Nonlinear Programming Applied to Calibrating Thermal and Fluid Models to Test Data

Jane Baumann, Brent Cullimore
C&R Technologies
Littleton, Colorado
303-971-0292

Abstract

Thermal modeling is fraught with uncertainties such as film coefficients, contact resistances, dissipation rates, and effective conductances and capacitances of complex components. Adjusting the values of uncertainties in a thermal/fluid model to achieve a better fit with test data is a necessary step; this procedure is even codified into military standards for electronic equipment design, for example.

Nonetheless, such "correlation" or "calibration" activities are typically done haphazardly and without any mathematical rigor, and are often impeded rather than aided by software.

This paper shows how readily available nonlinear programming (NLP) techniques that were developed for optimization problems have been successfully used to automate this critical but laborious calibration task. This paper briefly introduces NLP concepts, and then demonstrates their application both to a simplified curve-fitting exercise as well as a real case: a transient with a serpentine condenser plate.

Keywords: Calibration, correlation, validation, optimization, thermal analysis, parametric modeling, design automation.

Uncertainties in Thermal/Fluid Analysis

Variation can be classified in three categories:

1. uncertainties in performance parameters: contact resistances, film coefficients, dissipation levels, effective thermal capacitances and conductances of complex components, etc.
2. environmental or usage uncertainties: ambient temperature and humidity, duty cycle, etc., as well as degradations over the maintenance life of the product
3. unit-to-unit (manufacturing) variations: bonding, fan performance, filter resistance, etc.

Each category of variation is traditionally handled using different approaches. Because of differences between organizations, products, etc., the following attempt to describe "typical" approaches is necessarily a generalization.

Reference 1 describes the use of statistical design techniques for treating certain classes of uncertainties. This paper describes complementary techniques for reducing design uncertainty by calibrating some or all of the underlying thermal/fluid model to any available test data, perhaps by exploiting tests performed on previous versions of the vehicle or product. Such a calibrated model then can be used with greater confidence to predict design performance in untested or even untestable conditions. Model calibration is a necessary step in many industries and organizations, with both the model and its calibration requiring independent reviews.

Traditional Calibration "Techniques"

Values for performance uncertainties can be calculated from limited test data. Unfortunately, because of the system-level interactions of radiation and fluid flow, it rarely makes sense to perform thermal tests at low levels of assembly, and this means that the thermal/fluid model to be calibrated contains several (perhaps 5 to 30) simultaneous unknowns. Also, some unknowns (e.g., film coefficients) will vary over a range of test conditions (e.g., fan speeds).

When faced with many uncertainties and copious test data, engineers most often address each uncertainty serially: the parameter judged to be the most influential is corrected first, then left at a fixed value. The second parameter is subsequently adjusted, ad nauseam. Most analysis software makes it difficult to make sweeping changes in input values, even between runs. Therefore, because of the labor and tedium involved, rarely is the above cycle repeated: the initial value found for the first parameter is usually not rechecked once values for all of the other parameters have been determined.

In other words, current methods used for calibrating (or "correlating") models are primitive: repetitive analysis runs are made varying one parameter at a time. Worse, selection of best-fit values is most often based on a visual comparison of plots of test data versus predictions. The current "algorithm" for model calibration is then:

1. Choose the parameter with the most uncertainty and/or the parameter judged to have the greatest importance on the results.
2. Create a plot of the results based on a guessed value of the uncertain parameter, and make repeated runs until a better fit is visually evident. If allowed by the thermal/fluid analysis software, make a parametric sweep of the uncertain parameter and select the value that results in the best fit.
3. Choose the next most important/uncertain parameter on the list, and go to step 2. Continue through the list of uncertainties until either the desired match (e.g., error threshold) is achieved, or until the parameter list has been exhausted.

As will be described next, a superior calibration results by varying all parameters simultaneously and by using more mathematical rigor when making comparisons between test data and predictions. An important benefit of this new approach is that the laborious methods that were described in this section can be replaced by an automated search for the best fit.

Nonlinear Programming: Generalized Tasking

Nonlinear programming (NLP) techniques attempt to find the maxima and minima of an *objective function* in N dimensions, while obeying arbitrarily complex constraints. Many algorithms exist for solving such problems, as do several off-the-shelf software packages. For example, the Solver module in Microsoft's Excel® spreadsheet software is representative of this class of algorithm.

Formal mathematical descriptions of NLP techniques are not necessary to understand their importance to thermal/fluid model calibration and other automation tasks. Rather, it is important to understand the four parts of a optimization task, as listed below (and as depicted in Figure 1):

1. The objective function: an arbitrarily complex figure of merit to be maximized or minimized.
2. The design variables: the parameters whose values at the optimum point need to be determined.
3. Constraints: arbitrarily complex relationships that distinguish feasible design points (sets of values of design variables) from useless ones.
4. Evaluation procedures: the generation of current values of the objective function and the constraint functions given a current set of design values.

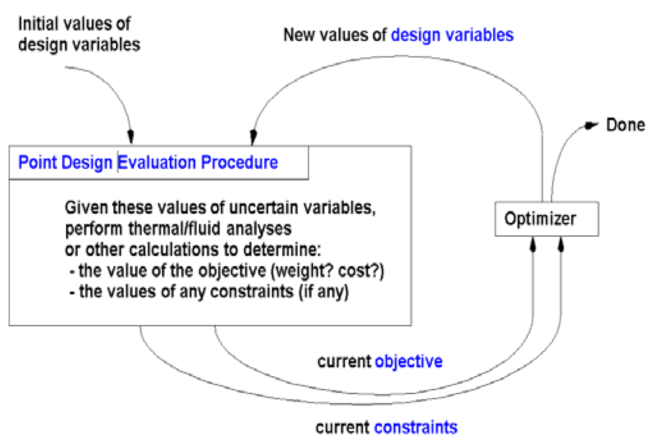

Figure 1: Four Concepts in Optimization

Many engineers have seen these algorithms applied to design optimization: the generation or synthesis of a design that minimizes weight or cost, or that maximizes performance. However, the math underlying NLP techniques can be applied to a wide variety of tasks: it is a generic means of defining a complex task or search.

For example, NLP algorithms can be applied to generating worst-case design scenarios, which represents yet another means of dealing with uncertainties in thermal/fluid design. In this case:

1. The objective function is the temperature of some component to be maximized ("find the hot case") or minimized ("find the cold case").
2. The design variables are the uncertainties. Especially common are environmental uncertainties such as ambient temperature, humidity, pressure (or altitude for avionics applications), and orientation (for aircraft and spacecraft applications).
3. The evaluation procedure might consist of a steady-state thermal/fluid analysis or a transient scenario that yields the temperature of critical components, given specific values of uncertainties ("design variables") as inputs.

Note that constraints are optional.

Applying NLP to Model Calibration Problems

In this paper, NLP techniques are described in relationship to model calibration tasks. One example of such a usage results in the following interpretations (Figure 2):

1. The objective function is the difference between tests and predictions, to be minimized. (There are many ways to define such a function, as will be described later.)
2. The design variables are the uncertainties: the bond resistance, the filter blockage or permeability, the fan efficiency, the film coefficient, etc.
3. The evaluation procedure might consist of a steady-state thermal/fluid analysis or a transient scenario that yields the temperature (or pressure etc.) of measured points, given specific values of uncertainties ("design variables") as inputs.

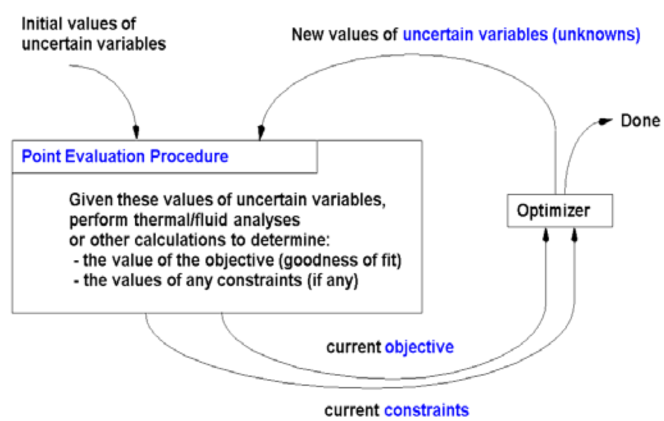

Figure 2: NLP Concepts Applied to Calibration

Again, constraints are optional. (Upper and lower limits on design variables are important, but are not true mathematical constraints and are often referred to as "side constraints.")

The reader will note that the above examples leave plenty of room for interpretation. This is an important feature: the engineer retains complete control over what is uncertain (and by how much), how to define a good fit, and how to minimize the computations required to find that fit. For example, it is possible to calibrate to temperature derivatives in time instead of temperatures, or to find the least cubes fit instead of the least squares fit, or to add weighting factors to critical measurements etc.

However, it is not the purpose of this paper to exhaustively list all of these possibilities. Instead, the basic concepts will be clarified via specific examples with the understanding that many, many more customizations are possible.

Example: Simple Curve Fitting

To illustrate the application of optimization concepts to calibration of models, an industry- and model-independent demonstration of a polynomial curve fit will be used.

Assume that 13 data points (depicted in Figure 3) are to be fitted to a simple third order polynomial:

$$Y_p = A + BX + CX^2 + DX^3$$

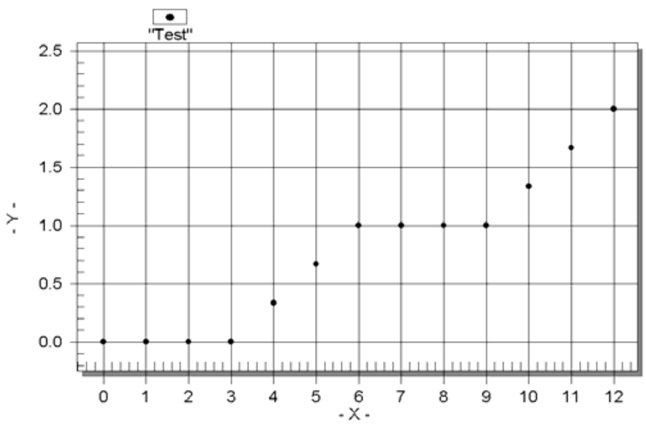

Figure 3: "Test Data" to be Curve Fit

For this example, the above equation is the "model" and the "uncertainties" are the four variables A, B, C, and D. To cast this into an optimization format requires that decisions be made regarding how to define a good fit. For example, using a root sum of squares (RSS) as the objective to be minimized is equivalent to a least squares curve fit. For each of the thirteen points:

$$\text{OBJECT} = \text{SQRT}[\ \sum_{i=1,13}(Y_{t,i} - Y_{p,i})^2\ ]$$

$Y_{t,i}$ is the test data and $Y_{p,i}$ is the prediction at the i^{th} point based on the "design variables" A, B, C, and D. OBJECT is the current value of the objective function, which is to be minimized. No constraints are needed, although upper and/or lower limits could be imposed on the design variables. (No such limits are applied in this simple example.)

The "evaluation procedure" consists simply of calculating the thirteen values of $Y_{p,i}$ given current values of A, B, C, and D, then computing the above objective.

The results of this exercise are shown in Figure 4.

Figure 4 also depicts the results of an alternative definition of a good fit: minimized maximum error ("Minimax"). The Minimax method often produces better fits to data, but is more sensitive to noise in the test data and often slower to solve. Also, for most NLP algorithms the simple replacement of an objective function with OBJECT = $|Y_{t,i} - Y_{p,i}|_{max}$ is unacceptable because it introduces discontinuities.

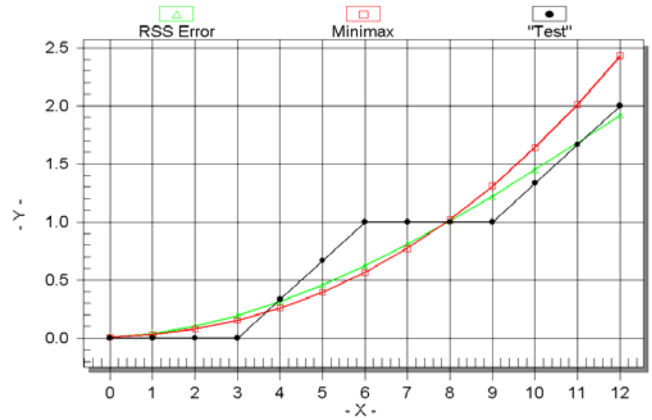

Figure 4: RSS and Minimax Curve Fits

To avoid discontinuities, a fifth design variable "E" is introduced and set equal to the objective function to be minimized (i.e., "OBJECT = E"). Then thirteen constraints are generated, one for each (i^{th}) data point/prediction pair:

$$-E < (Y_{t,i} - Y_{p,i}) < E$$

More details on the uses of Minimax methods, along with examples, are presented in Reference 2 (see Section 5 and Sample Problem E). The point of introducing this alternative here is to illustrate the flexibility available to the engineer in defining the calibration problem. Other possible objectives include minimizing cubic or quartic errors, standard deviations, and weighted error (i.e., make calibration at some points more important than at others).

Of course, the usefulness of the resulting calibrated model (in this case, the third order polynomial with four fitted values of the coefficients) is dependent on the model itself. A fourth order polynomial would have generated a better fit, as would many other functions. More critically, a poorly chosen predictive formula would always result in erroneous predictions of test data, no matter how well it was calibrated or fit. This is analogous to calibrating an inappropriate or error-ridden thermal/fluid model: *calibration can't fix a bad model*. This will be discussed further in a later section.

Example: Condenser Transient

This section demonstrates the application of automated model calibration techniques to an actual test of an ammonia condenser.

A thick aluminum plate (65kg) is bonded to a serpentine duct, as depicted in Figure 5 (the uneven spacing is intentional: the sketch is approximately to scale). The duct is not plain piping, but rather internally grooved to enhance condensation: it is a trapezoidally axially grooved (TAG) aluminum heat pipe extrusion, although it was not used as a heat pipe in this test.

The plate is attached to a cold sink via a malleable, conductive pad, but this pathway does not provide sufficient rejection for the heat load that will be supplied. Instead, the plate is initially cold and warms up over the course of a 42 minute transient event.

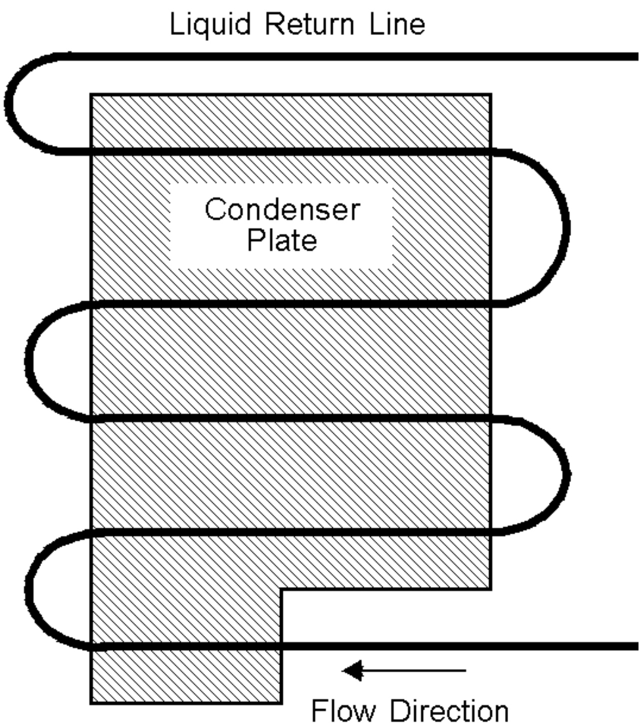

Figure 5: Geometry for Condenser Plate Transient

Initially, the entire system is quiescent at 6.6˚C: the ammonia within the condenser is stagnant liquid. At time zero, saturated ammonia vapor at 29˚C is supplied upstream at a rate corresponding to a heat input of 510W. As the plate warms, the condensation point progresses through the plate until it has reached the exit and the plate can no longer provide complete condensation.

Data at three points along the condenser was available as functions of time.

Assuming that the 29˚C saturation temperature and 510W evaporative input are both correct (they will be selected as uncertainties later), a simple SINDA/FLUINT (Ref 2) thermal/fluid model of the system was generated. Vendor-supplied data was used for the conductive pad, and the condensing film coefficient in the grooved tubing was estimated using a correlation generated for heat pipes.

The test data (black/solid) and the initial predictions (blue/dashed) for the transient event are provided as Figure 6. The top curves correspond to a point near the condenser inlet, whereas the bottom curves correspond to a point near the outlet.

The first step towards calibrating the model to the test data is to identify the key uncertainties. Based on engineering judgement, four quantities are identified along with limits on their reasonable range of variation:

1. The power input into the vaporizer. Although measured to be 510W, a measurement error of 5% is allowed. Also, heat leaks may cause less than the full amount of heat to flow into ammonia. The power input is therefore allowed to vary from 90% to 105% of the nominal value.
2. The saturation temperature of the ammonia system, which was measured in the test to be 29˚C at a reservoir. A 0.5˚C uncertainty is assigned to this value, plus an additional 0.5˚C on the upper end because the vapor in the reservoir can compress during start-up (and this effect might not be evidenced in the thermocouple). Thus, the saturation temperature is allowed to range from 28.5˚C to 31˚C.
3. The thermal resistance of the conductive pad is suspect since vendor data was used and might therefore be optimistic. Also, unit-to-unit variation exists due to clamp pressures and hysteresis (from previous clamp/release cycles). A large uncertainty is therefore allowed in this parameter: from 50% to 150% of the nominal conductance value. This correction factor is assumed constant throughout the pad.
4. The condensation coefficient. The film coefficient correlation used in the condenser may not be appropriate for forced flow. A single correction factor on the resulting coefficients is therefore applied throughout the condenser, with values of between 75% and 125% of nominal assumed.

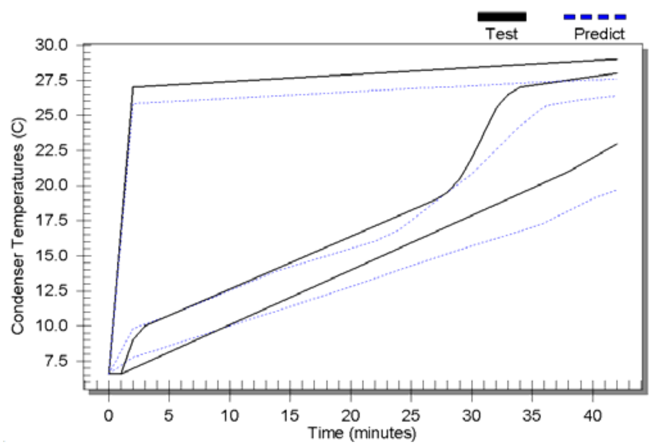

Figure 6: Test Data vs. Pre-calibration Predictions

The above uncertain parameters are applied as “design variables” to the NLP solver, with limits applied as side constraints. The evaluation procedure is to generate the transient temperature profiles using current values of the four parameters, and compare these with the test data to generate the objective function value.

As with the previous curve fitting example, two definitiions of the objective function are used: a “least squares” method (RMS or root-mean-square, which is equivalent to RSS since they have the same minumum) and “Minimax” (minimized maximum error) method. The results, generated using the built-in NLP “Solver” module in SINDA/FLUINT, are shown in Figure 7.

As can be seen in Figure 7, both methods return about the same predictions, resulting in very good agreement with the test data. The RMS method takes 37 evaluations (transient analyses) while the Minimax method requires almost twice as many. The RMS method returns a calibrated model with an RMS temperature error of about 1˚C, while the Minimax method returns a maximum error of about 2˚C.

However, the resulting values of the four uncertainties are not the same for both traces: multiple solutions exist. The RMS method resulted in 105% of the nominal input power (e.g., the limit) while the Minimax method required only 102%. Whenever such a limit is reached, then its selection must be questioned because the limit is influencing the answers. In other words, could a larger range of variation have been possible?

Both methods agreed that the saturation temperature was too low, but the RMS method returned a value of 30.2˚C while the Minimax method required a much smaller departure of 29.2˚C.

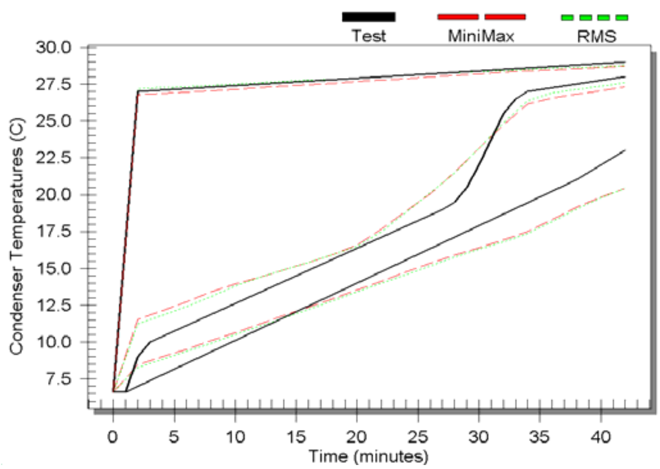

Figure 7: Test Data vs. Calibrated Predictions

However, the largest source of disagreement were the correction factors for the conductive pad and condensation heat transfer. The RMS method made hardly any change to the condensation heat transfer coefficient (101% of nominal) while the Minimax method lowered this factor to its lower limit (75% of nominal, again hitting a limit). Conversely, the RMS method dropped the pad conductance to 84% of the nominal value, while the Minimax method retained 95% of that conductance. At first, these disagreements might seem contradictory until one realizes that both factors are applied in parallel to the same heat flows: a reduction in one term results in about the same temperature predictions as does a reduction in the other term. The RMS method reduced one parameter and left the other alone, while the Minimax method reduced the complementary parameter, such that they both yeilded approximately the same temperature predictions (as evident in Fiture 7).

Also, both methods struggle to fit to the last (coldest) trace near the outlet towards the end of the transient. Although it is likely that this discrepancy is due to the simplicity of the model employed, it is also possible that some heat transfer pathway or physicial process was neglected. For example, spatial variations in the conductive pad performance are common but were neglected in the underlying model. This could have been accounted in the calibration procedure, but at the cost of a much greater number of uncertain parameters. For example, one could apply one adjustment factor for each of the 9 regions of the plate. Such an agumentation of uncertain parameters (from 4 to 13) would require even more transient evaluations (from 40 to about 200 in this case).

These difficulties have been elaborated in this example because they illustrate generalized points to be made in the next section. However, they should not detract from the fact that an *automatically* calibrated model resulted in a fit that was not only better than military standards, but was also better than was achieved using traditional (sequential, visual) techniques such as the labor-intensive ones that were described earlier.

Challenges for the Engineer

As was noted above, calibration can't fix a bad model. Despite all of the benefits of automated thermal/fluid model calibration techniques, analyst responsibility is not eliminated so much as shifted. The analyst retains the responsibility of building a sensible and complete model, with appropriate attention to the physics of each problem. In fact, because the model will be run parametrically many times, it might even have to be more robust and faster to execute than was tolerable using prior manual calibration techniques. Fortunately, these model preparations do not represent a departure from previous techniques or experience.

Challenges that might be new to the engineer using automated calibration techniques are listed below.

First, the choice of which parameters to declare as uncertain, and within which bounds, is critical. Failure to include a critical parameter or sufficient variation in a parameter can yield a false fit, yet too many parameters with bounds that are too liberal is inefficient.

Second, as was noted above, many different definitions of "best fit" can be mathematically specified. For example, a weighted least-squares is possible assigning more value to good correlation at critical components, or at critical simulation times, etc.

Third, although it is theoretically possible just to list all test cases with corresponding model runs as an "evaluation procedure" and activate all possible uncertain variables at once, huge efficiencies can be gained by a little preplanning and preparation of subset calibrations. For example, it is common practice to first calibrate thermal resistances/conductances to steady state test results, and then proceed to calibrating effective capacitances to transient test results.

Fourth, the engineer must accept or reject the resulting calibrated model, checking to see if limits in uncertainties have been reached. The engineer should consider improving the model or expanding the set of uncertain parameters as needed to achieve a reasonable fit.

This verification stage also includes searching for multiple solutions. The easiest way to check for the existence of multiple solutions is to rerun the problem using different initial values of the uncertain parameters, and see if either the same fit was achieved or if an equally good fit results using different final values of the uncertainies.

Challenges for Analysis Software

How does a thermal design engineer exploit the availability of these advanced techniques using their favorite thermal/fluid analyzer? Model calibration techniques involve a higher level of analysis beyond a traditional "point design simulation." Most engineering analysis software is set up to solve a deterministic set of equations, either steady state or transient, given a fixed set of inputs. In other words, these programs provide predictions of how a single point design performs under specific environments. Automated model calibration, on the other hand, requires either using or creating a software tool that can perform multiple iterative point design evaluations. This section describes three approaches toward achieving such a capability.

The first option uses an in-house development approach. First, engineers can write their own optimization engine or purchase one commercially. Then, a means of executing the thermal/fluid analyzer iteratively must be achieved, perhaps via an API (application programmer interface) if available, or perhaps simply by modifying and rewriting text input files and reading text output files. A script can be generated to iteratively run the thermal/fluid analyzer, driving the uncertain inputs with the optimization engine such that a best match is achieved between simulation predictions and test data. This option is cost effective only if software development labor is inexpensive or if an organization is large enough to recoup the investment of the development of a general-purpose utility. Otherwise, considerable effort will be spent rewriting the software every time a new calibration task arises.

As the second option, engineers can acquire a general purpose MDO (multidisciplinary optimization) environment. Examples of such software include Engineous' iSIGHT®, Phoenix Integration's ModelCenter®, MSC Software's RDCS, Synapse' Pointer®, VR&D's VisualDOC®, LMS' Optimus®, and Samtech's BossQuattro. To varying degrees, these programs enable the engineer to set up their favorite thermal/fluid simulation code as part of the evaluation of any one set of unknown or random inputs. The advantages are that these thermal/fluid simulation codes need not "know" that they are being used in such an iterative fashion: little to no modifications of the simulation codes and models are required. This approach also has the advantage of providing an infrastructure that reduces the time to create a new calibration or reliability estimation task. However, disadvantages of the MDO approach include the cost of acquiring and learning such codes, and the relatively slow speeds resulting from inefficiencies in running the simulation code in such a disconnected fashion. Nonetheless, such an approach is clearly better than the current "manual" and "serial" method of calibrating models.

A third choice is to use a thermal/fluid analyzer that already has these advanced features built-in. This avoids the overhead associated with the first choice, and the additional costs associated with the second choice, and is much faster to execute than either of those choices for various reasons.[1] However, choices are limited for two reasons. First and most important, few thermal/fluid analysts are aware of such capabilities, and hence they more typically demand additional detailed phenomenological modeling power rather than more help with design and calibration tasks. Forgivably, commercial vendors listen to them, and the demand for high-level decision support tools is therefore slack. Second, even after analysts discover these gains in productivity and begin to demand them, software suppliers will find it difficult to accommodate these requests without significant changes in their software. To accommodate high-level analyses such as model calibration and reliability estimation, the software must first become fully parametric instead of expecting single-valued ("hard-wired") design and environment specifications. There is hope, however: structural analysis and CAD software have increasingly emphasized such capabilities in their new releases over the last five years. It is hoped that thermal/fluid analysis tools can follow these examples and catch up once the user community has been educated and the demand for new capabilities is established.

Conclusions

Removal or reduction of uncertainty is an important if not required step in most thermal/fluid analyses. However, existing techniques are labor-intensive and faulty since they are rarely rigorous. This paper has shown how existing models built using existing software can be automatically rerun using NLP technology tasked with seeking a best fit. In software designed to include these capabilities as "native," application of automated calibration techniques is becoming commonplace.

The resulting techniques are not magic and still require a good model and an experienced engineer making sound decisions. However, a significant improvement in both productivity and predictability has been demonstrated and is in current active use.

References

1. B. Cullimore, "Dealing with Uncertainties and Variations in Thermal Design", IPACK2001-15516, InterPack '01 Proceedings, July 2001.
2. SINDA/FLUINT User's Manual, PDF available at www.crtech.com.

[1] In addition to avoiding interprocess communication and overhead associated with starting and restarting programs, a built-in capability can exploit the advantage that previous steady state solutions (which usually comprise the majority of calibration and reliability assessment tasks) in the search were close to the current solution, and can jump quickly to incremental answers.