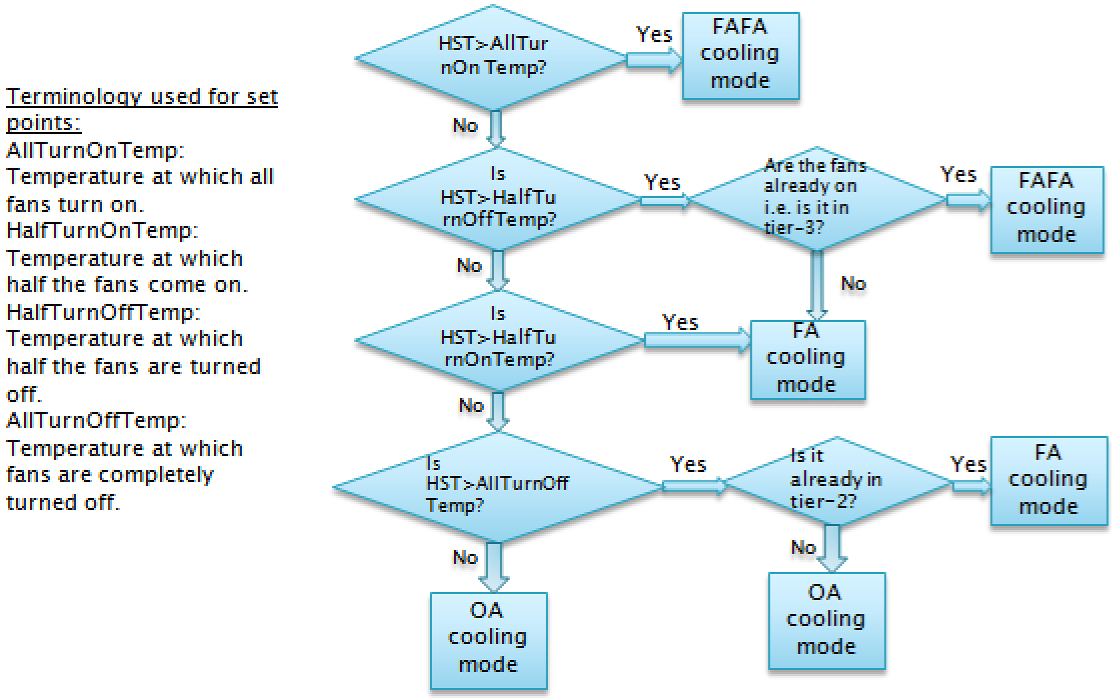


Figure 3.1 Flowchart for Cooling Mode Determination

1. When the HST is greater than the AllTurnOnTemp, all fans are on and the operating cooling mode is FAFA cooling mode.

2. When the HST is below AllTurnOnTemp, above the HalfTurnOffTemp and all fans are on i.e. if the transformer is already in FAFA cooling mode with the temperature dropping due to cooling, the operating cooling mode remains FAFA and all fans continue to remain on until the HST hits the HalfTurnOffTemp. However, if only half the fans are on and the AllTurnOnTemp is not reached yet, the operating cooling mode is FA.

3. If the HST is between HalfTurnOnTemp and HalfTurnOffTemp, the transformer is operating in FA cooling mode.

4. If the HST is between the HalfTurnOnTemp and AllTurnOffTemp and half the fans are on i.e. the transformer is operating in FA cooling mode, it will continue to be in

FA cooling mode. However if all fans are off and temperature has not yet reached the HalfTurnOnTemp, the transformer is operating in OA cooling mode.

5. If the HST is below the AllTurnOffTemp, the transformer is operating in OA cooling mode.

The foregoing discussion represents ideal operation; however, we have come across some issues in accurately determining the cooling mode of a transformer for the purposes of creating data sets for model building: the fan-switching does not always follow the set-points described above. For example, ideally, if all fans are on and the temperature drops below the AllTurnOnTemp due to cooling, the fans should remain on until the temperature hits the HalfTurnOffTemp. However we have found cases where half of the fans turn off sooner than the above algorithm would predict. This may occur due to the set-points not being followed. Similarly when the transformer is in FA cooling mode, if the HST drops below the HalfTurnOnTemp due to cooling, the fans are supposed to remain on until the AllTurnOffTemp is reached. In reality however, it is not certain that the fans actually remain on until the AllTurnOffTemp is reached. Thus the measured-data cooling-mode assignment may be in error. Since the fan status is not monitored, there is no way to know if our assumptions about the operating cooling mode are correct which is very important for both segregating data according to cooling mode and changing models when the dynamic loading calculations range over several cooling modes. If we know the fan status accurately, we can get a better idea about how accurate our cooling mode transition-point estimations are and how the fans are actually switching on and off. Since this irregularity in the fan status is a nonlinearity, we believed that it could be the cause of the “V-shaped” residual plots observed for some transformers. For the definition of residual plots and how they are used, refer to [13]. These details are also provided in section 4.1.2.

It was observed that some unreliable transformer thermal models had V-shaped residual plots even when those plots are taken from the training data sets as seen from the example in Figure 3.2.

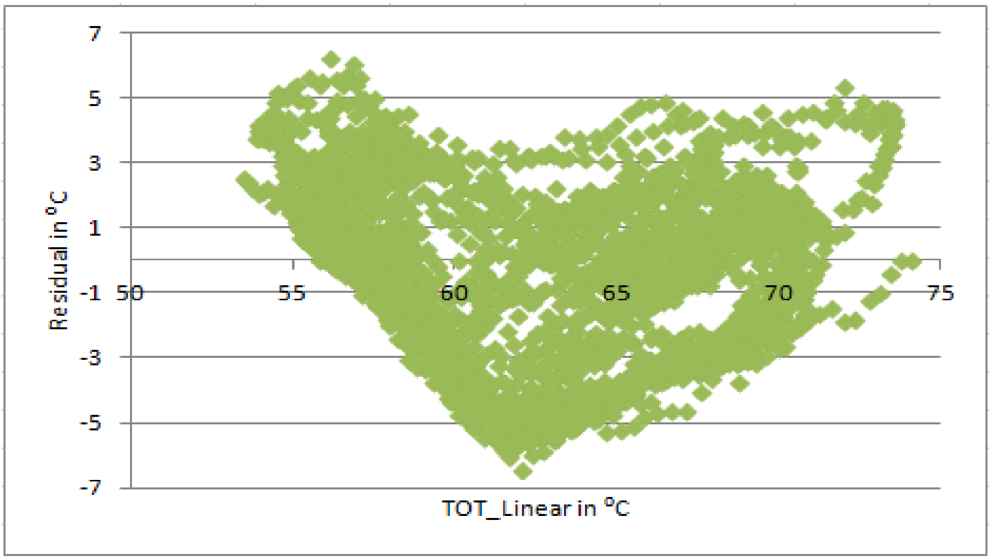


Figure 3.2 Residual Plot for an Unreliable Model for Transformer Highline-3

In order to investigate if inaccurate fan status was the reason behind the V-shaped residual plots, the cooling-mode-separation set-points were adjusted to exclude data near the cooling mode transition points. The resulting residual plot showed significant improvement as seen in Figure 3.3. It can be seen from this plot that the magnitudes of the errors have reduced and they now lie in a range of +/- 3^{o} C as compared to the range of +/- 7^{o} C observed earlier. Also the V-shape is greatly reduced. Similarly the RMS error is reduced from 1.63^{o} C to 0.4^{o} C. In an opposite experiment, the V-shape of some of the transformers' residual plots became more prominent when more FA cooling mode data was introduced in the FAFA model by changing our assumptions about the set-points, that is, assuming the set-points were lower.

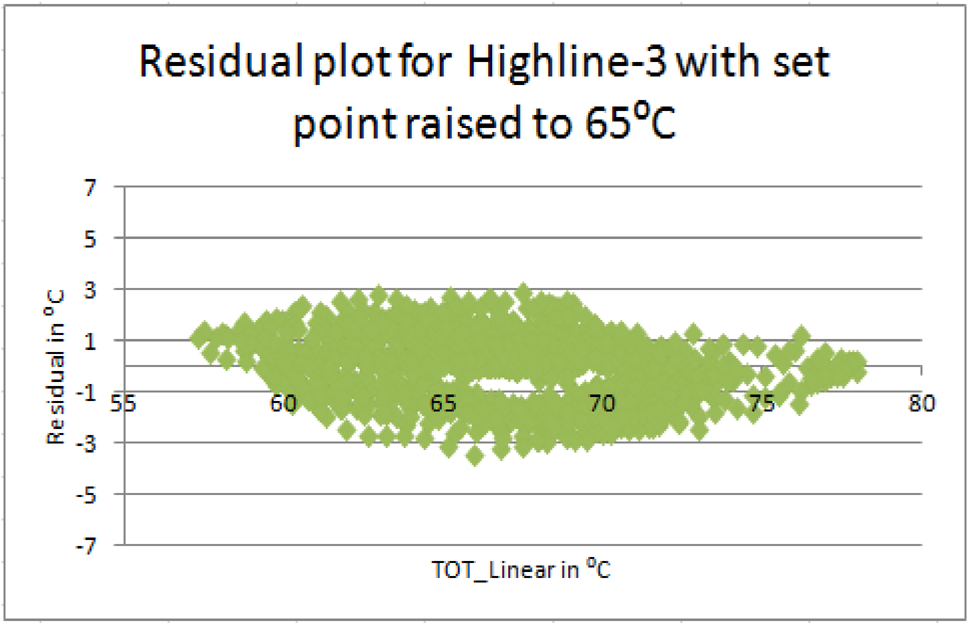


Figure 3.3 Residual Plot on Removing the Suspicious Data Points from the Input

Thus we strongly believed that the reason behind the nonlinearity indicated by the V-shaped residual plots was that fan-switching was not following the desired set-points which may be identified by detecting the V-shaped residual plots at the model building stage.

3.2 Detecting V-shaped Residual Plots

A number of methods were investigated to perform automatic identification of V-shaped residual plots.

3.2.1 Nonlinear Regression

It was postulated that if a V-shape was present in the residual plot then the residual behavior as a function of temperature would have a good fit to a quadratic curve.

Nonlinear regression was used to fit the lower boundary of the residual plots to a quadratic equation given by (3.2) and the coefficients of the model were studied.

$$f(x) = ax^2 + bx + c \tag{3.2}$$

It was observed that models with V-shaped residual plots had coefficients of the squared term greater than 0.02, coefficients of the linear term lower than -5 and the constant term greater than 100. This can be seen from Figure 3.4, Figure 3.5 and Figure 3.6 respectively.

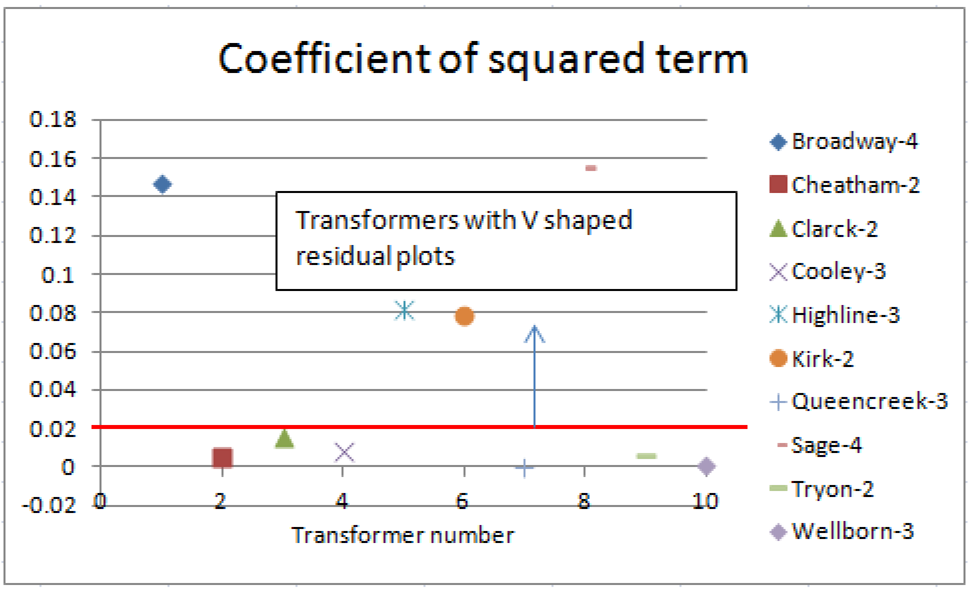


Figure 3.4 Coefficient of the Squared Term Obtained from Curve-Fitting

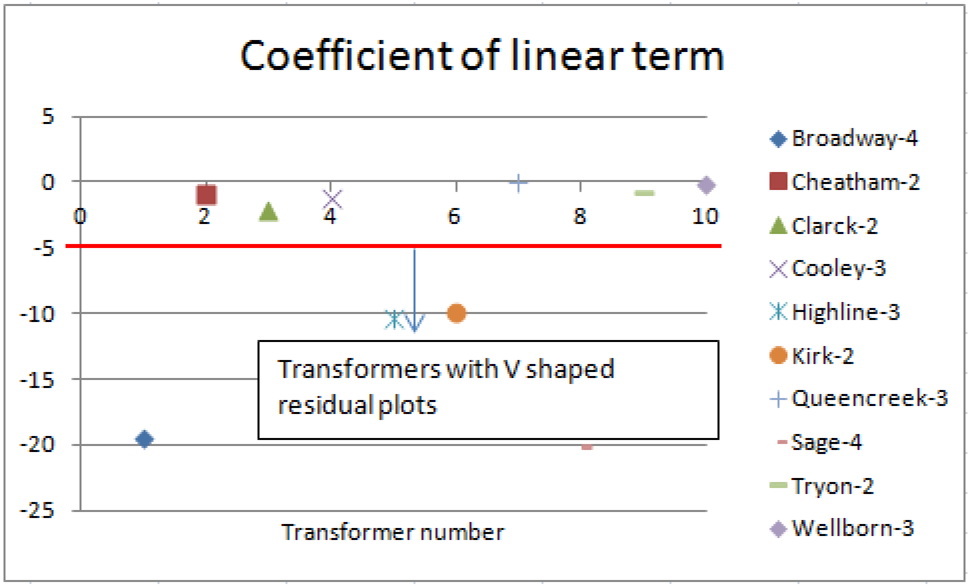


Figure 3.5 Coefficient of the Linear Term Obtained from Curve-Fitting

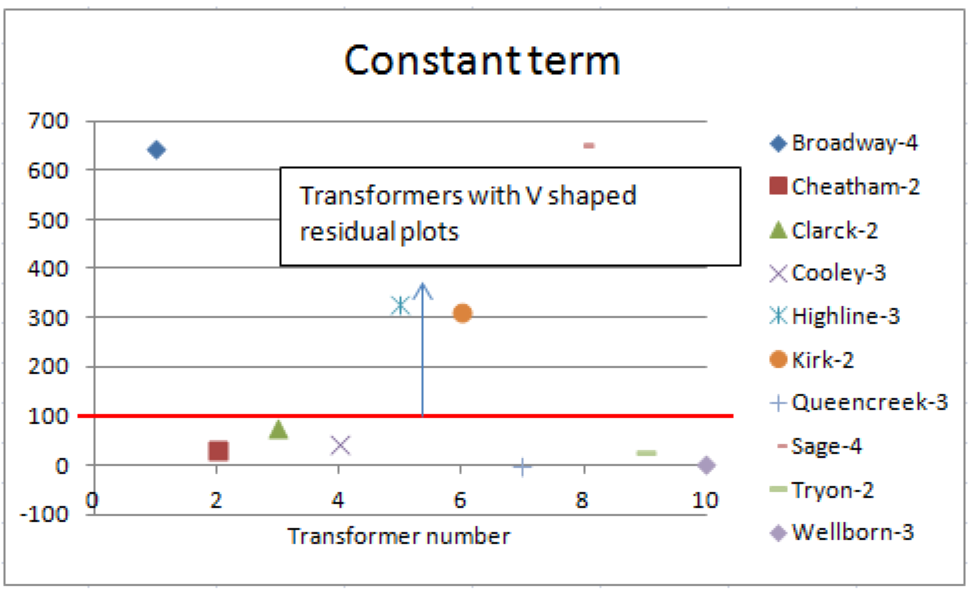


Figure 3.6 Coefficient of the Constant Term Obtained from Curve-Fitting

3.2.2 Standard Deviation of the Predicted Maximum Steady State Load

A second method postulated for detecting V-shaped residual plots was the reliability with which multiple models built using random sampling of the data, would predict the maximum steady state load, SSL_{Max}. Bootstrapping was used to construct thousands of samples from an independent and identically distributed population of the observed dataset and then the sampled datasets were used to build HST and TOT models of the transformers. Each of these models was used to predict the maximum steady-state load sustainable by the transformer and the standard deviations of these predictions were compared for the transformers in our data sets. It was observed that the standard deviations of the maximum predicted steady state loads for TOT from bootstrapping was greater than 1.0 p.u. for the transformers with V-shaped residual plots as seen from Figure 3.7.

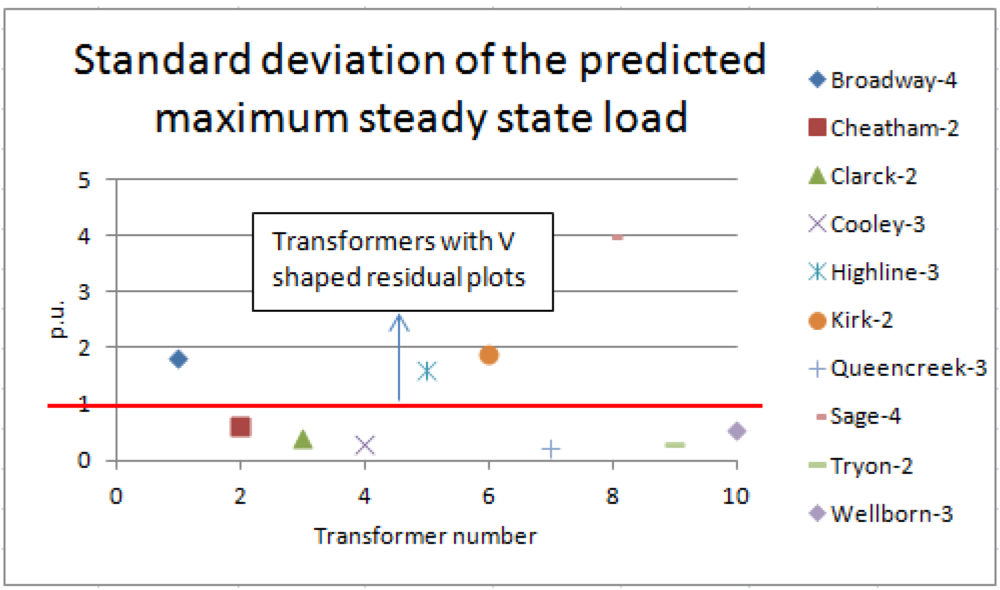


Figure 3.7 Standard Deviation of the predicted SSL_{Max}

3.2.3 Maximum Confidence Interval

The third metric used to identify V-shaped residual plots was the confidence interval, at a given confidence level, in the maximum steady-state load prediction calculated above. The confidence interval of the maximum steady state load defines a region around the mean value of the p.u. steady state loading predicted by the models built from the bootstrapped samples. For our purposes, it is interval centered at the mean predicted SSL_{Max}, such that there is a 95% probability that the true steady-state load rating is within that interval. It was observed that transformers with V-shaped residual plots had maximum confidence intervals (corresponding to the bootstrapped samples built with a data-set size of one day) greater than 0.3 p.u. as seen from Figure 3.8.

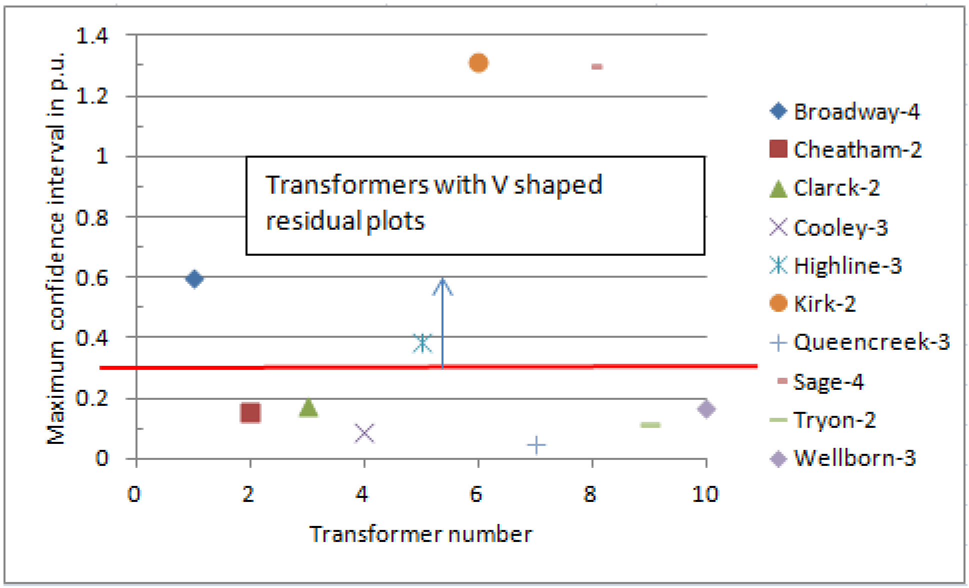


Figure 3.8 Maximum Confidence Interval Plot

3.2.4 Summary

All the three methods were able to successfully identify the models with nonlinearities such as improper fan-switching that were not captured by data segregation algorithm. The application being developed by the authors uses the nonlinear regression method and the standard deviation of the predicted SSL_{Max} since these methods provide a larger gap between the two categories of transformers i.e. with and without the V-shaped residual plots.

In the existing application, if the coefficient of the squared term for a transformer TOT model is greater than 0.02 or if the standard deviation of the predicted SSL_{Max} is greater than 1.0 p.u., the value assigned to its residual plot quality is 5, else it is 10. This residual plot metric is used to evaluate the 'Model Quality' described in Chapter 4.

3.3 Source of the V-shape

In this work, it was important not only to identify and classify as erroneous any models that had V-shaped residual plots but to also find the cause of the V-shape. Using the erroneous fan status as a working hypothesis, we examined the thermal behavior of the Highline-3 transformer. Figure 3.9 provides the measured and predicted TOT values for Highline-3 for the portions of five days when the transformer was in FAFA cooling mode.

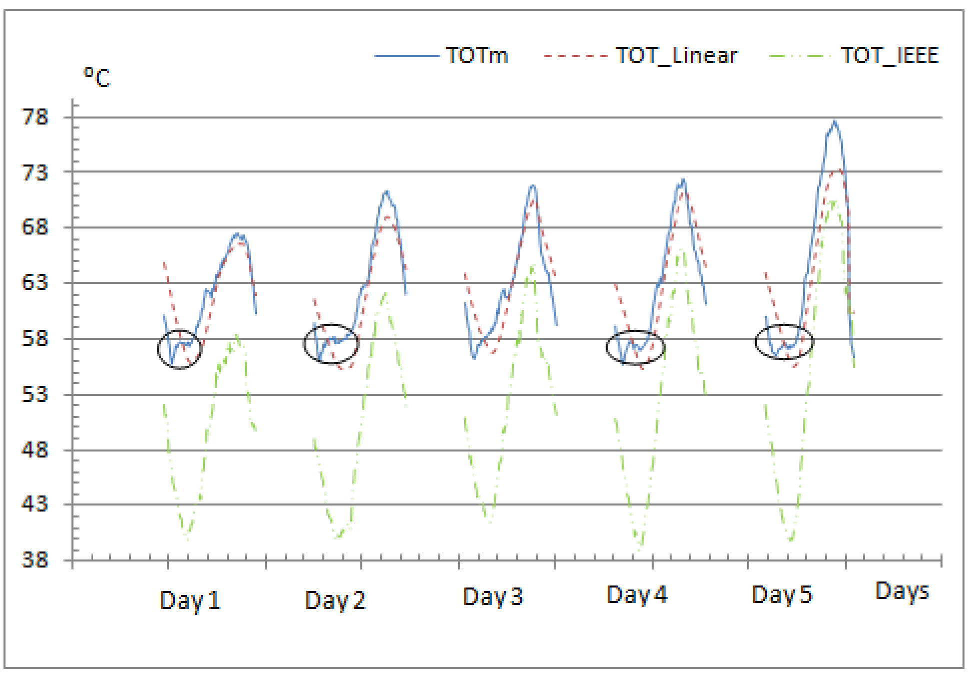


Figure 3.9 Plot of TOT versus Time for an Unreliable Model

The large errors at the peak and the phase shift of the predicted temperature response in this figure, shows that neither the IEEE nor the linear model are reliable. A sudden rise and fall was noticed at the troughs of the measured TOT as shown circled in Figure 3.9. This sudden rise/fall is inconsistent with the recorded load and weather behavior during those time intervals. One cause consistent with this behavior is a cooling mode change/error brought on by fan status that is inconsistent with the fan set points and simulated HST as calculated using (3.1). One of the days on which such a fall and rise was observed at the troughs was analyzed to find the cause of this behavior. The HST and load data for Highline-3 for 12/7/2009 are shown in Figure 3.10 and Figure 3.11, respectively. To understand the information this data contains, it is necessary to understand that Highline-3 does not have an FA cooling mode as per its set points; it has only FAFA and OA cooling modes and it uses simulated HST, rather than measured HST, as the input to its fan controller circuits. All fans are set to turn on when the HST rises above 75^{o} C and turn off when the HST drops below 60^{o} C.

As seen in Figure 3.10, on 12/7/2009, for the Highline-3 transformer, the simulated HST was observed decreasing from midnight for approximately two hours and then began increasing for approximately two hours, starting at point B in this figure even though the load was continuously decreasing as seen from Figure 3.11. The HST value resumed decreasing starting at point C for nearly three hours. Starting at point D, it began increasing for approximately the next ten hours. The relatively short duration for which the simulated HST was increasing (interval B-C) during a relatively steady or decreasing load period, is inconsistent with the transformer remaining in the FAFA cooling mode at point B; however, based on the simulated HST values (equation (3.1)) shown in Figure

3.10 and cooling mode-transition settings, the cooling mode should have remained in FAFA cooling. The most likely explanation of this behavior is that the transformer entered OA cooling mode at point B and then all of the cooling fans turned on (returned to FAFA cooling) at point C causing the temperature to decrease until, at point D, due to increasing load, the HST began increasing again and continued to do so throughout the afternoon which is the peak load period. This explanation of behavior is consistent with thermodynamic principles.

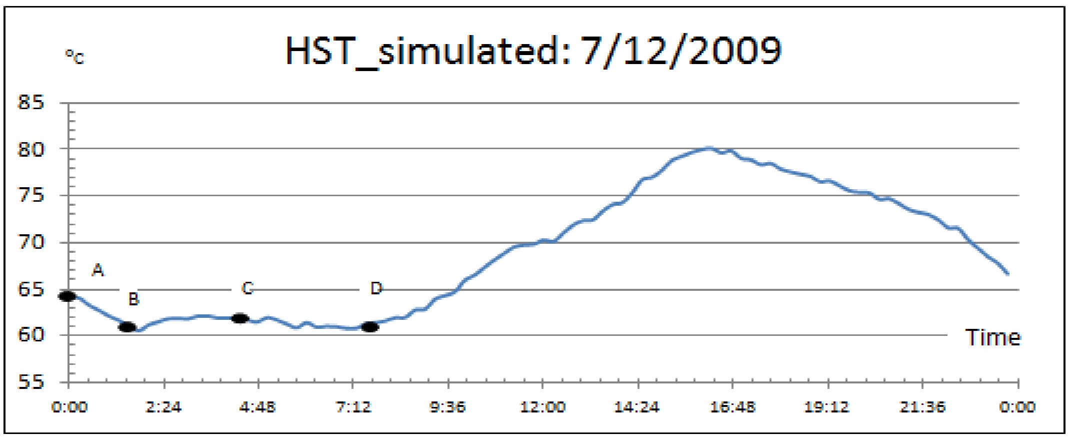


Figure 3.10 Simulated HST Data for Highline-3

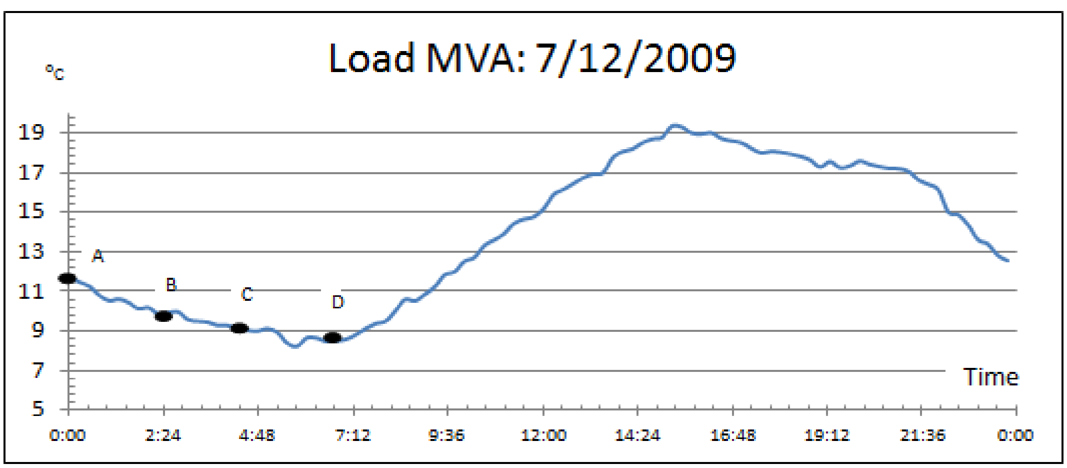


Figure 3.11 Load Data for Highline-3

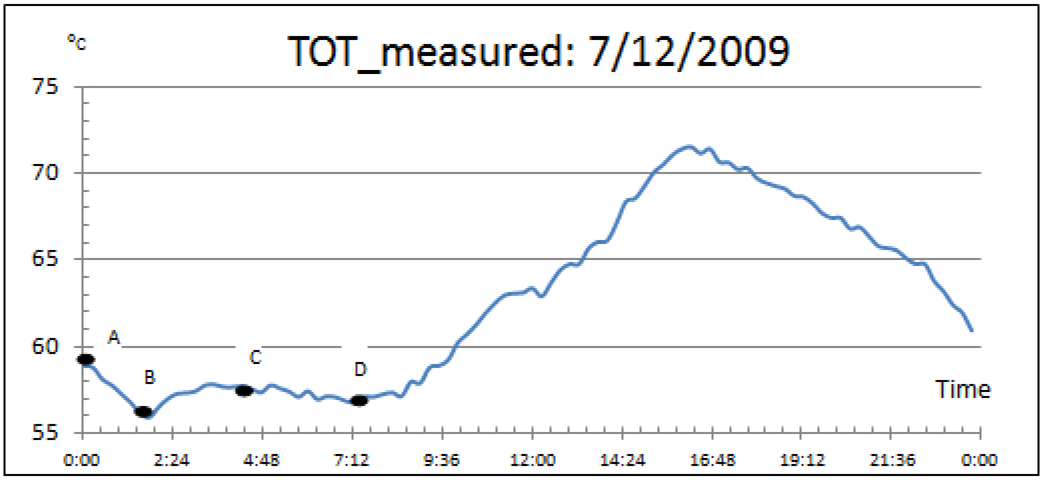


Figure 3.12 Measured FAFA TOT Data for Highline-3

Recorded pseudo-measurements of the HST (which are PMHST values “calculated” by the transformer’s controller) were obtained for Highline-3 for this period of time. A difference was observed between the simulated HST obtained using (3.1) and the pseudo-measured (simulated) HST recorded and shown in Figure 3.13. It can be seen that about 5 AM, the pseudo-measured HST drops below 60° C which is the fan turn-off set point for this transformer. The pseudo-measured HST did not cross the fan turn-on set point of 75°C until nearly 2 PM. Thus if we were to use the pseudo-measured HST, the fans would enter OA cooling mode at 5 AM and would not enter FAFA cooling mode until about 2 PM. This confirms that the transformer enters OA cooling mode, although the time during which it is in the OA cooling mode is not consistent with that deduced in the previous paragraph. However, since the fans may not follow the set-points exactly, it is quite plausible that the transformer entered OA cooling mode at point B and the rest of the thermal behavior is consistent with thermodynamic principles.

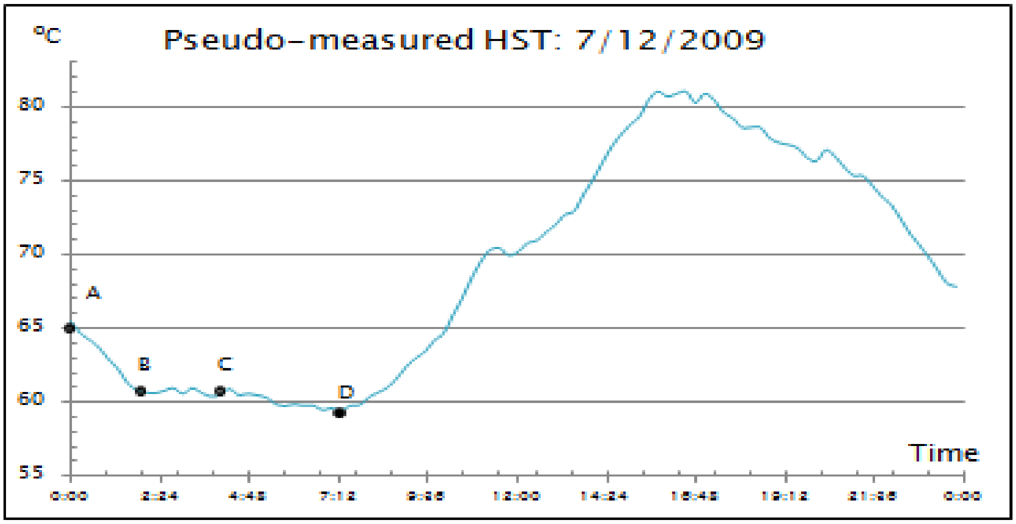


Figure 3.13 Pseudo-Measured HST Data for Highline-3

The sudden rise and fall at the troughs in the measured TOT data in Figure 3.12 can be seen to occur on a consistent basis for this transformer as shown in Figure 3.9. This means that the data contained in Figure 3.2 is a mixture of OA and FAFA cooling mode data. As stated previously, the reason that using this mixture of OA and FAFA cooling mode data produces an erroneous model is that each different thermodynamic condition must be modeled using a different set of model coefficients. By including data belonging to two different cooling modes in one thermal model, the linear regression model is being forced to fit two different thermodynamic conditions, which appears like a nonlinearity in the data. The linear thermal model is not able to capture this non-linearity and hence the thermal models are unreliable for either of the cooling modes. This understanding of the mixture of data in model building also explains why excluding data near the cooling-mode transition boundaries significantly reduced the V-shape in the residual plots. The linear TOT model was built assuming that the fans turned off when the simulated HST dropped below 65^{o} C, thus removing the set of data points from the FAFA model where the transformer entered the OA cooling mode for a few hours. With this data-screening mechanism in place, points in the time domain plot associated with the sudden rise and drop of measured TOT at the troughs were removed eliminating the effective nonlinearity.

Thus the source of the V-shape was believed to be inaccurate fan status.

3.4 Summary

It was observed that inaccuracy in determining the fan status during the cooling-mode-assignment of measured data points can lead to the thermal models whose predictions are inaccurate. The identification of the V-shaped residual plot is used as one of the measures to determine the level of reliability of the model. Possible methods to detect the linear TOT models with V-shaped residual plots were identified. Also the source of the V-shaped residual plot was analyzed. Inaccurate fan status was believed to be the cause of the V-shape.

CHAPTER 4

METRICS FOR MODEL SCREENING

The transformer thermal model building application (named TTeMP) developed as a planning tool at ASU in an earlier project, builds two kinds of models for TOT and HST for all transformers: the IEEE models and the ASU linear models. It analyzes the reliability of the models based on the metrics described in Chapter 2 and provides the linear model coefficients, parameters for the IEEE model, error duration curves, load duration curves, probability density plots and comments about model reliability for both models to the user for TOT as well as HST. However, the metrics discussed in Chapter 2 are insufficient to successfully discard all unreliable models. Some additional metrics for the linear TOT model have been evaluated that will be described in this chapter. These metrics have been developed in order to obtain an efficient model screening process and discard the linear TOT models with poor prediction accuracies.

4.1 Possible Metrics for Model Screening

In order to develop the metrics for the linear TOT models, the model quality was qualitatively and quantitatively assessed with the help of visual inspection and the quantitative metrics described in this section. Although a lot of the candidate screening metrics were good at discriminating between reliable and unreliable models, some were more efficient than others. A systematic method to detect unreliable models is described in section 4.2. The results discussed in this section, have been obtained for 28 MVA substation distribution transformer TOT models *only*.

4.1.1 Visual Inspection

Visual inspection was the first step to judge whether a model predicts the TOT accurately. Quantitative metrics were developed based on these observations. Visual inspection involved plotting the measured TOT, the linear model TOT prediction and the IEEE model TOT prediction on the same axes for the period during which the transformer was operating in FAFA cooling mode and then judging the accuracy of the predictions.

Figure 4.1 shows the TOT plots for five high-temperature days during which the transformer was in FAFA cooling-mode. In this figure, TOTm represents the measured TOT, TOT_Linear represents the linear model prediction and TOT_IEEE represents the IEEE model prediction. The predictions shown in this figure correspond to a thermal model we believe to be reliable, for the Queencreek-3 transformer. As seen from the plot, the linear model predictions track the measured TOT very accurately. The IEEE model predictions are not as accurate as the linear model but still considered acceptably accurate.

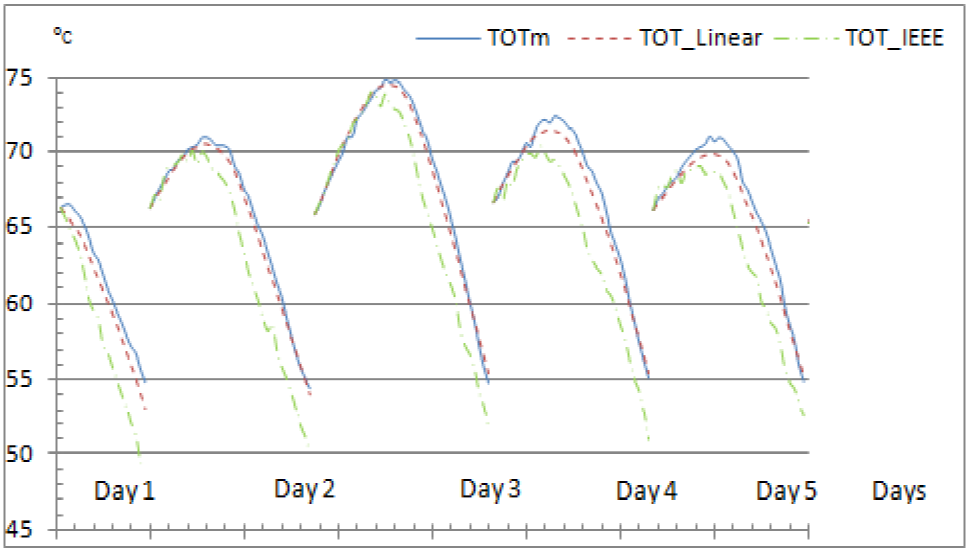


Figure 4.1 Plot of TOT versus Time for a Reliable Model

Figure 4.2 shows the measured and the linear model and IEEE model TOT predictions for five high-temperature days during which the transformer was in FAFA cooling-mode for the transformer Highline-3. The thermal models for this transformer are considered to be unreliable. It is observed that at higher temperatures, the linear model predictions are nearly 5^{o} C lower than the measured TOT. Additionally, there exists a significant phase shift between the measured and predicted TOT from both models. Neither the linear model nor the IEEE model is considered reliable. The IEEE model is less accurate as compared to the linear model and this has been observed in general for all transformers. The sudden rise and fall in the measured TOT at the troughs as shown circled in Figure 4.2 is worth noting. Its significance and possible cause has been explained in Chapter 3 where the difficulties in cooling mode determination are discussed.

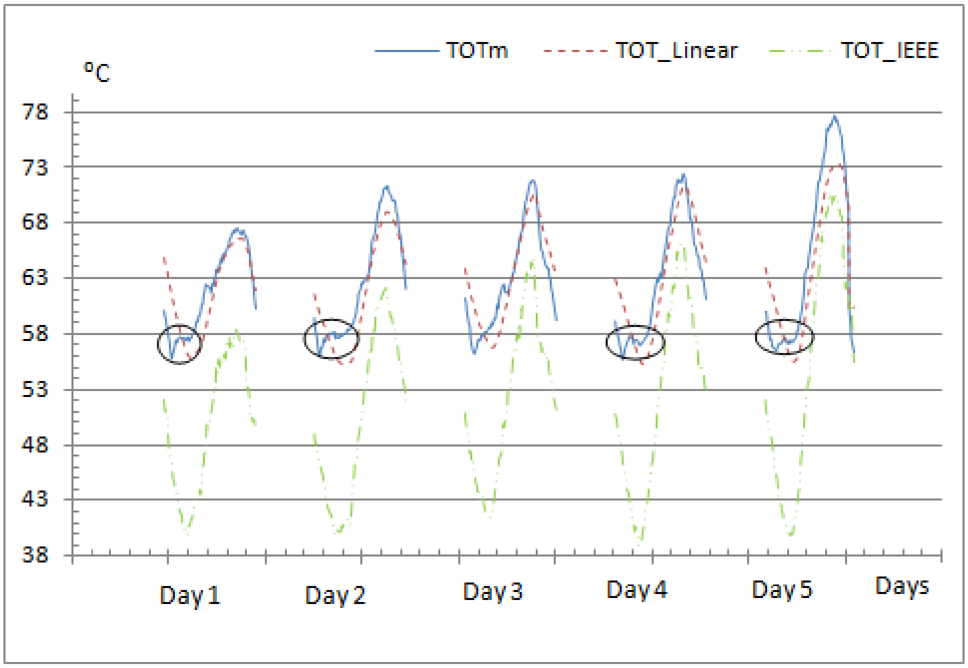


Figure 4.2 Plot of TOT versus Time for an Unreliable Model

Visual inspection was not used as a metric to discard unreliable models. Instead, it was used as a sanity check to validate the classification rules involving the metrics described below. The metrics have been developed only for linear models so far and similar research on IEEE models is anticipated as future work.

4.1.2 Residual Plots for Linear Model Acceptability Determination

An important metric for determining the acceptability of a model is a residual plot. A residual plot is a graph of the residual, e_i, versus the corresponding fitted (predicted) values $\hat{y}_i$. The residual, e_i, is given by (4.1)

$$e_i = y_i - \hat{y}_i, \quad i = 1,2,\ldots,n \tag{4.1}$$

where y_i represents measured values.

For a model to be considered reliable, the points in a residual plot should be confined to a horizontal band. Figure 4.3 shows the residual plot for the transformer Queencreek-3. Queencreek-3 has a reliable thermal model which can also be verified from the accurate predictions in the time domain plots given in Figure 4.1. However if the residual plot is V-shaped, this indicates that there is some non-linearity in the process which the model is unable to capture leading to inaccurate predictions. The residual plot for Highline-3, shown in Figure 4.4, has a V shape. It can be verified from the time domain plot for this transformer given in Figure 4.2 that the model predictions are not accurate.

The residual plot for Queencreek-3 shows that the errors remain in a bounded range of +/- 2^{o} C for all temperatures. Whereas the residual plot for Highline-3 shows that the linear model underpredicts the TOT at lower and higher temperatures whereas it overpredicts the TOT in the middle ranges of temperature. Thus the errors do not have a consistent behavior for this thermal model. Also the error magnitudes are very high and lie in the range of +/- 7^{o} C.

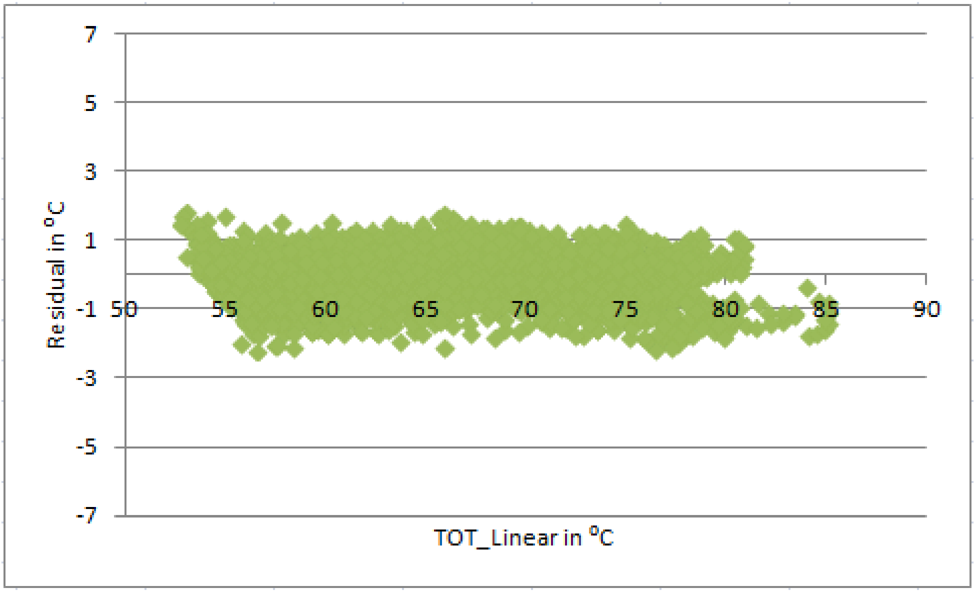


Figure 4.3 Residual Plot for a Reliable Model for Transformer Queencreek-3

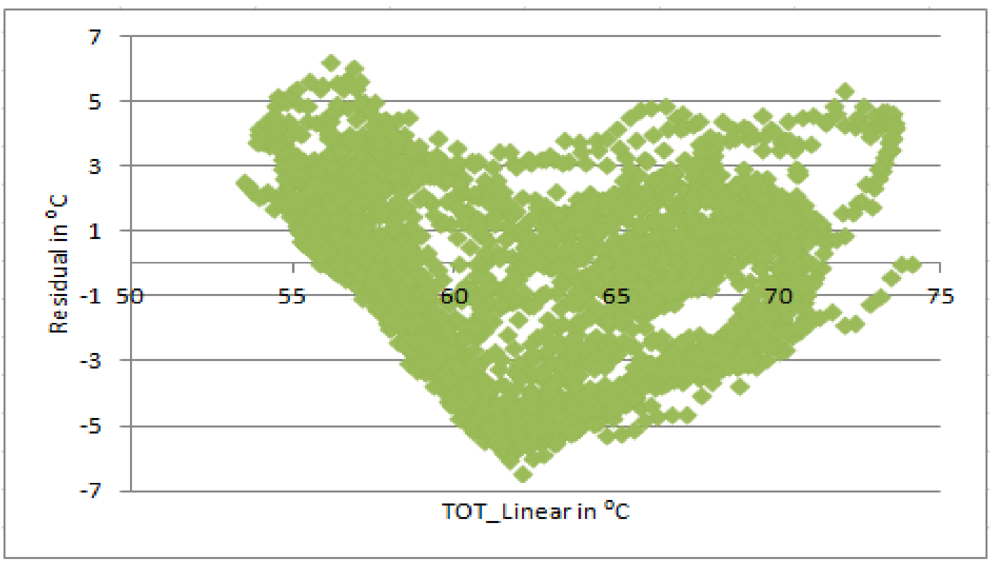


Figure 4.4 Residual Plot for an Unreliable Model for Transformer Highline-3

4.1.3 RMS Error

Root-mean-squared error, another important indicator whether a model is reliable or not, is defined as given by the equation (4.2)

$$RMSerror = \sqrt{(\Sigma(\hat{y}_i - y_i)^2)/n} \tag{4.2}$$

where n is the number of observations.

A high RMS error indicates that the model is unable to predict the TOT accurately. The RMS errors for the transformer thermal models trained on 2010 data are given in Figure 4.5 where the transformers with reliable thermal models are shown in boxes in the legend.

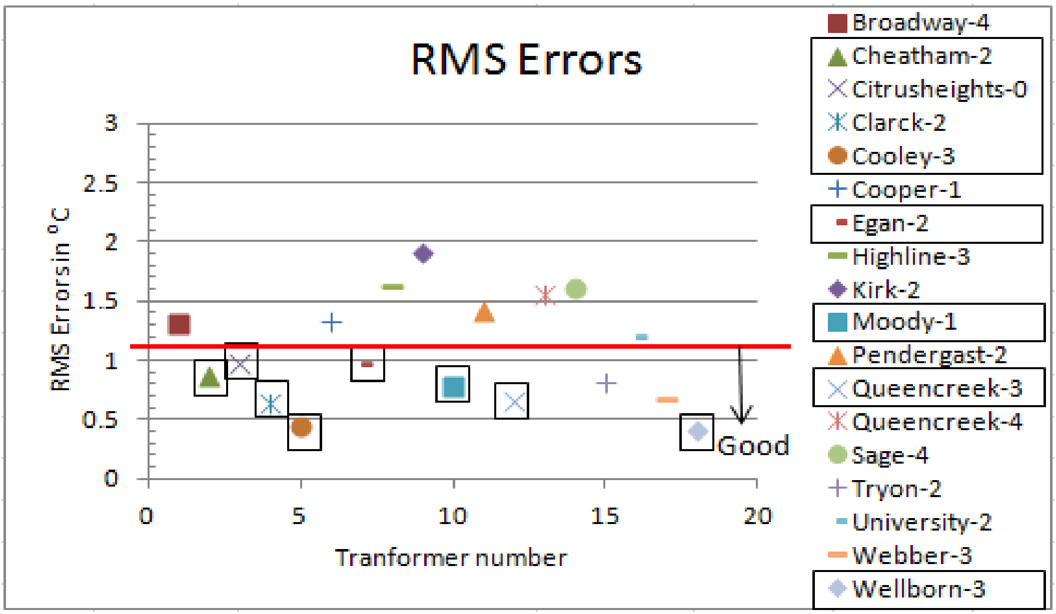


Figure 4.5 RMS Errors for Models Trained on 2010 Data

4.1.4 Time Constant

Another important metric is the model's predicted time constant. From equations (2.22) and (2.23), the time constant for the linear TOT model, in terms of model coefficients can be obtained as,

$$K_2 = \frac{\Delta t}{\Delta t + \tau} \tag{4. 3}$$

$$\tau = (\Delta t / K_2) - \Delta t \tag{4.4}$$

where the sampling period Δt is 15 minutes or 0.25 hours.

Typically, the time constant of the linear TOT model for 28 MVA transformers is in the neighborhood of 2.5 hours. Figure 4.6 gives the time constants for thermal models trained on 2010 data with the names of reliable models shown in boxes in the legend. It was observed that unreliable models had higher time constants.

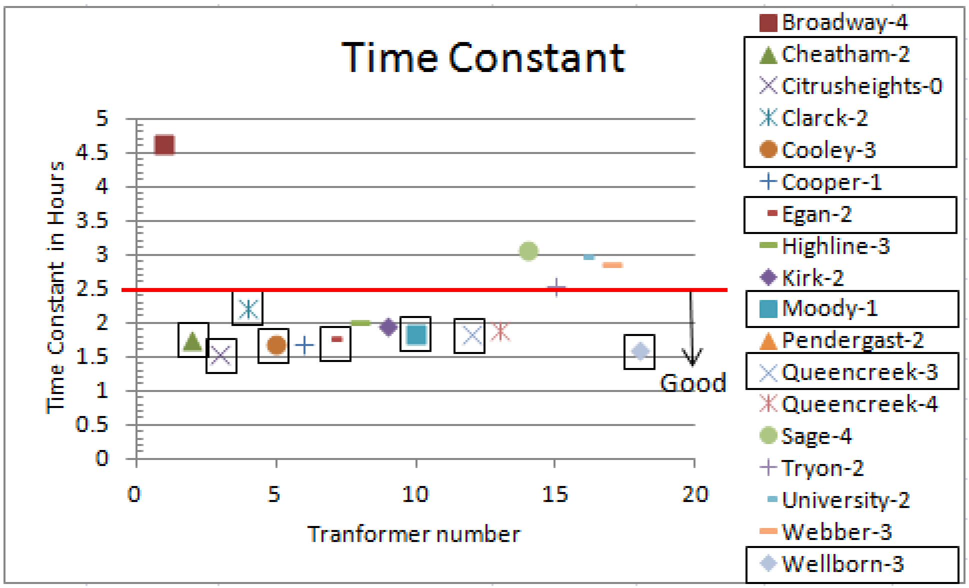


Figure 4.6 Time Constants for Models Trained on 2010 Data

4.1.5 Correlation Coefficient

Correlation coefficient also known as Pearson's product-moment correlation coefficient, is a measure of linear correlation between two variables. It can take values between -1.0 and +1.0. A value of zero implies no correlation, -1.0 implies a perfect negative correlation and +1.0 implies a perfect positive correlation. Therefore using the correlation coefficient we can estimate the level of correlation between measured TOT and predicted TOT. The correlation coefficient is given by the formula (4.5) where covariance is a measure of how two random variables change together.

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{4.5}$$

Figure 4.7 gives the correlation coefficients between measured TOT and predicted TOT for thermal models trained on 2010 data with the reliable models shown in boxes in the legend. It was observed that TOT models that had accurate predictions had correlation coefficients greater than 0.95.

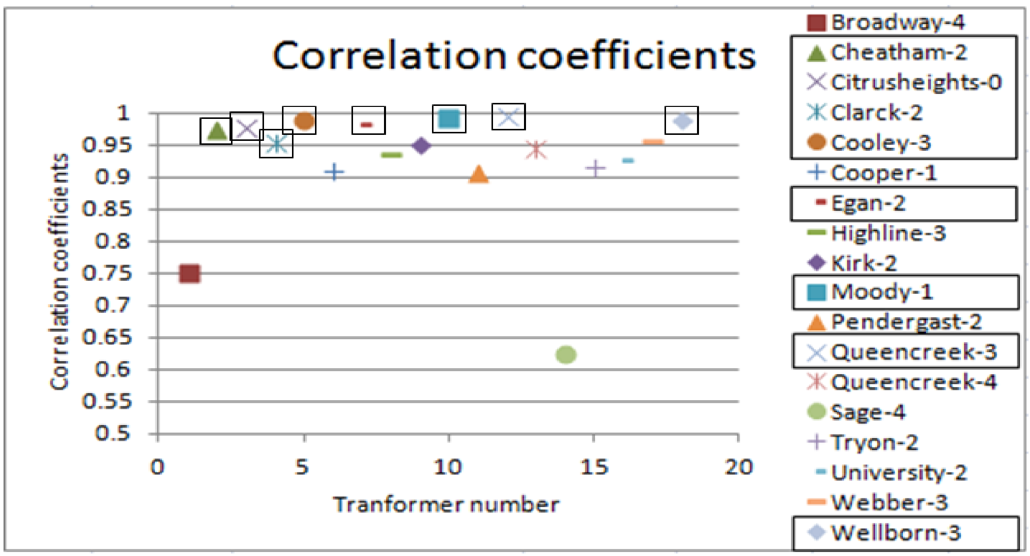


Figure 4.7 Correlation Coefficients for Models Trained on 2010 Data

4.1.6 Negative Model Coefficients

A model is considered unreliable if any of the model coefficients are negative. Since K_1 is the load coefficient, it cannot be negative as this would imply that the predicted TOT decreases with an increase in load. The coefficient K_2 is the coefficient of heat dissipated to the atmosphere and K_4 is the no-load top oil temperature. Based on these physical interpretations of the model coefficients, any model with negative coefficients is nonphysical.

4.1.7 Percentage Standard Deviation of the Model Coefficient K_2

Bootstrapping was used to construct thousands of samples from an independent and identically distributed population of the observed dataset and then the sampled datasets were used to build thermal HST and TOT models of the transformers. The standard deviations of the model coefficients were observed. It was observed that transformers whose models were unreliable were those whose set of bootstrapped models had large standard deviations (measured in percent) of the model coefficient K_2. Percentage standard deviation is defined as the ratio of standard deviation of K_2 and the model coefficient K_2 expressed in percent. Figure 4.8 shows a plot of the percentage standard deviations of K_2 for all transformers contained in the 2010 dataset used for this research. The names shown in boxes in the legend are the transformers with reliable thermal models.

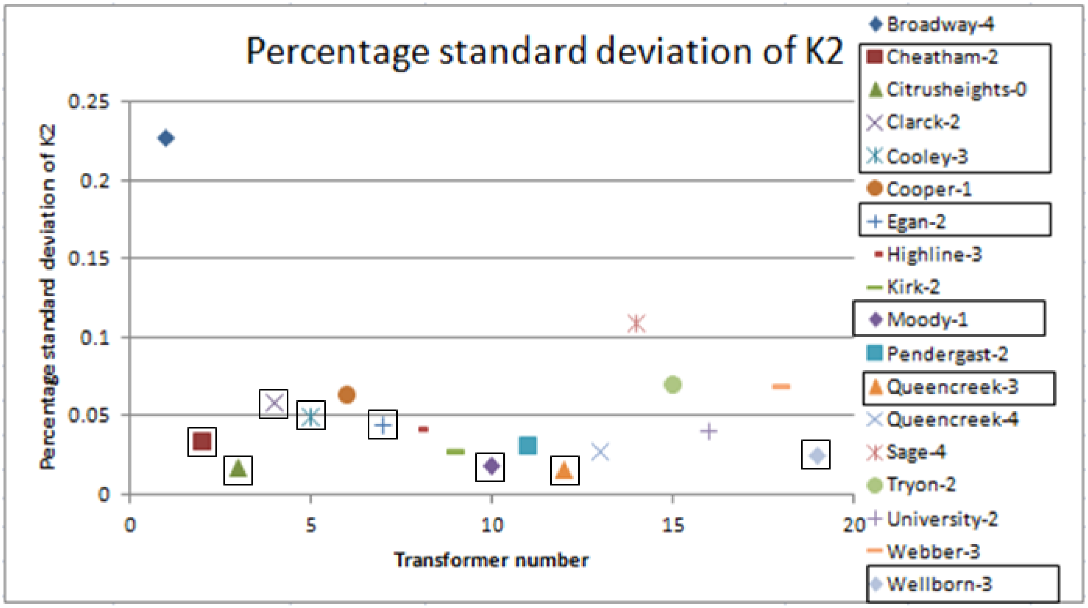


Figure 4.8 Percentage Standard Deviations of K_2 for Models Trained on 2010 Data

4.1.8 Intercept 'a' For Steady-State Load Confidence Interval vs. Sample Size Plot

The model building application TTeMP also provides the user with an analytical representation for the maximum steady-state-load confidence interval vs. sample size used for bootstrapping. The plot looks as shown in Figure 4.9.

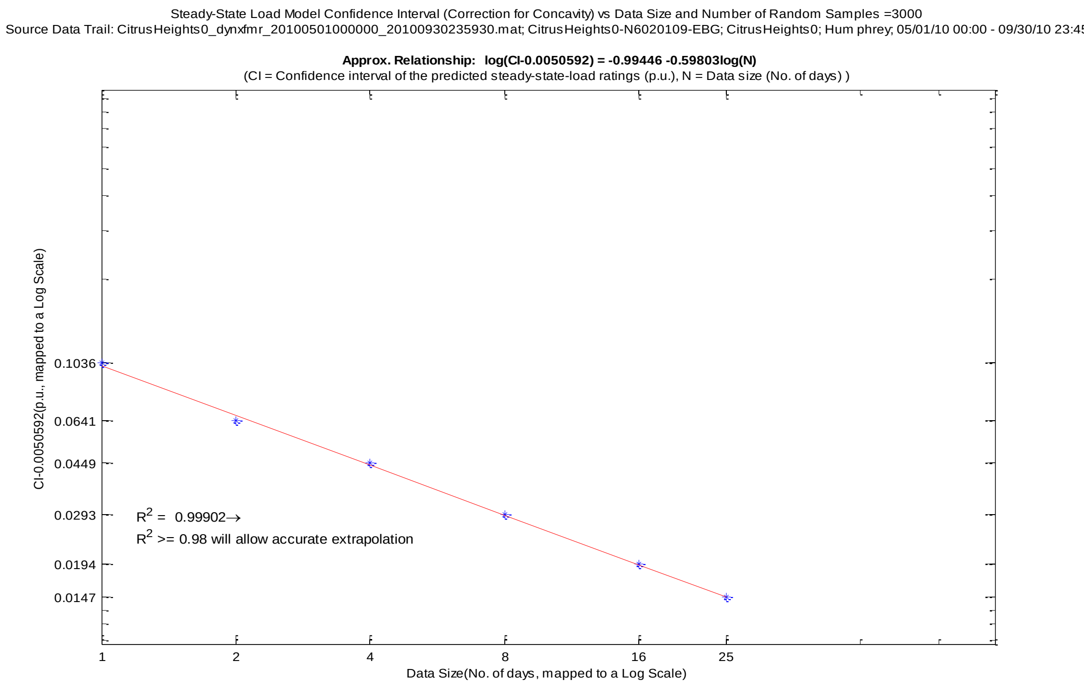


Figure 4.9 Plot of Maximum Steady-State Load Confidence Interval Plotted Against the Data Size

The equation for this plot is given by the following equation.

$$\log(\mathrm{CI} - \mathrm{alpha}) = \mathrm{a} + \mathrm{b} * \log\,(\mathrm{N}) \tag{4. 6}$$

where CI represents the confidence interval of the maximum steady-state load, alpha represents the minimum confidence interval attainable (which is a function of data noise level) and N represents the sample size.

The models which were not reliable had small magnitudes (less negative values) of the 'a' intercept which is a negative number. Since more reliable models would have smaller confidence intervals for any sample size, the intercept of the log(CI-alpha) curve with the log(N)=0 axis, i.e., a, for reliable models will be smaller, that is, more negative. Figure 4.10 shows the plot of the intercept 'a' for all thermal models built from 2010 data and demonstrates that the models with the more negative intercepts are the more reliable models.

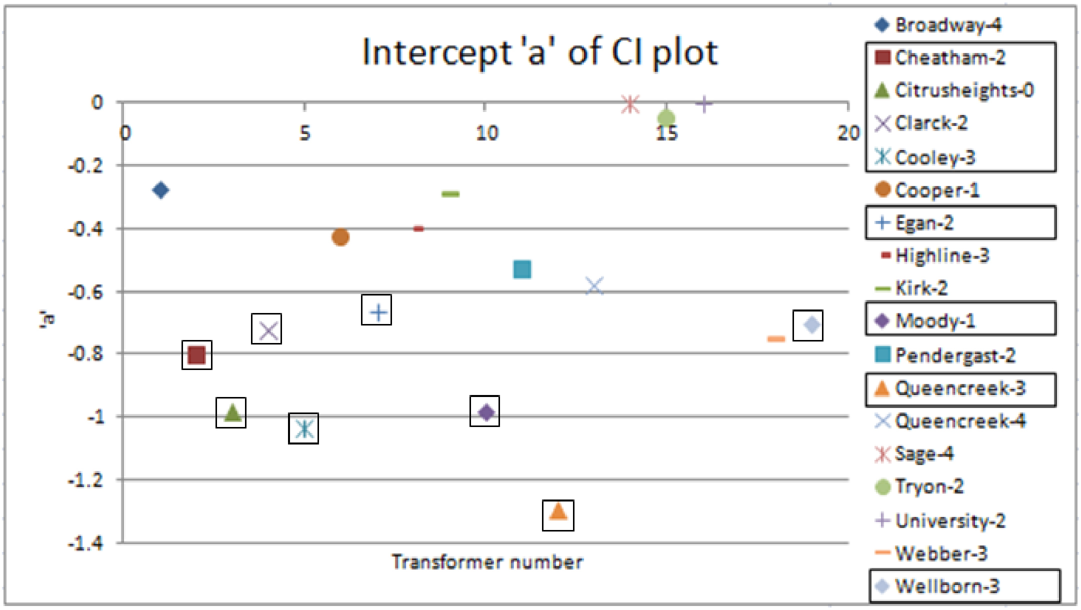


Figure 4.10 Intercept 'a' of CI Plot for All Models Trained on 2010 Data

4.1.9 The Load Coefficient K_1

Another important metric is the load coefficient, K_1. The variable K_1 is the load coefficient in the linear model given by (2.23). It is expected that the load should be a predominant factor in determining the temperature rise. If the K_1 is too low, it indicates that load is not a dominating factor in the thermal model which implies a nonphysical model. A very high value of K_1 will also indicate an unreliable model. It is desired that the K_1 be between certain upper and lower limits. From numerical experimentation and observation, it was determined that for a reliable model K_1 should be within a range of 3.0 to 4.5 for a 28 MVA substation distribution transformers. Figure 4.11 gives a plot of the load coefficients for thermal models built from 2010 data. The names shown in boxes in the legend indicate transformers with reliable thermal models.

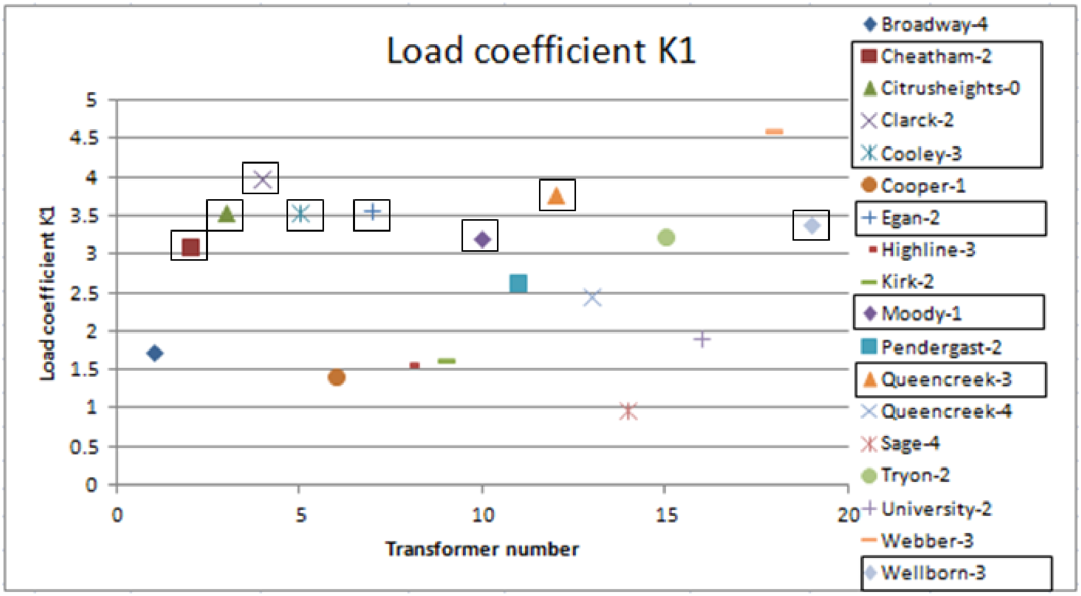


Figure 4.11 Load Coefficient K_1 for Models Trained on 2010 Data

4.1.10 Summary

After much data analysis, we have determined several metrics which are individually somewhat reliable in identifying bad models. Regardless of the rule-base procedure used to identify unreliable models, if any one of the aforementioned metrics taken alone is used, the rules result in rejecting some reliable models and accepting some unreliable models. By collectively using several of these metrics in a rule-based procedure, a highly reliable way of identifying unreliable models can be has been identified. It was found experimentally that not all of the metrics studied were needed for model screening and, in order to keep the model-screening process as simple as possible, all of the metrics described above were not used. Only the metrics that falsely rejected the fewest reliable models and falsely accepted the fewest unreliable were used as a part of the model screening process. The model screening process is described in the next section.

4.2 Identifying Bad Models for 28 MVA Transformers

The following metric values and the results of applying thresholds for model screening are based on the observations made for 19 transformers using data for the calendar years 2006, 2009 and 2010. Depending on the amount of data available for each year, independent thermal models were developed for each transformer for most of these years, giving 39 thermal models upon which the following observations were based.

4.2.1 Time Constant and RMS Error

Through visual inspection of the errors between measured and predicted TOT, and visual inspection of the residual plots, it was observed that all linear models with a time constant greater than 2.5 were unreliable. Also all models with an RMS error greater than 1.1 were unreliable.

Figure 4.12 shows the time constants of the models trained on 2010 data and Figure 4.13 shows the RMS errors for models trained on 2010 data, with the models considered as reliable using both visual inspection and residual plots shown in boxes in the legends of both figures. The horizontal lines in these figures indicate the approximate boundary where models transition from being acceptable to unacceptable.

However it was found that using these metrics alone had two drawbacks: sometimes reliable models were erroneously rejected and sometimes unreliable models were classified as good. To discriminate between reliable and unreliable models, the following rule was found to yield reliable results: a model was considered reliable *if and only if* its time constant was less than 2.5 *and* the RMS error was less than 1.1. Otherwise it was considered unreliable. Using this rule, all of the unreliable models were successfully discarded and only the reliable models were accepted. This method was successfully tested on all the models corresponding to different years and was found to work reliably.

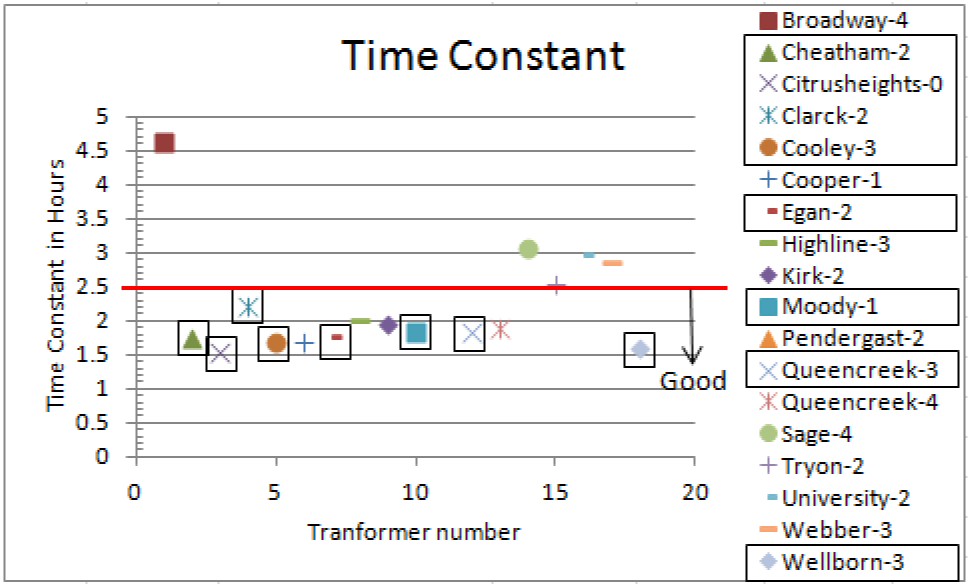


Figure 4.12 Time Constants for Models Trained on 2010 Data

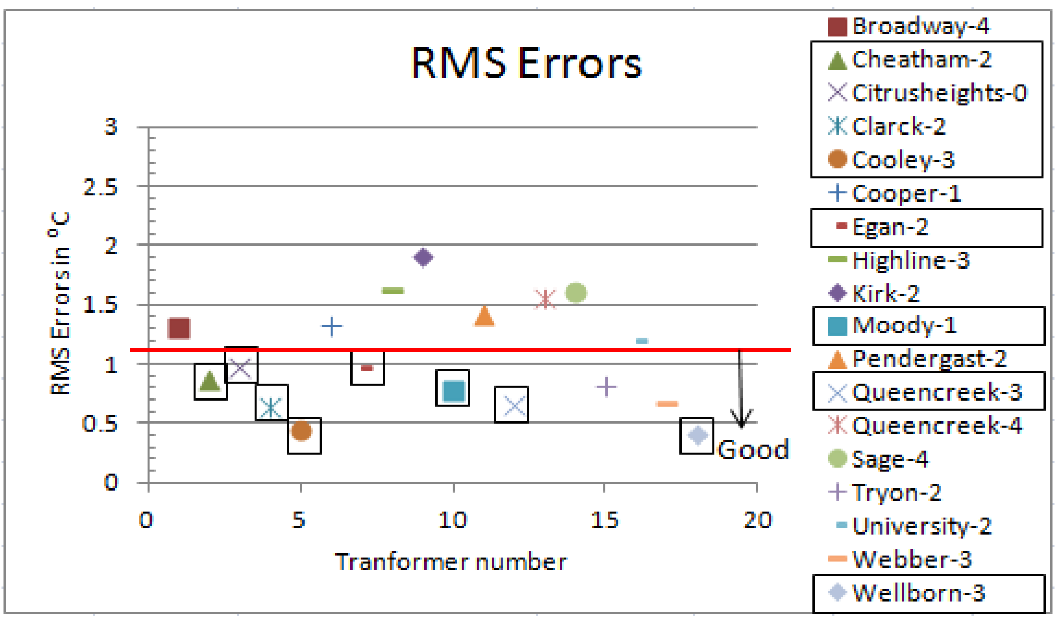


Figure 4.13 RMS Errors for Models Trained on 2010 Data

It is possible to identify unreliable transformer thermal models based on the following metrics: RMS error and time constant. Using the thresholds established for these metrics, the bad-model identification process was found to be 100% reliable for the 39 cases tested. The metrics also show some promise in being able to diagnose operational issues that lead to unreliable models.

The metrics presented in this section and corresponding threshold values were calculated *only* for 28 MVA substation distribution transformers. Acceptable ranges of these metrics, such as the time constant, vary for transformers of different ratings; and are therefore design specific. Determining appropriate ranges for transformers of different ratings is a matter for future work.

4.3 Model Reliability of Thermal Models

Using thermal models built from measured historical data, we found that transformers fell into certain arbitrary classes.

- High Quality Models: For these models, the TOT prediction performance, subjectively viewed, was good. These models were built using moderate transformer loads, which represented the highest recorded historical loads for these transformers.
- Poor Quality Models: After looking at model performance using several metrics and using visual inspection, we found that there are likely to be various causes that result in poor-quality models. We believe that some root causes of poor-quality models are:

- Fans do not turn on/off according to the set-point or the set-point in the field is different from the set-point value contained in the PTLoad files. Since these points are used to separate measure data into various cooling modes, erroneous data will be incorporated into our models if the fan turn on/off times are incorrect.
- Transformers have unbalanced loading. The unbalance is expected to affect the hot-spot models since hot-spot temperatures are recorded for each phase.

4.4 Model Quality

It was desired to go beyond the classifying of models into two bins, reliable/unreliable, and rank models according to their degree of reliability. In order to provide the user with an easily comprehensible model quality index, an attempt is made to classify thermal models in different categories such as 'Excellent', 'Good', 'Fair', 'Poor' and 'Unacceptable'. This is done using the correlation coefficient, time constant predicted by the model and the quality of the residual plot. A metric is determined for the model quality comprised of a weighted correlation metric, time constant metric and a metric assigned to the quality of the residual plot. This summative metric is used to determine the category of the models as described below.

4.4.1 Discussion of Model Quality

The correlation metric used was obtained using the following equation

$$\text{Correlation Metric} = 3.5 * \log(1/(1 - \text{Correlation Coefficient})) \tag{4.7}$$

Since the correlation coefficient is higher for reliable models, the reciprocal of (1-correlation coefficient) is higher for reliable models. The logarithmic scale provides a better distinction between reliable and unreliable models and the scaling constant 3.5 is used in order to make the correlation metric of the most accurate model close to 10. Thus the transformers' correlation metric were evaluated on a scale of zero to 10 where reliable models had higher correlation metric values.

Visual inspection of the time-domain plots was performed to judge the reliability of the models and then the time constants were observed. It was observed that the time constants of reliable models are in the range of 1.5 to 2.5. Thus the metric was assigned on the basis of Table 2.

Table 2 Time Constant Metric

Time Constant Range	Time Constant Metric
Time Constant>=3	4
2.5<=Time Constant<3	6
2<=Time Constant<2.5	8
1.5<=Time Constant<2	10
Time Constant<1.5	6

The metric for the residual plot is determined using non-linear regression to fit a curve to the lower boundary of the residual plot. The coefficients obtained by nonlinear regression indicate whether the residual plot has a V-shaped curve. Based on the thresholds for these coefficients, it is judged whether a residual plot has a V-shape or not.

Residual plot quality is assigned a value of five or ten based on the coefficient obtained from non-linear regression and the standard deviation of the predicted SSL_{Max}. This procedure is described in detail in Chapter 3.

All the above three metrics i.e. correlation metric, time-constant metric and residual-plot metric are averaged to obtain the summative metric. This cumulative metric is used to classify the model quality based on the classification given in Table 3.

Table 3 Model Quality Based On Cumulative Metric

Cumulative Metric	Model Quality
Metric>=9	Excellent
8<=Metric<9	Good
7<=Metric<8	Fair
6<=Metric<7	Poor
Metric<6	Unacceptable

Table 4 gives a list of transformers, which have been classified based on the accuracy of the thermal-model top-oil temperature (TOT) predictions using the summative metric. The transformers are in decreasing order of the model quality with the most reliable model being the first transformer in the table.

Table 4 Model Quality Assigned to Transformer TOT Models Based on a Weighted Metric

Transformer	Correlation Coefficient	Time Constant	Residual Plot Metric	Model Quality
Queencreek-3, 2010	0.995	1.82	10	Excellent
Moody-1, 2010	0.992	1.83	10	Excellent
Burton-3, 2006	0.991	1.86	10	Excellent
Webber-3, 2012	0.991	1.99	10	Excellent
Cheatham-2, 2009	0.989	1.52	10	Good
Citrusheights-0, 2009	0.990	1.55	10	Good
Wellborn3, 2010	0.989	1.60	10	Good
Cooley3, 2010	0.988	1.69	10	Good
Egan-2, 2009	0.983	1.76	10	Good
Webber-3, 2011	0.980	1.71	10	Good
CitrusHeights-0, 2010	0.976	1.54	10	Good
Clarck-2, 2009	0.975	1.87	10	Good
Cheatham-2, 2010	0.973	1.74	10	Good
Queencreek-4, 2009	0.944	1.89	10	Good
Clarck-2, 2010	0.954	2.21	10	Fair
Tryon-2, 2010	0.914	2.35	10	Fair
Webber-3, 2010	0.957	2.87	10	Fair
Kirk-2, 2010	0.952	1.96	5	Poor
Cooper-2, 2010	0.909	1.67	5	Poor
Highline-3, 2010	0.937	2.01	5	Unacceptable
University-2, 2010	0.926	2.99	5	Unacceptable
Highline-3, 2009	0.881	2.53	5	Unacceptable
Broadway-4, 2009	0.847	4.19	5	Unacceptable
Broadway-4, 2010	0.751	4.61	5	Unacceptable
Sage-4, 2010	0.626	3.08	5	Unacceptable

This 'Model Quality' index has been incorporated into DLTA application and is made available to the user on the interface of the 'Model Results' window which is a part of the model building application. It is believed that this classification gives a good estimate of the reliability of the TOT model. For instance, it is possible that a highly reliable model will be classified as 'Good' or 'Fair' or vice-versa. However the probability of accepting unreliable models is low.

4.5 Testing Model Reliability under Heavily-Loaded Conditions

In order to test/validate how well the IEEE and linear models predicted performance under high loads, it was desirable to observe the transformer’s TOT performance under heavily-loaded conditions and then compare the measured values to those predicted by the models built using moderate loads. From the historical data available, we provided the engineers at SRP with a list of transformers with thermal models we considered reliable and requested them to overload as many of those transformers as possible. The engineers at SRP agreed to overload two transformers 'Cooley-3' and 'Freestone-3' for two days upto 35 MVA and provided us the data to test our metrics. We obtained measured load, TOT and HST and ambient temperature data of these two heavily overloaded transformers for the summer of 2013 and were able to verify these results. In order to build a thermal model, the heavy load data was removed and then just enough data to build a model was used. i.e. it contained data equivalent to 120 hours of FAFA cooling mode data. This model was then tested to observe its predictions on the over-loaded days which were compared with the measured TOT on those days. It was observed that the linear model predictions were very accurate and this corresponded well with the conclusions we reached on the basis of the performance in interpolation.

4.6 Possible Causes Behind Unreliable Models

Once an unreliable model is detected, our goal is to identify potential causes for the poor predictions. Ultimately the cause of all bad models is either bad data or an inadequate model structure. We assumed that our linear model's structure was acceptable and identified some causes of bad models. The sophisticated input-data screening process we are using has been reported in [12]; however it is difficult to predict all of the ways in which data can be bad. For instance, it was observed that models trained for a certain set of transformers using 2006 data were unreliable whereas models trained on 2009 and 2010 data for the same transformers were considered reliable. For certain transformers the models trained on the 2006 data, the SSL_{Max} values were below 0.7 p.u. whereas, for the 2009 and 2010 models of these transformers, the SSL_{Max} values were between 1 p.u. and 1.3 p.u. Having inspected the 2006 data carefully, it was observed that the transformers with a low SSL_{Max} value had bad ambient temperature data. The ambient temperature for these transformers remained constant at 8^{o} C throughout many summer days in Phoenix which is improbable. While eliminating temperatures out of range for the year (above 122^{o} F or below 17^{o} F) was part of the original screening algorithm, screening for temperatures out of range for summer conditions alone had not been implemented. The conclusion for this study was that a low SSL_{Max} could be caused by faulty ambient temperature data. The outcome of this study was to pass this information on to SRP so they could investigate the cause of the faulty data.

In order to identify whether there were any systematic factors that affected both TOT and HST models, we tried to observe if there was any correlation between unreliable TOT and unreliable HST models. The RMS errors for HST models were plotted against the RMS errors for the TOT models for all transformers as shown in Figure 4.14. It was observed that the accuracy of the predictions of the TOT models had no correlation with the accuracy of the predictions of the HST models. The R squared value for this plot was as low as 0.0751. Thus there is no correlation between unreliable TOT and unreliable HST models and no such factors exist that affect both TOT and HST models.

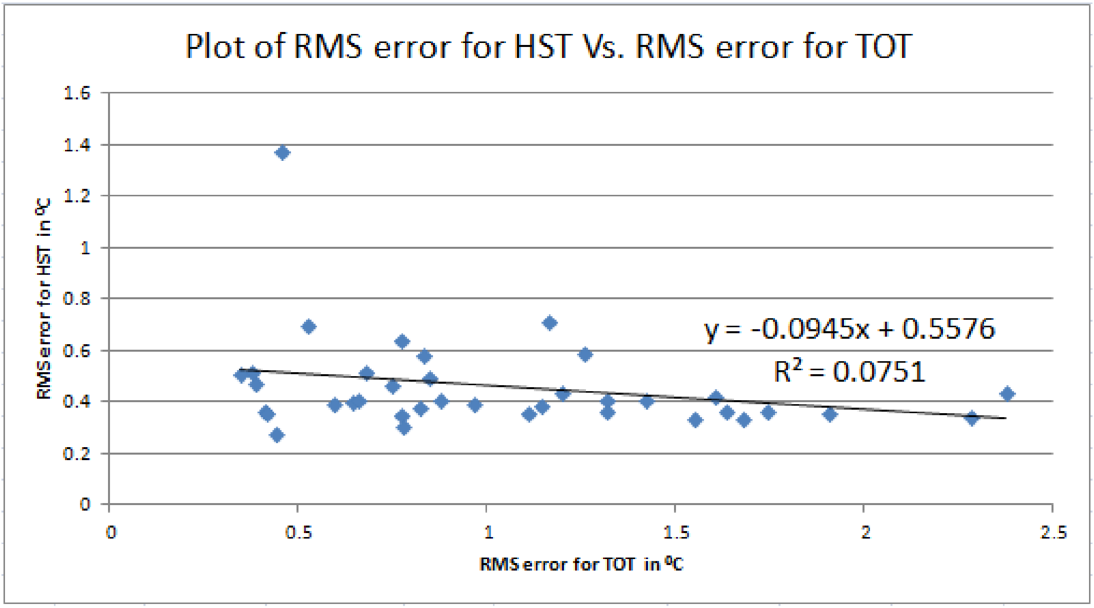


Figure 4.14 Plot of RMS Error for HST v/s RMS Error for TOT

4.7 Summary

Thus metrics have been identified to distinguish between reliable and unreliable thermal models. A method to obtain an easily comprehensible model quality index has been developed. Similar research needs to be conducted for IEEE models and the HST models in the future. It is very important to successfully discard an unreliable thermal model. If this is not done, the results provided by the dynamic loading application (DLTA) may be either non-conservative which may lead to over-heating of the insulation, thus reducing the transformer life. Or the predicted loading may be too conservative, which will lead to under-utilization of the available resources. Possible causes behind unreliable thermal models were identified and it was concluded that no such factors exist that affect both TOT and HST models.

CHAPTER 5

METHODS TO IMPROVE THE ACCURACY OF THE LINEAR MODEL PREDICTIONS

It can be seen from the discussion in the previous chapters that some transformers have unreliable thermal models. The possible causes behind the poor predictions by these models have also been explored. This chapter describes some methods to improve the accuracy of the temperature predictions of the linear models.

5.1 Linear Model for OA and FA Cooling Modes

From our experience with building thermal models from measured data, it has been observed that without significant forced cooling, such as in the OA and the FA cooling modes, factors such as minor changes in external conditions, (i.e. wind, solar radiation) and quantization of the measurements, lead to unreliable models. Thus, the IEEE models have been exclusively used for temperature predictions in the OA and FA cooling modes in the DLTA application and the linear models have been used for the FAFA cooling mode, provided the linear models are reliable. It is important to have accurate predictions in the OA and FA cooling modes because, as seen from (2.23) and (2.29), the TOT and HST predictions at a given observation point for the linear model depend on the previously predicted values. Thus even if an accurate model is used *only* for the FAFA cooling mode, if the error at the point of entry in the FAFA cooling mode itself is high, then the next prediction will similarly have a large error. The errors compound in the consecutive observation points resulting in temperature predictions at higher temperatures being very inaccurate. Thus, although the primary concern for accuracy is in the higher temperature ranges when the transformer is in FAFA cooling mode and the insulation is running close to is thermal limits, an attempt was made to improve the accuracy of the predictions in OA and FA cooling modes as well.

Due to the poor performance of the IEEE model (in the FAFA cooling mode,) we suspected that the linear model might predict better than the IEEE model in the OA and FA cooling modes as well. Hence a linear model was built for OA and FA cooling modes using linear regression techniques and measured load, TOT, HST and ambient temperature data to test such models' performance for the respective cooling modes. The equations used to build the linear models for these cooling modes were same as those used for the FAFA cooling mode given by (2.23) and (2.29). For the IEEE models as well, the equations were the same as those for the FAFA cooling mode, given by (2.12) and (2.19), with different exponents obtained from the transformer heat-run test reports used for each cooling mode, as recommended by IEEE Std C57.91-1995. The accuracy of the linear model predictions was then compared with that of the IEEE model predictions.

It is important to minimize the error at the FA-to-FAFA cooling-mode transition point in order to minimize the error in subsequent temperature predictions. Hence we conducted an experiment to compare the errors in the IEEE FA cooling mode model predictions with those of the linear FA cooling-mode-model predictions at the points at which the transformer was predicted to enter the FAFA cooling mode from the FA cooling mode. The error is given by (5.1)

$$e = Temperature_{predicted} - Temperature_{measured} \tag{5.1}$$

This experiment was conducted for the transformer Broadway-4 for the TOT model built from the data available for the year 2009 from the first day of May at midnight to the last day of September. The data was segregated during model-building into different cooling modes using the simulated HST calculated using (3.1), where measured TOT and load data were used to calculate the simulated HST. The simulations were then performed on the data-sets obtained, for the linear and the IEEE models of the respective cooling modes. Initialization to measured TOT was done at every point where the transformer was predicted to enter the OA and FA cooling modes respectively for both the linear as well as the IEEE models. The errors of the linear model predictions were then compared with the errors of the IEEE model predictions at the data points at which the transformer was predicted to enter the FAFA cooling mode where error is given by (5.1). Figure 5.1 gives a plot which contains the errors in TOT predictions by the linear and the IEEE models at the points at which the transformer was assumed to enter FAFA cooling mode from FA cooling mode for the transformer Broadway-4 trained on 2009 data. It can be seen that the linear model had an accuracy of +/- 2^0 C whereas the IEEE model always under-predicted the measured TOT by 2^0 C to 8^0 C at these data points.

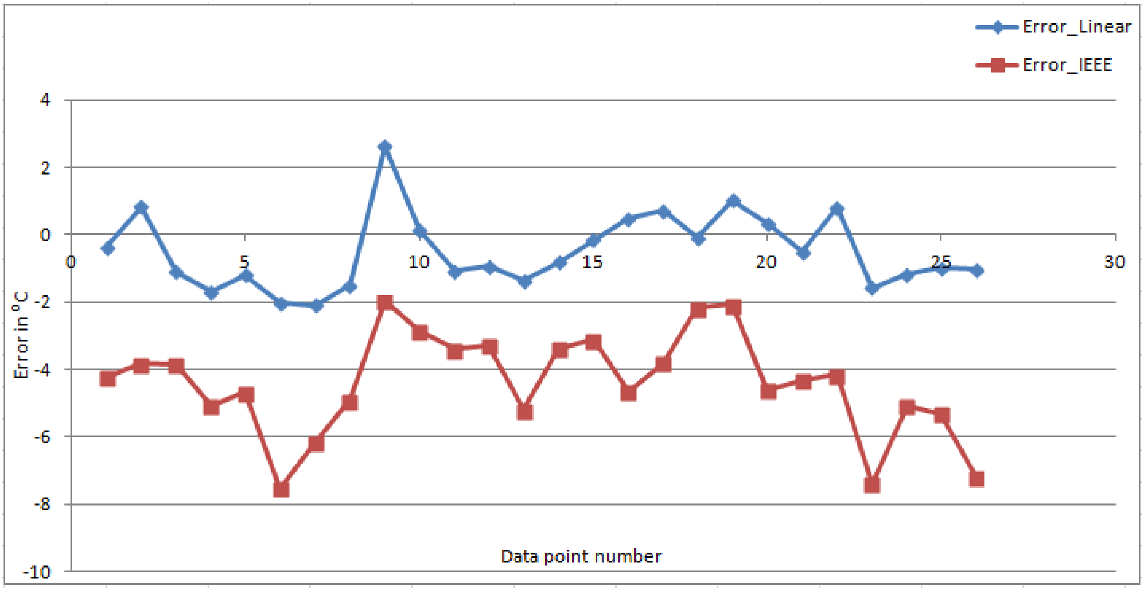


Figure 5.1 TOT Errors at the Points where the Transformer Enters FAFA Cooling Mode for Broadway-4

Thus it was observed that the IEEE TOT model prediction error at the point at which the transformer was predicted to enter FAFA cooling mode from FA cooling mode was greater than the linear TOT model prediction error at that point.

Since the errors were high for the IEEE model at the points at which the transformer entered the FAFA cooling mode, it was hypothesized that the linear model would have lower RMS errors in OA and FA cooling modes as compared to the IEEE model for these cooling modes. In order to test this hypothesis, linear models were built for OA and FA cooling modes using measured data which was obtained from different years for ten transformers, three of which did not have the FA cooling mode. The linear and IEEE models were used to predict the TOT as well as the HST for the entire training data-set which typically had three to five months of summer data. The predicted TOT and HST for the observation points at which the cooling mode transitions took place were initialized to the corresponding measured TOT and HST for both the linear and the IEEE models and then the RMS errors were calculated for the different cooling modes based on the training data set. The errors were calculated between the predicted TOT and measured TOT, and between the predicted HST and measured HST for the linear and the IEEE models for both OA and FA cooling modes. As seen from Figure 5.2, Figure 5.3, Figure 5.4 and Figure 5.5, the ratio of RMS error of the IEEE model to the RMS error of the linear model is greater than 1.0 for most of the transformers for TOT as well as HST in both OA and FA cooling modes. Thus it was observed that the RMS errors of the TOT and HST predictions in the OA and FA cooling modes for the corresponding linear models were lower as compared to the RMS errors for the IEEE models.

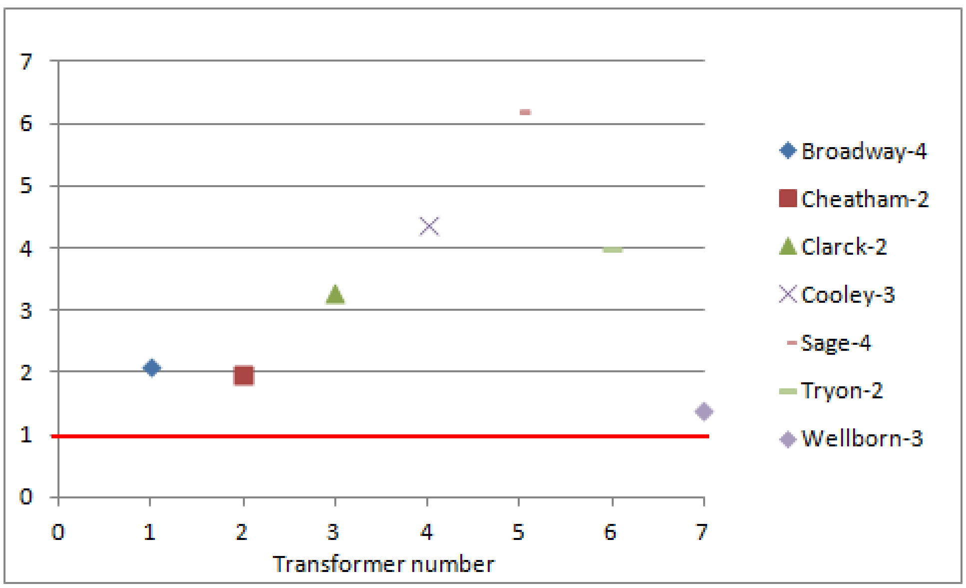


Figure 5.2 Ratio of RMS Errors of the IEEE TOT Model Prediction to the Linear TOT Model Prediction for FA Cooling Mode

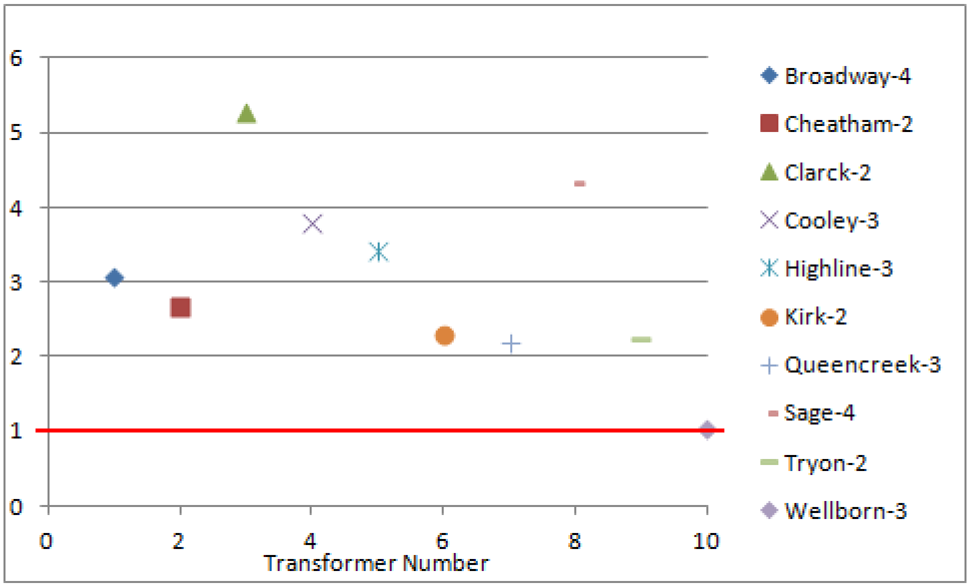


Figure 5.3 Ratio of RMS Errors of the IEEE TOT Model Prediction to the Linear TOT Model Prediction for OA Cooling Mode

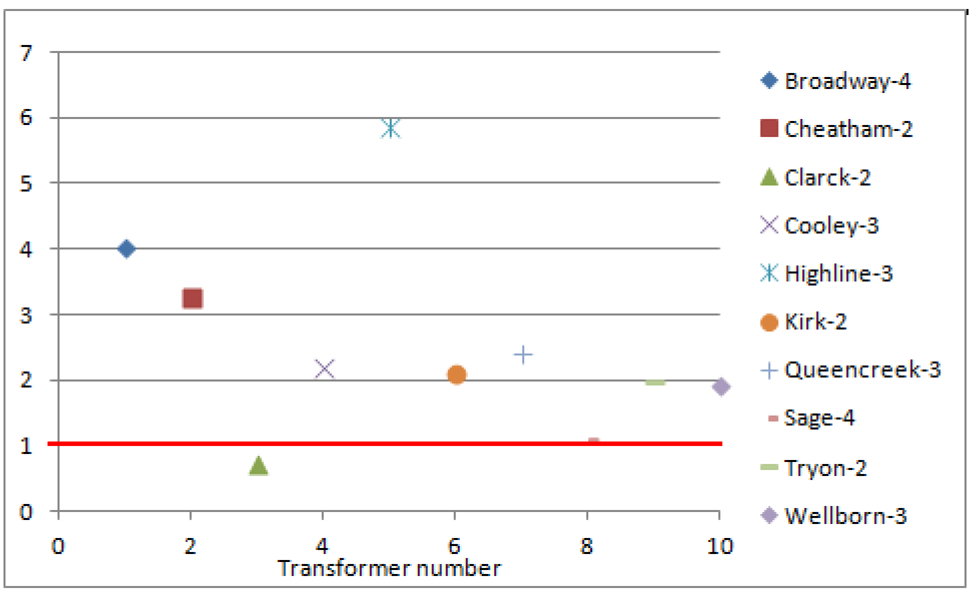


Figure 5.4 Ratio of RMS Errors of the IEEE HST Model Prediction to the Linear HST Model Prediction for OA Cooling Mode

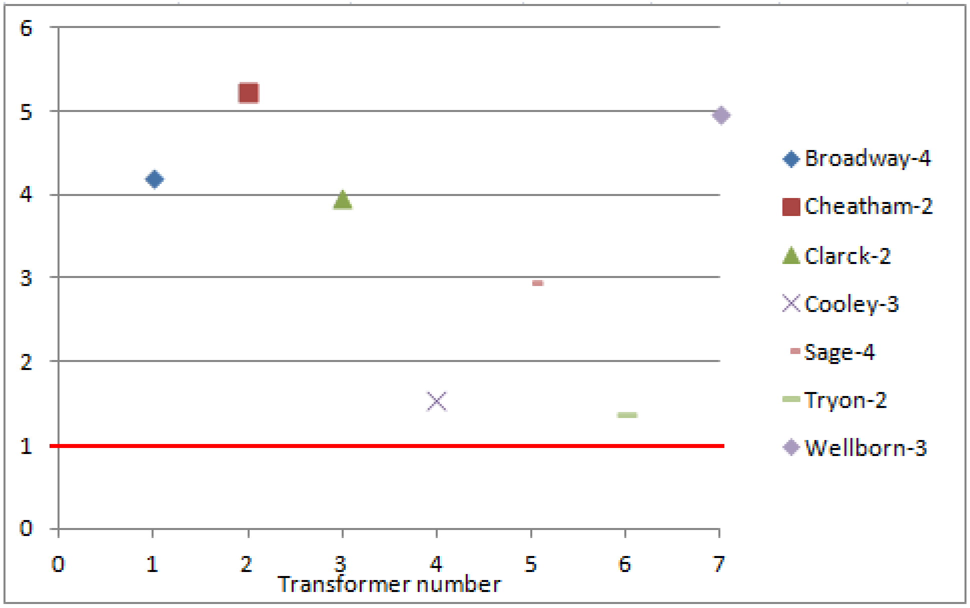


Figure 5.5 Ratio of RMS Errors of the IEEE HST Model Prediction to the Linear HST Model Prediction for FA Cooling Mode

In order to further observe if using the linear models in all cooling modes led to more accurate predictions than the IEEE models, the linear model and the IEEE model TOT predictions were plotted along with measured TOT for 72 continuous hours for a number of transformers with the first data point initialized to the measured TOT for both models. The data was segregated during model-building into different cooling modes using the simulated HST calculated using (3.1), where measured TOT and load data were used to calculate the simulated HST. The simulations were then performed on the data-sets obtained, for the linear and the IEEE models of the respective cooling modes. Figure 5.6 shows the continuously modeled TOT for the transformer Webber-3, starting at midnight, with the first data point initialized to the corresponding measured TOT for the linear as well as the IEEE models. In this figure, TOTm represents the measured TOT, TOT_Linear represents the linear model prediction and TOT_IEEE represents the IEEE model prediction. It can be seen that the linear model predictions followed the measured data more closely than the IEEE model predictions in OA as well as FA cooling modes. Thus it was again observed that the linear model predictions were more accurate than the IEEE model predictions in all cooling modes. This behavior was observed for both reliable and unreliable models with the classification between reliable and unreliable models being made based on the FAFA model reliability, since the reliability metrics as described in Chapter 4 for the FAFA cooling mode have not yet been developed for OA and FA cooling modes. This is a matter of future work.

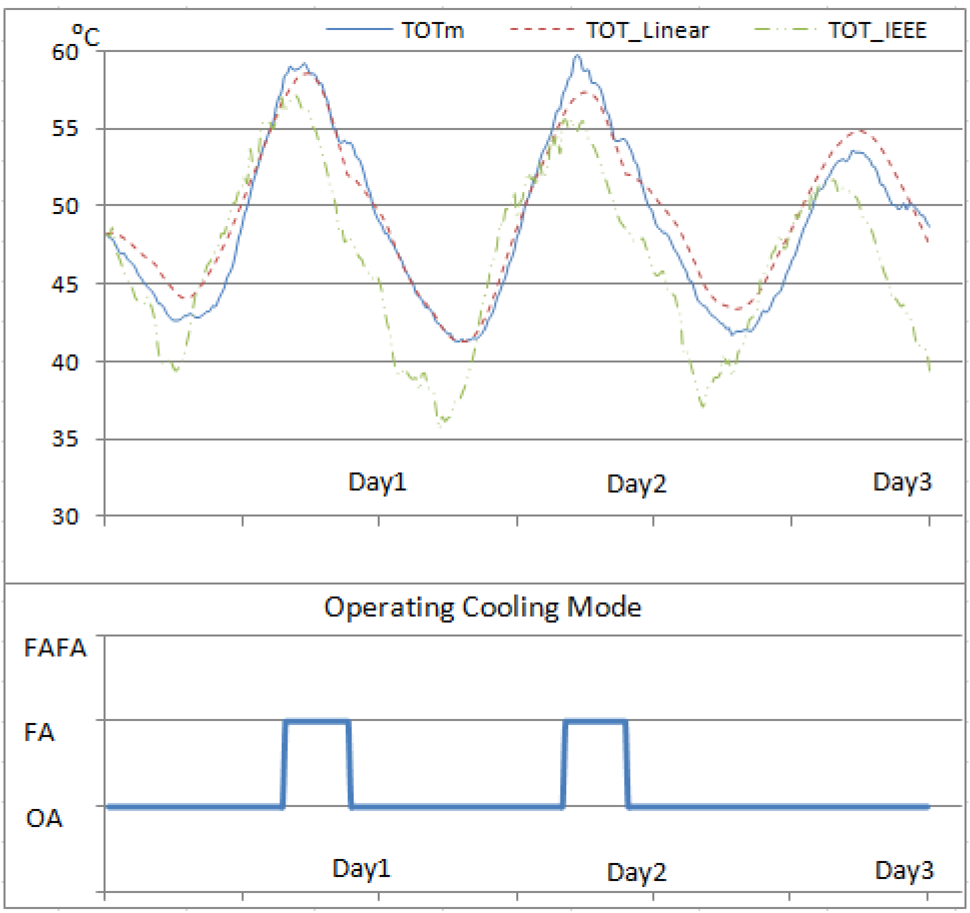


Figure 5.6 Continuously Modeled TOT Predictions for Webber-3 Model Trained on 2010 Data and the Corresponding Operating Cooling Modes

Hence the results showed that the linear model predictions in the OA and FA cooling modes are generally more accurate than the IEEE model predictions in OA and FA cooling modes respectively. Since the linear model predictions are highly dependent on the predictions in the previous time step, it is important that the model predictions be accurate at the points at which the transformer is assumed to enter the FAFA cooling mode.

5.2 Linear Regression Model using Least Absolute Value Method

The parameters of the linear model used in the work thus far have been calculated using the least squares (LS) method as described in Chapter 2. In order to improve bad data rejection, we considered using the least absolute value (LAV) method to estimate the model coefficients. The LAV method simultaneously detects and rejects bad data thus providing accurate model coefficients. Thus the LAV helps build more accurate regression models if there are many outliers in the measured data. The LAV regression coefficients are chosen to minimize the sum of the absolute values of the residuals. By minimizing sums of absolute values rather than sums of squares, the effect of outliers on the coefficient estimates is diminished. The LAV is governed by (5.2),

$$\min \sum_{i=1}^{n} |e_i| = \min \sum_{i=1}^{n} \left| y_i - \sum x_{ij} \beta_j \right| \tag{5.2}$$

The formulation used to implement the LAV based regression is given by (5.3).

$$\min \sum_{i=1}^{n} e_i \tag{5.3}$$

Subject to $\beta_j x_{ij} - y_i \leq e_i \; and - \beta_j x_{ij} + y_i \leq e_i$

where β_j is the model coefficient of the j^{th} independent variable, x_{ij} is the i^{th} measurement of the j^{th} independent variable and y_i is the i^{th} measured value of TOT or HST.

It was desired to observe whether the LAV models yield superior results as compared to the LS models. To that end, linear TOT models were built whose coefficients were obtained for the FAFA cooling mode using both the LAV and LS method. The data to build the models was obtained from different years for ten transformers and was typically available for three to five months of the summer. The data was segregated during model-building into different cooling modes using either simulated HST or measured HST data (if it was available for a given transformer). The simulations were then performed on the FAFA data-set for the linear and the IEEE models. Initialization to measured TOT was done at every point where the transformer was predicted to enter FAFA cooling mode for both the linear as well as the IEEE models. The correlation coefficients between the measured and predicted TOT values based on the training data were compared with those obtained when the model was built using the LS method. Table 5 provides the correlation coefficients for the TOT models built using the LAV and LS methods. In order to highlight the transformers for whom the LAV method yields better results, when the correlation coefficient for a transformer model built using LAV method is greater than the correlation coefficient for the same transformer with the model built using LS method, the coefficients are written in green font, otherwise they are in red font. It can be seen that for most transformers, the correlation coefficient was higher when the model was built using LAV method.

Table 5 Comparison of Correlation Coefficients by LAV and LS Methods

Transformer	Correlation Coefficient by LAV	Correlation Coefficient by LS
Broadway-4	0.860087	0.846771
Cheatham-2	0.990232	0.989899
Clarck-2	0.974809	0.975209
Cooley-3	0.988108	0.987886
Highline-3	0.882594	0.881317
Kirk-2	0.955133	0.951515
QueenCreek-3	0.994953	0.994889
Sage-4	0.635617	0.625612
Tryon-2	0.934866	0.931673
Wellborn-3	0.989103	0.989271

5.3 Incremental and Decremental Models

In the present model building application, a single thermal model is built for the entire time that the transformer is in FAFA cooling mode, including both, the time when the transformer is heating up and when it is cooling down. The oil-circulation behavior is different when the transformer is heating up versus when it is cooling down. Assuming the transformer is operating in FAFA cooling mode, when a transformer is heating up, the oil viscosity is low and thus it circulates faster causing the insulation to cool down faster by the forced air cooling. Whereas, when the transformer is cooling down, the oil viscosity is increasing and thus its circulation rate reduces and the insulation cools at a slower rate than it does when the insulation temperature is increasing. It was desired to determine whether temperature prediction could be improved if separate models were built to simulate thermal performance under increasing and decreasing temperature conditions. Thus a numerical experiment was conducted for some transformers where different sets of transformer coefficients were determined for increasing and decreasing temperature periods. The data used to build the models as well as to test the accuracy of predictions belonged to different years for seven transformers and was typically available for three to five months of the summer. The equation used for building the linear models for the increasing and decreasing temperature periods, was same as that used for building a unified model, given by (2.23), with the measured data used to train the models belonging to the corresponding time periods. The incremental model was used when the measured TOT was increasing and the decremental model was used when the measured TOT was reducing in order to test the predictions. The data was segregated during model-building into different cooling modes using either simulated HST or measured HST data (if it was available for a given transformer). The simulations were then performed on the FAFA data-set for the linear and the IEEE models. Initialization to measured TOT was done at every point where the transformer was predicted to enter FAFA cooling mode for both the linear as well as the IEEE models. Figure 5.7 gives a plot of the ratio of RMS error when a single thermal model was used to the RMS error when the incremental and decremental models were used for the transformers for which this experiment was conducted. It can be seen that this ratio is greater than one for most transformers. Hence by using the incremental and decremental model, the RMS error is reduced for most transformer models.

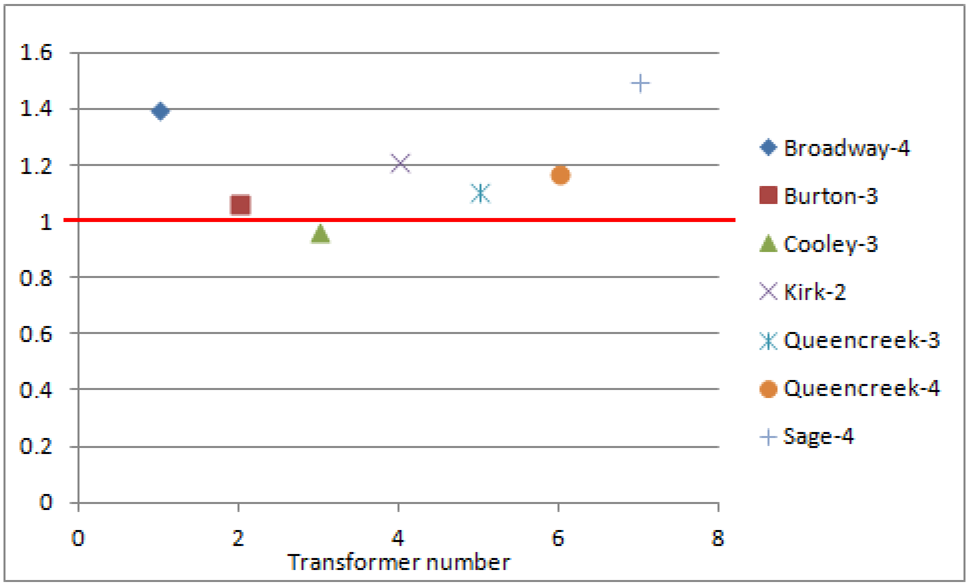


Figure 5.7 Ratio of RMS Errors of Single Model to the Incremental & Decremental Model

Thus using two different models for the time periods when the insulation temperature is increasing and when the insulation temperature is decreasing may improve the accuracy of the thermal model predictions.

5.4 Using the Hottest Winding at each Data Point for Cooling Mode Determination

At present, if measured HST data is used to determine the cooling mode, we select one of the three windings with the highest mean HST from the given training data to determine the cooling modes for the entire training data-set during model building. Another possible way to improve the accuracy of the model predictions is to use the hottest winding at each time step while determining the cooling modes during model building, for those transformers which use measured HST to determine the cooling mode. However, very few transformers have the fiber-optic instruments needed to measure the

HST for all windings. Hence most of the transformers whose data was available to us, used simulated HST to determine the cooling mode. Only one of the transformers, Cooley-3, used measured HST to determine the cooling modes. Thus we conducted a numerical experiment using data from the Cooley-3 transformer where the cooling modes were determined by the winding with the highest measured HST at each time-step and the linear LS TOT model was built using data segregated in this way. (For this experiment, we used the measured data obtained from SRP where Cooley-3 was heavily loaded for two days. This is the same dataset that was used to test the performance in extrapolation as described in Chapter 4.) Initialization to the measured TOT was done at every point at which the transformer was predicted to enter FAFA cooling mode respectively for both linear as well as the IEEE models. The RMS errors for the TOT linear model when the model was built using the above method and when the model was built using a single winding to determine the cooling modes, for the two over-loaded days are given in Table 6. The experiment showed an improvement in predictions with more than 10% reduction in the RMS error in extrapolation when the winding to be used for cooling mode determination was determined at each time step as compared to when the same winding was used for the entire data-set.

Table 6 RMS Errors for the TOT Linear Model in Extrapolation

Date on which the transformer was over-loaded	RMS error with a single winding used for cooling mode determination	RMS error with different windings used for cooling mode determination	Percentage improvement
6/29/2013	1.32877	1.1864	10.71442
6/30/2013	1.803	1.5585	13.56073

As the new transformers being installed have the fiber-optic instruments necessary to measure the winding temperature, this method may help improve the predictions of the linear models for these transformers.

5.5 Summary

It was shown that the linear model predictions are more accurate than the IEEE model predictions in OA and FA cooling modes. Since the linear model predictions at a given data point depend on the predictions in the previous data-point, using the more accurate linear models for OA and FA cooling modes can help improve the thermal model predictions in FAFA cooling mode due to the more accurate initialization used at the point at which the transformer is assumed to enter the FAFA cooling mode. Also the LAV method can be used instead of the LS method to obtain better linear models. The accuracy of predictions may also be improved by using two different models for the time periods when the insulation temperature is increasing and the insulation temperature is decreasing. Also the hottest winding can be selected at each data point while determining the cooling modes during model building for those transformers that have measured HST data available, in order to obtain more accurate data-to-cooling-mode assignment and consequently increasing the accuracy of temperature predictions.

One or more of the above methods can be used to improve the accuracy of the thermal model predictions and thus yield more reliable dynamic loading results.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

The research reported was directed towards detecting unreliable thermal models and improving the accuracy of the thermal model predictions for substation distribution transformers. The three aspects investigated in detail in this research are:

- Identifying reliable and unreliable thermal models based on various metrics. This involved determining the quality of a thermal model in terms of easily comprehensible terminology such as 'Excellent', 'Good', 'Fair', 'Poor' and 'Unacceptable'. Metrics such as correlation coefficients, RMS errors, time constants, residual plot quality and model coefficients were used to arrive at the 'Model Quality'.
- Identifying the possible causes behind unreliable thermal models. This involved judging if the actual fan turn-on and turn-off points are compliant with the desired set points based on the residual plot shape and identifying if a transformer thermal model has a "V-shaped" residual plot.
- Improving the accuracy of the thermal model predictions by:
    - Using linear regression model for OA and FA cooling modes.
    - Using the least absolute value (LAV) method instead of least squares (LS) method to build a linear regression based thermal model.
    - Using two different models for the modes when the insulation temperature is increasing and decreasing.
    - Selecting the hottest winding at each data point while determining the cooling modes during model building for those transformers that have measured HST

data available, in order to obtain more accurate data-to-cooling-mode assignment.

Based on the above research, we have been able to refine the process of building thermal models and thus increase the reliability of the results obtained for the dynamic loading calculations.

Also a different application has been developed in order to automate the process of building both the IEEE and linear thermal models for transformers. The application uses historical measured data of load, TOT, HST and the ambient temperature to build the linear models, while the IEEE model requires the data from the PTLoad files. With the automatic model building application, the user needs to specify the folders which contain the transformer historical data files and the IEEE parameter files. The application builds the thermal models for all the transformers in the specified folder and provides the results of the model building process and comments about the model reliability to the user for each transformer. A .csv file, called 'XfmrModelParameters.csv', is generated which contains the data necessary for the dynamic loading application, DLTA, also known as the operator tool. Also a summary of the model building process is provided in a .csv file named ' Report_Xfmr_Thermal_Models.csv'. Further details about this application are provided in the appendix.

6.2 Future Work

While we have been using three-phase load to build our models, hottest-spot temperature is dependent on individual phase loads. Since there is unbalance in the phase loading of all transformers, we are hoping to record transformer phase currents to build better hottest-spot models. We also hope to be able to predict the amount of transformer unbalance by looking at certain metrics and plots we obtain for model screening

We are investigating whether introducing some nonlinearities into our model, such as a more complex time constant, for example that given by the IEEE model, or an oil-viscosity model, such as that proposed by Susa et al., may yield improved results.

Finally, research into applying similar metrics to those proposed here to the HST model for unreliable-model determination is ongoing. Also similar metrics need to be developed for the IEEE models.

REFERENCES

[1] IEEE Guide for Loading Mineral-Oil-Immersed Power Transformers, IEEE C57.91-1995.

[2] D. J. Tylavsky, Q. He, G. A. McCulla and J. R. Hunt, “Sources of Error and Quantization on Transformer Dynamic Thermal Modeling,” IEEE Transactions On Power Delivery, Vol. 15, No. 1, Jan. 2000, pp. 178-185.

[3] B. C. Lesieutre, W. H. Hagman and J. L. Kirtley Jr., “An Improved Transformer Top Oil Temperature Model for use in an On-line Monitoring and Diagnostic System,” IEEE Transactions On Power Delivery, Vol. 12, No. 1, Jan. 1997.

[4] Ali Abur and Antonio Exposito, “Power System State Estimation: Theory and Implementation,” Taylor & Francis, New York, 2004.

[5] G. Swift, T.S. Molinski and W. Lehn, “A Fundamental Approach to Transformer Thermal Modeling—Part I: Theory and Equivalent Circuit,” IEEE Transactions on Power Delivery, Vol. 16, No. 2, April 2001, pp. 171-175.

[6] G. Swift, T. S. Molinski and W. Lehn, “A Fundamental Approach to Transformer Thermal Modeling—Part II: Field Verification,” IEEE Transactions on Power Delivery, Vol. 16, No. 2, April 2001, pp. 176-180.

[7] L. W. Pierce, “An Investigation of the Thermal Performance of an Oil Filled Transformer Winding,” IEEE Transactions on Power Delivery, Vol. 7, No. 3, July 1992, pp. 1347-1356.

[8] L. W. Pierce, “Predicting Liquid Filled Transformer Loading Capability,” IEEE Transactions on Industry Applications, Vol. 30, No. 1, January/February 1994, pp. 170-178.

[9] D. Susa, M. Lehtonen and H. Nordman, “Dynamic Thermal Modeling of Power Transformers,” IEEE Transactions on Power Delivery, Vol. 20, No.1, January 2005, pp. 197-204.

[10] L. Jauregui-Rivera and D. J. Tylavsky, “Acceptability of Four Transformer Top-Oil Thermal Models—Part 1: Defining Metrics,” IEEE Transactions on Power Delivery, Vol. 23, No. 2, April 2008, pp. 860-865.

[11] L. Jauregui-Rivera and D. J. Tylavsky, “Acceptability of Four Transformer Top-Oil Thermal Models—Part 2: Comparing Metrics,” IEEE Transactions on Power Delivery, Vol. 23, No. 2, April 2008, pp. 866-872.

[12] D. J. Tylavsky, Xiaolin Mao and G.A.McCulla, “Data Screening to Improve Transformer Thermal Model Reliability,” Proceedings of the 37th Annual North American Power Symposium, pp.560-568, 23-25 Oct. 2005.

[13] Douglas C. Montgomery, Elizabeth A. Peck and G. Geoffrey Vining, “Introduction to Linear Regression Analysis,” Wiley Series in Probability and Statistics: Texts, References, and Pocketbooks Section, New York, 2001

[14] Lida. Jauregui-Rivera, “Reliability Assessment of Transformer Thermal Models,” Ph.D Dissertation, Arizona State Univ., Tempe, 2006.

[15] Oluwaseun Adeyemi Amoda, “Evaluation of Models for Predicting Hottest-spot Temperature in Substation Distribution Transformers”, Ph.D Dissertation, Arizona State Univ., Tempe, 2009.

[16] Ming Zhang, “Dynamic Loading of Substation Distribution Transformers: An Application for use in a Production Grade Environment”, M.S. Thesis Dissertation, Arizona State Univ., Tempe, 2013.

[17] Shruti Rao, Ming Zhang and Daniel Tylavsky, “Data Screening to Improve Transformer Thermal Model Reliability,” Proceedings of the 37th Annual North American Power Symposium, pp. 560-568, 23-25 Oct. 2013.

[18] Lida. Jauregui-Rivera, Xiaolin Mao and Daniel Tylavsky, “Improving Reliability Assessment of Transformer Thermal Top-Oil Model Parameters Estimated From Measured Data,” IEEE Transactions on Power Delivery, vol. 24, January 2009.

APPENDIX A

USER'S MANUAL FOR THE MODEL BUILDING APPLICATION

The dynamic loading application (DLTA) requires transformer thermal models in order to calculate the maximum dynamic load that it can sustain without violating the thermal limits of the insulation. Those thermal models are produced by the Model Building Tool (MBT) described below and output as the .csv file: "XfmrModelParameters.csv".

Overview:

Once the Model Building Tool (MBT) is started, the user will select the directories which contain the historical TOT, and HST data files and the IEEE parameter files and then the application will build the models for all the transformers that have sufficient data in those directories. All of the results, such as the model coefficients and comments about the model reliability, are stored in excel and .MAT files which are needed by DLTA and can be used by engineers to scrutinize the models built.

1 System Requirements

CPU: Intel Pentium 4 or above

Memory: At least 1 GB, recommended 2 GB

Disk Space: At least 500 MB

Operating System: Microsoft Windows 7, 64 bit

2 Other Prerequisites for Running the Model Building Tool (MBT)

The MBT must be run in a MATLAB environment. Verify that the MATLAB Compiler Runtime (MCR) is installed and ensure you have installed version 8 or later. If the MCR is not installed, run “MCRInstaller.exe”, which is provided on the CD under the folder "MCRInstaller" in the root directory "Model Building Application". The folder contains the "MCRInstaller.exe" files for 32-bit and 64-bit Windows computers. For more information about the MCR and the

MCR Installer, see “Working with the MCR?” in the MATLAB Compiler User Guide found at http://www.mathworks.com/help/compiler/working-with-the-mcr.html.

The MCRInstaller can also be downloaded from:

http://www.mathworks.com/products/compiler/mcr/

The version downloaded must be version 8.0 or later.

Note: You will need administrator rights to run MCRInstaller.

3 *Install the MBT*

Copy the entire "Model Building Application" directory from the CD to the location where you want the program to be installed, e.g. C:\myprogram.

The program to be run is “MBT.exe”. You may want to create a shortcut for it on your desktop. However the original executable file must be in the folder "Model Building Application". This folder contains some data files necessary for the application to run. Please note that if you want to run it from the desktop, it is necessary to create a shortcut for it. Copying the file to desktop will not suffice since the folder that contains "MBT.exe" must also contain the other data files necessary for it to run successfully.

4 *Run the MBT*

Double click the "MBT.exe" file. The main interface of the program will appear as shown in Figure 1.

Note: A pop-up window may appear and ask for administrator rights to run the application. You should click on ‘Yes’ in order to give the authority to run the application.

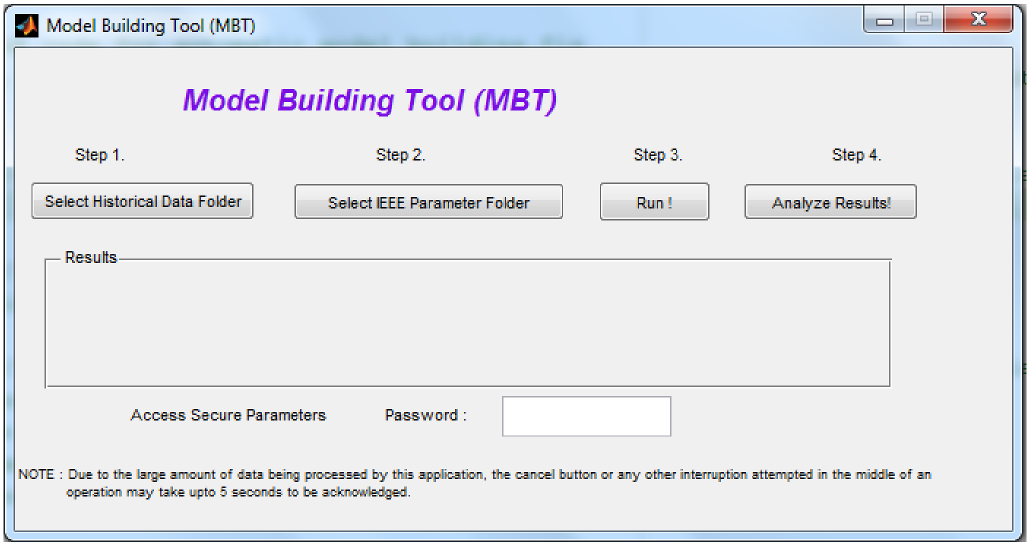


Figure 1 Main Interface of the MBT

Step 1: Select the folder containing the historical TOT and HST files.

The main interface of the model building application is shown in Figure 1. A pop-up window will appear when you click on the push button labeled "Select Historical Data Folder" on the main interface. The interface of the pop-up window is as shown in Figure 2. Browse to the folder which contains the historical data files and click on the button labeled "Select Folder". (The required format for the historical data files is given in Section 5.2.) Also it is important to note that each transformer should have <u>*only*</u> <u>*one*</u> data file in that folder. (If multiple files for the same transformer are present, the results for the file(s) processed first will be over-written by the results for the file processed last.)

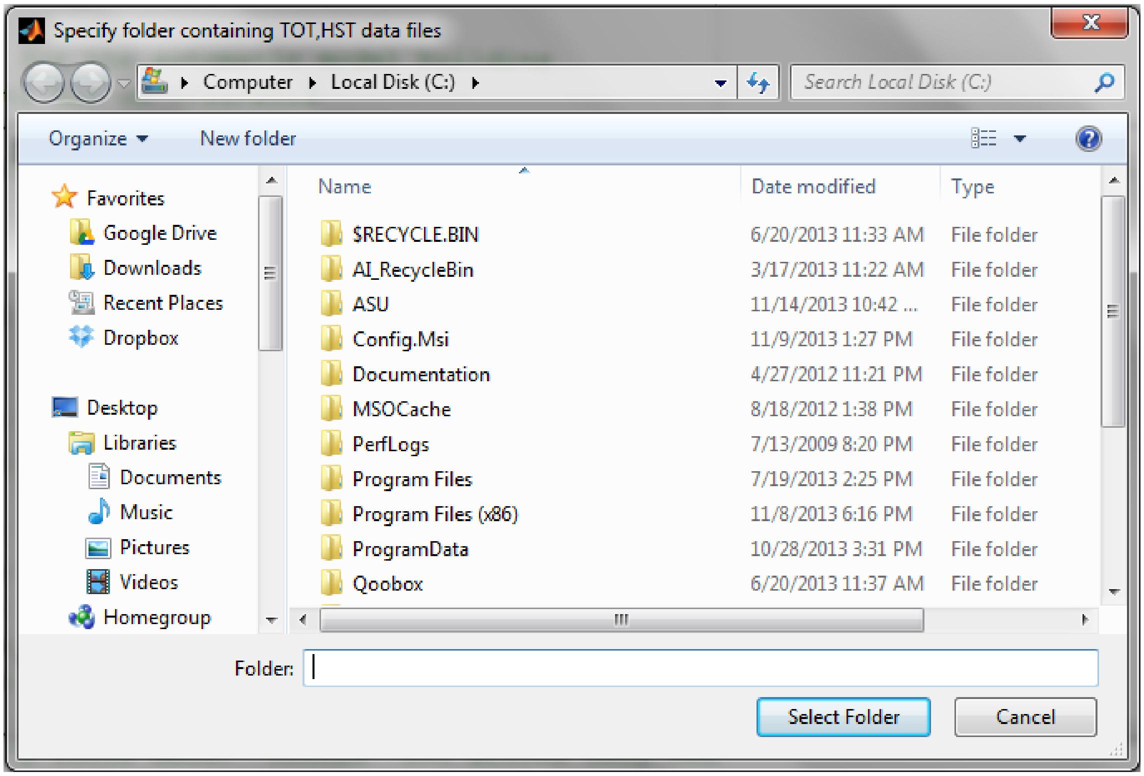


Figure 2 Interface to Select the Directory Containing TOT, HST Historical Data Files

Step 2: Select the folder containing the IEEE parameter data files.

A pop-up window will appear when you click on the push button labeled "Select IEEE Parameter Folder" on the main interface of MBT. The interface of the pop-up window is similar to the figure as shown in Figure 2. Browse to the folder containing the IEEE parameter files and click on "Select Folder". The IEEE parameter files are the "*.run" files obtained from EPRI software PTLoad.

Optional step: Modify secure parameters used for the model building process:

The Access Secure Parameters panel in the MBT window displays the default parameters used in the program which can be modified by the authorized user if he wishes to do so by typing “SRP” in the password field. These parameters include thresholds used in the model building process such as the maximum TOT and HST for steady-state load rating calculation, desired R squared and VIF values for the model screening. The explanation of these variables is provided in Table 7. The window that pops up when the correct password is entered looks as shown in Figure 3. The user can modify the parameters as desired and then click on "OK" to save the changes. If the user wishes, he can revert to the default parameters by clicking on "Restore".

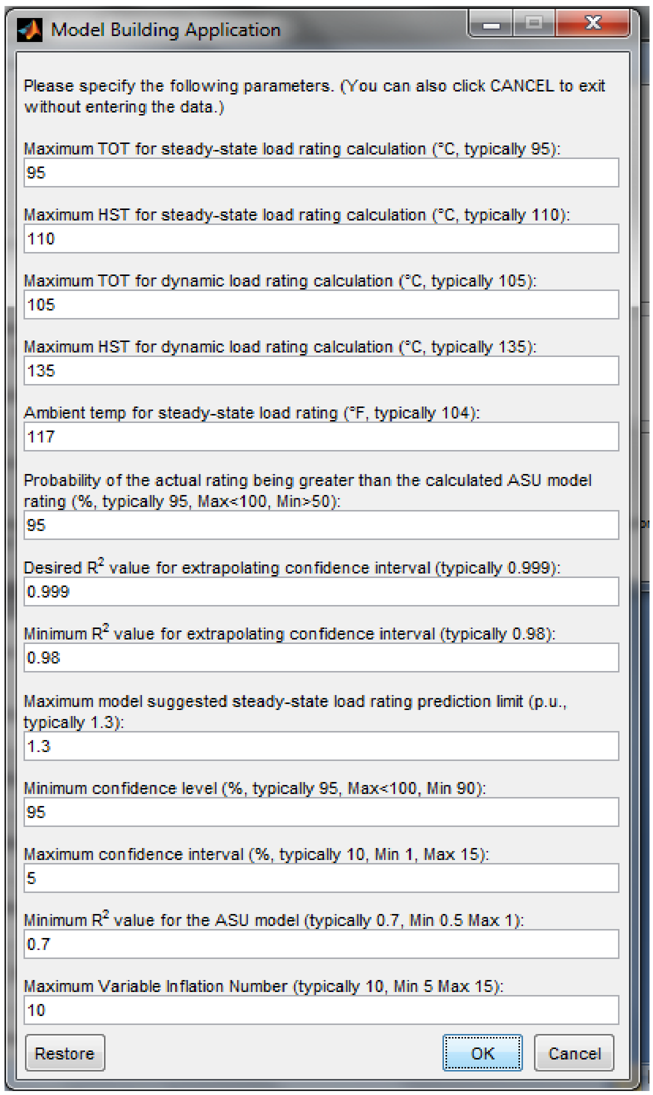


Figure 3 Access Secure Parameters Window

Table 7 Explanation of the variables in "Access Secure Parameters" Window

Maximum Steady state load rating (SSL_{Max})	The maximum load to which a transformer can be subjected, without violating the defined TOT or HST limits under steady state conditions
R^2 (Coefficient of determination)	Determines how well the data fits the linear regression model
VIF (Variance Inflation Factor)	Indicates if the predictor variables in the linear regression models are correlated with each other

Step 3[1]: Run!

Once the first two steps are completed, click on the push button "Run!". A pop-up message will appear similar to as shown in Figure 4. If the user selects "Yes", the model building process will start. A progress-bar will appear as shown in Figure 5. The progress-bar is provided with a Cancel button. Due to the large amount of data being processed during model building, the application may take a long time to respond to any interruption that may be attempted during the execution of the model building process. Please allow up to 5 seconds for the Cancel button or any other interruption to be acknowledged. If the user clicks on the small "x" on the top right corner of the progress-bar, the progress-bar window will close temporarily. When the program updates it again, the window will reappear.

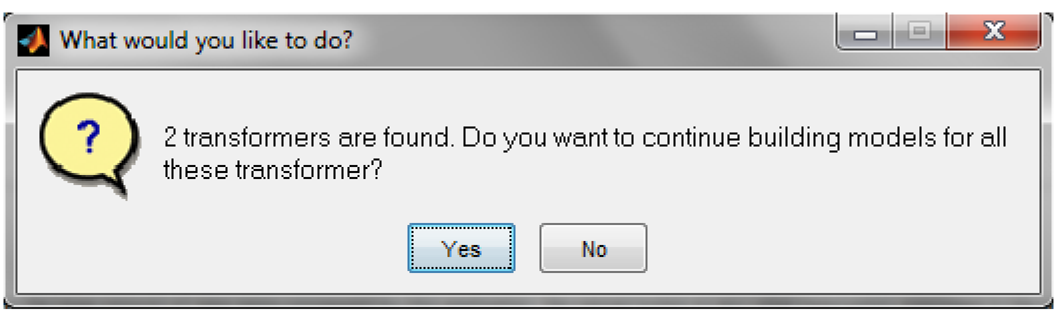


Figure 4 Pop-Up Message that Appears on Clicking "Run!"

[1] Be sure that you have selected the folders containing IEEE parameter and historical data files. If these steps have not been executed before starting the model building process, the application will give an error message, requesting that you select the folders.

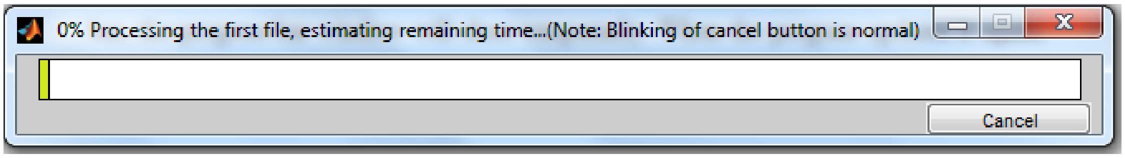


Figure 5 Progress-Bar

Once the program acknowledges the cancel request, a warning message will appear as shown in Figure 6. If the user selects "No", the message will disappear and the program will resume operation. If the user clicks on "Yes", the process will be cancelled and the results for the transformer data file that were completely processed will be available for the user to analyze.

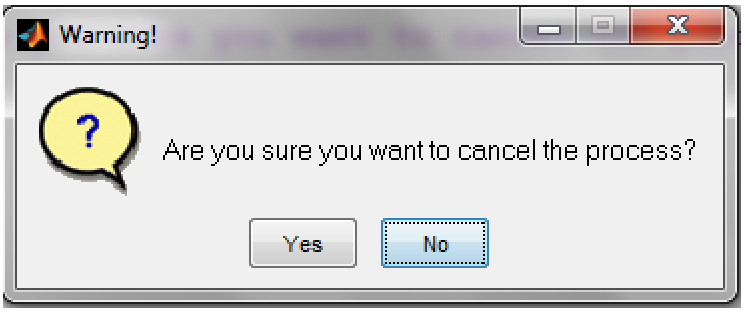


Figure 6 Warning Message on Clicking "Cancel"

If the user attempts to interrupt the program in any other way, for instance if he clicks on "Select Historical Data folder" during the execution of model building process, the program will complete the current task and then acknowledge the interruption. An error message will pop-up as shown in Figure 7.

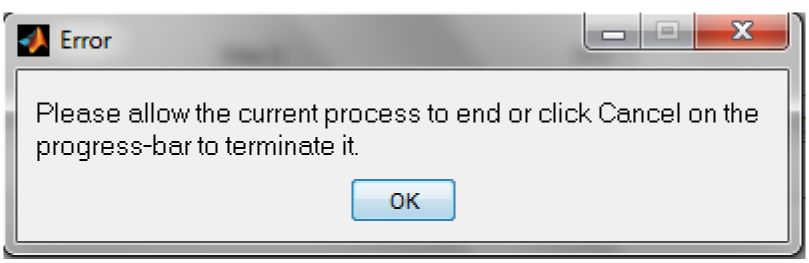


Figure 7 Error Message

Once the model building process has started, the "Results" panel in Figure 1 is updated as and when each file is processed. This is shown in Figure 8. Once all the data files have been processed, the "Results" panel displays the message "Model building process is complete."

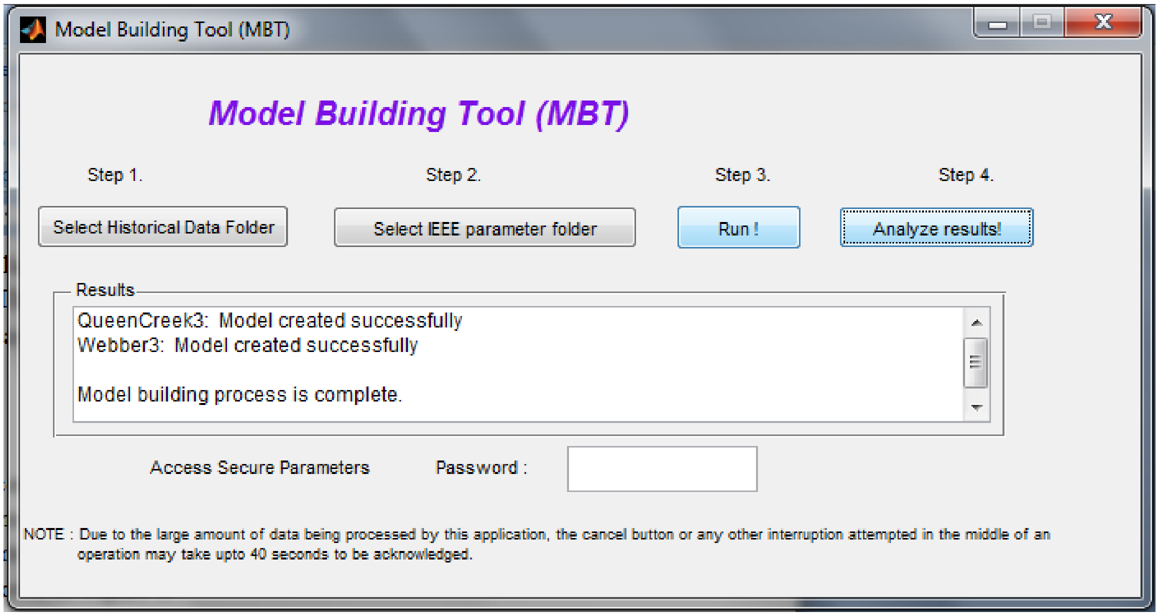


Figure 8 Updated Results Window as Files are Processed

If the user clicks on the "x" on the top right corner of the main interface, a warning message as shown in Figure 9 will appear. If the user clicks on "Yes", the application will terminate its operation and close the window. If the user clicks on "No", the warning message will disappear and the application will resume its operation.

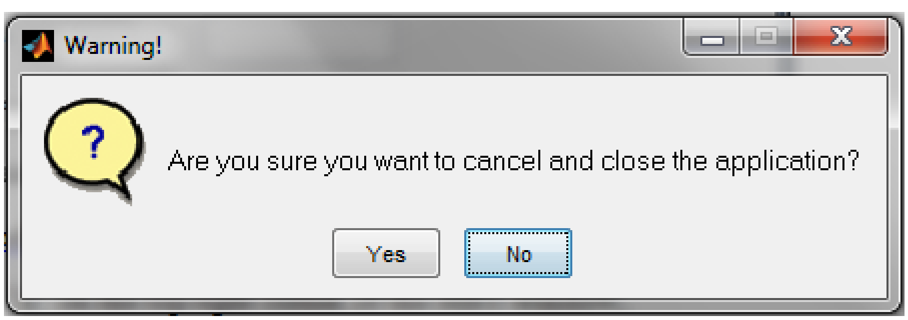


Figure 9 Warning Message on Attempting to Close the MBT Window

Step 4: Analyze Results!

In this step, the user can view the detailed information about the transformer TOT and HST models built and compare the performances of the linear and IEEE models.

If the user clicks on "Analyze Results!", a new window titled "Analyze results" will appear as shown in Figure 10.

Note: This button will automatically be disabled whenever the model building process is in progress. The user can click on it only before starting a new process or after the termination of the present process. If the window of "Analyze results" is already open and the user starts the model building process, he will get a warning message requesting him to wait until the completion of the process before trying to view the model-building reports.

In order to view the model-building reports for a particular transformer, the user needs to first select the desired transformer from the drop-down list provided in Step 1 of "Analyze results" window. The drop-down list contains a list of names of all transformer data files processed. The user then needs to select the results that he wishes to see. They can select one or more options from the following:

1. A window which contains the summary of the TOT and HST model coefficients, the load duration curve of the load data used to build this models, prediction error duration curves for the data sets used to train the data and the normal probability distribution plots for SSL_{Max} predicted (assuming TOT and HST values as the limiting criteria.).

2. A window containing a plot of the measured TOT and the TOT values predicted by the linear-model and the IEEE-model.

3. A window containing a plot of the measured HST and the HST values predicted by the linear-model and the IEEE-model.

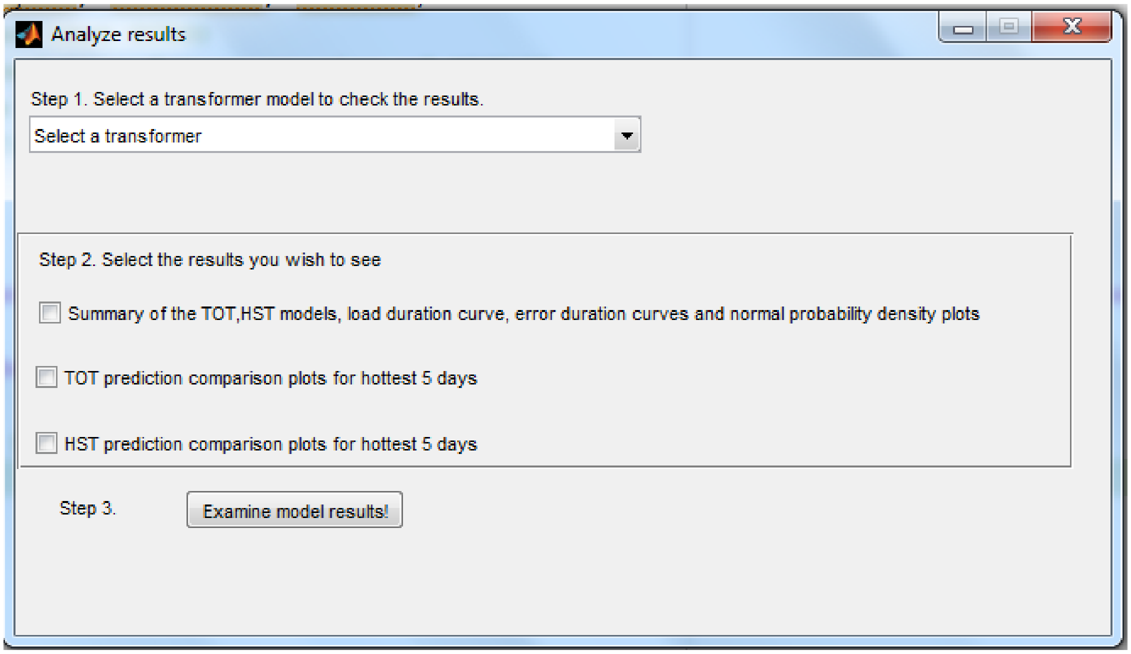


Figure 10 Analyze Results Window

If either of the two steps is skipped, an error message will appear, requesting the user to complete it.

The window that has the summary of the model details, load duration curve, error duration curves and probability density plots is shown in Figure 11. If the measured input data is corrupt or insufficient to build a reliable linear model, the application mentions that in the comments section. (In this case the IEEE models built by this application will be used by DLTA for dynamic loading provided these models are acceptable.)

The TOT prediction comparison and HST prediction comparison plots[2] are shown in Figure 12 and Figure 13.

[2] In order to generate these plots, the corresponding results for all transformer models are stored in .MAT files in a folder named "Model Results" in the same directory as the "MBT.exe" file.

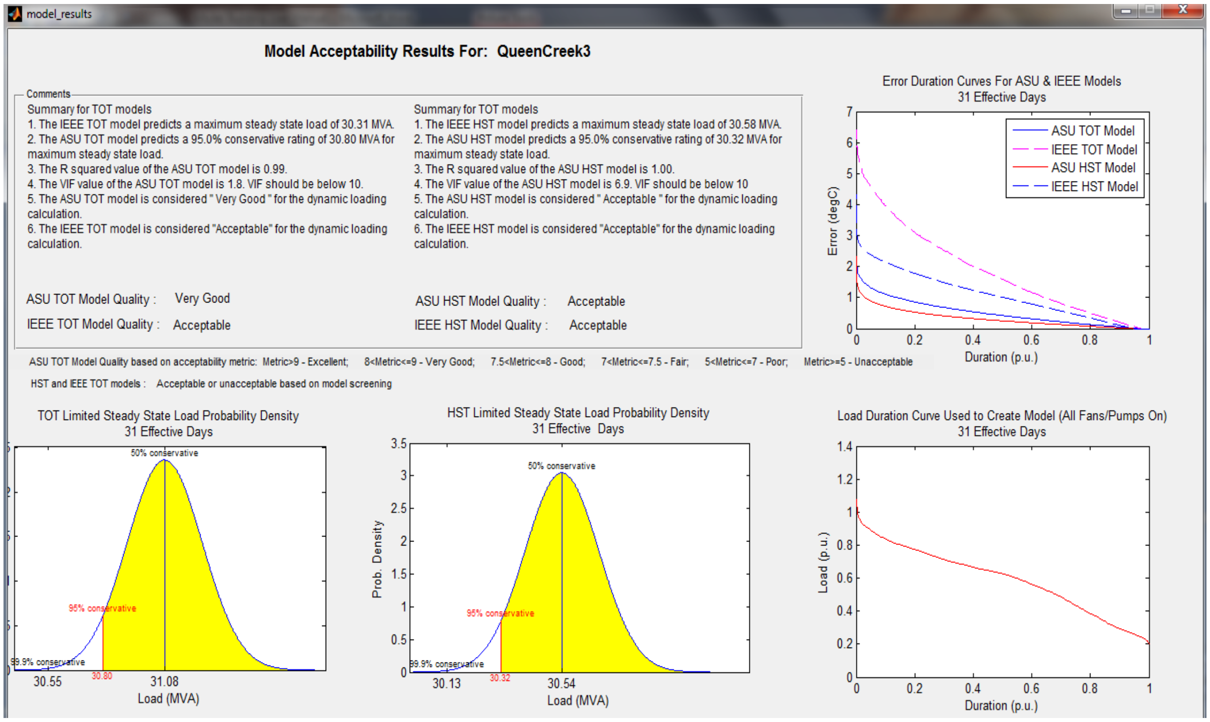


Figure 11 Model Results Window

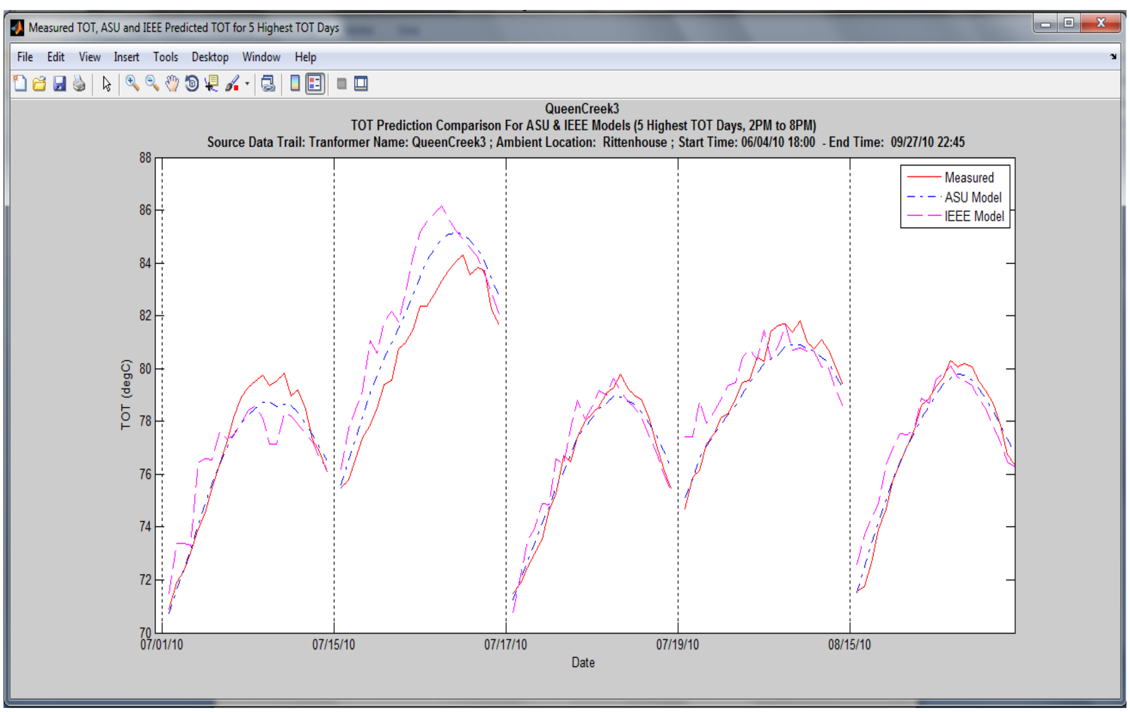


Figure 12 TOT Prediction Comparison Plot for Hottest Five Days

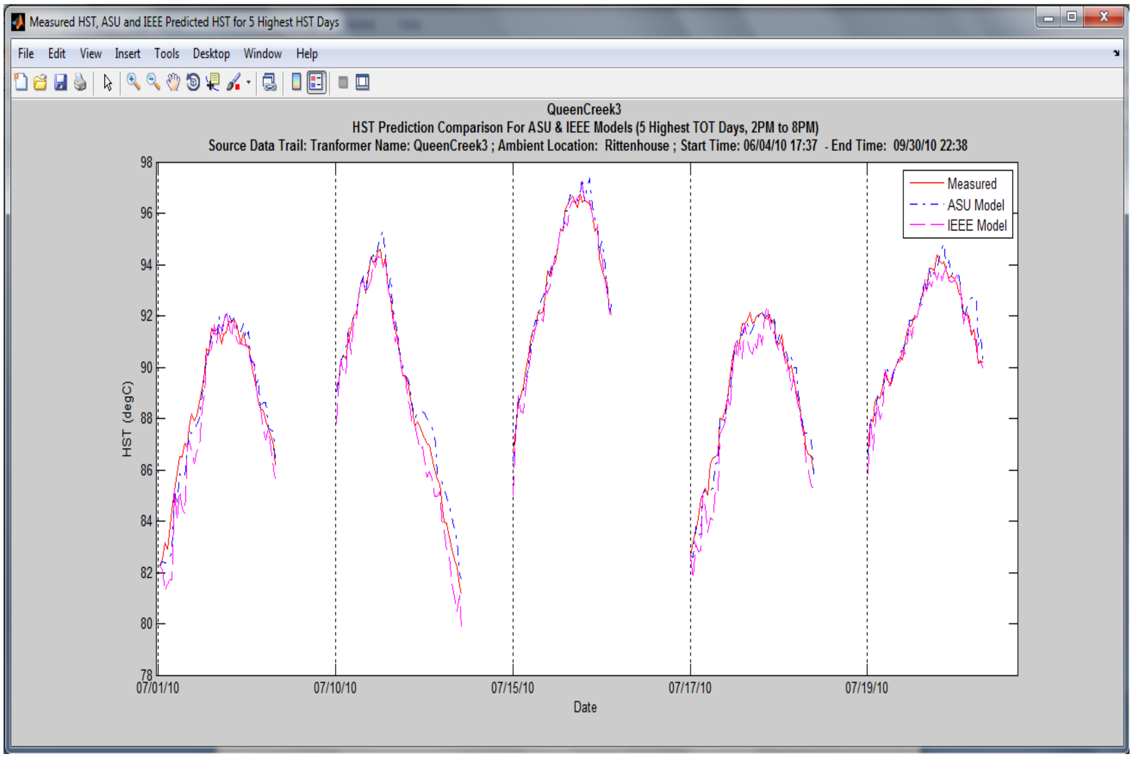


Figure 13 HST Prediction Comparison Plot for Hottest Five Days

In addition to these results, two excel files are generated in the folder "Model Results". The first file "Report_Xfmr_Thermal_Models.csv" contains a report of the entire model building process, including the number of data files processed, the number of failed and successful models and a summary of all the transformer models. If a model cannot be built successfully for any transformer, the cause for the failure is given in the summary. If this file or any of the other .MAT files are modified by reading them into a text editor and editing them or are deleted by the user, the user may not be able to view the model building results using the “Analyze Results!” option within the MBT main interface.

The second excel file "XfmrModelParameters.csv" contains all the linear and IEEE model parameters which will be required for dynamic loading. It is important that this file is not modified as it is required for the operator tool, DLTA, to function. See the DLTA user's manual for information on how this file is to be used. The file contains the detailed descriptions of the contents of all the columns in the file.

5 *Required Formats*

5.1 *Installation Directory*

The "Model Building Application" folder in Figure 14 is the installation folder that may be located anywhere on the hard drive.

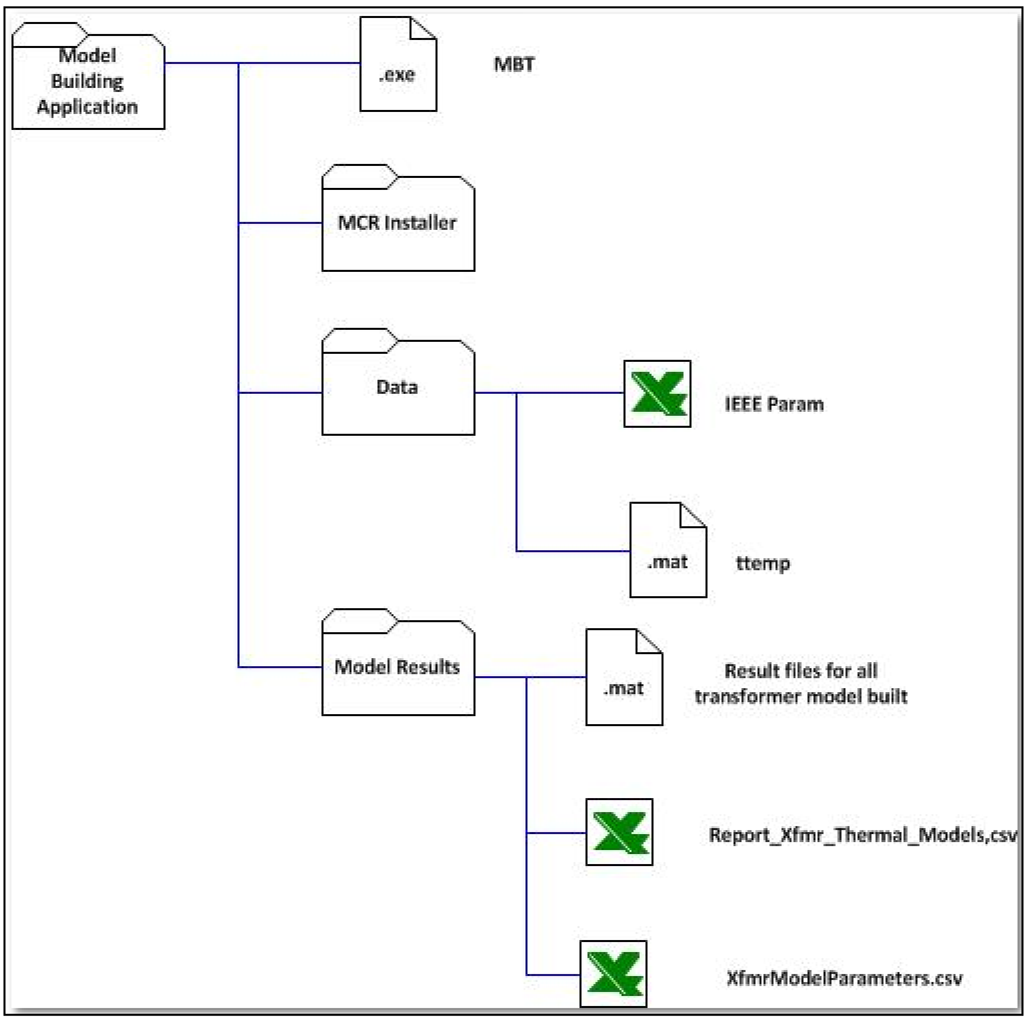


Figure 14 Directory Structure of the MBT

In Figure 14, "MCR Installer" is a folder that contains the runtime engine, MATLAB Compiler Runtime (MCR) installation files. The ".exe" is the executable file to launch the application. "Data" is a folder that contains the files IEEE param.csv and ttemp.mat that are necessary for the application to run. The folder "Model Results" contains the .MAT files and excel files that contain the results of the model building process.

Note: All the folders and file names should be strictly in accordance with what is shown in Figure 14. The user should not change the name of the folders or files.

5.2 Historical TOT and HST Data Files (.txt Files)*

The historical TOT and HST data files (*.txt files) are the files that contain measured data which are supplied by the user. The requirements for the format of real-time data files are as follows:

a. The data file should be in Text (Tab Delimited) format.
b. There should be one data file that contains all measured data associated with any given transformer (including ambient temperature data.)
c. First nonblank line should contain transformer name. The next line should contain a text string specifying the name of the ambient temperature location.
d. The column heading title line should immediately follow the ambient temperature location line. The first column heading title must be “Date” (no quotes in data file). Other column heading titles are:

   “MVA” (Load at the sample time, MVA) (mandatory)

   “AIR” (Primary ambient temperature, ºF) (mandatory)

   “OIL” (Top oil temperature, ºC) (mandatory)

“WindingA” (Hottest spot temperature at WindingA, ºC) (mandatory)

“WindingB” (Hottest spot temperature at WindingB, ºC) (mandatory)

“WindingC” (Hottest spot temperature at WindingC, ºC) (mandatory)

“QC” (Quality Control Index) (mandatory)

“Tamb1” (Secondary ambient temperature, ºF) (optional)

e. Immediately following the column heading title line should be the first row of data and the data should start from midnight (00:00:00).

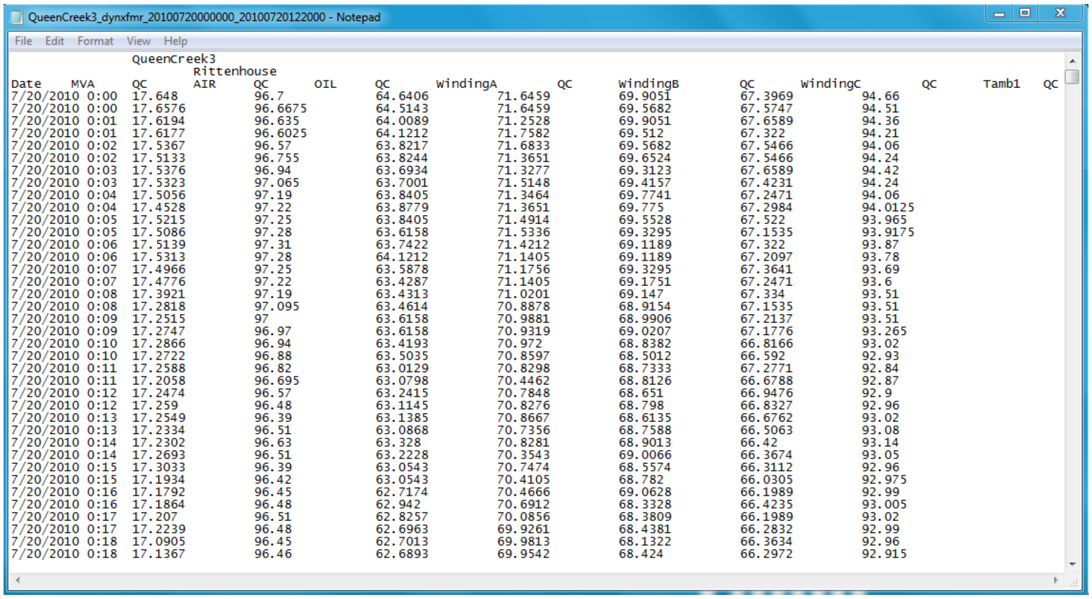
f. Column heading titles should match data.

g. A “QC” column is considered as the data quality check for the data in the immediately preceding column. The data file may contain multiple “QC” columns. If one or more “QC” values are found to be “?” in a row, that whole row of data will be discarded.

h. The secondary ambient temperature is the temperature that would be used as back up in dynamic loading calculation if AIR has missing data.

Note: To create a *.txt file, you can create a data file in an Excel format and then save it as Text (Tab delimited) format.

A sample of the historical data file format is provided in Figure 15.



QueenCreek3_dynxfmr_20100720000000_20100720122000 - Notepad

File Edit Format View Help

		QueenCreek3												
			Rittenhouse											
Date	MVA	QC	AIR	QC	OIL	QC	WindingA	QC	WindingB	QC	WindingC	QC	Tamb1	QC
7/20/2010 0:00	17.648		96.7		64.6406		71.6459		69.9051		67.3969		94.66	
7/20/2010 0:00	17.6576		96.6675		64.5143		71.6459		69.5682		67.5747		94.51	
7/20/2010 0:01	17.6194		96.635		64.0089		71.2528		69.9051		67.6589		94.36	
7/20/2010 0:01	17.6177		96.6025		64.1212		71.7582		69.512		67.322		94.21	
7/20/2010 0:02	17.5367		96.57		63.8217		71.6833		69.5682		67.5466		94.06	
7/20/2010 0:02	17.5133		96.755		63.8244		71.3651		69.6524		67.5466		94.24	
7/20/2010 0:03	17.5376		96.94		63.6934		71.3277		69.3123		67.6589		94.42	
7/20/2010 0:03	17.5323		97.065		63.7001		71.5148		69.4157		67.4231		94.24	
7/20/2010 0:04	17.5056		97.19		63.8405		71.3464		69.7741		67.2471		94.06	
7/20/2010 0:04	17.4528		97.22		63.8779		71.3651		69.775		67.2984		94.0125	
7/20/2010 0:05	17.5215		97.25		63.8405		71.4914		69.5528		67.522		93.965	
7/20/2010 0:05	17.5086		97.28		63.6158		71.5336		69.3295		67.1535		93.9175	
7/20/2010 0:06	17.5139		97.31		63.7422		71.4212		69.1189		67.322		93.87	
7/20/2010 0:06	17.5313		97.28		64.1212		71.1405		69.1189		67.2097		93.78	
7/20/2010 0:07	17.4966		97.25		63.5878		71.1756		69.3295		67.3641		93.69	
7/20/2010 0:07	17.4776		97.22		63.4287		71.1405		69.1751		67.2471		93.6	
7/20/2010 0:08	17.3921		97.19		63.4313		71.0201		69.147		67.334		93.51	
7/20/2010 0:08	17.2818		97.095		63.4614		70.8878		68.9154		67.1535		93.51	
7/20/2010 0:09	17.2515		97		63.6158		70.9881		68.9906		67.2137		93.51	
7/20/2010 0:09	17.2747		96.97		63.6158		70.9319		69.0207		67.1776		93.265	
7/20/2010 0:10	17.2866		96.94		63.4193		70.972		68.8382		66.8166		93.02	
7/20/2010 0:10	17.2722		96.88		63.5035		70.8597		68.5012		66.592		92.93	
7/20/2010 0:11	17.2588		96.82		63.0129		70.8298		68.7333		67.2771		92.84	
7/20/2010 0:11	17.2058		96.695		63.0798		70.4462		68.8126		66.6788		92.87	
7/20/2010 0:12	17.2474		96.57		63.2415		70.7848		68.651		66.9476		92.9	
7/20/2010 0:12	17.259		96.48		63.1145		70.8276		68.798		66.8327		92.96	
7/20/2010 0:13	17.2549		96.39		63.1385		70.8667		68.6135		66.6762		93.02	
7/20/2010 0:13	17.2334		96.51		63.0868		70.7356		68.7588		66.5063		93.08	
7/20/2010 0:14	17.2302		96.63		63.328		70.8281		68.9013		66.42		93.14	
7/20/2010 0:14	17.2693		96.51		63.2228		70.3543		69.0066		66.3674		93.05	
7/20/2010 0:15	17.3033		96.39		63.0543		70.7474		68.5574		66.3112		92.96	
7/20/2010 0:15	17.1934		96.42		63.0543		70.4105		68.782		66.0305		92.975	
7/20/2010 0:16	17.1792		96.45		62.7174		70.4666		69.0628		66.1989		92.99	
7/20/2010 0:16	17.1864		96.48		62.942		70.6912		68.3328		66.4235		93.005	
7/20/2010 0:17	17.207		96.51		62.8257		70.0856		68.3809		66.1989		93.02	
7/20/2010 0:17	17.2239		96.48		62.6963		69.9261		68.4381		66.2832		92.99	
7/20/2010 0:18	17.0905		96.45		62.7013		69.9813		68.1322		66.3634		92.96	
7/20/2010 0:18	17.1367		96.46		62.6893		69.9542		68.424		66.2972		92.915	

Figure 15 A Sample of the Real-Time Data File Format