Assessing the reliability of linear dynamic transformer thermal modelling

X. Mao, D.J. Tylavsky and G.A. McCulla

Abstract: Improving the utilisation of transformers requires that the hot-spot and top-oil temperatures be predicted accurately. Using measured (noisy) data to derive equivalent linear dynamic thermal models yields performance that is superior to the ANSI standard model, but the reliability of these model coefficients must be assessed if the user is to have confidence in the model. By adding arbitrarily large amounts of data in the modelling process it was expected to make the reliability measures of these models arbitrarily small. When this did not happen, an investigation began that showed why there is a limitation to the accuracy of models derived from noisy data. It is also shown that a standard technique for assessing the reliability of model coefficients is invalid because of the absence of unmeasured driving variables. An alternative method for assessing transformer model reliability is provided.

1 Introduction

The maximally efficient dynamic loading of transformers requires a model that can accurately predict both top-oil and hot-spot temperatures (HSTs and TOTs). We know that the traditional top-oil model [1] does not accurately model dynamic behaviour [2, 3], but yields accurate steady-state behaviour. Nevertheless, this model [1] is used ubiquitously in the industry for three primary reasons. First, since this model has been the industry standard for many years, most utilities have developed in-house software that implements the model equations and have had reasonable success using it. Secondly, the model yields accurate results, provided only long-term and/or steady-state behaviour is needed. Thirdly, this model requires only parameters that can be distilled from measurements recorded on the transformer's test report, a document that most utility engineers can find in their archives.

The competitive concerns of deregulation have caused many utilities to look at improving the efficiency of their transmission and distribution system. Accurate dynamic loading, which is one way of making this improvement, requires that the dynamic performance limitation of what is known as the top-oil rise model (ANSI model) [1] be overcome. The model developed by Pierce [5, 6] overcomes these limitations of the top-oil-rise model but requires model parameters that are not often available to a planning or operations engineer in a production environment.

An alternative approach that reliably yields more accurate temperature prediction is to develop the transformer parameters from transformer performance *in situ* [2, 4, 8]. We have observed that using a linear transformer thermal model obtained through system identification techniques yields results that are superior to the nonlinear ANSI model [1]. Figure 1 shows the typical errors we observe when predicting transformer top-oil temperature using a linear model (i.e. top-oil model) constructed from data measured *in situ* against a nonlinear ANSI model constructed from transformer test report coefficients. Other authors have also observed that models created from field data perform well [9, 10].

The model created from field data is superior because it accurately represents what is in the field which is often different from OEM equipment. For example, failed cooling fans are sometimes replaced with fans of ratings different from the OEM, or sometimes not replaced at all. The model created from field data is superior also because it neither relies on a single transformer test, which may be inaccurate, nor does it rely on the existence of transformer test reports, which are sometimes lost, making determining the parameters of the transformer a guessing game.

As we experimented with building models from measured data [2, 4, 8] we noticed a larger than expected variation in

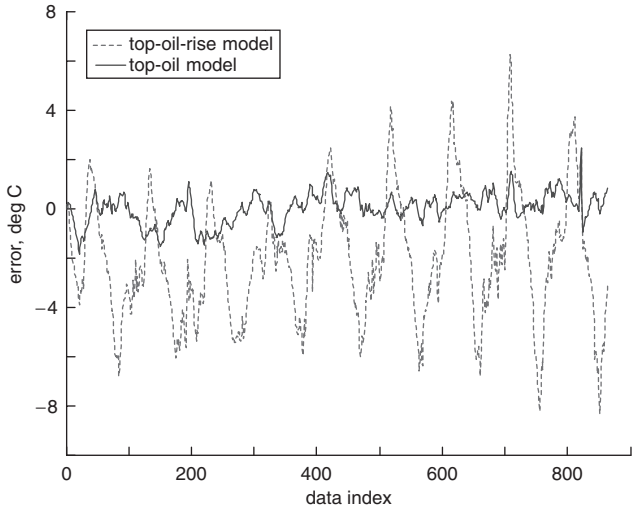

Fig. 1 *Comparison of top-oil model and top-oil-rise model*

IEE Proceedings online no. 20050172

doi:10.1049/ip-gtd:20050172

Paper first received 11th May and in final revised form 7th December 2005

X. Mao is with Department of Electrical Engineering, Arizona State University, USA

D.J. Tylavsky is with Department of Electrical Engineering, Arizona State University, USA

G.A. McCulla is with Salt River Project, Arizona, USA

E-mail: xiaolin.mao@asu.edu

the coefficients of the models built using data taken from the same transformer, data that was, to the eye, similar. While these models exhibited very similar performance, as measured by the top-oil temperature, the variability cast some doubt on the reliability of the models. The objective of this paper is to explain the source of that variability and provide one way of measuring the reliability of the model. To that end we derive a mathematical model to simulate the effect of noise on parameter calculations and show that mathematical model is consistent with simulation results. This derivation shows why the variations in the model coefficients cannot be made arbitrarily small by using an arbitrarily large number of measured data points to construct a model. We show through traditional reliability analysis, and by using the FFT as a low-pass filter to eliminate high-frequency noise, that there is a difference between the performance of measured data and simulated data and, further, that this difference is due to unmeasured driving variables and/or unmodelled nonlinearities. Finally we show that traditional reliability analysis yields erroneous reliability results and propose a sample-based approach for predicting model reliability.

In the following Section we establish a common notation for transformer thermal model development and develop a notation for linear regression analysis.

2 Fundamental model

The traditional top-oil-rise model [1] is governed by the differential equation

$$T_o \frac{d\theta_o}{dt} = -\theta_o + \theta_u \tag{1}$$

which has the solution

$$\theta_o = (\theta_u - \theta_i)(1 - e^{-(t/T_o)}) + \theta_i \tag{2}$$

where

$$\theta_u = \theta_{fl}\left(\frac{K^2 * R + 1}{R + 1}\right)^n \tag{3}$$

$$T_o = \frac{C\theta_{fl}}{P_{fl}} \tag{4}$$

and

θ_o	top-oil rise over ambient temperature (°C)
θ_{fl}	top-oil rise over ambient temperature at rated load (°C)
θ_u	ultimate top-oil rise for load L (°C)
θ_i	initial top-oil rise for $t = 0$ (°C)
θ_{amp}	ambient air temperature (°C)
T_o	time constant
C	thermal capacity (Wh/°C)
P_{fl}	total loss at rated load (W)
n	oil exponent
K	ratio of load L to rated load
R	ratio of load loss to no-load loss at rated load.

The TOT is then given by,

$$\theta_{top} = \theta_o + \theta_{amb} = (\theta_u - \theta_i)(1 - e^{-(t/T_o)}) + \theta_i + \theta_{amb} \tag{5}$$

The top-oil model [3] (*cf.* top-oil-rise model) corrects the dynamic limitations of the top-oil-rise model by including in (1) the dependence of the time rate of change of θ_{top} on ambient temperature θ_{amb}

$$T_o \frac{d\theta_{top}}{dt} = -\theta_{top} + \theta_{amb} + \theta_u \tag{6}$$

This equation has the solution

$$\theta_{top} = (\theta_u + \theta_{amb} - \theta_i)(1 - e^{-(t/T_o)}) + \theta_i \tag{7}$$

To obtain a discrete-time model we discretise (6) by applying the forward Euler discretisation rule

$$\frac{d\theta_{top}[k\Delta t]}{dt} = \frac{\theta_{top}[k\Delta t] - \theta_{top}[(k-1)\Delta t]}{\Delta t} \tag{8}$$

to yield (with the assumption that $n = 1$)

$$\theta_{top}[k] = K_1 I[k]^2 + K_2\theta_{amb}[k] + (1 - K_2)\theta_{top}[k-1] + K_3 \tag{9}$$

where K_1–K_3 are functions of the differential equation coefficients [3], and $I[k]$ is the per-unit transformer current (based on the rated value of the transformer) at time-step k. To obtain a model based on measured field data we choose the coefficients that best fit the measured data (rather than using the formulas for the K_x s from test report data). We have examined many optimisation techniques of finding the best K_x s and have observed linear regression (least-squares method) to be among the best and easiest to use.

To use the least-squares method to obtain K_1–K_3, (9) is reformed as

$$Y[k] = K_1 X_1[k] + K_2 X_2[k] + K_3 \tag{10}$$

where

$$Y[k] = \theta_{top}[k] - \theta_{top}[k-1]$$
$$X_1[k] = I[k]^2$$
$$X_2[k] = \theta_{amb}[k] - \theta_{top}[k-1]$$

Assuming m sets of independent X measurements, (10) can be rewritten in matrix format as

$$Y = K_1 X_1 + K_2 X_2 + K_3 [\,1 \quad 1 \quad \cdots \quad 1\,]^T_{1\times m} \tag{11}$$

where

$$Y = [\,Y[1] \quad Y[2] \quad \cdots \quad Y[m]\,]^T$$

$$X_i = [\,X_i[1] \quad X_i[2] \quad \cdots \quad X_i[m]\,]^T, \quad i = 1, 2$$

Averaging both sides of (11) over time-step index k yields

$$\overline{Y} = K_1\overline{X}_1 + K_2\overline{X}_2 + K_3 \tag{12}$$

where the over-bar represents variables averaged over time. For example

$$\overline{X}_1 = \frac{1}{m}\sum_{k=1}^{m} X_1[k]$$

Subtracting (12) from each equation in (11) gives

$$\tilde{Y} = K_1\tilde{X}_1 + K_2\tilde{X}_2 \tag{13}$$

where the tilde represents variables of zero mean and

$$\tilde{Y} = \left[\,Y[1] - \overline{Y} \quad Y[2] - \overline{Y} \quad \cdots \quad Y[m] - \overline{Y}\,\right]^T$$

$$\tilde{X}_i = \left[\,X_i[1] - \overline{X}_i \quad X_i[2] - \overline{X}_i \quad \cdots \quad X_i[m] - X_i\,\right]^T \quad i = 1, 2$$

Equation (13) can be rewritten as

$$\tilde{Y} = \tilde{X}\,K \tag{14}$$

where

$$\tilde{X} = \lfloor\,\tilde{X}_1 \quad \tilde{X}_2\,\rfloor$$

$$K = [\,K_1 \quad K_2\,]^T$$

With the least-squares method the formula to calculate these coefficients is

$$K = (\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T\tilde{Y} \tag{15}$$

After K_1 and K_2 are obtained, K_3 can be calculated from (12) as

$$K_3 = \overline{Y} - K_1\overline{X}_1 - K_2\overline{X}_2 \tag{16}$$

Using these equations for K_x we can calculate the transformer thermal time constant and top-oil rise at full load.

3 Cause of model variability

When using (15) and (16) to calculate the K_x coefficients we found our models were not as consistent as we expected, even when derived from load and temperature data that seem, to the eye, to be consistent. For example, we constructed eight models using data measured from one transformer over a period of two summer months in Arizona. During this time the daily variation of load and ambient temperature was relatively uniform; however the coefficients resulting from our model building, listed in Table 1, show variations as high as 19% (STD% in Table 1 represents the standard deviation of the coefficient divided by the sample mean in percent). We suspected that these variations were due to noise in the input data. We expected that if we put sufficient data into the modelling procedure, the random effects of measurement noise would average out and that we could bring the K_x coefficients into an arbitrarily narrow band. To test this hypothesis we designed experiments to duplicate the effect of measurement noise using simulated (rather than field) data. Using simulated data allowed us to eliminate the effects of nonlinearities that may be present in the physical process and to eliminate the effects of any unknown missing driving variables. Using simulated data also provided us with a process for which we knew the theoretically correct models.

Table 1: Variability of transformer thermal model coefficients

Data set	K_1	K_2	K_3
1	3.1068	0.0961	0.5277
2	3.0589	0.0905	0.3901
3	3.1632	0.0911	0.4277
4	2.8474	0.0797	0.2667
5	2.7130	0.0814	0.4367
6	2.9740	0.0843	0.3706
7	2.8518	0.0864	0.4621
8	2.8679	0.0936	0.4860
Mean	2.9479	0.0879	0.4210
STD	0.1538	0.0059	0.0801
STD%	5.22%	6.66%	19.02%

We created simulated data sets by first using ambient temperature and load field data for a transformer from data provided by the Salt River Project to serve as typical driving variable data. Next, we assigned to K_1, K_2, and K_3 typical values for the transformer from which these data came (i.e. $K_1 = 2.1882$, $K_2 = 0.08269$, $K_3 = 0.70$, which correspond to $T_0 = 2.77$ h, $R = 3.13$ $\theta_{fl} = 34.9$°C). We then used (9) to generate the simulated TOT data, which we call 'true' TOT. Then we added random noise (first gaussian and then uniformly distributed) to the 'true' TOT, load and ambient temperature, one at a time. Finally we preformed the linear regression to obtain the new coefficients K_1–K_3.

3.1 Observation in coefficients

Figure 2 shows K_1 against the standard deviation of the gaussian noise added to the TOT data. It can be seen that K_1 becomes larger than its true value when the TOT data is noisy. Further, it increases monotonically as the magnitude of the noise increases. Results show that plots of K_2 and K_3 against the standard deviation of the noise have a similar pattern.

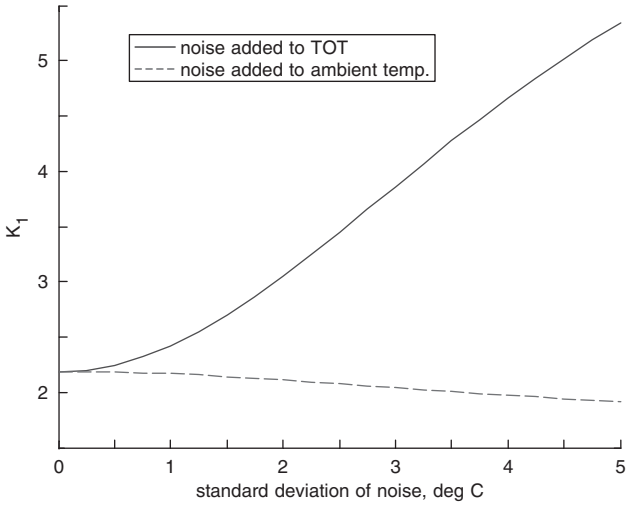

Fig. 2 *K_1 against the standard deviation of gaussian noise added to the TOT data, and the ambient temperature data*

We obtained similar results when we applied uniformly distributed random noise to TOT data.

We found that all the coefficients were much less sensitive to noise added to the ambient temperature data or load data than noise is added to TOT as shown in Fig. 2 and that regardless of the amount of data we put in our modelling procedure, we could not bring the K_x values to within an arbitrarily narrow range.

3.2 Discussion

3.2.1 Ill-conditioned matrix: To discover why the modelling process is sensitive to noise we first looked at the condition number of the coefficient matrix. Consider the equation used to calculate the coefficients (15), written in an equivalent form

$$(\tilde{X}^T\tilde{X})K = \tilde{X}^T\tilde{Y} \tag{17}$$

It is well known that the matrix $\tilde{X}^T\tilde{X}$ has a larger condition number than the $\tilde{X}$ matrix. (The condition number of a matrix may be defined as the ratio of the largest singular value of the matrix to the smallest.) A large condition number indicates that the matrix is nearly singular. For example, using the simulated data described earlier (without noise) in (17) yields as expected

$$\tilde{X}^T\tilde{X} = \begin{bmatrix} 32.654 & -240.404 \\ -240.404 & 24402.0 \end{bmatrix}$$

$$\tilde{X}^T\tilde{Y} = \begin{bmatrix} 51.5744 \\ 1491.76 \end{bmatrix}$$

$$K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} 2.1882 \\ 0.08269 \end{bmatrix}$$

The condition number of $\tilde{X}^T\tilde{X}$ matrix is $cond(\tilde{X}^T\tilde{X}) = 806.0$, a number which is large for a 2×2

matrix. This indicates the matrix is closer to singular than desired, which means that slight noise in $\tilde{X}^T\tilde{X}$ or in $\tilde{X}^T\tilde{Y}$ can lead to large changes in the K_x coefficients. In our example, a noise level of 0.7% added to TOT yields changes in K_1 and K_2 of 1.0% and 3.5%, respectively, a five-fold increase over the input noise level. Further, we've observed that noise that is more highly correlated (which we find in practice) will lead to a more exaggerated response in the coefficients.

To show the effect of a large condition number graphically we expand (17) as

$$\begin{bmatrix} \tilde{X}_1^T\tilde{X}_1 & \tilde{X}_1^T\tilde{X}_2 \\ \tilde{X}_2^T\tilde{X}_1 & \tilde{X}_2^T\tilde{X}_2 \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} \tilde{X}_1^T\tilde{Y} \\ \tilde{X}_2^T\tilde{Y} \end{bmatrix} \tag{18}$$

which can be expressed as

$$\begin{bmatrix} (m-1)\mathrm{var}(X_1) & (m-1)\mathrm{cov}(X_1,X_2) \\ (m-1)\mathrm{cov}(X_1,X_2) & (m-1)\mathrm{var}(X_2) \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} (m-1)\mathrm{cov}(X_1,Y) \\ (m-1)\mathrm{cov}(X_2,Y) \end{bmatrix} \tag{19}$$

where *var* is the variance operator, and *cov* is the covariance operator. Equation (19) represents two lines in the K_2 against K_1 co-ordinate plane, the intersection of which is the solution of (19)

$$\text{line1:} \quad K_2 = \frac{\mathrm{cov}(X_1,Y) - \mathrm{var}(X_1)\cdot K_1}{\mathrm{cov}(X_1,X_2)} \tag{20}$$

$$\text{line2:} \quad K_2 = \frac{\mathrm{cov}(X_2,Y) - \mathrm{cov}(X_1,X_2)\cdot K_1}{\mathrm{var}(X_2)} \tag{21}$$

When the relatively small amount of uniformly distributed random noise in TOT, ± 0.5°C, is applied to measured TOT data, $\tilde{X}^T\tilde{X}$ and $\tilde{X}^T\tilde{Y}$ become

$$\tilde{X}^T\tilde{X} = \begin{bmatrix} 32.654 & -240.650 \\ -240.650 & 24490.0 \end{bmatrix}$$

$$\tilde{X}^T\tilde{Y} = \begin{bmatrix} 51.5844 \\ 1565.61 \end{bmatrix}$$

And the solution of the K_x coefficients becomes

$$K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} 2.2110 \\ 0.08565 \end{bmatrix}$$

Figure 3 shows the two lines whose intersection is the solution to (17) before and after applying a ± 0.5°C uniformly distributed noise in TOT. This Figure shows several relevant characteristics of the modelling process. It shows that the lines that characterise coefficient matrix are nearly collinear, which explains the sensitivity of the modeling process to noise. It shows that even though the presence of noise in TOT has little impact on line 1, noise in TOT causes line 2 to shift upward slightly which, because of the colinearity problem, causes the solution point for K_1 and K_2 to shift considerably. This Figure does not explain why adding a large amount of data to the modelling procedure does not cause the effect of noise to be made arbitrarily small.

3.2.2 Analysing coefficient matrix when noise is present: In this Appendix (Section 8) we

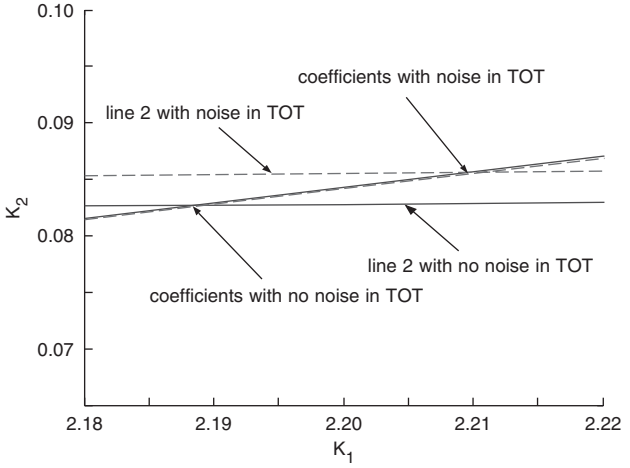

Fig. 3 *Effect on coefficients with slight noise in TOT*

analyse the coefficient matrix after introducing a noise model. The results of this analysis, shown in approximations (22) and (23), shows that line 1 is (approximately) unaffected by noise, while the approximation for line 2 differs from (21) by the presence of an additional term var(N_{-1}). This new term the variance of the added noise to the simulated TOT.

$$\text{line1:} \quad K_2 \approx \frac{\mathrm{cov}(X_1,Y) - \mathrm{var}(X_1)\cdot K_1}{\mathrm{cov}(X_1,X_2)} \tag{22}$$

$$\text{line2:} \quad K_2 \approx \frac{\mathrm{cov}(X_2,Y) + \mathrm{var}(N_{-1}) - \mathrm{cov}(X_1,X_2)\cdot K_1}{\mathrm{var}(X_2)} \tag{23}$$

This analysis shows that K_1 and K_2 will always increase when noise in TOT is added, which is consistent with the numerical results we observed and shown in Fig. 2. Further, it shows that regardless of the number of points in the data set, the coefficients obtained from data sets with different amounts of noise cannot be brought into an arbitrarily small range unless either the variance of the applied noise is reduced or the variances and covariances shown in (22) and (23) are made invariant, something that is not within our control with field data.

To eliminate or reduce the value of var(N_{-1}) in (23), we investigated using the FFT as a low-pass filter to reduce noise in the TOT input data. We chose to investigate noise in TOT because this source of noise has the most severe effect on the variance of the coefficients. We first applied this filter to simulated data and then to field data.

4 Using FFT to eliminate TOT noise

4.1 Appling FFT to filter TOT noise in simulated data

Because the time constant associated with a transformer's TOT is about 3 to 5 hours, only the relatively low frequencies of the TOT have information that comes from the TOT heating process. Measurement noise introduced into the TOT manifests itself mainly as high-frequency components. Our hypothesis was: we can reduce the var(N_{-1}) in (23) by eliminating the high-frequency components.

Using the simulated data constructed earlier, normally-distributed zero-mean random noise with 1.5°C standard deviation was first added to the TOT and then the FFT was applied to the noisy TOT with different cutoff frequencies.

Table 2: Applying FFT to one simulated data set

Cutoff freq. (cycle/day)	K_1
3	2.0741
5	2.0992
8	2.1156
10	2.1266
17	2.1546
21	2.1903
30	2.3187
40	2.5286
no FFT	2.7000
true value	2.1882

Typical results for the experiments we conducted are shown in Table 2.

It can be seen from Table 2 that when the cutoff frequency is too low, the K_1 coefficient strays far from its true value: too many frequency components are eliminated which eliminates signal in addition to noise. On the other hand, when the cutoff frequency is too high, K_1 strays far from its true value also: too much noise is preserved. Our experiments showed that the optimum filter involved retaining about 21 cycles per day.

In another test we repeated the experiment on many independent data sets and used the variation of the model coefficients as the evaluation of the effectiveness of the FFT. The assumption here is that if the FFT is removing noise, the models derived should have a narrower range of coefficients. Table 3 shows results obtained by using the FFT to filter TOT data with different cutoff frequencies applied to 17 sets of simulated data. It can be seen from Table 3 that the FFT reduces the variation in the model coefficients, and the swift reduction continues until the FFT cutoff frequency drops below about 21 cycles per day.

Table 3: Applying FFT to many independent simulated data sets

Cutoff freq (cycle/day)	STD% of K_1
3	1.55
5	1.40
8	1.31
10	1.42
17	1.55
21	1.79
30	3.74
40	6.65
no FFT	8.88

4.2 Appling FFT to filter TOT noise in measured data

Since there is no guarantee that the cutoff frequency determined will be the same for measured data, the experimental approach to determining the cutoff frequency, by minimising the variance of the K_x values, was performed using measured data. The experiment determined the K_x values using 17 independent data sets, and used the variation of the model coefficients as the evaluation of the effectiveness of the FFT. Table 4 shows the results of using the FFT to filter TOT data with different cutoff frequencies. While the variation in these coefficients is acceptable for modelling, it can be seen from Table 4 that the FFT essentially does not reduce the variation in the model coefficients. The lack of change in the field-data STD near 21 cycles a day implies that variability of the coefficients is not caused by random measurement noise, but by either missing driving variables or unmodelled nonlinearities. Note that both of these confounding modelling issues (nonlinearities and missing driving variable) will appear as virtual input–data noise leading to unwanted variation in our coefficients.

Table 4: Applying FFT to field data

Cutoff freq. (cycle/day)	STD%		
	K_1	K_2	K_3
3	4.57	4.44	6.60
5	4.76	4.59	6.40
8	4.83	4.72	6.53
10	4.86	4.79	6.63
17	4.90	4.91	6.86
21	4.94	4.98	6.95
30	5.00	5.12	7.16
40	5.03	5.22	7.37
no FFT	5.06	5.29	7.49

Since there is no way to eliminate the variation in coefficients it is necessary to quantify the reliability of the model we produce.

5 Reliability analysis

5.1 Traditional reliability analysis

For a given confidence level it is possible to calculate the confidence interval of each of the coefficients that result from linear regression [7]

$$CI_{Ki} = [K_i - \sigma\sqrt{(\tilde{X}^T\tilde{X})^{-1}_{ii}} \cdot t, \quad K_i + \sigma\sqrt{(\tilde{X}^T\tilde{X})^{-1}_{ii}} \cdot t], \quad i = 1, 2 \tag{24}$$

where CI_{Ki} is the confidence interval of K_i ($i = 1, 2$); σ is the standard deviation of the residuals; $(\tilde{X}^T\tilde{X})^{-1}_{ii}$ is the ith diagonal element of the inverse of the matrix $\tilde{X}^T\tilde{X}$; and t is the number of standard deviations corresponding to a given confidence level.

Assuming a desired confidence *level* of 95% for K_1, and then after calculating the corresponding confidence *interval* we can then say we are confident that the true value of K_1 for our model lies within the calculated confidence interval 95% of the time. That is, we can expect that for each coefficient we calculate, the 95% confidence interval surrounding that coefficient will enclose the true value 95% of time, or 19 out of 20 times [7].

We first applied the confidence interval approach to the simulated data with TOT corrupted by normally-distributed zero-mean random noise with 0.5°C standard deviation. Figure 4 shows the calculated K_1 coefficient with confidence intervals corresponding to a 95% confidence level for data samples taken from various times of the year. This Figure shows that the real value of K_1 as well as the mean value and the median value of calculated K_1 stay in the confidence interval 100% of the time. This is slightly more often that we would expect; however, because our sample

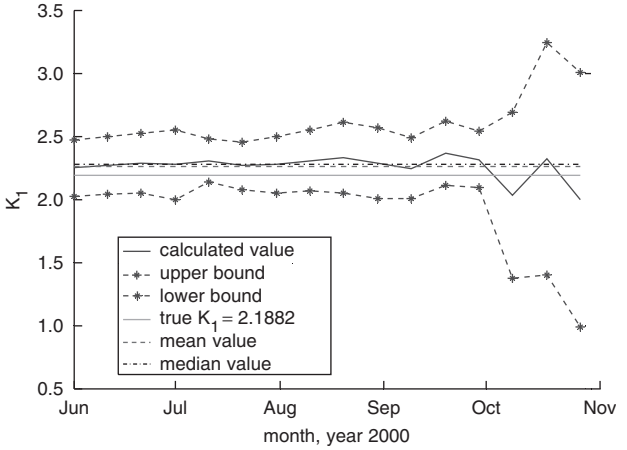

Fig. 4 *Applying confidence interval to simulated data with normally distributed noise*

was so low (we had only 17 sets of coefficients), the results shown do not violate the theory.

We applied this confidence interval approach to the field data, calculating the confidence interval based on a 95% confidence level. Inspection of Fig. 5 (which corresponds to the numerical confidence intervals shown in the middle column of Table 5) shows that there is no way to pick a 'true K_1' that is within the calculated confidence intervals 95% of the time. This apparent paradox shows that the inconsistency of the model coefficients is not mainly due to random noises, but to either the incompleteness of the model, e.g. missing driving variables or the nonlinearity in the transformer thermal process. It is impossible with a linear regression approach to distinguish whether the virtual noise in our data comes from unmodelled nonlinearities or unmodelled (and unmeasured) driving variables. Both of these sources of virtual noise are perceived identically by the linear regression process. To show that either of these sources of virtual noise can cause the effect on confidence interval and confidence level observed in Fig. 5, we generated simulated transformer TOT values using a linear model which was modified by adding an extra driving variable to the model. This extra variable accounted for the increase in heat caused by solar radiation model. Then we calculated the model coefficients and confidence intervals based on the original model (without a solar-radiation variable). The results of Fig. 6 show that the behaviour of the coefficient's confidence interval is similar to that of field

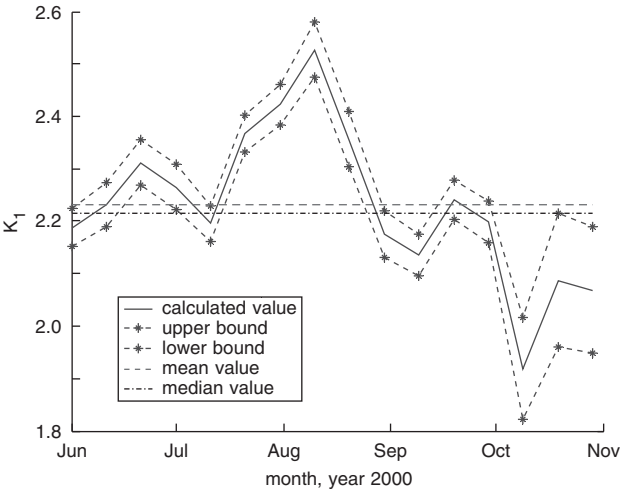

Fig. 5 *Applying confidence interval to field data*

Table 5: Comparison of linear-regression-based and sample-based confidence intervals

Start date of sample	Magnitude of linear-regression-based confidence interval for K_1	Magnitude of sample-based confidence interval for K_1
Jun 01	0.073	
Jun 11	0.0839	
Jun 21	0.0863	
Jul 01	0.0877	
Jul 11	0.0667	
Jul 21	0.0703	
Jul 31	0.0783	
Aug 10	0.1051	
Aug 20	0.1053	0.5781
Aug 30	0.0879	
Sep 09	0.0789	
Sep 19	0.0728	
Sep 29	0.0789	
Oct 09	0.1961	
Oct 19	0.2538	
Oct 29	0.24	

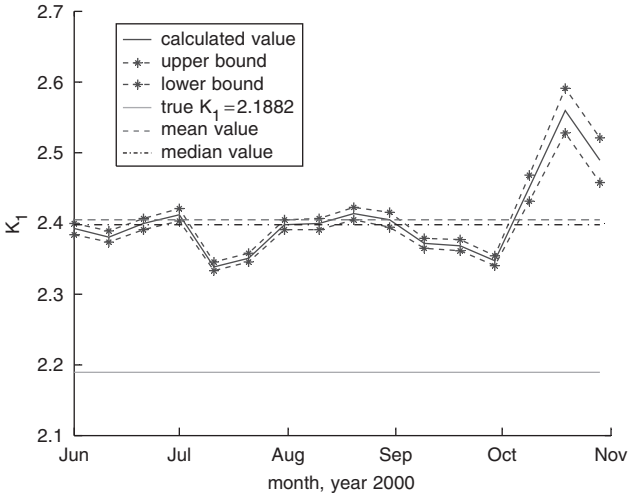

Fig. 6 *Applying confidence intervals to simulated data modelling missing driving variable*

data; there is no way to pick a 'true K_1' that it is within the calculated confidence intervals 95% of the time.

The failure of the traditional approach to yield consistent confidence levels and intervals also explains why using a low-pass filter does not increase the reliability of our model: the noise in our data is virtual noise rather than actual measurement noise.

To estimate the reliability of our model we needed a different approach.

5.2 Sample-based reliability analysis

Sample-based reliability calculations involve calculating many models from different data samples and then calculating the standard deviation of, for example, K_1. This sample standard deviation can be used to estimate the standard deviation of the population. Using this standard deviation the confidence intervals and confidences levels can

be calculated. Specifically, assuming we have m data sets with K_1 values: $K_{11}, K_{12}, \ldots, K_{1m}$, the standard deviation is defined as

$$\sigma_{K1} = \sqrt{\frac{1}{m-1}\sum_{j=1}^{m}\left(K_{1j} - \overline{K}_1\right)^2}$$

Based on the number of samples Student's T-distribution can be used to calculate the confidence interval for a given confidence level. For sample sizes larger than 15, the confidence interval calculated using Student's T-distribution is close to that obtained when assuming a gaussian distribution, and a Gaussian assumption may be used. Assuming Student's T-distribution, the confidence interval can be calculated as

$$CI_{Ki} = [K_i - \sigma_{Ki} \cdot t, K_i + \sigma_{Ki} \cdot t], \quad i = 1, 2 \qquad (25)$$

where σ_{Ki} is the sample-based standard deviation of K_i and t is the point on the scale of the Students T-distribution corresponding to a given confidence level. Figure 7 shows the calculated K_1 coefficients taken from field data with confidence intervals corresponding to a 95% confidence level for data samples taken from various times of the year. Numerical confidence-interval values corresponding to Fig. 7 are shown in the right-most column of Table 5. Inspection of Fig. 7 shows that both the mean and median K_1 values are within 15 of 16 confidence intervals, which is consistent with a 95% confidence level.

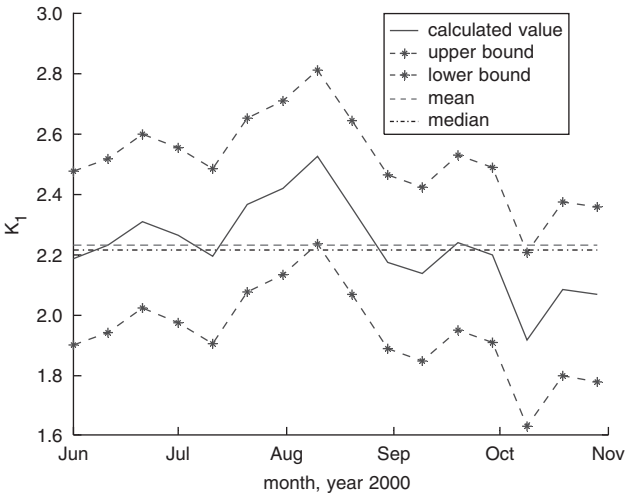

Fig. 7 *Applying sample-based confidence intervals to K_1 derived from field data*

To show that this same behaviour exists for other transformers, we examined two transformers. We calculated the linear-regression-based confidence intervals for several samples and the sample–based confidence intervals for the same set of samples. This data is shown in Table 6, which clearly shows that the relationship between the linear-regression-based and sample–based confidence intervals is similar to what we've observed in Figs. 5 and 7.

We observed that regardless of the reliability metric we chose the reliability index we calculated varied, depending on which K_x we used in the assessment. To estimate the reliability of the model it is necessary to develop a measure which includes all coefficients. While many such measures can be justified we chose to use steady-state load as a measure, since the ultimate use of these models will be to predict loading capability. Our goal is to use these models to perform dynamic loading. Defining a measure of dynamic

Table 6: Comparison of linear-regression-based and sample-based confidence intervals for two other transformers

Transformer name sample number	Magnitude of linear-regression-based confidence intervals		Magnitude of sample-based confidence interval	
Transformer 1	K_1	K_2	K_1	K_2
1	0.4432	0.0112		
2	0.2764	0.0069		
3	0.2550	0.0062		
4	0.3594	0.0090		
5	0.3570	0.0090		
6	0.2279	0.0059		
7	0.3575	0.0080	1.1465	0.0402
8	0.3700	0.0089		
9	0.2491	0.0059		
10	0.2860	0.0060		
11	0.3035	0.0068		
12	0.5344	0.0065		
Transformer 2	K_1	K_2	K_1	K_2
1	0.2415	0.0074		
2	0.2535	0.0063		
3	0.3033	0.0088		
4	0.2336	0.0057		
5	0.1787	0.0046		
6	0.1490	0.0038	1.0026	0.0285
7	0.2266	0.0059		
8	0.2724	0.0075		
9	0.1822	0.0051		
10	0.2309	0.0053		
11	0.1687	0.0040		
12	0.4712	0.0064		

loading requires many arbitrary assumptions, such as daily load shape, and daily ambient temperature profile and amplitude. We've observed that steady-state loading may be defined with many fewer arbitrary assumptions and further, the changes in predicted steady-state loading level correlate well with changes in predicted dynamic loading level. Steady-state loading for a fixed ambient temperature of 37°C can be calculated by assuming $\theta_{top}[k] = \theta_{top}[k-1]$ and setting this quantity to the maximum TOT allowed for the transformer in (9) and then solving for load I to get

$$I_{SS} = \sqrt{\frac{K_2(TOT_{\max} - 37) - K_3}{K_1}} \qquad (25)$$

where I_{SS} stands for the steady-state loading. A similar result can be obtained if hot-spot temperature is the limiting criteria.

We observed that the confidence intervals in Fig. 8 for steady-state load (using field data) are always much smaller than the corresponding intervals (in percentage of respective parameters) for the K_x coefficients. This seeming contradiction can be explained. We observed that K_1 and K_2 tend to rise and fall in unison from model to model. The variable K_1 is proportional to the heat generated per unit load during each time-step. The variable K_2 is proportional to

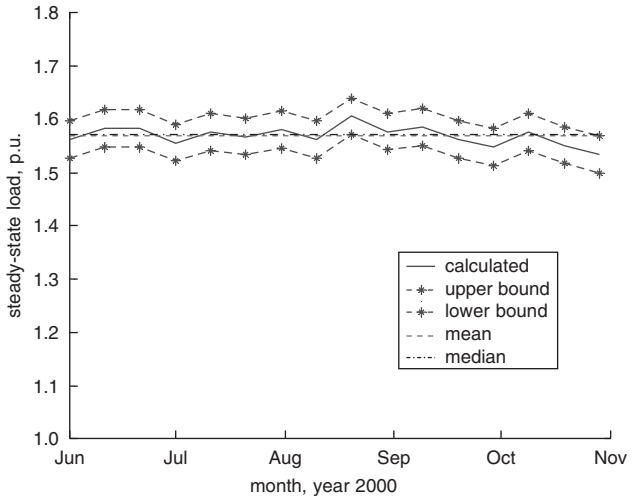

Fig. 8 *Sample-based confidence intervals for steady-state load predictions made from field data*

the heat lost to air during each time-step. If both K_1 and K_2 change in synchronism, their effect on loading capability offset, leading to much more consistent predictions of loading level than the variability of K_1 or K_2 would suggest.

6 Conclusions

We have shown that as long as there is noise in the measurement process, inaccuracies in the dynamic thermal models from measured data will persist. Luckily, with most of the data we have looked at, the noise is relatively small and the models derived are 'good.' The goal of this work was to provide a method for assessing the reliability of these models or, equivalently, to define in a quantitative way, what 'good' means. We showed, using confidence intervals as a measure, that the behaviour of real data is very different from the behaviour of data contaminated with random noise and that the standard technique for assessing reliability is invalid when applied to real data. Through numerical experimentation we provided a plausibility argument that this difference is caused by the absence of unmeasured driving variables in the model. We provided a valid method for assessing model reliability. Rather than using model coefficients as a measure of model reliability, we proposed using loading capability as a more relevant measure of model performance and explain why the reliability of predicted loading is much higher than the reliability of the coefficients of the model.

Being able to calculate the reliability of a model derived from measured data is another advantage of using system identification techniques. Reliability assessment gives the user some measure of confidence in the model they are using. At present there is no method for assessing the reliability of models derived from transformer test reports. From the data of Fig. 1, it is likely that such models are much less reliable than the models we propose here.

7 References

1 IEEE Guide for Loading Mineral-Oil-Immersed Power Transformers, IEEE C57.91-1995
2 Tylavsky, D.J., He, Q., McCulla, G.A., and Hunt, J.R.: 'Sources of error and quantization on transformer dynamic thermal modeling', *IEEE Trans. Power Deliv.*, 2000, **15**, (1), pp. 178–185
3 Lesieutre, B.C., Hagman, W.H., and Kirtley Jr, J.L.: 'An improved transformer top oil temperature model for use in an on-line monitoring and diagnostic system', *IEEE Trans. Power Del.*, 1997, **12**, (1)
4 He, Q., Si, J., and Tylavsky, D.J.: 'Prediction of top-oil temperature for transformers using neural networks', *IEEE Trans. Power Del.*, 2000, **15**, (4), pp. 1205–1211
5 Pierce, L.W.: 'Predicting liquid filled transformer loading capability', *IEEE Trans. Ind. App.*, 1994, **30**, (1)
6 Pierce, L.W.: 'An investigation of the thermal performance of an oil filled transformer winding', *IEEE Trans. Power Del.*, 1992, **7**, (3)
7 Stoll, H.G.: 'Least-cost electric utility planning' (Wiley, New York, 1989), pp. 177–192
8 Tylavsky, D.J., He, Q., Si, J., McCulla, G.A., and Hunt, J.R.: 'Transformer top-oil temperature modeling and simulation', *IEEE Trans. Ind. Appl.*, 2000, **36**, (5), pp. 1219–1225
9 Swift, G., Molinski, T.S., and Lehn, W.: 'A fundamental approach to transformer thermal modeling—Part I: theory and equivalent circuit', *IEEE Trans. Power Deliv.*, 2001, **16**, (2), pp. 171–175
10 Swift, G., Molinski, T.S., and Lehn, W.: 'A fundamental approach to transformer thermal modeling—Part I: field verification', *IEEE Trans. Power Deliv.*, 2001, **16**, (2), pp. 176–180

8 Appendix

8.1 Analysis of why K_x increases when noise is added to the TOT

When random noise with zero mean N is added to θ_{top}, the variables Y, X_1, and X_2 become

$$\begin{aligned} Y'[k] &= \theta_{top}[k] + N[k] - \theta_{top}[k-1] - N[k-1] \\ &= Y[k] + N[k] - N[k-1] \end{aligned}$$

$$X_1'[k] = I[k]^2 = X[k]$$

$$\begin{aligned} X_2'[k] &= \theta_{amb}[k] - \theta_{top}[k-1] - N[k-1] \\ &= X_2[k] - N[k-1] \end{aligned}$$

$$k \geqslant 1$$

Accordingly (19) becomes

$$\begin{bmatrix} \mathrm{var}(X_1') & \mathrm{cov}(X_1', X_2') \\ \mathrm{cov}(X_1', X_2') & \mathrm{var}(X_2') \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} \mathrm{cov}(X_1', Y') \\ \mathrm{cov}(X_2', Y') \end{bmatrix} \tag{26}$$

where for an m length sample

$$Y' = [\, Y'[1] \quad Y'[2] \quad \cdots \quad Y'[m] \,]^T$$

$$X_i' = [\, X_i'[1] \quad X_i'[2] \quad \cdots \quad X_i'[m] \,]^T, \quad i = 1, 2$$

Apparently we have

$$\mathrm{var}(X_1') = \mathrm{var}(X_1)$$

$$\mathrm{cov}(X_1', X_2') = \mathrm{cov}(X_1, X_2) - \mathrm{cov}(X_1, N_{-1})$$

$$\mathrm{var}(X_2') = \mathrm{var}(X_2) - 2\mathrm{cov}(X_2, N_{-1}) + \mathrm{var}(N_{-1})$$

$$\mathrm{cov}(X_1', Y') = \mathrm{cov}(X_1, Y) + \mathrm{cov}(X_1, N) - \mathrm{cov}(X_1, N_{-1})$$

$$\begin{aligned} \mathrm{cov}(X_2', Y') = {} & \mathrm{cov}(X_2, Y) + \mathrm{cov}(X_2, N) - \mathrm{cov}(X_2, N_{-1}) \\ & - \mathrm{cov}(N_{-1}, Y) - \mathrm{cov}(N_{-1}, N) + \mathrm{var}(N_{-1}) \end{aligned}$$

where

$$N = [\, N[1] \quad N[2] \quad \cdots \quad N[m] \,]^T$$

$$N_{-1} = [\, N[0] \quad N[1] \quad \cdots \quad N[m-1] \,]^T$$

If the data length is sufficiently long, the covariance between any of the deterministic variables X_1, X_2 and Y, and any of the noises N and N_{-1} will approach to zero. Also since N is a random noise, $\mathrm{cov}(N_{-1}, N) = 0$. Thus

$$\mathrm{var}(X_1') = \mathrm{var}(X_1)$$

$$\mathrm{cov}(X_1', X_2') \approx \mathrm{cov}(X_1, X_2)$$

$$\mathrm{var}(X_2') \approx \mathrm{var}(X_2) + \mathrm{var}(N_{-1})$$

$$\mathrm{cov}(X_1', Y') \approx \mathrm{cov}(X_1, Y)$$

$$\mathrm{cov}(X_2', Y') \approx \mathrm{cov}(X_2, Y) + \mathrm{var}(N_{-1})$$

Substituting these approximations into (26) and solving for K_2 yields the same equation for line 1 (20) in K_2 against K_1 co-ordinate plane, (as shown in Fig. 2) while line 2 with noise becomes

$$\text{line2: } K_2 \approx \frac{\mathrm{cov}(X_2, Y) + \mathrm{var}(N_{-1}) - \mathrm{cov}(X_1, X_2) \cdot K_1}{\mathrm{var}(X_2) + \mathrm{var}(N_{-1})} \tag{27}$$

From the definition of $X_2 : X_2[k] = \theta_{amb}[k] - \theta_{top}[k-1]$, $\mathrm{var}(X_2)$ is the variance of the top-oil-rise. Therefore for a normal magnitude of noise, it may be assumed that $\mathrm{var}(X_2) >> \mathrm{var}(N_{-1})$. Thus (27) becomes approximately

$$\text{line2: } K_2 \approx \frac{\mathrm{cov}(X_2, Y) + \mathrm{var}(N_{-1}) - \mathrm{cov}(X_1, X_2) \cdot K_1}{\mathrm{var}(X_2)} \tag{28}$$

Equation (28) indicates that, after noise is added to TOT, the slope of line 2 has changed little, whereas its intercept will increase. Also notice that normally the slopes of both lines are positive since usually $\mathrm{cov}(X_1, X_2) < 0$ due to the positive correlation between load and top-oil-rise. The increase of the intercept of line 2 given, for the example in Fig. 2, by $\mathrm{var}(N_{-1})/\mathrm{var}(X_2) = 0.08/24402/(960-1) \approx 0.003$ for $m = 960$) will then shift the intersection of the two lines to the upper right, causing both K_1 and K_2 to increase, as shown in Fig. 2.