A Pass-Fail Algorithm for Generating Hesterman's Transformer Model

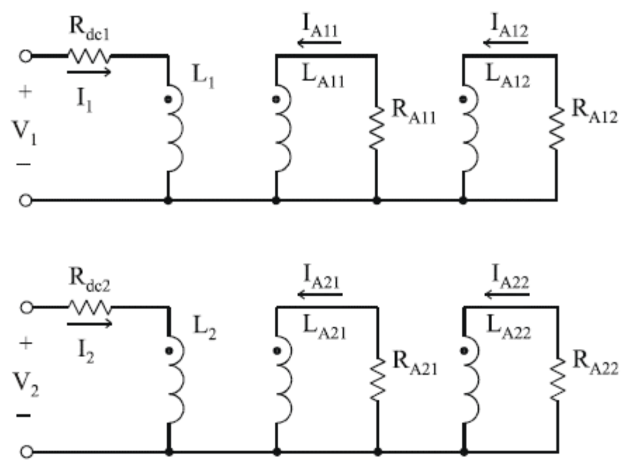
by Norman J. Elias

October 23, 2020

Introduction:

In January 2000 Mombello and Möller [1] published a transformer model that accounts for frequency-dependent impedances, losses, and resonances. In subsequent discussions with the original authors, Bryce Hesterman [2] proposed several refinements including revised curve-fitting routines. Curve-fitting automates parameter selection but must be initialized by a reasonable choice of values for those parameters. In his comments, Hesterman noted that he was still seeking "…a method for finding a good initial guess" to use in the curve-fitting routines.

Earlier this year, I became aware of this need, when I agreed to port Bryce's Mathcad parameter extraction routines [3] to Excel. After implementing his algorithm for the two-winding transformer model I began to explore a solution to the initial guess problem. My goal was to find a reliable method that executes in a time-frame comparable to or better than the curve-fitting routine itself. This report presents the results of that study and introduces the prospects it opens for further refinement of the curve-fitting routine.



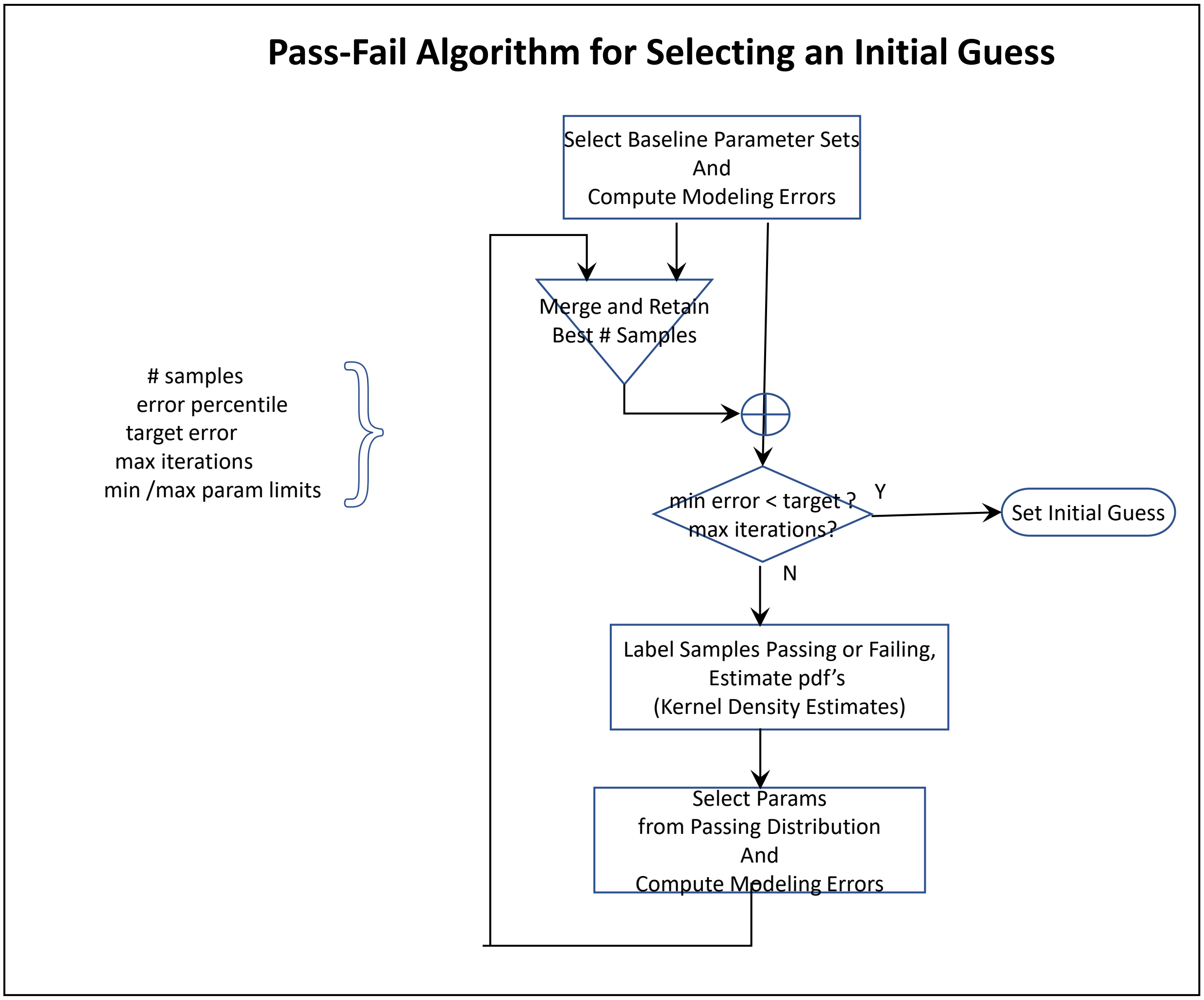
Coupling Coefficients:

k_{11} L_1 L_{A11}	k_{21} L_2 L_{A11}	k_{ps} L_1 L_2
k_{12} L_1 L_{A12}	k_{22} L_2 L_{A12}	
k_{13} L_1 L_{A21}	k_{23} L_2 L_{A21}	
k_{14} L_1 L_{A22}	k_{24} L_2 L_{A22}	

Hesterman's two-winding transformer model.

My solution to this problem is an adaptation of a statistical inference methodology I originally developed to optimize component tolerances for maximum circuit yield [4]. In that context I referred to that methodology as Pass-Fail analysis. In the following paragraphs, I describe the Pass-Fail approach as I've applied it to initializing Bryce's curve-fitting routines, I exemplify its operation by applying it to the two-winding transformer model. In anticipation of follow-up investigations, I introduce an additional feature of Pass-Fail analysis that provides important insights into the algorithm.

Solution By Pass-Fail Analysis:



Pass-Fail methodology uses statistical inference to provide a computationally efficient Monte Carlo search for the initial parameter set. A range of parameter values is much easier to provide than a single initial guess and may even be programmed into the software. In principle, random sampling can ultimately identify an acceptable initial set of model parameters or, even, a final selection, if the number of samples is large enough. However, the modeling error figure of merit for each parameter set provides information that can guide the search toward a solution with much fewer samples.

The flow chart above outlines an iterative search that uses computed errors to grade each sample on a pass-fail basis. To control the number of samples in each group, the passing level is set at a fixed (usually 20th but modifiable by the user) percentile level. After an initial random baseline, iterations select new parameter values from the distribution of passing values. At each iteration, the minimum error is identified and compared to a target value to terminate the iterations. A limit on the final number of iterations prevents the process from running indefinitely. Since the minimum error of a set of samples is, itself, a random variable, half of the samples used at any iteration are retained from previous iterations. Consequently, the minimum error at each iteration is a non-increasing function of the iteration count.

Appendix A derives and explains some of the mathematics employed by this algorithm.

In addition to search routines and tolerance (yield) optimization, Pass/Fail methodology has potential application to a variety of topics including …

- Tolerance/yield optimization:
- Failure Diagnostics:
- Design/product Reliability, MTTF:
- Cost reduction:
- Thermal Design:
- Statistical sensitivity analysis:
- Analog Testing:

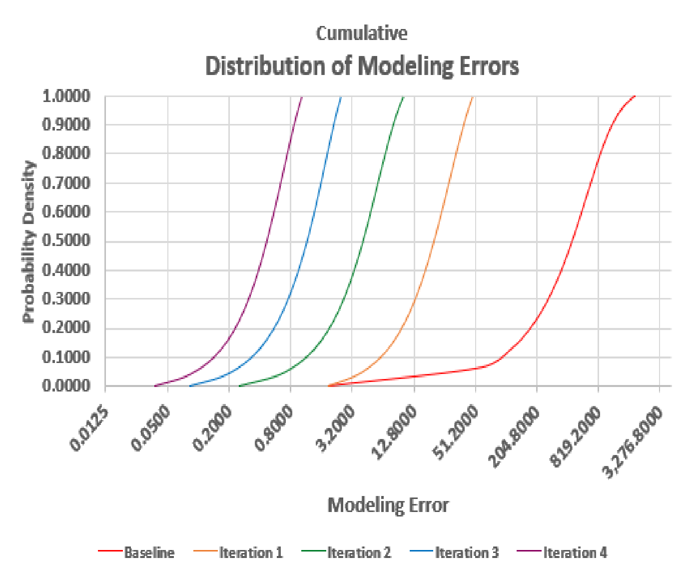
A Two-Winding Transformer Example:

I tested the algorithm on a two-winding model using measurements provided by Bryce Hesterman [3]. This is the same transformer I used to develop the spreadsheet. The goal was to demonstrate convergence to an acceptable error measure with a run time comparable to or better than the curve fitting solver. Tests were run with two different ranges of resistor and coupling coefficient values. To view progress through four iterations, I set a low error target of 0.01. To limit the run time, I held the number of samples at 64 (a power of 2).

The adjacent figure shows the results for the first range of parameter values tested. The family of curves plotted represent the cumulative distribution of modeling errors calculated for the baseline and four successive iterations. Each curve ranges from zero at the lowest error up to one at the highest error. Points along the curve represent the probability of obtaining a modeling error at or below the x-axis value.

Search Range	min	max
RA11,RA21	100	1000
RA12,RA22	2000	10000
k	0	0.5

Iteration	Cumulative Execution Time (sec)	min(error)
Baseline	5.61	1.858551E+00
Iteration 1	26.96	1.858551E+00
Iteration 2	47.96	2.576789E-01
Iteration 3	69.13	8.250658E-02
Iteration 4	90.24	3.803676E-02

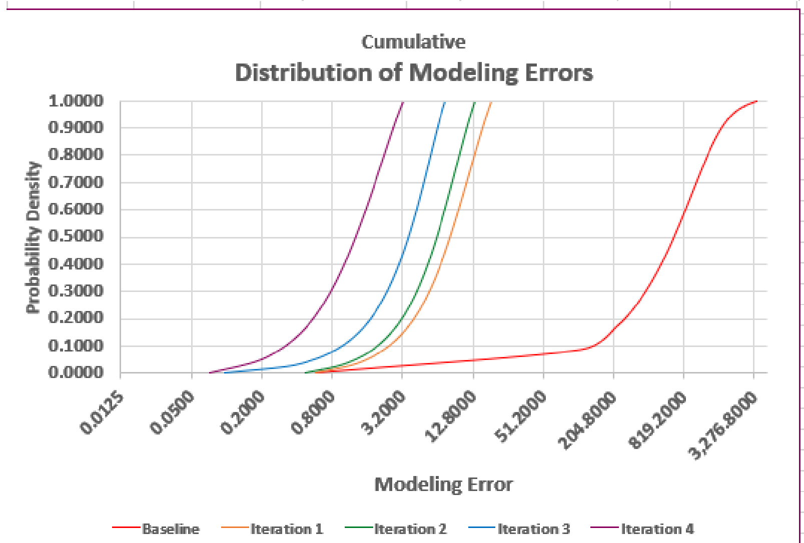


The first important observation is that these curves shift consistently to the left (lower error values) from one iteration to the next. This is proof that the Pass-Fail approach is successful in driving down the modeling error. The second significant observation is that these curves are confined to an increasingly narrow range of error values. This is indicative of a convergent process. The downside of this observation is that the lowest error in the distribution is decreasing more slowly than any other value along the curve. The bottom line on this observation is that this approach would be very inefficient at locating the absolute minimum error.

The upside of these results is that the algorithm is proven capable of providing the desired initial guess. Minimum error values from each iteration are tabulated above the distribution curves along with actual elapsed run time measured from the start of the baseline computation. After completing the Baseline phase in under 6 seconds, the Pass-Fail algorithm averaged about 20 seconds per iteration, completing four iterations in 90 seconds. This compares favorably to the block solver for which typical elapsed time varies between 50 and 200 seconds depending on how far it can reduce the modeling error.

Search Range		
	min	max
RA11,RA21	100	2000
RA12,RA22	2000	40000
k	0	0.95

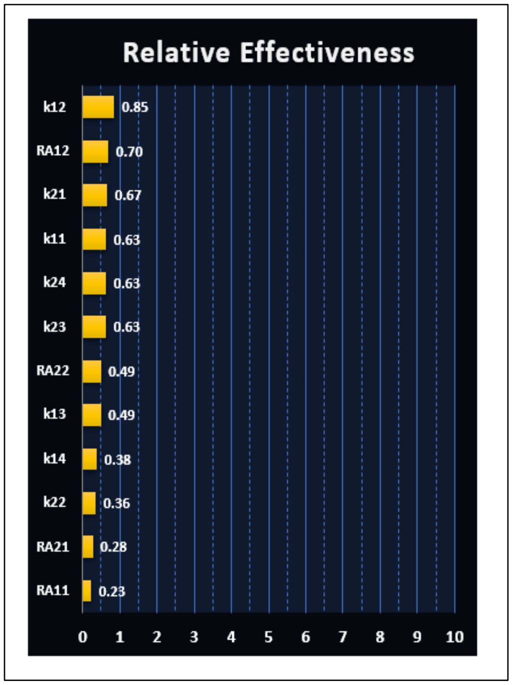
Iteration	Cumulative Execution Time (sec)	min(error)
Baseline	2.86	5.709580E-01
Iteration 1	23.47	5.709580E-01
Iteration 2	43.82	4.670602E-01
Iteration 3	65.23	9.723732E-02
Iteration 4	88.98	7.111427E-02



A second test of the algorithm confirmed the previous observations as can be seen from the adjacent figure. Substantially similar performance was obtained for a significantly wider range of parameter values.

The Relative Effectiveness Metric (RE):
An important by-product of the Pass-Fail algorithm is a statistical measure that can be used to quantify the contribution each of the model parameters makes towards determining the modeling error. The bar chart below presents RE data for the first of the two examples

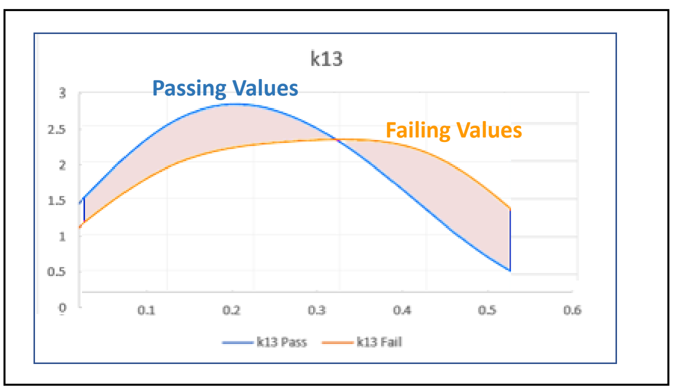
reported above. It ranks each parameter on a scale from 0.0 to 10.0 with 0.0 indicating that the parameter has no effect on the modeling error and 10.0 indicating that the parameter completely determines the modeling error. The significant results are that the numbers are all similar and that they are all relatively small. I'll explain the RE metric after discussing these results and the consequences they have toward follow-up studies of the curve fitting operation.



Similar values indicate a degree of balance that is probably a desirable attribute of the model. If the Pass-Fail analysis were repeated several times using modeling errors at specific frequencies, it would show us whether, for example, any particular pairing of auxiliary winding resistance and coupling coefficient such as RAij/kij dominates the error at some specific frequency range. It would be interesting to see if frequency dependencies such as this could be used to optimize the curve fitting operation either in terms of reducing the modeling error or speeding up the computation of that result.

The relatively small RE results suggest further exploration to optimize the curve fitting by substituting some alternative to the model parameters. For example, there might be a more effective set of expressions from which we can calculate the resistance and coupling coefficient parameters. These two observations are left to be explored in future studies.

The relative effectiveness metric is best explained in terms of the adjacent Pass-Fail diagram. The diagram shows the passing and failing probability distributions for a single parameter. The shaded region is the area between these two curves. That area goes to zero if the distributions overlap completely, to two if the distributions do not overlap anywhere and to an intermediate value in all other cases. Scaling that area by a factor of five produces an RE metric ranging from zero to ten.



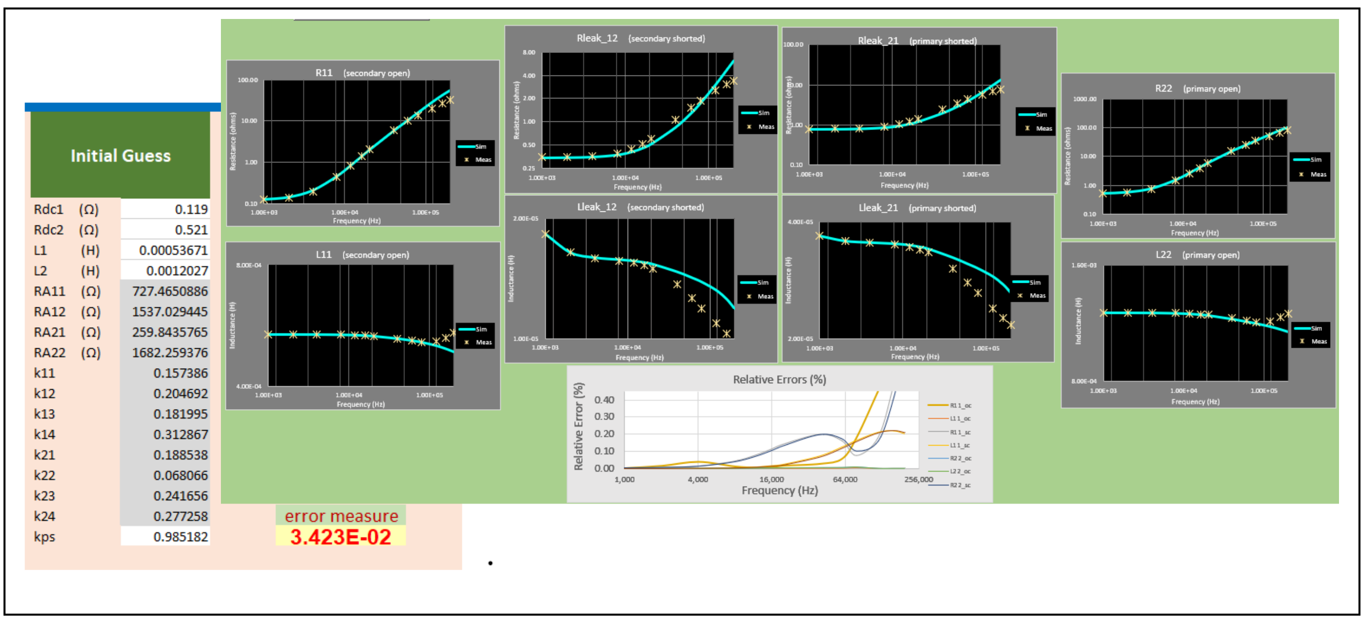
To understand how this metric reflects the effectiveness of the parameter, consider the two extreme cases mentioned above. Overlapping pass and fail distributions mean that any parameter value is equally likely to appear in a passing (low error) model or in a failing (higher error) model. This can only happen if the pass-fail distinction is completely determined by some other parameter or combination of parameters. RE=0 means that the parameter has no effect on the modeling error.

At the other extreme, the passing and failing distributions range over completely separate values. Every value of that parameter is always associated with the same outcome, either passing (low error) or failing (higher error) regardless of the values of all other parameters. This can only happen if the pass-fail distinction is completely determined by the parameter in question. RE=10 means that the parameter has 100% effect on the modeling error. Hence RE values for each parameter reduce the comparison of parameters to a list of single numbers.

It is important to bear in mind, however, that the RE measures are statistics, random variables that can vary over a range of values. Values for various parameters must be statistically different for any conclusions to be drawn. Appendix B analyzes the statistics of the relative effectiveness measure.

Summary and Conclusion: The bottom line is that the Pass-Fail algorithm solves the problem of initialing the curve-fitting algorithm for Hesterman's mutual impedance transformer model. The function of the Pass-Fail initialization is to replace the initial guess with a, more easily chosen, initial range of parameter values. The Pass-Fail analysis is iterative and, in theory, capable of completing the curve-fitting operation. However, the directed search of the block solver is much more efficient at that task. For that reason, it is best to restrict the Pass-Fail algorithm to the initialization task with a modest target error on the order of 0.05 to 0.10 and then apply the block solver.

In the following figure an initial guess based is compared to measurements. The parameter values were computed in a single iteration with a target error of 0.05. At that error level the



parameters match the measurements reasonably well except for the leakage inductances at frequencies above 50KHz.

In addition to the main topic of curve fitting, the report introduces a relative effectiveness (RE) metric, which grades the parameters in terms of the degree to which it influences the modeling error. Example results demonstrate a relatively uniform sharing of influence among the parameters at a consistently low level. These observations suggest follow-up studies into prospects for either speeding-up the curve-fitting operation or reducing the modeling errors even further. Possibilities mentioned include resolving the modeling error into low, medium, and high frequency components to see if the added structure is beneficial to the curve-fitting operation. A second suggestion would be to introduce an equal number of new variables that have a one-to-one dependence on the original set (e.g., translations and rotations within the space of all parameters). These ideas are under consideration.

Acknowledgements

My original interest in this aspect of transformer modeling came at the suggestion of Ray Ridley who directed my attention to proximity effects. He notes that frequency-dependent winding losses are commonly overlooked in power supply design. I am especially indebted to Bryce Hesterman who introduced me to the mutual-impedance transformer model and provided the algorithm that I implemented in Excel as well as the transformer measurements used in my studies. It is my intent to make the Excel available to anyone interested and to share credit with those who have guided my work.

References

[1] E. E. Mombello and K. Moller, "New power transformer model for the calculation of electromagnetic resonant transient phenomena including frequency-dependent losses," IEEE Transactions on Power Delivery, vol. 15, No. 1, January 2000, pp. 167-174.**https://ieeexplore.ieee.org/document/847246/**

[2] B. L. Hesterman, E. E. Mombello and K. Moller, "Discussion of "New power transformer model for the calculation of electromagnetic resonant transient phenomena including frequency-dependent losses" [Closure to discussion]," in IEEE Transactions on Power Delivery, vol. 15, no. 4, pp. 1320-1323, Oct. 2000.**https://ieeexplore.ieee.org/document/891529**

[3] Bryce Hesterman, " Mutual Impedance Transformer Model (Rev 6)" in http://verimod.com/pdf_files/Mutual_Impedance_Transformer_Model_rev_6.zip , March 5, 2020

[4] N. J. Elias, "New statistical methods for assigning device tolerances", *Proc. Int. Symp. Circuits and Systems*, pp. 329-332, 1975. https://ci.nii.ac.jp/naid/80013145034/

Appendix A – Constructing the Passing Distribution and Selecting Model Parameters from It

Constructing the Passing Distribution:

Given the formula for the passing pdf, the next challenge is to pick a random value of x_i and in This appendix covers both aspects of the parameter selection process. The passing distribution must be constructed reliably before an effort is made to select from that distribution. Looking at the numbers, 10% to 20% of a 64-sample data set is a very sparse population. An estimate drawn from 80% to 90% provides at least 50 samples to work with. For that reason, the passing distribution is derived from the failing samples as follows.

Let **A** be the event that the model lies within the 20th percentile (or whatever value we choose). If P(**A**) is the probability that a model chosen at random will be a member of the group that meet that test, then by definition of percentiles, P(**A**) = 0.2 and 1 - P(**A**) = 0.8.

If x_i represents the ith of 12 parameters selected at random, it will have a statistical distribution over the range of values allowed for selecting parameters. Probability theory defines two representations for the distribution of x_i. The cumulative distribution (CDF) $P_i(x_i)$ is the probability that any value of this parameter lies at or below the numerical value chosen for this parameter. If I use $\underline{x}_i$ to denote an arbitrary choice of the parameter and if I stick to x_i as a test level, then $P_i(x_i)$ = Probability($\underline{x}_i \le x_i$). Probability theory also defines a probability density function (or pdf) $p_i(x_i)$ such that $p_i(x_i)\,dx_i$ represents the probability that a chosen value $\underline{x}_i$ will lie between x_i and $x_i + dx_i$. The pdf and CDF are related by

$$P_i(x_i) = \int_{-\infty}^{x_i} p_i(\xi)\,d\xi$$

where ξ is a dummy variable.

Probability theory also defines conditional probabilities, distributions, and densities. For example, $P_i(x_i|\mathbf{A})$ and $p_i(x_i|\mathbf{A})$ represent, respectively the cumulative distribution and probability density functions of x_i values for which event **A** applies meaning for values of parameter x_i for which the associated model lies in the 20th percentile of modeling errors. In other words, $p_i(x_i|\mathbf{A})$ is the passing pdf of x_i and $p_i(x_i|\textbf{not A})$ is the failing pdf.

The law of total probability relates all these pdf's by the relation

$$p_i(x_i) = p_i(x_i|\mathbf{A})P(\mathbf{A}) + p_i(x_i|\textbf{not A})(1-P(\mathbf{A}))$$

which simply states that the density is a weighted sum of the only two possibilities. Either x_i is a sample from a model in the 20th percentile, which happens 20% of the time or it's from a model from the rest of the population , which happens 80% of the time. If we solve the above equation for $p_i(x_i|\mathbf{A})$ we get a relation that lets us compute the passing distribution from the failing distribution and the total distribution as

$$\begin{aligned} \mathrm{p_i}(\mathrm{x_i}|\mathbf{A}) &= \frac{\mathrm{p_i}(\mathrm{x_i})\text{-}\mathrm{p_i}(\mathrm{x_i}|\mathbf{not\ A})(1\text{-}\mathrm{P}(\mathbf{A}))}{\mathrm{P}(\mathbf{A})} \\ &= 5\ \mathrm{p_i}(\mathrm{x_i})\text{-}4\ \mathrm{p_i}(\mathrm{x_i}|\mathbf{not\ A}) \end{aligned}$$

Note that this relation also means

$$\begin{aligned} \mathrm{p_i}(\mathrm{x_i}|\mathbf{A})\text{-}\mathrm{p_i}(\mathrm{x_i}|\mathbf{not\ A}) &= \frac{\mathrm{p_i}(\mathrm{x_i})\text{-}\mathrm{p_i}(\mathrm{x_i}|\mathbf{not\ A})}{\mathrm{P}(\mathbf{A})} \\ &= 5\left(\mathrm{p_i}(\mathrm{x_i})\text{-}\mathrm{p_i}(\mathrm{x_i}|\mathbf{not\ A})\right) \end{aligned}$$

Selecting Model Parameters from the passing distribution:

Given the formula for the passing pdf, the next challenge is to pick a random value of x_i in such a way that the pdf of all such x_i will equal the passing pdf. The most computationally efficient way to do that involves the cumulative distribution $P_i(x_i|\mathbf{A})$, which we can approximate from pdf values at discrete values of the parameter. If parameter x_i ranges from $x_{i,min}$ to $x_{i,max}$, we can partition that range into N equal intervals of width $\Delta_i = (x_{i,max} - x_{i,min})/N$. At each interval, the CDF would be approximated as

$$\mathrm{P}_i\left(\mathrm{x}_{i,K}|\mathbf{A}\right) = \mathrm{P}_i\left(\mathrm{x}_{i,min} + K\ \Delta_i|\mathbf{A}\right) = \sum_{k=0}^{K} \mathrm{p_i}\left(\mathrm{x}_{i,min} + \mathrm{k}\Delta_i|\mathbf{A}\right)\Delta_i$$

Intermediate values would be calculated by interpolation. For any given value of P_i, the x_i would be computed by searching through the discrete values for upper and lower estimates and interpolating between them. By this mechanism we can select x_i by selecting a random value from a uniform distribution between 0 and 1 and using that selection as a value of P_i.

By this mechanism, the probability of selecting a value of x_i at or below any $\underline{x}_i$ would equal the probability of selecting a uniform random variable at or below P_i and that probability is equal to P_i. By that reasoning, the values of x_i selected would have P_i as their CDF as required.

Appendix B – Statistics of The Relative Effectiveness (RE) Metric:

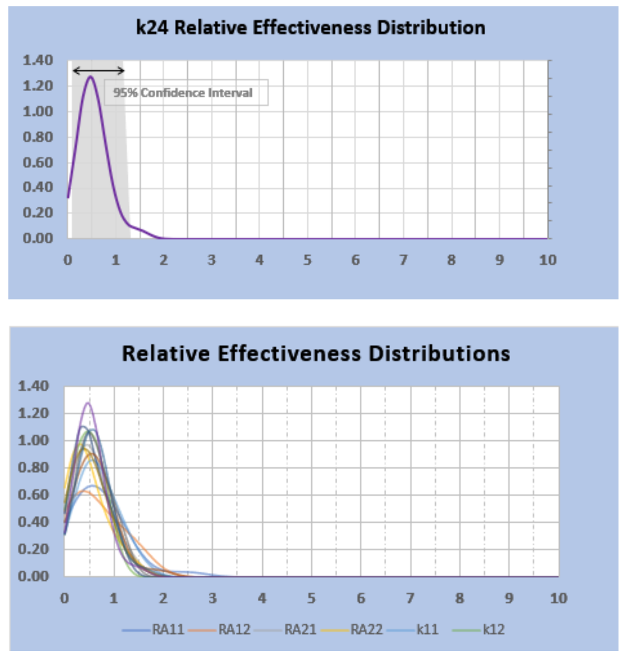
Using the notation introduced in Appendix A, we can express the area between the passing and failing distributions of the i^{th} model parameter as:

$$\mathrm{RE_i} = 5\int_{-\infty}^{\infty} |\ \mathrm{p_i}(\mathrm{x_i}|\mathbf{A})\ \text{-}\ \mathrm{p_i}(\mathrm{x_i}|\mathbf{not\ A})\ |\,\mathrm{d\,x_i}$$

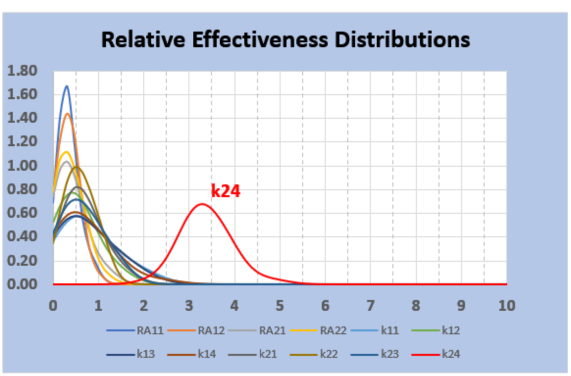
As described in the report, variations from parameter to parameter measure the relative effect of each parameter on the model error. But there are also statistical variations because

1. The passing and failing distributions are estimates generated from a finite number of samples and
2. The above is estimated from a finite sum.

These statistical variations are the noise that obscures small parameter to parameter variations. The figure below shows distributions of the RE metric computed from 25 repetitions of the baseline analysis for the first example reported above. The RE distribution for coupling coefficient k24 shows that 95% of the statistical variations lie within a window a little more than ±0.5 wide. Distributions of all parameters are of comparable width. These results suggest a rule of thumb judgement that parameter to parameter variations less than 1.0 be overlooked.



As an illustrative example I ran another set of 25 RE computations with k24 values selected over the full range of 0.0 to 1.0 but with all other parameters confined within a narrow range. The figure below demonstrates how the RE metric reveals the dominant parameter.



Composite Plot Comparing RE Values with one parameter dominant