Acceptability of Four Transformer Top-Oil Thermal Models—Part II: Comparing Metrics

Lida Jauregui-Rivera, *Member, IEEE*, and Daniel J. Tylavsky, *Senior Member, IEEE*

***Abstract*—The acceptability of four transformer top-oil thermal models is examined vis-a-vis training with measured data. Acceptability is defined as having the qualities of adequacy, consistency, and accuracy. Metrics are used to characterize the likely deficiencies of these models. It is shown that the classical IEEE/ANSI standard model is unacceptable for model identification purposes. The linear top-oil model is acceptable for FOFA transformers but not NOFA. The models by Susa *et al.* [7] and Swift *et al.* [5], [6], may be useful for FOFA transformers and NOFA transformers. Further research with larger training data sets is warranted.**

***Index Terms*—ANSI C57.91, system identification, top-oil temperature, transformer, transformer thermal modeling.**

I. Introduction

IN a companion paper, we defined four different transformer top-oil models and metrics to determine the acceptability of these models [2]. In this paper, we apply these metrics to models constructed for two different transformers: one is rated at 167 MVA (FOFA) and the other is at 28 MVA (NOFA). We chose transformers with very different ratings and cooling modes to see how the various models will handle a wider range of conditions. The heat-run values and basic data for these transformers are provided in the Appendix.

II. Data Collection

Top-oil temperature (TOT), load, and ambient temperature were sampled every 15 min. The data were filtered to eliminate bad data and divided into separate data files for each of the three different cooling modes as described in [4]. The models built in this work use only the highest cooling mode—NOFA or FOFA. The data sets for training each model included 30 effective days of data, that is, 2880 data points.

III. Constructing the X Matrix

For the linear model, the X matrix contains measured values of I^2, θ_{amb}, and, as a lagged regressor θ_{top}, as shown in (1), where K_3 corresponds to β_0 in the companion paper. For the nonlinear problem, the iteration procedure is defined by (2), where $\boldsymbol{\beta} = \lfloor T_0 \quad \theta_{fl} \quad n \quad R \rfloor$ is a 4×1 vector of the regression coefficients (model parameters.) The X matrix for the nonlinear problem is that given by (3). The derivatives for all models are too lengthy to include here but may be found in [9]

$$\begin{bmatrix} I(1)^2 & \theta_{\mathrm{amb}}(1) - \theta_{\mathrm{top}}(0) \\ I(2)^2 & \theta_{\mathrm{amb}}(2) - \theta_{\mathrm{top}}(1) \\ \vdots & \vdots \\ I(n)^2 & \theta_{\mathrm{amb}}(n) - \theta_{\mathrm{top}}(n-1) \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} + \begin{bmatrix} K_3 \\ K_3 \\ \vdots \\ K_3 \end{bmatrix} = \begin{bmatrix} \theta_{\mathrm{top}}(1) \\ \theta_{\mathrm{top}}(2) \\ \vdots \\ \theta_{\mathrm{top}}(n) \end{bmatrix} \tag{1}$$

$$\hat{\boldsymbol{\beta}}^{\mathbf{k+1}} = \left(\mathbf{X^T}\left(\hat{\boldsymbol{\beta}}^{\mathbf{k}}\right)\mathbf{X}\left(\hat{\boldsymbol{\beta}}^{\mathbf{k}}\right)\right)^{-1}\mathbf{X^T}\left(\hat{\boldsymbol{\beta}}^{\mathbf{k}}\right)\left(\mathbf{y} - \boldsymbol{h}\left(\hat{\boldsymbol{\beta}}^{\mathbf{k}}\right)\right) + \hat{\boldsymbol{\beta}}^{\mathbf{k}} \tag{2}$$

$$\boldsymbol{X}\left(\beta^k\right) = \begin{bmatrix} \frac{\partial \boldsymbol{h_1}(\boldsymbol{\beta})}{\partial T_0} & \frac{\partial \boldsymbol{h_1}(\boldsymbol{\beta})}{\partial \theta_{fl}} & \frac{\partial \boldsymbol{h_1}(\boldsymbol{\beta})}{\partial n} & \frac{\partial \boldsymbol{h_1}(\boldsymbol{\beta})}{\partial R} \\ \frac{\partial \boldsymbol{h_2}(\boldsymbol{\beta})}{\partial \mathbf{T}_0} & \frac{\partial \boldsymbol{h_2}(\boldsymbol{\beta})}{\partial \theta_{fl}} & \frac{\partial \boldsymbol{h_2}(\boldsymbol{\beta})}{\partial n} & \frac{\partial \boldsymbol{h_2}(\boldsymbol{\beta})}{\partial R} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \boldsymbol{h_n}(\boldsymbol{\beta})}{\partial \mathbf{T}_0} & \frac{\partial \boldsymbol{h_n}(\boldsymbol{\beta})}{\partial \theta_{fl}} & \frac{\partial \boldsymbol{h_n}(\boldsymbol{\beta})}{\partial n} & \frac{\partial \boldsymbol{h_n}(\boldsymbol{\beta})}{\partial R} \end{bmatrix}_{\boldsymbol{\beta}^{\mathbf{k}}}. \tag{3}$$

IV. Consistency Metrics

The first metrics we look at are the consistency metrics, those measures that tell us how reliably we will be able to obtain the same model parameters from similar data. If a model and method yield inconsistent results, then there is little use in pursuing that model further.

A. Eigenvalues and Singular Values

We are interested in two different sets of eigenvalues and condition numbers for these transformers: 1) values calculated using nominal parameters and 2) values calculated at the solution point. The condition number $\boldsymbol{k}\left(\boldsymbol{X^T X}\right)$, evaluated using nominal parameters predicts whether an iterative algorithm has an initial direction which will reliably point toward the optimum. At the solution point $\boldsymbol{\beta}^*$, the eigenvalues and condition number $\boldsymbol{k}^*\left(\boldsymbol{X^T X}\right)$ measure how sensitive the solution point is to noise in the input data. The condition numbers at nominal and optimal parameter values for the Corbell3 (28 MVA) and DV6 (167 MVA) transformers are shown in Tables I and II,

Manuscript received July 26, 2006. This work was supported in part by the Salt River Project, in part by Arizona Public Service, and in part by PSERC. Paper no. TPWRD-00417-2006.

The authors are with Arizona State University, Tempe, AZ 85287-5706 USA (e-mail: Lida.Jauregui-Rivera@aps.com; tylavsky@asu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRD.2007.905576

TABLE I
EIGENVALUES AND CONDITION NUMBERS FOR CORBELL3 FOR NOMINAL PARAMETERS

Corbell3				
Quantity	**LTOP**	**NTOP**	**Susa**	**Swift**
λ_1	1.36	4.00E+04	4.52E+04	4.87E+04
λ_2	0.64	1.50E+02	1.89E+01	2.77E+02
λ_3	---	7.80E-13	8.24E-01	1.36
λ_4	---	-9.06E-15	3.35E-03	6.23E-04
$\kappa(X^TX)$	**2.13**	**2.98E+20**	**1.35E+07**	**7.78E+07**
$\kappa^*(X^TX)$	**2.13**	**6.54E+18**	**3.56E+07**	**9.20E+07**

TABLE II
EIGENVALUES AND CONDITION NUMBERS FOR DV6 FOR NOMINAL PARAMETERS

DV6				
Quantity	**LTOP**	**NTOP**	**Susa**	**Swift**
λ_1	1.79	3.85E+04	2.85E+04	9.76E+03
λ_2	0.21	2.75E+02	9.06E+01	6.66E+01
λ_3	---	3.37E-09	2.01E+00	2.26
λ_4	---	-9.87E-14	4.99E-02	7.07E-02
$\kappa(X^TX)$	**8.66**	**6.25E+19**	**5.69E+05**	**1.38E+06**
$\kappa^*(X^TX)$	**8.66**	**3.66E+24**	**1.44E+05**	**1.78E+06**

respectively. The eigenvalues of $\mathbf{X^TX}$ in Tables I and II use nominal parameters. The nomenclature used in these tables is: LTOP is the linearized Clause 7 model [3], NTOP is the Clause 7 model [1], Susa is the Susa *et al.* model [7], and Swift is the Swift *et al.* model [5], [6].

Notice that the eigenvalues and condition numbers in these tables are similar for both transformers, indicating that this metric is determined more by the model than by the transformer type.

The condition number of the LTOP model is much better than that of any other model for both transformers. For DV6, the nominal value of n is 1.0 for all models except the Susa *et al.* [7] model. For $n = 1.0$, the NTOP and Swift models are identical to LTOP. Then why in Table II is there such a large difference in the condition numbers between the LTOP, NTOP, and Swift models? The difference is not caused by the input data since all models share the same data. The large difference arises because the LTOP model is linear in the parameters (K_1, K_2, K_3) while the nonlinear models are nonlinear in the parameters $\boldsymbol{\beta} = \lfloor T_0 \quad \theta_{fl} \quad n \quad R \rfloor$, and it is the derivatives with respect to those parameters that are entries into the $\mathbf{X}$ matrix.

Due to the large condition number of the NTOP, the iteration matrix at the nameplate (nominal) values is numerically singular. (For nominal values for these transformers, refer to Tables XII and XIII.) The practical consequence of this is that the iteration scheme defined by (2) does not converge, except for the trivial case of the linear model and in certain cases for the Swift *et al.* [5], [6] model. Where the Gauss–Newton scheme did not converge for us, we used a more robust MatLab optimization routine, FMINSEARCH.

TABLE III
PARAMETER SENSITIVITIES FOR CORBELL3 FOR NOMINAL PARAMETERS VALUES

Corbell3				
Parameter	**LTOP**	**NTOP**	**Susa**	**Swift**
R or K_1	0.019	1.1E+05	5.58	40.00
T_0 or K_2	0.019	2.3E-02	0.09	1.32
θ_{fl} or K_3	0.018	5.0E+05	1.22	1.61
n	---	5.8E+03	0.39	0.12

The large condition number of the NTOP model indicates the model is not well posed. References [5]–[7] have shown through first-principle derivations that the n exponent in the NTOP model is misplaced. The ANSI/IEEE standard [1] states that the placement of the NTOP-model exponent was ad hoc, derived, "empirically" to account for the "variation of the (top-oil temperature rise above ambient caused by the) … change in resistance with change in load." Thus, it is not surprising that the NTOP model may be inappropriate.

Another question arises: The Susa and Swift models are very similar to the NTOP model. Why are their condition numbers so much better? The eigenvalues (and condition numbers) are sensitive to small changes in entries of the X matrix since $\mathbf{X^TX}$ is close to singular. The small changes in these models are enough to move $\mathbf{X^TX}$ further from a singular matrix. This also suggests that the Susa and Swift models more accurately capture the features of the thermal process than the NTOP model.

When the $\mathbf{X^TX}$ matrix is well conditioned, the parameters derived from the model will be more consistent. When consistency is seen in the solution (investigated later), then the model (of necessity) must match the process it is modeling more closely. Since the $\mathbf{X^TX}$ matrix of the Swift and Susa models are also ill-conditioned, we would expect them to have consistency and convergence issue, as indeed they have.

Notice that the condition numbers evaluated using optimal parameters values $\boldsymbol{k}^*(\boldsymbol{X}^T\boldsymbol{X})$ in Tables I and II are consistent with $\boldsymbol{k}(\boldsymbol{X}^T\boldsymbol{X})$, indicating that $\boldsymbol{k}(\boldsymbol{X}^T\boldsymbol{X})$ is a good measure of the condition number at the solution point.

B. *Parameter Sensitivities*

The objective of evaluating parameter sensitivities is to identify which parameters are most likely to be skewed by the noise in the input data and/or model insufficiencies. The sensitivities of the parameters, calculated using nominal values, are shown in Tables III and IV. These sensitivities give relative values of the change in parameter for a change in input/measurement vector $\mathbf{y}$. The first three rows have double labels to account for the different parameters K_1, K_2, and K_3, used in the LTOP.

These tables show that R is the parameter that is most sensitive to input parameters. Neglecting the NTOP model, the Swift model, in general, shows higher sensitivities than the Susa model (consistent with their condition numbers) and the LTOP model shows the smallest sensitivities. Sensitivities calculated for optimal parameters are very similar to the ones shown in these tables with R remaining as the most sensitive.

TABLE IV
PARAMETER SENSITIVITIES FOR DV6 FOR NOMINAL PARAMETERS VALUES

DV6				
Parameter	LTOP	NTOP	Susa	Swift
R or K_1	0.031	1.4E+05	4.46	11.80
T_0 or K_2	0.031	2.7E-02	0.15	1.54
θ_{fl} or K_3	0.019	9.3E+05	0.57	0.81
n	---	2.0E+04	0.08	0.16

TABLE V
VIF'S FOR CORBELL3 FOR OPTIMAL PARAMETERS

Corbell3				
Regressor for	LTOP	NTOP	Susa	Swift
R or K_1	1.1	511.56	313.90	1602
T_0 or K_2	1.1	0.0005	0.008	1.754
θ_{fl} or K_3	---	11.151	3.902	2.614
n	---	0.034	0.007	0.02

TABLE VI
VIF'S FOR DV6 FOR OPTIMAL PARAMETERS

DV6				
Parameter for	LTOP	NTOP	Susa	Swift
R or K_1	2.7	34.694	16.069	138.7
T_0 or K_2	2.7	0.0007	0.010	2.36
θ_{fl} or K_3	---	3.707	1.30	0.67
n	---	0.0425	0.004	0.03

The R parameter occurs in the nonlinear models in the term

$$\frac{I^2R+1}{R+1}. \tag{4}$$

This quantity is relatively insensitive to R values of $I^2 > 0.5$ and $R > 4$. Hence, for optimization purposes, a wide range in R values yields similar results.

C. Variance Inflation Factors (VIFs)

The VIF for each parameter in the model measures the combined effect of the dependencies among the regressors on the variance (uncertainty) of that parameter. One or more VIFs that are larger than 10 indicates that the associated regression coefficients are poorly estimated because of multicollinearity. Tables V and VI show the calculated VIFs for Corbell3 and DV6 data sets, respectively. As expected, these VIFs show that the variance of the parameters of the LTOP model is the smallest and that the R parameter will be the most uncertain of all parameters. This is consistent with the observations made regarding (4).

D. Variation of Model Coefficients and $\mathrm{SSL}_{\max}$

The mean of $\mathrm{SSL}_{\max}$ and its coefficient of variation for Corbell3 and DV6, taken from six five-day data-set samples are shown in Tables VII and VIII. As expected, there is a lack of consistency in $\mathrm{SSL}_{\max}$ between the models. In terms of prediction variation, LTOP does the best, followed by the Susa and the Swift models. While the CVs of these latter models appear reasonable, these values are based on parameters which have large CVs and are sometimes negative, as shown in Tables IX and X. The low CVs of the nonlinear models (Tables VII and VIII) show that despite wide variations in the coefficients, the optimization algorithm works appropriately, but multicollinearity causes model parameters to be unrealistic.

TABLE VII
VARIATION IN $\mathrm{SSL}_{\max}$ FOR CORBELL3

Corbell3				
Parameter	LTOP	NTOP	Susa	Swift
$\mathrm{SSL}_{\max}$ (pu)	1.25	---	1.08	1.21
CV (%)	1.28%	---	1.67%	2.48%

TABLE VIII
VARIATION IN $\mathrm{SSL}_{\max}$ FOR DV6

DV6				
Parameter	LTOP	NTOP	Susa	Swift
Load (pu)	1.17	---	1.17	1.20
CV (%)	0.96%	---	2.54%	5.59%

TABLE IX
R PARAMETERS FOR CORBELL3

Corbell3				
Data Case	LTOP	NTOP	Susa	Swift
1	3.77	---	9.37	18.18
2	3.72	---	5.99	8.06
3	3.65	---	27.39	12.87
4	3.95	---	-17.48	30.55
5	3.87	---	-25.49	16.54
6	4.18	---	38.33	11.93
Avg	3.85		6.35	16.36

TABLE X
R PARAMETERS FOR DV6

DV6				
Data Case	LTOP	NTOP	Susa	Swift
1	8.20	---	8.95	3.46
2	6.75	---	8.32	22.2
3	7.65	---	15.1	7.46
4	7.71	---	7.67	18.6
5	9.77	---	16.01	2.20
6	5.76	---	8.31	11.5
Avg	7.64		10.71	10.90

V. ADEQUACY METRICS

Adequacy measures whether the model has the appropriate structure to capture the features of the process being modeled.

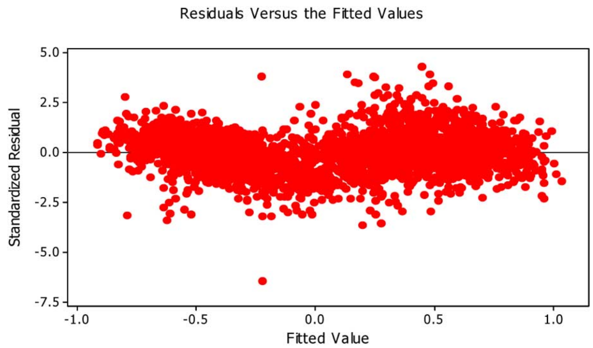

Fig. 1. Plot of residuals versus fitted values for LTOP applied to Corbell3.

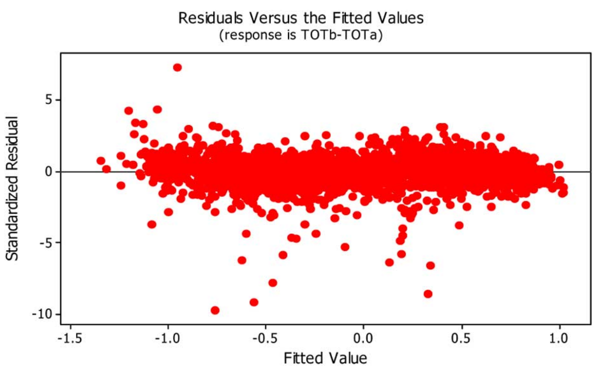

Fig. 2. Plot of residuals versus fitted values for LTOP applied to DV6.

TABLE XI
R^2_{Pred} VALUES

	R^2_{Pred}			
Parameter	**LTOP**	**NTOP**	**Susa**	**Swift**
Corbell3	0.956	---	0.947	0.944
DV6	0.886	---	0.875	0.873

TABLE XII
OPTIMIZED PARAMETER VALUES FOR CORBELL3

Corbell3					
Parameter	**Nominal**	**LTOP**	**NTOP**	**Susa**	**Swift**
Nominal *n*		**1.0**	**0.9**	**0.8**	**1.0**
R	**17.23**	4.77	-275.87	24.42	15.61
T_0	**2.14**	2.68	68.24	2.54	13.41
θ_{fl}	**43.90**	49.36	-1.17	44.71	44.82
n		1.00	-2.78	0.800	0.705

TABLE XIII
OPTIMIZED PARAMETER VALUES FOR DV6

DV6					
Parameter	**Nominal**	**LTOP**	**NTOP**	**Susa**	**Swift**
Nominal *n*		**1.0**	**1.0**	**0.8**	**1.0**
R	**7.03**	7.08	-83.24	5.03	7.73
T_0	**2.30**	1.81	85.59	1.81	1.98
θ_{fl}	**39.90**	43.32	0.00	43.35	43.16
n		1.00	-10.14	0.994	0.979

A. Plots of Residuals Versus Fitted Values

1) LTOP Model: A plot of studentized residuals r_i versus the fitted values $\hat{y}_i$ (calculated using the optimal coefficients) is helpful for detecting common types of model inadequacies. The plot of residuals versus fitter values for the LTOP model is shown for Corbell3 and DV6, in Figs. 1 and 2, respectively. The residual-versus-fitted-values plots show that the linear model fails to capture the mild nonlinearity in Corbell3 (where the load range from 0.08 to 0.64 p.u. and TOT ranging from 36.5 °C to 73.01 °C in the training data set). This mild nonlinearity is not sufficient to cause the model to be inadequate.

For DV6, an FOFA-cooled transformer, the residual plots show that the heating process is quite linear. (The load and TOT values in the training data set for these plots ranged from 0.12 to 0.88 p.u. and from 33.37 °C to 80.76 °C, respectively.) This linearity is consistent with the $n = 1$ exponent recommended by Swift *et al.* [5], [6] and the LTOP models. Other residual plots versus load and TOT for both Corbell3 and DV6 show no nonlinearity [9].

2) Susa et al. [7] and Swift et al. [5], [6] Models: All of the residual plots (versus fitted values, load, or TOT) for the Susa and Swift models give similar results to that of Fig. 2 for DV6. These results indicate that the extra heating due to a change in resistance with load is not significant at these load levels. These results suggest again that the placement of the n exponent in these models (compared to the NTOP model) captures the mild nonlinearity of Fig. 1.

B. R^2_{Pred}

R^2_{Pred} measures the ability of the model, to accurately predict each data point, when that data point is not used in building the model. The R^2_{Pred} statistics for each model for both transformers are contained in Table XI. These numbers show that under FOFA operation, all models have similar performance. Likewise, under NOFA operation, all models have similar performance. This implies that the linear model captures a linear model feature in NOFA operation not evident in our analysis.

VI. ACCURACY

The accuracy metrics assess the models' accuracy for predicting quantities of interest by interpolating (predicting values from within the data-set range) and extrapolating (predicting values outside the data-set range).

A. Comparison With Nominal Values

Since a wealth of positive experience exists with transformer thermal models using parameters derived from the test report, it is expected that the optimal parameters will be reasonably close to the nominal parameters. Optimal model parameters that vary widely from classical calculations indicate that the model is faulty and/or ill conditioned.

The results of optimizing the parameters for all models using 30-day training data sets taken from Corbell3 and DV6 are contained in Tables XII and XIII. The row labeled "Nominal n" in these tables represents the recommended nominal values (as we best understood them by reading [5]–[7].)

Table XII shows that none of the models accurately represents the NOFA transformer. While the Susa model comes closest, we see in Table IX (for smaller data sets—5 day versus 30 day) that there is a wide range of variability in the R parameter from the Susa model. The conclusion that none of the models can accurately model NOFA transformers is premature. We have observed that larger data sets yield more consistent models for LTOP and expect the same behavior for the nonlinear models. With a sufficiently large data set, the Swift or Susa models may be accurate. Table IX does show that the LTOP gives a consistently incorrect R parameter for the NOFA transformer indicating that the LTOP is unacceptable for NOFA transformers.

For the FOFA transformer, Table XIII shows the LTOP model is most accurate and Table X shows that the results are reasonably consistent. (Since R is the most sensitive parameter, it has the most variation. Other parameters for all models of DV6 are much closer to nominal values and have much smaller variations.) The Swift model may be acceptable for the FOFA cooling mode, even though Table X shows inconsistency in R parameter estimates. Again, a data set that is larger than 60 days may improve the performance of all models.

The optimized parameters for the NTOP model are far from nominal for both transformers for all coefficients. These unreasonable results, along with the variability of the Susa and Swift models, are consistent with the large condition number of the $\mathbf{X^TX}$ matrix contained in Tables I and II.

It is noteworthy that the Swift *et al.* model, [shown in (5) at the bottom of the page] performs better for the DV6 transformer (which has oil pumps) than for Corbell3 (which has none.) With oil pumps, oil circulation is almost independent of oil viscosity. Indeed, the Susa model results for the DV6 transformer given an n coefficient which is nearly 1.0. This means that the Susa model [shown in (6) at the bottom of the page] is nearly independent of oil.

The sensitivity results of Tables III and IV indicate that the parameter R is most sensitive to changes in input data and, therefore, contributes to large condition numbers of the $\mathbf{X^TX}$ matrix of all but the linear model. We test this hypothesis by fixing R at its nominal value and evaluating the remaining parameters of each model. The results of this optimization are shown in Tables XIV and XV. We observe that this somewhat improves both the Swift and Susa model performance. This is not surprising since by fixing the value of R, we improve the condition number of the $\mathbf{X^TX}$ matrix.

There is evidence that the oil viscosity term in the Susa model is an important feature. Compare the performance in Table XIV of both the Swift and Susa models for Corbell3 (which is without oil pumps.) The Swift *et al.* [5], [6] model shows worse optimized parameters than the Susa *et al.* model.

TABLE XIV
OPTIMIZED PARAMETER VALUES FOR CORBELL3 WITH R FIXED

Corbell3					
Parameter	**Nominal**	**LTOP**	**NTOP**	**Susa**	**Swift**
Nominal *n*		**1.0**	**0.9**	**0.8**	**1.0**
R	**17.23**	4.77	Nom.	Nom.	Nom.
T_0	**2.14**	2.68	118.8	2.46	5.17
θ_{fl}	**43.90**	49.36	$\approx -10^5$	43.20	47.12
n		1.00	23.23	0.988	0.856

TABLE XV
OPTIMIZED PARAMETER VALUES FOR DV6 WITH R FIXED

DV6					
Parameter	**Nominal**	**LTOP**	**NTOP**	**Susa**	**Swift**
Nominal *n*		**1.0**	**1.0**	**0.8**	**1.0**
R	**7.03**	7.08	Nom.	Nom.	Nom.
T_0	**2.30**	1.81	79.15	1.81	1.81
θ_{fl}	**39.90**	43.32	0.00	43.28	43.26
n		1.00	-17.50	0.99	1.00

Swift et al. Model

$$\begin{aligned}\boldsymbol{\theta}_{\rm top}(\mathbf{k}) &= \boldsymbol{\theta}_{\rm top}(\mathbf{k}-\mathbf{1}) + \frac{\boldsymbol{\Delta}\mathbf{t}}{\mathbf{T_0}} \\ &\quad \times \left[\left[\frac{\mathbf{I^2(k)R+1}}{\mathbf{R+1}}\right]\boldsymbol{\theta}_{\mathbf{fl}}^{\mathbf{1/n}} - [\boldsymbol{\theta}_{\rm top}(\mathbf{k}) - \boldsymbol{\theta}_{\mathbf{amp}}(\mathbf{k})]^{\mathbf{1/n}}\right] \end{aligned} \tag{5}$$

Susa et al. Model

$$\begin{aligned}\boldsymbol{\theta}_{\rm top}(\mathbf{k}) &= \boldsymbol{\theta}_{\rm top}(\mathbf{k}-\mathbf{1}) \\ &\quad + \frac{\boldsymbol{\Delta}\mathbf{t}}{\mathbf{T_0}}\left[\left[\frac{\mathbf{I^2(k)R+1}}{\mathbf{R+1}}\right]\boldsymbol{\theta}_{\mathbf{fl}} - \frac{\mathbf{1}}{[\boldsymbol{\mu}(\mathbf{k})\boldsymbol{\theta}_{\mathbf{fl}}]^{[\mathbf{1-n/n}]}}[\boldsymbol{\theta}_{\rm top}(\mathbf{k}) - \boldsymbol{\theta}_{\rm amb}(\mathbf{k})]^{\mathbf{1/n}}\right] \end{aligned} \tag{6}$$

$$\boldsymbol{\mu}(\mathbf{k}) = \frac{\mathbf{1.3573 \bullet 10^{-6} \bullet e}^{\left[\frac{\mathbf{2797.3}}{\boldsymbol{\theta}_{\rm top}(\mathbf{k})+\mathbf{273}}\right]}}{\boldsymbol{\mu}_{\mathbf{rated}}} \tag{7}$$

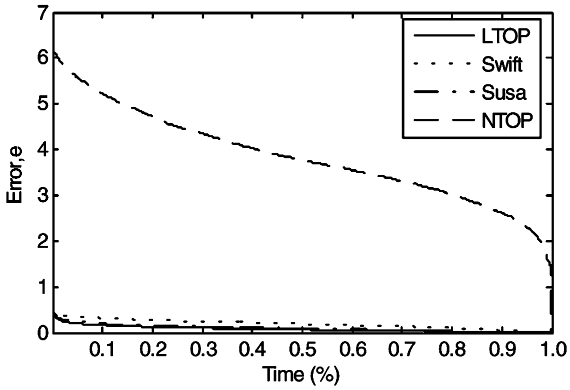

Fig. 3. Error duration curves for Corbell3.

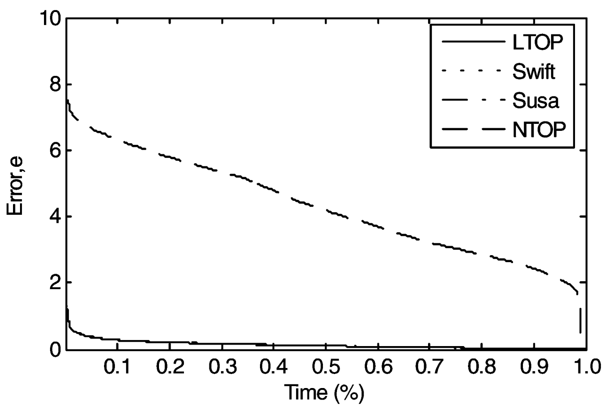

Fig. 4. Error duration curves for DV6.

When no oil pumps are present, heat transfer relies on natural convection, which is dependent on oil viscosity.

B. Residual Duration Curves

Plots of the residual (or error) duration curves, give a measure of how well each model predicts TOT compared to actual values. Fig. 3 shows the Corbell3 error duration curves for all models using optimized coefficients, except for the NTOP model. Since we could not obtain reasonable coefficients by training the NTOP model, we show it plotted using the nominal coefficients. A plot of the error duration curves for DV6 is shown in Fig. 4. Figs. 3 and 4 show that the linear model is either equal to or better than the other nonlinear models. All models are considerably better than the NTOP (Clause 7) model.

VII. Summary

Whether the model is NOFA or FOFA affects which, if any, model is appropriate. Of the four transformer top-oil models examined here, we can draw the following conclusions regarding acceptability of models for training using measured data.

- The NTOP (Clause 7) model is unacceptable.
- The LTOP model is unacceptable for NOFA cooling.
- The LTOP model is acceptable for FOFA cooling.
- The Susa *et al.* [7] and Swift *et al.* [5], [6] models may be acceptable for FOFA and NOFA cooling. These models should be further investigated by using larger data sets.

TABLE XVI
Transformer Heat-Run Data

Quantity (Units)	Corbell3	DV6
Rating (MVA)	28	167
No Load Loss (W)	10,740	75,300
Load Loss (W)	185,000	529,054
Top Oil Rise (°C)	43.9	39.3
Hot Spot Rise (°C)	49.5	39.1
Core/Coil Weight (lbs)	47,875	114,000
Tank Weight (lbs)	20,501	75,000
Oil Volume (gal)	4,404	11600

Regarding the metrics we used, the following obtained results showed:

- Condition numbers, sensitivity, VIF"s variation give *a priori* information about model consistency. These conclusions are reinforced by inspecting the variation of model parameters and $\mathrm{SSL}_{\max}$.
- Residuals plots and prediction R^2 help identify potential misspecification of the model.
- Comparison with nominal values and residual duration curves provide useful information about the accuracy and prediction capability of the model.
- The linear model takes into account an unknown feature that the nonlinear models do not.
- The $\mathbf{X}^{\mathbf{T}}\mathbf{X}$ condition number is the best single predictor of a model's consistency and adequacy.

In conclusion, model consistency, accuracy, and adequacy can be used to determine acceptability of transformer thermal models.

Appendix
Corbell3 and DV6 Transformer Descriptions

The heat-run data and basic data for these transformers are contained in Table XVI.

References

[1] "IEEE guide for loading mineral-oil-immersed power transformers," IEEE C57.91–1995, 1995.
[2] L. Jauregui-Rivera and D. J. Tylavsky, "Acceptability of four transformer top-oil thermal models—Part I: Defining models and metrics for system identification purposes," *IEEE Trans. Power Del.*, vol. 23, no. 1, Jan. 2008.
[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 Trans. Power Del.*, vol. 12, no. 1, pp. 249–256, Jan. 1997.
[4] D. J. Tylavsky, X. Mao, and G. A. McCulla, "Transformer thermal modeling: Improving reliability using data quality control," *IEEE Trans. Power Del.*, vol. 21, no. 3, pp. 1357–1366, Jul. 2006.
[5] G. Swift, T. S. Molinski, and W. Lehn, "A fundamental approach to transformer thermal modeling—Part I: Theory and equivalent circuit," *IEEE Trans. Power Del.*, vol. 16, no. 2, pp. 171–175, Apr. 2001.
[6] G. Swift, T. S. Molinski, R. Bray, and R. Menzies, "A fundamental approach to transformer thermal modeling—Part II: Field verification," *IEEE Trans. Power Del.*, vol. 16, no. 2, pp. 176–180, Apr. 2001.
[7] D. Susa, M. Lehtonen, and H. Nordman, "Dynamic thermal modeling of power transformers," *IEEE Trans. Power Del.*, vol. 20, no. 1, pp. 197–204, Jan. 2005.
[8] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 3 ed. New York: Wiley, 2001.
[9] J. Lida, "Reliability assessment of transformer thermal models," Ph.D. diseertation, Arizona State Univ., Tempe, 2006.

Lida Jauregui-Rivera (M'08) received the B.S. degree in electrical engineering from the University of San Antonio Abad of Cusco, Cusco, Perú, in 1998, the M.S. degree in electrical engineering from the University of Puerto Rico at Mayaguez, Mayaguez, Puerto Rico, in 2002, and the Ph.D. degree in electrical engineering from the Arizona State University, Tempe, in 2006.

From 1998 to 1999, she worked in the areas of generation, transmission, and distribution of energy utility industries in Peru. In 2007, she joined the marketing and trading group of the Arizona Public Service (APS) utility company. Her research interests include modeling and simulation of power systems and components, parameter estimation strategies, algorithms for the solution of parameter estimation in electrical energy processing systems, and stability and control of power systems.

Daniel J. Tylavsky (SM'88) received the B.S., M.S.E.E., and Ph.D. degrees in engineering science from Pennsylvania State University, University Park, PA, in 1974, 1978, and 1982, respectively.

From 1974 to 1976, he was with Basic Technology, Inc. and from 1978 to 1980, he was an Instructor of Electrical Engineering at Penn State. In 1982, he joined the Faculty in the Electrical Engineering Department of Arizona State University, Tempe, where he is currently Associate Professor of Engineering.

Dr. Tylavsky is an RCA Fellow, NASA Fellow, and member of Phi Eta Sigma, Eta Kappa Nu, Tau Beta Pi, and Phi Kappa Phi.