JOURNAL OF THERMOPHYSICS AND HEAT TRANSFER
Vol. 24, No. 4, October–December 2010

Application of a Versatile "Real-Space" Validation Methodology to a Fire Model

Vicente Romero,* Anay Luketa, and Martin Sherman
Sandia National Laboratories, Albuquerque, New Mexico 87185-0828

DOI: 10.2514/1.46358

This paper applies a pragmatic approach to validation of a fire-dynamics model involving computational fluid dynamics, combustion, participating-media radiation, and heat transfer. The validation problem involves experimental and predicted steady-state temperatures of a calorimeter in a wind-driven hydrocarbon pool fire. Significant aleatory and epistemic sources of uncertainty in the experiments and simulations exist and are transformed to a common basis of interval uncertainty for aggregation and comparison purposes. The validation comparison of experimental and simulation results, and corresponding criteria and procedures for model substantiation or refutation, take place in "real space" as opposed to "transform space" where various transform measures of discrepancy between experiment and simulation results are calculated and assessed. The versatile model validation approach handles difficulties associated with representing and aggregating aleatory and epistemic uncertainties (discrete and continuous) from multiple correlated and uncorrelated source types, including 1) experimental variability from multiple repeat experiments, 2) uncertainty of experimental inputs, 3) experimental output measurement uncertainties, 4) uncertainties that arise in data processing and inference from raw simulation and experiment outputs, 5) parameter and model-form uncertainties intrinsic to the model, and 6) numerical solution uncertainty from model discretization effects.

I. Introduction

ONE aspect of the work that Sandia National Laboratories performs for the U.S. Department of Energy is the design and assessment of safing systems that keep nuclear weapon firing systems inert in accidents and abnormal environments such as fires. (See, e.g., [1] for a synopsis of weapon risk considerations and assessment methodology for fire accidents.) In a supporting activity, the model validation effort described in this paper centers around the steady-state temperature response of a fire-heated cone-shaped calorimeter. The experimental setup, fire tests, and raw experimental data and results have been extensively documented [2]. Reference [3] discusses the geometry and experimental conditions important to modeling the fire tests. Also documented are the physics and discretization models in the Fuego [4] computational fluid dynamics, combustion, and heat transfer simulations and results. The massively parallel 3-D code involves 33 degrees of freedom at each node, representing x-y-z momentum, participating-media radiative transport directions or ordinates, chemical reactions/combustion, and turbulence, energy, pressure, and soot and chemical species conservation and transport.

The subject of this paper is validation of the Fuego fire-dynamics model. The following one-sentence contemporary definition of model validation, and close variants of it, is the accepted standard definition in [5–10]: *Model validation is the process of determining the degree to which a computer model is an accurate representation of the real world from the perspective of an intended use of the model*. Despite broad agreement on this one-sentence notional definition, at a detailed interpretational and implementational level there is room within the definition for considerable debate among validation methodology developers and practitioners concerning the specific procedures, steps, and products of model validation.

Hence, various modeling communities are still working out the technical procedures and criteria for assessing and deciding whether a model is considered adequate with respect to a particular intended use. To be sure, several model validation paradigms and methodologies exist in the literature. Many of these are considered in [11–15], in which it is determined that none of those considered appear to have the full set of enabling features, yet are pragmatically simple enough, to reasonably handle the difficult model validation problem pursued in this paper. Indeed, a novel model validation framework and methodology, developed in the course of model validation practice on a variety of very challenging application problems (see [14] for a description), is applied in this paper to the difficult fire-dynamics model validation problem.

A large variety of mathematical transforms exist in the literature to characterize discrepancy between experiment and simulation results (e.g., the simple subtraction transform in [9]). The transform measures in the literature can get relatively sophisticated and involved, with varying transparency or interpretability of the physical and decision-making significance of the numerical values yielded by the discrepancy measures. Indeed, workable criteria to demarcate adequacy of model/experiment agreement in the "transform space" of the various mathematical measures are found to be relatively elusive.

An important aspect of the present methodology is that model accuracy and adequacy relative to experimental data (in both calibration and validation settings) are *not* posed in terms of transform measures and acceptance criteria in a transform space. Instead, a simple and fundamentally different "real-space" approach has been developed and articulated in [11–15] and is applied in this paper. Indeed, project requirements of pragmatism and demonstrated functional usability have led to the present paradigm.

Another key feature of the model validation methodology in this paper is that it appropriately treats both types of uncertainty (epistemic and aleatory) often present in model validation activities. A fully general and viable validation framework must adequately handle not only probabilistic uncertainty, but also the types of epistemic uncertainty often found in experiments and models, such as interval uncertainties and model-form uncertainties among discrete (not parametrically connected) alternate plausible models. All these types of uncertainty are confronted and treated in the model validation problem in this paper.

The general algorithm/recipe of the model validation methodology is presented in [15]. The general algorithm cannot be fully

Presented as Paper 2009-2279 at the 11th AIAA Non-Deterministic Approaches Conference, Palm Springs, CA, 4–7 May 2009; received 15 July 2009; revision received 25 November 2009; accepted for publication 23 January 2010.

*vjromer@sandia.gov. Senior Member AIAA (Corresponding Author).

inferred from the specific application problem here, but a step-by-step presentation of the general algorithm would lengthen this paper too much. Nevertheless, the example here confronts and treats a large subset of circumstances found in many validation problems, so by analogy with the presented example, the reader should be able to treat a variety of model validation problems that they might confront.

In the following sections the fire experiments, simulations, and results are briefly summarized. The majority of the paper is devoted to presentation of a validation-relevant subset of the experiment and simulation results and uncertainties, processing of these into a form suitable for the particular type of model validation comparisons employed, and description of the comparisons along with interpretation within the model validation framework employed. For brevity here, only a representative subset of experimental and simulation results and model validation processing is presented. The full set of data, processing, and results from the validation activity can be found in [3]. It should be noted that this paper is a substantial revision of its predecessor [16]. In particular, the presentation here on the model validation paradigm and methodology employed is substantially reorganized and extended compared to [3,16], and various uncertainty quantification aspects are treated with greater rigor.

II. Experiments and Experimental Conditions

As Fig. 1 shows, the experiments involved a flow-through partial enclosure that serves the purpose of imitating foreseeable conditions in which a fire engulfs a weapon and heats the walls and roof of a weapon storage or transportation room or container. The walls and roof then radiate heat to the weapon, imparting more heat and thus embodying a more severe heating condition than if no enclosure were present.

The walls and roof of the enclosure were composed of 3/16-in.-thick steel sheet stock on the inside. These were insulated on the outside with blanket insulation to provide an easily modeled boundary condition for the validation activity. Therefore, an adiabatic boundary condition was modeled on the exterior of the enclosure walls and roof in the Fuego simulation. The floor is thermally massive and is assumed not to heat up appreciably. The cone-shaped calorimeter outside shell is made of 3/16-in.-thick steel sheet. For ease of modeling, the inside volume is filled with blanket insulation to prevent convective and radiative heat exchange among the inside walls of the cone. Hence, an adiabatic boundary condition is applied on the interior of the steel cone shell modeled in the Fuego simulations. The nearly adiabatic walls (with insulated backsides) of the calorimeter and enclosure also strongly promote a quicker arrival to steady-state wall temperatures and reradiation conditions in the experiments.

Thermocouples (TCs) were attached to the inside of the cone shell at 12 height levels that include the two levels shown in Fig. 2. TCs

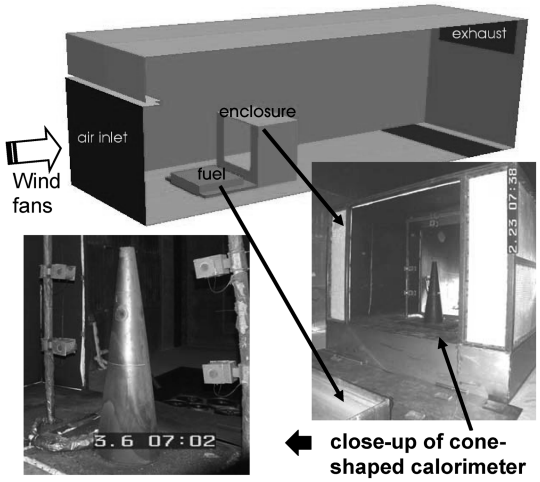

Fig. 1 Cone-shaped calorimeter inside flow-through enclosure with fuel pan in front. Stands at sides of calorimeter hold thermocouples and (visible) flux gauges.

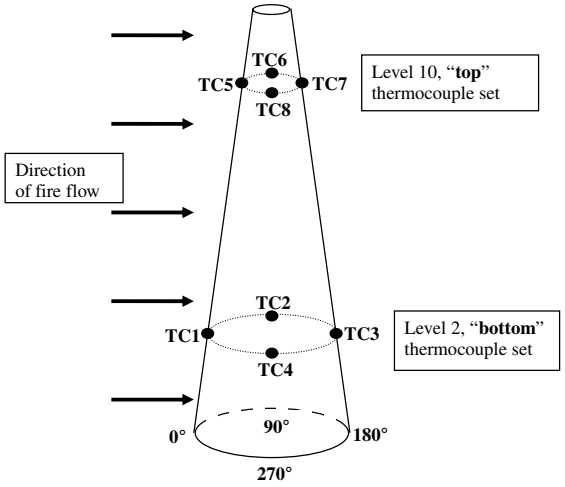

Fig. 2 Thermocouple locations on cone-shaped calorimeter for comparison of experimental results versus Fuego predicted results.

and flux gauges were also placed in the flowfield of the fire (on the stands shown at the sides of the calorimeter in Fig. 1) and on the walls/roof of the flow-through enclosure in which the calorimeter is at the center. The steel interior walls, roof, and floor of the enclosure and the exterior of the calorimeter were painted with black Pyromark high-emissivity paint so that they would readily exchange radiative heat with the fire and with each other, therefore more quickly reaching steady-state temperature and reradiation conditions. The Pyromark-painted surfaces have an initial emissivity ϵ of nominally 0.86 upon application to the room-temperature surfaces, but the emissivity after experimental burn-in of the paint at high temperature is a significant uncertainty in the experiments. Later it will be described how this important uncertainty is accounted for in the model validation analysis. Other uncertain inputs to the experiments will also be discussed.

Details of the fuel and fuel-pan, facility and structure geometries, experiment flow conditions such as the air inflow and exhaust outflow conditions, and all thermocouple and flux-gauge locations can be found in [3].

III. More Model Validation Considerations

Although a wealth of experimental data from the tests is available to compare against Fuego predictions for model validation purposes, time and resource considerations dictated that comparisons in the project be confined to only calorimeter temperature response. Such response is most directly associated with the validation question driving the experimental and simulation work here: how well does Fuego calculate surface heating conditions on objects (weapons) in wind-driven hydrocarbon fires?

To bring the scope of the validation activity into line with project resources, the validation comparisons are further limited to eight diverse and representative TC locations on the calorimeter, although many more locations were monitored in the experiments. Figure 2 shows four TC locations (at 0, 90, 180, and 270°) on the upper section of the cone above a field joint (internal bolted attachment) at midheight, and another four at the same angular positions on the lower section of the cone below the attachment joint.

On the upper section of the cone, the level 10 set of TCs in Fig. 2 was chosen in order to minimize local edge or anomaly effects on conduction behavior, and therefore temperature, of the calorimeter shell. level 10 is about halfway between the top edge of the cone and the height at which flux gauges were inserted into holes (discernible in Fig. 1) of 1.5–2 in. diameter in the shell. These holes create local anomalies in the shell's conduction and temperature fields. The level 10 TCs are sufficiently far from these anomalies.

On the lower section of the cone, unknown and possibly highly varying contact resistance around the bolted attachments at the midheight field joint and the base of the cone could produce local

effects on calorimeter temperature. The level 2 height in Fig. 2 was chosen in order to minimize local effects due to these bolted joints.

Calorimeter temperature response is not the only quantity relevant to Fuego validation. For example, enclosure temperature response is just as indicative of how well Fuego calculates object surface heating. Certainly, it would be significantly more revealing to compare Fuego predictions against experimental response for both the calorimeter and enclosure, instead of just one or the other. Both are time-space integrating sensors of the experimental and simulated fires, and corroboration (validation) based on two diverse such sensors is of course stronger than corroboration from one alone. However, corroboration based on two diverse integrating sensors also cannot be said to be *absolutely sufficient* to settle the model validation question with regard to the present experiments. For instance, a further check against the TCs and flux gauges on the stand in the flowfield of the fire (see Fig. 1) could indicate differently, and refute the fire model's predictive capability locally there. (Such further checks are essentially infinitely prescribable; see the discussion in [13] on field validation versus the validation of only certain *resultant effects* of the field important to an engineering purpose, effect validation.)

In fact, absolute sufficiency is probably only reachable in theoretical limits and it would appear that statistical procedures and arguments would have to be invoked to establish validation sufficiency in practical terms (e.g., odds or significance level associated with a validation conclusion). This is beyond the scope of the present treatment. In the present circumstances we can only address whether a *necessary* test of model adequacy is met by Fuego with respect to the narrow program-relevant quantity of calorimeter steady-state temperature. Indeed, it is shown in the following that a significant validation test is passed in this regard, lending credibility to Fuego's predictive abilities in the present experimental circumstances, although not sufficient to absolutely confirm its predictive capability, even for the present experimental conditions.

A relaxation of the strength of the validation case also comes from the fact that we compare only steady-state temperature of the calorimeter, as opposed to transient temperature response which would more stringently test Fuego's predictive capability. This is forced by several reasons. First, capabilities for realistic modeling of transient startup and growth of the pool fire were not yet available in Fuego. Second, even if available, time-accurate simulation from startup to stabilization to steady-state behavior of the fire would be prohibitively expensive and long to compute.

Fortunately, the predominant project driver (assessment of weapon safety risk in fires) is concerned with weapon internal responses that take tens of minutes to develop and are the result of sustained fire heating of the weapon and are fairly insensitive to the early-time initial flare-up and stabilization transients of the fire. Therefore, a validation test regarding steady-state heating of the calorimeter adequately serves the project objectives concerned with predicting sustained steady-state heating loads for weapon risk analysis. Indeed, there are many instances in engineering and science in which validation of steady-state behavior suits the validation project needs very well.

IV. Plan of Fuego Simulations for Uncertainty Quantification in the Model Validation Procedure

Fuego simulations were used in the validation procedure to quantify modeling uncertainty. Resource constraints dictated that only five or six steady-state Fuego simulations could be run at the medium-resolution model described in [3]. (The simulations took several weeks on 256 processors.)

Here, model resolution involves spatial-discretization cell size, number of discrete-ordinate directions for solution of participating-media radiation transport, and solver numerical-error tolerance parameters employed in the steady-state Fuego computations. The resolution level was surmised from past Fuego modeling projects to be sufficient for the validation purpose here, i.e., to yield simulation results that would change with further resolution refinement only relatively little compared to the effects of other physical modeling uncertainties and bias errors in the validation activity.

Table 1 Intrinsic sources of uncertainty in the fire-dynamics model

Source	Values
Heat of combustion (HOC)	44.66 kJ/mol ± 10%
Soot extinction coefficient (SEC)	7 ± 10%
Convection coefficient (CC) calculated at object surface	−50% to +100%
Flame value coefficient (FVC)	2.13 ± 30%
Flame loading coefficient (FLC)	0.41 ± 30%
Turbulence model form	TFNS (nominal) versus BVG

To check that the medium-mesh discretization was indeed sufficient, a simulation was planned at a significantly higher resolution (see [3]). However, the high-resolution simulation never completed, because of difficulties with the required multithousand processor jobs. Thus, in assuming that the calculation results here are sufficiently numerically resolved, there is only previous experience to cite from [17] which compared results from a similar medium resolution to results from a substantially higher resolution. Nonetheless, it is explained later how the validation framework used in this paper would handle any characterized uncertainty (e.g., by methods presented in [18]) due to numerical resolution in the model and simulations.

Regarding fire-physics modeling uncertainty, a consensus of Sandia fire-modeling expert judgment concluded that the physics submodel forms and coefficients listed in Table 1 were the largest sources of *intrinsic* modeling uncertainty in the validation activity. Nominal values and associated uncertainties are listed in Table 1. Here, *intrinsic* signifies a category of modeling uncertainty that is innately associated with the *traveling* portion of the model of the validation experiment. The latter model is termed the *validation experiment model* or *VE model*. Only the traveling portion of the VE model goes on to new predictions beyond the validation exercise. That is, certain elements or aspects of the VE model will be carried forward (will travel) to new predictions, and certain aspects are specific only to the validation setting. The intrinsic uncertainties in Table 1 all exist in the VE model, but also travel to new predictions as part of the Fuego fire-dynamics model, as uncertainties that are propagated to simulation results in the new prediction settings as well. Other uncertainties in the experiment and in the VE model, e.g., emissive uncertainty of the painted calorimeter and enclosure walls, are confined to the validation experiments alone. In general, predictions after the validation activity will involve heated objects and enclosures with different emissive properties/uncertainties than the specially prepared Pyromark-painted surfaces in the validation experiments. Thus, these emissivity uncertainties will not travel to new predictions: they are not traveling uncertainties. As explained in [15] and demonstrated in this paper, the validation framework handles traveling uncertainties differently from nontraveling uncertainties.

Some of the modeling entities in Table 1 require further description. Brief descriptions are given here. Fuller descriptions are given in [3]. The heat of combustion is the lower heating value equal to the total heat released when one unit of fuel is completely oxidized minus the heat of vaporization for water vapor in the products.

The soot extinction coefficient is a parameter required for the calculation of the emittance of a spherical carbon soot particle layer at a given temperature and volume fraction [19]. The dependence is such that an increase (decrease) in the soot extinction coefficient results in a lower (higher) emittance.

The flame volume coefficient is a proportionality constant for the ratio of the maximum reaction-zone volume per eddy volume. This ratio was defined in [20] to determine the volume fractions of the reaction zone and the surroundings zones. The basic premise is that the overall heat release rate is controlled by the mass transport into the reaction zone. Thus, the flame volume coefficient will have an effect on the overall heat release rate.

The flame loading coefficient is a proportionality constant in the relation to determine the residence time in the reaction-zone volume.

Table 2 Fire-model input variations for the six Fuego simulations run

Fuego run	Simulation	Description	Turbulence model	HOC	SEC	CC	FVC	FLC	$\epsilon_{\text{Pyromark}}$
Set 1	1	Baseline case, TFNS with nominal parameter values in Table 1	TFNS	44.66	7	100%	2.13	0.41	0.86
Set 2	2	High-heating-parameter combination with TFNS	TFNS	44.66 + 10%	7	100%	2.13	0.41 − 30%	0.86
Set 3	3	Low-heating-parameter combination with TFNS	TFNS	44.66 − 10%	5	200%	2.13 − 30%	0.41 + 30%	0.86
Set 4	4	Same as high set 2, except with BVG turbulence model	BVG	44.66 + 10%	7	100%	2.13	0.41 − 30%	0.86
Set 5	5	Same as low set 3, except with BVG turbulence model	BVG	44.66 − 10%	5	200%	2.13 − 30%	0.41 + 30%	0.86
Set 6	6	Same as set 2, except for emissivity value	TFNS	44.66 + 10%	7	100%	2.13	0.41 − 30%	0.96

This residence time is compared to a minimum residence time based on precalculations using a perfectly stirred reactor model and chemistry. If the residence time within a computational cell is less than this minimum value, then flame extinction occurs. Thus, this parameter affects the heat release by allowing for combustion to either cease or continue within a cell.

Fuego currently has two different turbulence models considered to be plausible choices for modeling wind-driven fires of the nature here. These are the buoyant vorticity generation (BVG) turbulence model [21] and the time-filtered Navier–Stokes (TFNS) turbulence model [22]. These are implemented within a Reynolds-averaged Navier–Stokes (RANS) computational fluid dynamics formulation.

The following strategy was employed to estimate the uncertainty contributed by the six modeling factors in Table 1. These uncertainties are all epistemic in nature (not random or aleatory), and are of either interval or discrete form. The nature of the uncertainty and the small available budget of only five or six Fuego simulations dictated that the goal should be simply to define reasonable upper and lower bounds on the joint uncertainty that these factors contribute to prediction of calorimeter heating. This requires running a simulation at the combination of factor values within their uncertainty ranges that gives the highest heating of the calorimeter, and then at a combination of factor values that gives the lowest heating of the calorimeter.

In pursuing this goal, the severe limit on the allowable number of simulations dictated that previous information about model trends with respect to the uncertain factors would have to be extrapolated to the current modeling setting. In particular, it is known that the inclusion of a baroclinic torque source-term in the k equation enhances mixing (over the standard k-ε turbulence model) and therefore promotes enhanced combustion in the BVG formulation. Indeed, in Fuego simulations the BVG-RANS formulation generally results in larger and hotter predicted fires (and greater heating of objects) than the TFNS-RANS formulation does.

Regarding the other five factors in Table 1, the results of a previous uncertainty/sensitivity study [23] were leveraged. The study involved a blocked experimental design of 16 simulations to explore the above five-factor uncertainty space plus other factors such as spatial discretization and computational solver resolution. Sixteen simulations were run for airplane fire and 16 more for truck fire wind-driven fire scenarios. The sensitivity of weapon heating to the uncertainties was analyzed. From the sensitivity results, the parameter combinations for high and low weapon heating (as a global spatial average over the heated object, but not necessarily locally at all points on its surface) in each accident scenario were inferred. These are referred to, respectively, as the hot-fire and cool-fire parameter combinations.

With these hot-fire and cool-fire parameter sets and the knowledge that the BVG-RANS turbulence model usually results in greater heating than TFNS-RANS, combinations of all six factors can be formed to yield high and low heating bounds for the fire-dynamics model, as could reasonably be determined under the resource constraints. Just one Fuego simulation with the TFNS model and the cool-fire parameter set (case 3 below) and a run with the BVG model and the hot-fire parameter set (case 4 below) were necessary to ostensibly bound the uncertainty associated with the fire-dynamics model. Nevertheless, to more thoroughly examine the effect of turbulence model form on predicted heating in the validation experiments, both model forms were run with the hot-fire and cool-fire parameter sets. Thus, the following simulations were run. Table 2 lists the specific factor settings in the run matrix.

It should be mentioned here that the soot extinction coefficient, the convection coefficient, and the flame volume coefficient are all concave-down in their main-effects sensitivities (effect on object-heating over the uncertainty ranges in Table 1). Thus, the nominal or central values of these parameters also correspond to the maxima at which object heating is greatest over their uncertainty ranges. Hence, the parameter values of SEC, CC, and FVC do not change from set 1 (baseline case with all parameters at nominal or central values) to sets 2 and 4, in which the high-heating hot-fire parameter values are invoked.

In addition to the intrinsic uncertainties of the fire-dynamics (traveling) portion of the VE model, there are uncertainties in the nontraveling portion of the VE model that must be considered in the validation activity. These consist of uncertain inputs to the validation experiment(s) that are nontraveling, such as uncertainties in calorimeter/enclosure/facility geometry; in emissivity of the Pyromark-painted surfaces; in fuel regression rates and airflow rates (wind velocity) in the experiments, etc. With the following rationale, our plan of Fuego UQ simulations only treats systematic uncertainties in the experimental conditions.

In principle, we can also account for random (aleatory, non-systematic) uncertainty in the experimental inputs if they are known. For example, with very little measurement uncertainty the fuel regression rates in the validation-relevant experiments 6 and 7 (described in [3]) were recorded as 4.0 and 4.16 mm/ min. This

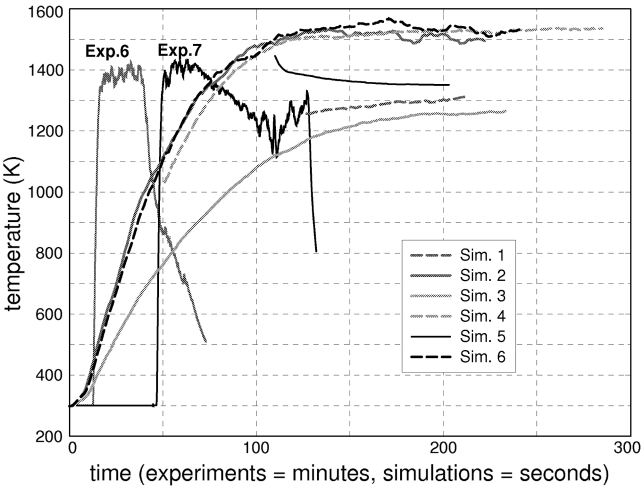

Fig. 3 Calorimeter temperature at TC5, facing the fire and near the top of the calorimeter.

random (uncontrolled) change or variation in the fuel regression rates in the two nominal repeat experiments, along with any other experimental inputs that varied from experiment 6 to experiment 7, ostensibly show up as a difference in the experiment output results. Indeed, Fig. 3 shows a small but noticeable difference of about 20 K between the ultimate steady-state calorimeter temperatures in experiments 6 and 7. Analogously on the simulation side of things, we could run two separate simulations with the known differences in experimental inputs applied in the two simulations. Simulations of experiments 6 and 7 would assign input fuel regression rates of, respectively, 4.0 and 4.16 mm/ min. Other inputs known to differ between the two experiments such as facility airflow rates and wall temperatures would also be faithfully assigned in the simulations. Thus, these known experimental variations would be propagated to simulation outputs, just as the real experimental system (the fire dynamical system) propagated them to experimental outputs. Then a cleaner basis of comparison would exist for compared simulation and experimental outputs.

Even so, the comparison basis would still not be completely fair, because other variations between experiments (such as calorimeter and enclosure emissivities) were not characterized and available for propagation through the simulations. It is generally the case in model validation activities that some experimental variations are either not known/measured or are not easily or affordably propagated through the simulation model. Hence, frequently in practice, a somewhat skewed basis of comparison between experiment and simulation results cannot be avoided.

If experimental input variabilities are neglected in the validation simulations it is normally harder to meet the threshold for claiming a validation-substantiated model [15]. Therefore, such neglect likely errs to the conservative side in model validation assessments. It is also easier and less expensive computationally to ignore (in the modeling) the variance contributed by known input variations in the experiments. Therefore, the decision to ignore is sometimes made as a simplifying and likely conservative choice in the model validation assessment.

In the present validation problem the experimental variability in thermocouple temperature due to all random variations (known/measured and unknown) in the experimental conditions of experiments 6 and 7 is derived in the next section to be nominally ± 39 K (two-sigma range). A significant amount of this uncertainty could have been modeled by propagating known experimental input variations, thereby neutralizing some of the disparity in random uncertainty in the experimental and simulation outputs. However, a wealth of previous experiments and simulations in other settings indicated that a particular source of systematic uncertainty in the experimental data due to epistemic uncertainty in the emissivity of the calorimeter and enclosure surfaces would be more substantial and therefore more important to estimate with the remaining budget of Fuego simulations.

As established below, emissivity ϵ of these fire-tarnished surfaces can reasonably be said to lie within a range 0.76 to 0.96. Although emissivity of these surfaces would be expected to vary to a small degree from experiment 6 to experiment 7, such random variation is thought to be rather small, given that these experiments were nominally identical in nature and the surfaces had already been well burned-in via the five previous preparatory experiments. Hence, the uncertainty is predominantly systematic in nature: ϵ is nominally the same within a small perturbation in the two experiments, but this common value is unknown to within a relatively large systematic uncertainty of 0.76 to 0.96. Again, treating the slightly varying emissivity in the experiments to be nonvarying (nonrandom) in the simulations is a likely conservative approach in the validation analysis, so is an acceptable tactic if needed to facilitate the assessment.

Most of the epistemic uncertainty associated with calorimeter steady-state temperature is deemed to come from just this one source of systematic uncertainty in the experiments (emissivity). All other systematic uncertainties are judged to impact steady-state temperature to a vastly lesser degree, so were ignored in view of the limited number of Fuego simulations that could be run.

The emissivity uncertainty, a nontraveling systematic uncertainty, is accounted for in the following way. First, a simulation (no. 6, below) is run to quantify the effect of a substantial perturbation from a nominal emissivity value $\epsilon_{\text{nominal}} = 0.86$, with all other model inputs held constant. From this, a first-order finite difference approximation to the derivative $\mathrm{d}T/\mathrm{d}\epsilon$ for change in *local* (at a given location) steady-state temperature at the cone surface with respect to change in surface emissivity is formed. Then the linearized relationship $\Delta T \cong [\mathrm{d}T/\mathrm{d}\epsilon] \bullet (\Delta\epsilon)$ can be used to estimate the steady-state temperature change at a given TC location on the calorimeter surface for any value of emissivity different from $\epsilon_{\text{nominal}}$. For certain reasons not evident in what is presented in this paper, simulation 6 is a perturbation from parameter set 2. In what follows, the assumption is made that the derivative $\mathrm{d}T/\mathrm{d}\epsilon$ at a given thermocouple location would be approximately the same for a similar emissivity perturbation from *any* of the simulations 1–5. Although the derivative might actually change nonnegligibly if computed by perturbations from the other simulations, this sensitivity could not be investigated because of the limit on the number of Fuego simulations that could be run. Hence, the local ΔT effect from varying ϵ over its applicable uncertainty range can be investigated analytically with the above relation, where $\Delta\epsilon = \epsilon - \epsilon_{\text{nominal}}$. This will be applied later in the paper. In fact, it will impact the calculation of aggregate uncertainties for both simulated and experimental steady-state temperatures.

It is straightforward that the emissive uncertainty in the experiments is propagated through the model to express the effect that this uncertainty has on predicted steady-state temperatures. It may be less straightforward to the reader that the uncertainty in ϵ should also be propagated to an uncertainty on the experimental data itself (of measured steady-state temperatures). This is because it is a systematic uncertainty in the experiments, so its effects will not be present in the observed variability of the experimental output data like will the effects of any randomly varying experimental factors. At a very simple and familiar level, this is analogous to propagating any systematic measurement uncertainty to the experimental data readings taken in an experiment. The actual readings are deterministic values coming from the measurement instruments. To the deterministic reading a measurement uncertainty is applied, often in the form of a $\pm\delta$ uncertainty range about the reading value. This is a simple form of *data conditioning* of the experimental data in order to account for measurement uncertainty on experimental outputs. Another form of data conditioning is applied in Sec. V.C to account for systematic uncertainty of the experimental input ϵ in the experiments.

To condition the experimental data as mentioned earlier, $\mathrm{d}T/\mathrm{d}\epsilon$ must be explicitly calculated. This required the extra Fuego simulation 6. If the emissivity uncertainty only had to be propagated to an effect on prediction uncertainty (but not also to an uncertainty on the experimental steady-state results), then $\mathrm{d}T/\mathrm{d}\epsilon$ would not need to be calculated. Instead, the high and low emissivity bounds of 0.76 and 0.96 could simply be folded into the joint bounding parameter combinations in simulations 1–5 for high heating and low heating of the calorimeter. However, the emissivity uncertainty is a nontraveling uncertainty, so as mentioned earlier, there are differences in how it is treated versus the traveling uncertainties (such as the model-intrinsic uncertainties) in the validation assessment.

The uncertainty range for emissivity of the calorimeter exterior surface and the enclosure interior surfaces (walls, roof, and floor) is taken to be 0.76 to 0.96 based on measured values [24–27] for both heavily oxidized steel surfaces and burned-in Pyromark-painted steel surfaces such as in the present experiments. Set 6 employs an emissivity perturbation that goes to the high extreme of the emissivity uncertainty range, 0.96. Set 6 (simulation 6) is the same as set 2, except $\epsilon = 0.96$.

Another nontraveling systematic uncertainty worth noting in the experiments is the heat transfer contact resistance between the bottom of the steel cone and the thermally massive steel floor to which it was bolted. This was judged to be substantially less significant to steady-state temperatures than was the uncertainty in ϵ, and it was also not practical to model contact conductance at this

interface. Instead, a simple adiabatic boundary condition was applied at the bottom of the cone. The consequent error (and uncertainty thereof) remains unquantified in the present effort. Nevertheless, strong arguments are made in Sec. VI that neglecting this factor probably does not materially change the final validation conclusions.

Geometry uncertainties were also systematic to experiments 6 and 7 but were judged likely to be of second-tier importance for steady-state temperatures. Finally, any systematic measurement uncertainties associated with fuel regression rate, airflow rate, and facility wall temperatures were also judged to cause only very small bias errors in calculated steady-state temperatures. Uncertainties of steel and insulation thermal properties of the calorimeter and enclosure were deemed to have relatively little effect based on applicable sensitivity studies from [17].

V. Experiment and Simulation Results and Uncertainty Processing for Model Validation Comparisons

Here, we consider only a representative subset of validation-relevant experiment and simulation results and uncertainties. We focus on TC5, about one-fifth of the way down from the top of the calorimeter and facing the oncoming wind-driven fire. The experiment and simulation results and validation processing for the other seven TCs in Fig. 2 can be found in [3].

Figure 3 shows experimental and predicted temperature response curves for TC5. The data for two nominal repeat experiments, 6 and 7, are shown. The previous five experiments were preparatory experiments needed to refine the geometry configuration and experimental conditions, and to burn in the Pyromark-painted radiating surfaces. It should be mentioned that in experiment 7 at about 70 min, a slow die-off in the TC5 temperature trace begins. According to the experimentalist [2], this is not because the fire begins to die out, but because shunting occurred in the thermocouple. Shunting is a partially shorted junction caused by degradation of the insulation covering the TC wires.

Also plotted on the Fig. 3 are the results of all six simulations. The simulation results will be discussed in detail later, but it is mentioned here that these were false-transient simulations to reach steady-state temperatures as quickly as possible. Therefore, the transient response for the simulations is plotted on a scale of seconds to reach steady state (whereas the experimental responses are plotted on a scale of minutes to reach steady state). It is also informative to mention here that results for simulation 5 are falling rather than rising to a steady state, because a mistake in one of the parameter values was noticed and corrected midway through the weeks-long calculation. Only the late-time portion of the simulation 1 results were saved for plotting.

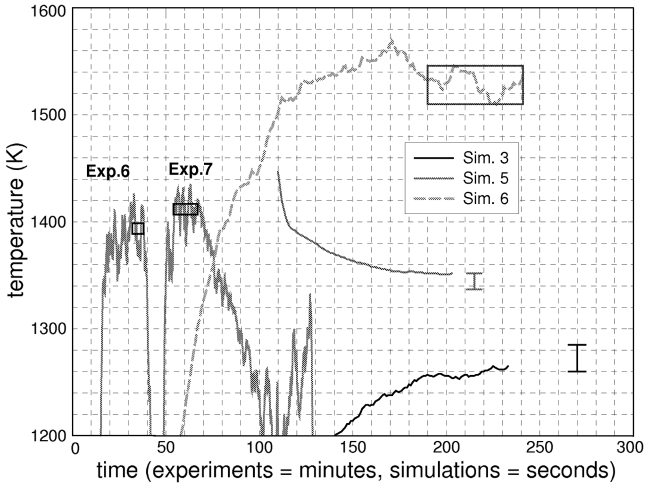

Fig. 4 Uncertainty on steady-state experimental and simulation results at TC5.

Table 3 Mean, maximum, and minimum temperatures over time windows in Fig. 4

Experiment no.	Time-window mean, μ	Max over time window	Min over time window
6	1394 K or 1121°C	$\mu 6 + 33$ K or °C	$\mu 6 - 31$ K or °C
7	1412 K or 1139°C	$\mu 7 + 24$ K or °C	$\mu 7 - 34$ K or °C

A. Characterization of Raw Experimental Results

For the experimental results in Fig. 3 it is obvious that it is important to use appropriate time windows over which to average TC5 temperature responses in order to get representative steady-state values for model validation comparisons. The time windows used are depicted graphically in Fig. 4. These correspond to a 6 min window from 32–38 min in experiment 6 and a 13 min window from 54–67 min in experiment 7. Various other reasonable time windows for steady-state temperature averaging are considered in Appendix A. It is determined that model validation results arrived at later in this paper are not narrowly dependent on the particular time windows shown in Fig. 4 and used in the model validation analysis.

The mean temperature over each time window is listed in Table 3. The upper and lower horizontal lines that define the top/bottom of the box are given, respectively, by the mean ± 5 K. The rationale is that other time windows than those used here, with different width and/or shifted later or sooner in time, could be just as appropriate or representative, but will yield different estimates for steady-state temperature. Several alternative reasonable time windows are considered in Appendix A. These result in time-window means that generally differ from those in Table 3 by a magnitude of less than 5 K. Accordingly, a reasonable uncertainty band of ± 5 K about the steady-state point estimates in Table 3 is employed in the ensuing analyses. The table also lists maximum and minimum *instantaneous* temperatures over the time windows. These will be used in later analysis and should not be mistaken for uncertainty bounds on the steady-state means.

B. Characterization of Raw Simulation Results

Representative steady-state values are actually *less* definite in many instances for the simulation results than for the experimental results. In particular, Figs. 4 and 5 show that the temperature predictions have not yet stabilized completely. The simulations had to be terminated due to resource limitations in the project. (Each simulation took several weeks on 256 processors.) Consequently, the trends were (initially) visually extrapolated to estimated steady-state values. For example, for simulation 3 it is visually judged that the steady-state asymptote of the temperature curve is within the uncertainty bar plotted in Fig. 4 that runs from 5 K below the last

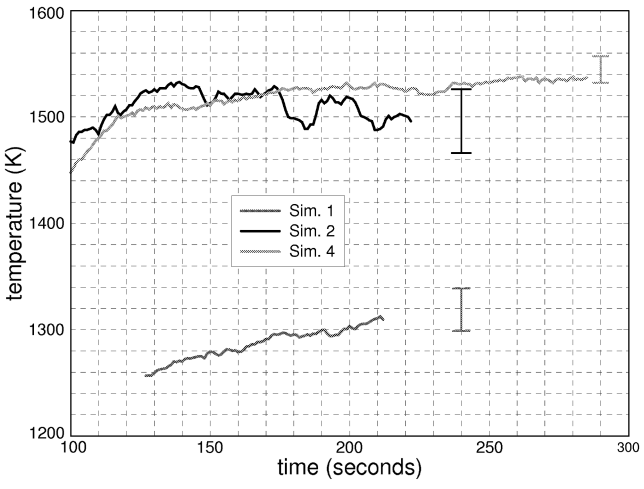

Fig. 5 Uncertainty on steady-state simulation results at TC5.

Table 4 Visually projected limits (in brackets) to anticipated change from end-of-simulation temperatures to asymptotic steady-state values

Simulation no.	Value
1	[−10, +30] K
2	[−30, +30] K
3	[−5, +20] K
4	[−5, +20] K
5	[−15, 0] K
6[a]	190–241 s

[a]For simulation 6 the averaging time window is listed from which a mean estimate for steady-state temperature is determined.

Table 5 End-of-simulation temperatures at TC5 location, with mean or extrapolated-asymptotic steady-state values and uncertainty bounds

Fuego run	Ending T, K	Max, K	Min, K	Central estimate, K
Simulation 1	1309	1339	1299	1319 (midpoint)
Simulation 2	1496	1526	1466	1496 (midpoint)
Simulation 3	1265	1285	1260	1273 (midpoint)
Simulation 4	1537	1557	1532	1545 (midpoint)
Simulation 5	1352	1352	1337	1345 (midpoint)
Simulation 6	1534	1546	1510	1529 (mean)

plotted temperature from the simulation to 20 K above the last plotted temperature. This information is recorded in Table 4. These values of −5 and +20 K are added to simulation 3's last plotted temperature (listed in Table 5) to get the associated minimum and maximum values in the table. The minimum and maximum values correspond to the lower and upper extents of simulation 3's uncertainty bar shown in Fig. 4. The midpoint of the uncertainty bar for simulation 3 is also reported in Table 5. Similar methodology was employed to obtain the simulations 1, 2, 4, and 5 results listed in Table 5 and the associated uncertainty bars shown in Figs. 4 and 5. Quantitative regression procedures applied in Appendix B corroborate these results. There, regression curves are fit to the simulation data and asymptotic steady-state temperature estimates are determined where the curves come to a zero slope.

For simulation 6, an estimate of steady-state temperature was obtained by averaging over the time window shown in Fig. 4 and listed in Table 4. However, unlike for the experimental data, no empirical investigation was performed to determine sensitivity of calculated mean temperature to perturbations of the time window by shifting it and/or altering its width. Instead, uncertainty bounds listed in Table 5 are obtained from the maximum and minimum instantaneous temperatures over the time window. These are also depicted by the upper and lower extents of the time box in Fig. 4.

The latter approach is an easy way to get uncertainty bounds for results that have reached a noisy plateau, but is thought to give considerably exaggerated uncertainty estimates. Indeed, the time-window maximum and minimum for the experimental data are found in Table 3 to be roughly 25 to 35 K higher or lower than the time-window mean. Yet a much smaller range of ±5 K about the time-window mean was found to be a reasonable uncertainty tolerance by experimentation with various plausible averaging windows in Appendix A. In comparison, the upward and downward instantaneous differences from the mean for simulation 6 in Table 5 are roughly 20 K: considerably less than the 25 to 35 K differences from experimental time-window means. Therefore, if the experimental and simulation results have proportionate sensitivities of their time-window means to different plausible time windows that could reasonably have been chosen, then something less than ±5 K would be more reasonable to assign for the uncertainty of steady-state temperature for simulation 6.

Indeed, for experiments 6 and 7 and for simulation 6, the scale of temperature oscillation over their time windows in Fig. 4 is roughly the same. Yet, the vertical extents of the simulation 6 time-window box are of clearly larger scale than the vertical extents of the boxes for experiments 6 and 7. Thus, the uncertainty ascribed to the simulation 6 point estimate of steady-state temperature is relatively oversized. Hence, it appears that an easy and conservative approach to bounding the uncertainty on time-window means (of plateaued results), as an alternative to the labor-intensive approach of sampling different reasonable time windows (recall that we had seven other TCs to process, times six simulations), is to use instantaneous maximum and minimum temperature over the time window. Unfortunately, this appears to exaggerate by a considerable amount the uncertainty that should reasonably be attributed to steady-state point estimates given by time-window means.

From the various error bars and time-window boxes shown in the figures, readers can judge for themselves whether the visually determined uncertainty ranges on steady-state temperatures are reasonable. Indications that these ranges are reasonable are given in Appendix B. Estimates from the quantitative regression procedures in Appendix B all lie within the uncertainty bars in Figs. 4 and 5 with minimum to maximum ranges listed in Table 5.

These uncertainties are referred to as *graphical processing uncertainties* in the following. Both the experimental and simulation results possess nonnegligible graphical processing uncertainties.

C. Uncertainty Processing of Experimental Data for Model Validation Comparisons

The experimental data is processed here in a specific manner for model validation comparison to simulation results according to the validation paradigm and procedures detailed in [14]. There are many other model validation approaches of various types that one could pursue (see [14] for a sampling of references on this subject). However, the one applied here appears to uniquely have the required features to handle all the difficult attributes of the current Fuego validation problem.

Figure 6 presents the perceived dominant experimental uncertainties (from a model validation perspective) concerning steady-state temperature at the location of TC5. The sizes of the error bars in the figure are approximately to scale for the numerical magnitudes denoted in the figure.

We first consider experiment 6. From Table 3 it has a steady-state mean temperature of 1394 K over the applicable time window. It was previously established that this mean is subject to a ±5 K graphical uncertainty; thus, the designation 1394 K ±5 K in the figure.

The ±11 K measurement uncertainty indicated in the figure is an amalgamation of several sources. First, manufacturing variability of the 1/16-in.-diam ungrounded-junction mineral-insulated metal-sheath type-K TCs used in the experiments is specified by the manufacturer to result in a two-sigma accuracy of 2.2°C or 0.75% of reading in degrees Celsius, whichever is greater (see [28]). This accuracy level is said to exist over a temperature measurement range from 200 to 1000°C. The wording is generally interpreted (e.g., [28]) to imply that, although a given TC's particular transducing error ($\varepsilon = T_{\text{indicated}} - T_{\text{true}}$) is not known, the relative likelihood of what the

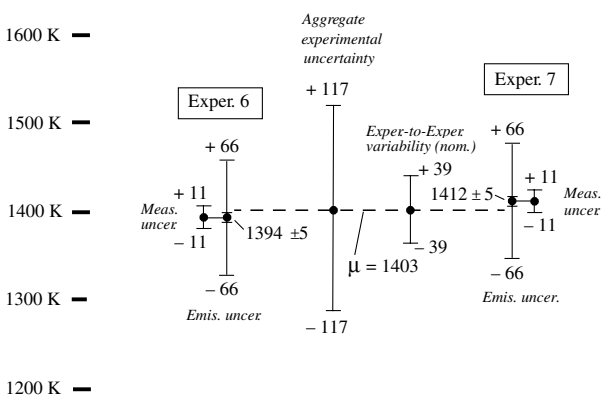

Fig. 6 Experimental uncertainties and rollup to aggregate experimental uncertainty at TC5. All temperatures are in degrees Kelvin.

error might be is governed by a normal distribution with standard deviation $\sigma = \frac{1}{2}$ (2.2°C or 0.75% of reading in degrees Celsius, whichever is greater). (See the shaded distribution associated with note 2 in Fig. 7.) Hence, the probability that the absolute value of the error is less than 2σ is given by integrating the shaded portion of the normal distribution that lies within $\pm 2\sigma$ of $T_{\text{indicated}}$. This integration yields a value slightly larger than 0.95. Accordingly, greater than 95% probability exists that the given TC's error has a magnitude less than 2σ; $|T_{\text{indicated}} - T_{\text{true}}| < 2\sigma$. Unfurling this inequality yields

$$T_{\text{indicated}} - 2\sigma < T_{\text{true}} < T_{\text{indicated}} + 2\sigma$$

Hence, greater than 95% probability exists that the true bead temperature lies within $\pm 2\sigma$ of the indicated bead temperature.

The above characterization is presumably obtained with a measurement system (particular data acquisition system, length of TC leads, length of electrical signal cables, etc.) that is properly calibrated for accuracy. Bias error is then introduced by the different particulars of the Sandia measurement system relative to the manufacturer's system. Characterization of the Sandia measurement TC channels in, e.g., [28], has determined that channel accuracies usually range well within 1°C, even for high-temperature calibration signals (standards) emulating 1000°C (1273 K), in the neighborhood of calorimeter temperatures in the present tests. With the error characterization from these activities, individual TC readings in the experiments could be bias-corrected, but usually are not because of the relatively small errors involved. Instead, a blanket uncertainty is assigned to the reported data to cover any error due to the Sandia measurement system. We assign an interval uncertainty here of ±0.25% of degrees Celsius reading. This equates to greater than ±2.5°C in the present case, representing a conservative upper bound on measurement-system errors traditionally seen in assessments against calibration standards at these elevated temperatures.

Hence, two significant sources of uncertainty are surmised to exist with regard to face-value temperature readings in the present experiments. Figure 7 helps illustrate the uncertainty sources and their implications. The probability density functions (PDFs) in the figure represent the uncertainty in true temperature due to TC-to-TC variable error (aleatory uncertainty) characterized by the manufacturer, assuming channel bias in their measurement system is corrected using a reliable calibration standard. The depicted potential shifts in the PDF (and therefore where the true temperature is expected to lie) are caused by potential differences between the manufacturer and Sandia measurement systems. Uncertainty in the magnitude of the systematic differences is taken here to be an interval (epistemic) uncertainty of ±0.25% of reading in degrees Celsius.

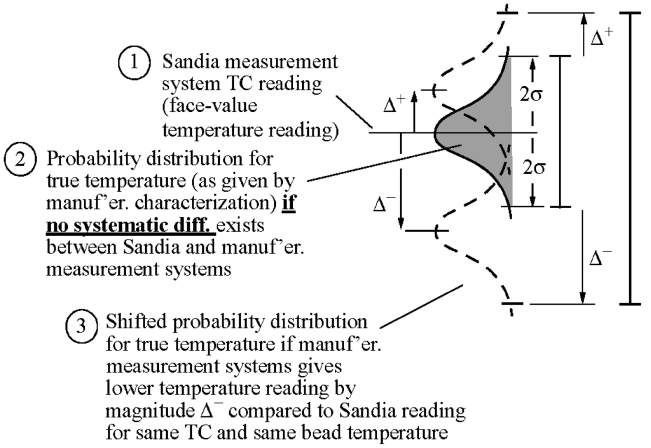

Fig. 7 Uncertainty of TC reading error is governed by probabilistic uncertainty (PDF) and potential systematic shifts in the PDF. Aggregate uncertainty interval shown at right is given by shifting either PDF or its interval representation shown, through the uncertainty range $[-\Delta^-, +\Delta^+]$ of the potential error of Sandia-measured system relative to TC manufacturer-measured system (to which the manufacturer-supplied PDF is referenced).

Accordingly, if no Sandia measurement-system bias exists, the true TC bead temperature would be expected with greater than 95% probability to lie within the range marked by the smaller error bar in the figure. However, when Sandia measurement-system bias does exist, and its value is characterized to lie, e.g., within an uncertainty range of $-\Delta^-$ to $+\Delta^+$, as depicted in Fig. 7, then the said PDF uncertainty (and therefore where the true temperature is expected to lie) might be shifted up or down by any value within the extremes shown. Under such shifts, the superposition spoken of here is a simplifying approximation, but is reasonable for the present circumstances and validation assessment. True superposition requires that the PDF in Fig. 7 is independent of the second uncertainty that precipitates the shifting and that the distribution remain unchanged or invariant as it shifts. The independence condition is met here and in many other real settings. However, the invariance condition is not strictly met here and is probably difficult to strictly meet in general. Here, the standard deviation (thus width) of the distribution is posed as a percentage of temperature reading. Since temperature magnitude changes as the distribution shifts vertically over the $[-\Delta^-, +\Delta^+]$ range in Fig. 7, the width of the distribution would change as well, violating the strict superposition conditions. However, this noninvariance effect is not large enough to be material to the validation conclusions here. If the noninvariance effect *is* significant and therefore important to capture, a Monte Carlo procedure such as that demonstrated in [29] can be employed.

Thus, the region within which the true temperature is expected to lie (at the 95% probability level) is given by the large error bar at right in Fig. 7. The lower and upper extremes of the large error bar are formed by simply adding the respective extremes of the PDF and interval source uncertainties: $[L, U] = [(-\Delta^- + -2\sigma), (\Delta^+ + 2\sigma)]$.

For the present case we get $[L, U] = [(-0.25\% + -0.75\%), (0.25\% + 0.75\%)]$ of degrees Celsius reading $= [-1\%, 1\%]$ of degrees Celsius reading. This uncertainty range can also be expressed conventionally as ±1% of degrees Celsius reading. Consulting Table 3 for degrees Celsius reading for experiments 6 and 7, ±1% yields the numerically rounded [−11 K, +11 K] error bars in Fig. 6 (analogous to the large error bar in Fig. 7).

Usually, a third (and dominant) source of temperature indication error exists, due to TC attachment effects. That is, the TC bead is not usually at the same temperature as the surface it is attached to, because of contact resistance effects and heat transfer to or from the bead. See [30] for a detailed presentation of the issues and quantification of the effects via finite-element modeling. Since it is really the surface temperature that is desired from the measurement, and not actually the bead temperature, any difference between the two is usually considered to be an error in the measurement. Such error has been indicated in Sandia investigations (e.g., [24,28]) to be as much of 6% of degrees Celsius reading under certain conditions. Such temperature differences or lags are not considered in the present activity, because steady-state temperature is the quantity of interest. For this quantity, positive and negative temperature lags are presumed to cancel out in the calculation of mean temperature as the instantaneous temperature oscillates noisily about the steady-state mean. Depending on the actual oscillation history, nonsymmetric time-weighting of positive and negative lags could lead to incomplete cancellation, but any such effect is likely to be small to negligible. It seems reasonable to assume here that any such effects are small enough to be covered by the significant margin of conservatism in the assigned 0.25% calibration uncertainty discussed previously.

Another source of validation uncertainty in the indicated experimental steady-state temperature is represented by the [−66 K, +66 K] error bars labeled *Emis. uncer.* in Fig. 6. It is evident that these are dominant uncertainties in the individual experiments. These arise from an uncertain *input* to the experiments (calorimeter and enclosure emissivity) as opposed to the aforementioned uncertainties in the measurement and processing of the experimental *outputs* (TC temperatures).

In general, even if experimental outputs (TC temperatures) are known with zero uncertainty, systematic uncertainty in the input conditions of the experiment implies that we cannot precisely resolve how the tested real system transforms experimental inputs to outputs.

Since the validation paradigm in this paper assesses how closely the experimental and modeled systems transform experimental inputs to output results, it is important to account for any resolution uncertainty in how the real system transforms inputs to outputs. Thus, as part of the data conditioning mentioned earlier, such resolution uncertainty is mapped to an equivalent amount of uncertainty applied to the output results of the experiment(s). This reduces the chances of committing *type X* model validation error. The validation methodology employed here is skewed toward preventing a type X model validation error [14] of an incorrect conclusion of no significant model bias when in fact significant bias does exist but is hidden by systematic uncertainty in nontraveling input factors in the experiments. A drawback is that the framework likely exaggerates the uncertainty resolution level within which it can be established that the model emulates the real system. The tradeoff is analogous to the situation in statistical hypothesis testing, where the more one chooses to decrease the odds of incorrectly rejecting a true hypothesis (type I error), the more likely it is that a type II error will be committed of incorrectly accepting a false hypothesis. As in hypothesis testing, the presence of uncertainty here forces a decision of which is the least undesirable outcome among two undesirable outcomes. The methodology here favors incurring a type Y error of exaggerating the range of possible model bias relative to what it is likely to actually be, and accepts this tradeoff in preference to incurring a type X error of underestimating the model bias. This conservative choice is argued in [14] to best support the objectives of best-estimate-with-uncertainty extrapolative prediction with the model.

In the present problem, emissivity ϵ of the calorimeter and enclosure surfaces is thought to be the only nontraveling systematic experimental input uncertainty to contribute significant uncertainty concerning the mapping of inputs to outputs by the tested fire/calorimeter system. Hence, the uncertainty in ϵ is mapped to an equivalent uncertainty in the experimental steady-state temperatures via the following procedure.

First, the emissivity uncertainty is propagated through the model. The uncertainty range for emissivity of the calorimeter exterior surface and the enclosure interior surfaces (walls, roof, and floor) is 0.76 to 0.96 as established in Sec. IV. The upward ΔT perturbation at TC5, corresponding to the upper-bound perturbation $\epsilon = 0.96$ from $\epsilon_{\text{nominal}} = 0.86$, is given directly by subtracting the central estimate for simulation 2 steady-state temperature in Table 5 from the central estimate for simulation 6 steady-state temperature. The result is $\Delta T = +33$ K. From the fact that a downward uncertainty perturbation $\epsilon = 0.76$ from the nominal value 0.86 is the same magnitude as the upward perturbation from 0.86 to 0.96, the downward temperature perturbation is -33 K by linearity (approximately) in this input factor. Therefore, the full uncertainty range of predicted temperature at TC5 due to an emissive uncertainty [0.76, 0.96] is projected to be 66 K, given by the uncertainty interval [-33 K, $+33$ K].

In the next section this interval [-33 K, $+33$ K] contributes to the *prediction* uncertainty for steady-state temperature at TC5 (see Fig. 8). It turns out here that the data-conditioning uncertainty to be added to the experimental results is twice that added to the *simulation* results, or [-66 K, $+66$ K], as shown in Fig. 6. This multiplier of two is derived in [14] and strictly applies to cases in which an experimental determination of the gradient of system response with respect to the systematically uncertain input factor(s) is used in the data-conditioning procedure. In most cases (such as the one here having ϵ as the input factor) the model must be used to get an approximation for the experimental gradient. Since the present case concerns only one input factor, we use slope instead of gradient in the following.

Because the model-indicated slope only approximates the true experimental slope, the sensitivity of validation conclusions to slope approximation error should be investigated as part of the validation procedure. It is calculated that the experimental slope can be as much as 50% greater than the modeled slope before the experimental uncertainty bar in Fig. 8 would extend outside the range of prediction uncertainty and change the validation conclusions arrived at later. The actual experimental slope is expected to be well less than 50%

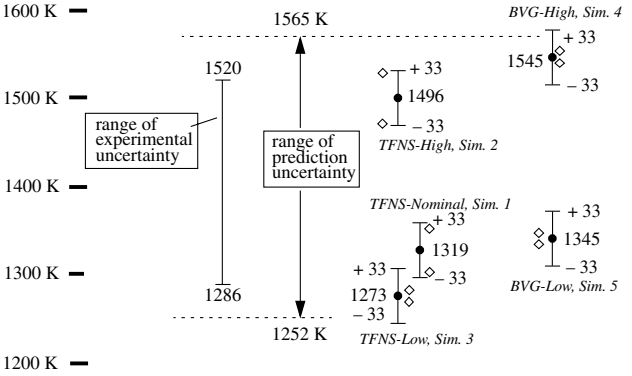

Fig. 8 Simulation results and uncertainties at TC5, with range of prediction uncertainty compared against range of aggregate experimental uncertainty. All temperatures are in degrees Kelvin.

greater than the modeled slope. If, in the other direction, the experimental slope is less than the modeled slope, the [-66 K, $+66$ K] uncertainty used for the data conditioning here is automatically conservative in that this direction of error makes it harder, not easier, to substantiate the model than if the true experimental slope were used for the data conditioning.

Finally, we address the test-to-test variability of the experimental results. Consider the uncertainty in Fig. 6 associated with the nominal point results of experiments 6 and 7. First note that the measurement and emissivity-related uncertainties are completely correlated. That is, since the same TC and data acquisition system and channel was used for TC5 in the two experiments, any associated bias errors in reading vs true temperature are essentially the same in the two experiments. The same is true of the uncertainty due to emissivity in the experiments. Although ϵ in the experiments is unknown to within a relatively large range of 0.76 to 0.96, it is reasonable to postulate the experiment-to-experiment differences in emissivity are small comparatively. Hence, the associated uncertainty in experiments 6 and 7 results is closely correlated (systematic over the two experiments).

Conversely, the ± 5 K graphical uncertainties associated with the steady-state temperature averaging windows for the two experiments can be considered to be independent and uncorrelated among the two experiments. This dictates that this source of uncertainty, for the purposes of characterizing experiment-to-experiment variability, be treated differently than the ones above.

First, however, consider the instructional case of only one experiment (either 6 or 7). The aggregate uncertainty for the single experiment would be constructed as follows. The [-11 K, $+11$ K] measurement uncertainty in Fig. 6 can be conceived as being subject to vertical shifting over the range [-66 K, $+66$ K] due to the emissive uncertainty. Such shifting is already familiar from the previous discussion pertaining to Fig. 7. Adding or superposing the uncertainty ranges according to the assumptions explained previously would yield

$$[(-11\text{ K} + -66\text{ K}), (11\text{ K} + 66\text{ K})] = [-77\text{ K}, +77\text{ K}]$$

This uncertainty bar about the nominal measured temperature would in turn be subject to vertical shifting of ± 5 K associated with the steady-state temperature graphical processing uncertainty. An aggregate uncertainty of [-82 K, $+82$ K] would result. This would be the case for either experiment 6 or 7 alone.

Now consider experiments 6 and 7 together. With reference to Fig. 6, the nominal results are, respectively, 1394 and 1412 K. The 18 K difference between these results cannot be explained by the ± 5 K uncorrelated processing uncertainties in each result. (Recall that the other uncertainties are effectively correlated or systematic between the two experiments, so cannot explain or contribute to any relative differences in the two experimental results.) Therefore, some other explanation lies behind the experimental differences. Certainly, things varied between the two experiments that we could not

characterize or did not explicitly treat in this project due to practical limitations.

In any case, we can reasonably posit that if many other repeat experiments were run, the results would vary according to a normal distribution, as is often the case with complex experimental systems. We can get estimates of what the mean and variance of the normal distribution would be by calculating these from the 1394 and 1412 K nominal experimental results. The mean of these is 1403 K as denoted in the figure. The standard deviation S, times two, is $2S = 25$ K. It must be kept in mind that this two-sigma magnitude of 25 K only *nominally* corresponds to 95% included probability in the postulated normal distribution of experiment-to-experiment steady-state temperature at TC5. The actual standard deviation of a large number of experimental repeats could be very different from the $S = (25\text{ K})/2$ calculated from just the two experiments. The small-sample uncertainty in the standard deviation value, and also in the mean value, is not accounted for in this paper. Hence, we cannot state with reasonable statistical confidence that the mean $\pm 2S =$ 1403 K $\pm$ 25 K defines an interval that encompasses 95% of the postulated normal distribution for experiment-to-experiment steady-state temperature variability. We can only state that this interval gives a nominal quantification of the physical variability. In fact, if the small-sample uncertainty on the apparent experimental variability is taken into account, this uncertainty on the variability or epistemic uncertainty on the aleatory uncertainty is large enough that it can potentially overturn our validation affirmations at some of the eight TC locations. Nonetheless, within the sampling uncertainty, it is also possible that an even stronger substantiation of model validity could occur, depending on the direction of the actual errors in the calculated values $\tilde{\mu} = 1403$ K and $S = (25\text{ K})/2$.

The nominal treatment also ignores the ± 5 K graphical processing uncertainties in the two steady-state temperatures. Because this uncertainty is uncorrelated over the two experiments, a worst-case (largest variance) treatment of this uncertainty involves decreasing the lower nominal result, 1394 K, by the maximum possibility over the applicable ± 5 K uncertainty range; and increasing the upper nominal result, 1412 K, by the maximum possibility over its independent ± 5 K uncertainty range. This yields adjusted results of 1394 K $-$ 5 K $=$ 1389 K, and 1412 K $+$ 5 K$=$ 1417 K. The accompanying two-sigma magnitude is $2S_{\text{high}} = 39.1$. This compares to the two-sigma value of 25 obtained in the previous paragraph, which did not account for the ± 5 K graphical processing uncertainties. A rounded value of 39 (shown in Fig. 6) is used in the following.

The effect of the correlation treatment is very significant here. If the ± 5 K uncertainties were perfectly correlated among the two experiments, then these would constitute a systematic uncertainty over the two experiments. Then, consistent with Fig. 7, the aggregate uncertainty from these two factors (graphical processing and experiment-to-experiment variability) would be to shift the (two-sigma value of 25 K) normal distribution over a ± 5 K range. The result would be

$$[(-5\text{ K} + -25\text{ K}), (5\text{ K} + 25\text{ K})] = [-30\text{ K}, +30\text{ K}]$$

by the linear superposition approximation.

Instead, if the ± 5 K graphical processing uncertainties are treated as independent and uncorrelated over the two experiments, the result is

$$[-2S_{\text{high}}, +2S_{\text{high}}] = [-39\text{ K}, +39\text{ K}]$$

as already determined. Therefore, the added effect of the uncorrelated ± 5 K processing uncertainties is a ± 14 K increment to the standing $\pm 2S = \pm 25.0$ nominal uncertainty from experimental variability. The ± 14 K is nearly 3 times the added effect as when the ± 5 K processing uncertainties are treated as perfectly correlated (systematic) over the two experiments.

To close out this subsection, we combine the experimental measurement and emissivity-related interval uncertainties with the normal PDF for experiment-to-experiment variability (mean 1403 K and standard deviation $S_{\text{high}} = (39\text{ K})/2$). Following the paradigm of Fig. 7, the [-11 K, $+11$ K] measurement uncertainty (interval) is superposed/added with the PDF uncertainty to get

$$[(-11\text{ K} + -39\text{ K}), (11\text{ K} + 39\text{ K})] = [-50\text{ K}, +50\text{ K}]$$

The interval uncertainty [-66 K, $+66$ K] due to uncertain emissivity is combined in by further superposition to get

$$[(-66\text{ K} + -50\text{ K}), (66\text{ K} + 50\text{ K})] = [-116\text{ K}, +116\text{ K}]$$

By using the numbers from each term with more decimal-place precision, we get the more accurate result [-117 K, $+117$ K] depicted by the uncertainty bar labeled Aggregate experimental uncertainty in Fig. 6. Analogous quantities and results to those presented in Fig. 6 for TC5 are presented in [3] for the other seven TCs.

D. Uncertainty Processing of Simulation Results for Model Validation Comparisons

Figure 8 shows Fuego simulation results at TC5 location for simulations 1–5. These were all run with the nominal value of emissivity, $\epsilon_{\text{nominal}} = 0.86$. The range of aggregate experimental uncertainty from Fig. 6 is plotted in Fig. 8 to lend a sense of scale and location relative to the simulation uncertainty. The central-estimate values of steady-state temperature from Table 5 are printed in the figure beside the plotted filled dots. The maximum and minimum reasonable steady-state temperature bounds from Table 5 are also plotted in Fig. 8 as unfilled diamond symbols.

The *nominal* steady-state temperature values indicate that the BVG turbulence model yields greater object heating than the TFNS model, all other simulation parameters being equal. This agrees with expectations (see Sec. IV).

When the emissivity uncertainty is accounted for, the expectations still hold up. For instance, consider the TFNS-high simulation (no. 2) and the BVG-high simulation (no. 4). The emissivity-related [-33 K, $+33$ K] uncertainty bars of these two simulations overlap some in Fig. 8. However, recall that these uncertainty bars represent *correlated* or systematic uncertainty over the set of simulations. That is, whatever the true value of ϵ is, it is the same for all simulations. If the true value (or any value) were input to the simulations, it would not yield, e.g., a value on the upper portion of the uncertainty bar of simulation 2 and a value on the lower portion of the uncertainty bar of simulation 4. Rather, the results would be correlated such they would both lie at closely the same position vertically within each error bar. Therefore, although the error bars of simulation 2 and simulation 4 overlap some, giving the *appearance* that it is possible to get a temperature realization from BVG-high simulation 4 that is lower than a corresponding realization from TFNS-high simulation 2, this is not really the situation here.

The expected ordering holds even when accounting for the graphical processing uncertainty. In contrast to the emissivity-related uncertainty, the graphical processing uncertainties (minimum and maximum bounds in Table 5) are *uncorrelated* across the various simulation results, and *can* significantly shift simulation results relative to each other. For example, in the extreme worst combination, the graphical uncertainty could allow 1) a shift of simulation 4's uncertainty bar downward until its midpoint (filled dot) aligns with the local lower diamond and 2) a shift of simulation 2's uncertainty bar upward until its midpoint aligns with the local upper diamond. Even in this worst case, the BVG-high simulation 4 uncertainty bar remains higher in vertical position than the TFNS-high simulation 2 bar. This implies a higher simulated temperature for BVG-high than for TFNS-high. Recall also that this is under a worst-case realization of the uncorrelated graphical uncertainties and that the graphical uncertainty magnitudes are likely highly exaggerated for the simulation results, as discussed earlier.

It is otherwise observed that, for a given turbulence model form (TFNS or BVG), the prescribed parameter sets for low, nominal, and high heating give consecutively hotter fires, or at least consecutively greater heating of the calorimeter as indicated at TC5 and the other

seven TC locations (see [3]). This ordering is retained at all TC locations under all possible realizations of the graphical and emissivity uncertainties.

These results support the proposition that for the model validation purposes here (and for extrapolative predictions in general) it is not necessary to expend simulations at the BVG-low and TFNS-high and medium parameter sets. These parameter sets (5, 2, and 1, respectively) routinely yield results that lie between the heating extremes of parameter sets 3 (TFNS-low) and 4 (BVG-high).

Thus, the effects of the six dominant sources of intrinsic modeling uncertainty in the fire-dynamics model (Table 1) can be effectively bounded by running Fuego simulations at just the two extreme parameter sets 3 and 4. Of course, other simulations have to be run to assess the effects of other sources of uncertainty (such as uncertain emissivity, numerical discretization, etc.) in a given prediction.

We now turn to the model validation objective of our analysis. Ultimately the aggregate experimental uncertainty in Fig. 8 will be compared against an aggregate simulation uncertainty that is appropriate to a validation assessment. The criterion for the model to be *substantiated* in the sense adopted here is that the model predictions, with uncertainties properly accounted for, yield an uncertainty band that encompasses the aggregate experimental uncertainty.

Otherwise, the processed experimental uncertainty, as presumably the best empirical evidence of where reality resides, lies outside the range predicted by the model. If the predictions fall short of spanning and capturing reality in the validation setting, it is reasonable to project that a similar relationship will hold in extrapolative predictions; reality will also lie outside of model predictions in extrapolation settings. Clearly, this is not what most designers, analysts, and decision-makers would want from their model predictions. They would like to be reasonably assured that reality lies *within* the uncertainty of their model predictions: the case shown in Fig. 8. Hence, as well as extrapolation behavior can be anticipated, risk mitigation in extrapolation appears to be best served by a validation objective that the uncertainty bars of the prediction in the model validation setting should effectively encompass reality. This supports an objective of best-estimate-with-uncertainty modeling and prediction.

In the present framework, the preliminary aggregate prediction uncertainty in Fig. 8 (before accounting for graphical uncertainties) ranges from the low end of the lowest uncertainty bar (given by TFNS-low simulation 3), to the high end of the highest uncertainty bar (from BVG-high simulation 4). That is, the emissive uncertainty and the six uncertainties in Table 1 have possible combinations or realizations over their joint uncertainty space (where these seven factors are justifiably assumed to be independent of each other) that can yield model predictions which vary from the low end to the high end of the said uncertainty range. This preliminary range of prediction uncertainty is impacted by the graphical uncertainties as explained next.

The graphical processing uncertainty for the BVG-high simulation 4 results is given by the minimum and maximum bounds in Table 5. The graphical uncertainty allows that simulation 4's uncertainty bar in Fig. 8 can really lie anywhere within an upward or downward shift, where its midpoint (filled dot) remains between the upper and lower unfilled diamonds to the side of the uncertainty bar. Analogous freedoms are allowed for the simulation 3 uncertainty bar to be shifted between the upper and lower diamonds at its side. Recall that the graphical uncertainties are not correlated with each other, so the allowable shifting of simulation 4's uncertainty bar is independent of simulation 3's.

At the upper end of the simulated temperature range, treatment of the graphical uncertainty for simulation 4 figures into the validation determination relative to the high end (1520 K) of the experimental temperature uncertainty range shown in Fig. 8. Analogous considerations hold for the graphical uncertainty for simulation 3 and the validation determination at the low end (1286 K) of the experimental uncertainty range. Within the graphical uncertainty, the simulation 4 uncertainty bar could be translated upward, and the simulation 3 results could be translated downward, such that validation margins are greater at both the upper and lower ends. The opposite extreme possible combination is a downward shift in simulation 4 results and an upward shift in simulation 3 results, such that validation margins decrease at both the upper and lower ends.

A treatment that increases validation margins here, or which tends to create a closer comparison when positive margins such as those in Fig. 8 do not exist, is said to be nonconservative. Accounting for simulation uncertainty of the intrinsic modeling factors of Table 1 and for the emissive uncertainty have the effect of expanding the prediction uncertainty as well. Yet, this expansion is not considered to be nonconservative in the validation formulation. From a model validation perspective (in the context of extrapolative predictions and hierarchical modeling), these factors are different in nature from the graphical uncertainty, so are handled differently. The intrinsic modeling factors of Table 1 proceed to any new predictions with the (validated) model, so their uncertainty is transported inherently to new prediction results. This is not the case with the graphical uncertainties being discussed here. New graphical uncertainties/magnitudes will be present in new simulation results. For example, if the new simulations are terminated after arriving at a smooth, perfectly flat asymptotic steady-state, no graphical uncertainty will be present at all. In terms of the emissive uncertainty, such as the uncertainties in Table 1 this is an uncertain *input* to the model, not a (graphical) uncertainty from processing of the *outputs* of the model. For the emissive uncertainty, it is not a traveling model uncertainty such as the model-intrinsic uncertainties, but through a different mechanism avoids any nonconservative effects. It propagates to simulation results with a 1X effect in expanding the prediction uncertainty, but propagates to the experimental results via data conditioning with a conservative 2X uncertainty effect as previously explained in association with type X error. This type of treatment could enable the validation criterion to be (falsely) met or approached closer, while the actual value of the factor (for no graphical processing error) might correspond to a more biased model than the validation assessment with graphical processing error may lead one to believe. Hence, this is one possible way to commit type X model validation error. Type X error can arise from many other sources, such as model discretization uncertainties, systematic uncertainties of nontraveling experimental inputs (here, emissivity as already discussed), and other sources catalogued in [14].

To guard against the potential for type X error that the graphical uncertainties pose, the framework takes a conservative approach of attempting to eliminate the risk entirely. To do this, the most extreme possible combination is invoked by shifting the simulation 4 uncertainty bar downward the full allowable amount until its midpoint (filled dot) is beside the lower diamond at its side and by shifting the simulation 3 uncertainty bar upward the full amount until its midpoint is beside the upper diamond at its side. This maximally decreases the validation margins at both the upper and lower ends of the data range. Hence, this likely causes a type Y model validation error of the framework exaggerating the perceived extent of potential model bias. Unfortunately, eliminating the risk of type X validation error comes with a tradeoff of likely committing a type Y error. An alternative is to just simply ignore the graphical processing uncertainties. However, in closer cases than the one here, this can incur a substantial risk of type X error in the validation conclusions.

Hence, the validation-conservative lower temperature limit of prediction uncertainty is obtained as follows. The simulation 3 uncertainty bar in Fig. 8 is centered on the upper diamond to its right, which has a temperature of 1285 K (maximum value for simulation 3 in Table 5). The temperature at the bottom of this uncertainty bar is therefore $1285\ \text{K} - 33\ \text{K} = 1252\ \text{K}$. This temperature is marked by the lower horizontal dashed line in the figure.

The validation-conservative upper temperature limit is obtained by a mirrored procedure. The simulation 4 nominal 1545 K result is shifted downward to the lower diamond at its right, at 1532 K (minimum value for simulation 4 in Table 5). The temperature at the top of this uncertainty bar is therefore $1532\ \text{K} + 33\ \text{K} = 1565\ \text{K}$. This temperature is marked by the upper horizontal dashed line in the figure. Accordingly, the final aggregate range of prediction uncertainty is 1252 to 1565 K.

VI. Analysis and Discussion of Validation Comparisons

Analogous quantities and results to those plotted in Figs. 6–8 for TC5 can be found in [3] for all eight TCs. Speaking to the TC5 results here, there is considerable margin for error in the uncertainty estimates and processing of experimental and simulation results before the model validation conclusions would be overturned. The upper and lower simulation bounds in Fig. 8 are seen to fairly spaciously encompass the range of the aggregate experimental uncertainty within which steady-state experimental temperatures are provisionally expected to lie. Errors would have to conspire (i.e., coordinate in enough antagonistic directions and magnitudes) to overcome the existing margin for error of 45 K (1565 K − 1520 K) at the top. At the bottom, a margin of 34 K (1286 K − 1252 K) exists against errors that might conspire. The existence of errors is a given. However, it is much less likely that they would be sufficiently large and sufficiently conspiring that the net result would exceed the indicated margins for error.

At the other seven TC locations the margins for error are even greater at 13 of the 14 upper and lower ends. In no case does the range of experimental uncertainty extend outside the prediction uncertainty. Therefore, the model is provisionally substantiated based on the validation results at the eight diversely spaced TCs on the calorimeter.

Provisionally signifies here that large error may exist in our quantification of experiment-to-experiment variability. Nominal mean and standard-deviation results based on only two experimental repeats were used. The uncertainty on these small-sample statistics, and also in the assumption itself of normally distributed variability, could result in errors in an antagonistic direction and large enough to overturn the nominal affirmations at most TC locations. However, note that antagonistic error from this source could be partially offset or even completely overwhelmed by the substantial conservatisms introduced at several points in our experimental and simulation results processing. Moreover, the small-sample errors could go the other way, to more strongly affirm the validation conclusions at any and all TC locations. In any case, uncertainty due to experiment-to-experiment variability is at least nominally accounted for in our validation analysis, results, and conclusions.

Finally, a mitigating factor is pointed out concerning contact resistance between the bottom of the steel cone and the thermally massive steel floor to which it was bolted. It was not practical in this project to model the contact resistance at this interface, so an adiabatic boundary condition was applied at the bottom of the cone. This modeling error could be substantial, depending on the particular location considered on the calorimeter. The error causes higher temperatures to be predicted in the calorimeter than would occur if heat transfer from the cone to the steel floor was modeled. It turns out that the experimental uncertainty bars at all TC locations are skewed toward the lower ranges of the prediction bars. That is, modeling the contact conductance at the interface would shift the predicted temperatures downward, especially for the bottom four TCs 1–4. This would move things in the right direction toward better central agreement between the experimental and predicted temperatures. Since the corrections would be expected to be only tens of degrees Kelvin (not hundreds), there is sufficient margin at the top of the simulation uncertainty bars to absorb such corrections and still encompass the experimental data at the upper end.

In fact, the smallest upper margin for such a correction is 45 K at TC5. From [3], all other TCs have a significantly larger upper margin for correction. Note that TC5 is from the upper set of TCs (level 10 in Fig. 2), in which the effect of the correction would be expected to be fairly small. Among the lower set of TCs, where the correction would be expected to be much larger, perhaps tens of degrees Kelvin, the smallest available margin for correction at the upper end of the uncertainty bars of all eight TCs is 168 K. This is well in excess of what the correction is reasonably expected to be.

Correcting for what is thought to be easily the largest unaccounted-for source of physics modeling error in the validation comparisons would be expected to move things in a direction that strengthens the validation conclusions, but not by a magnitude so large that it goes too far in this direction such that the experimental results are no longer bounded above by the tops of the simulation uncertainty bars.

Finally, the issue of calculation verification is considered. As previously explained, project constraints required the use of medium spatial discretization cell sizes and a medium number of discrete-ordinate directions for resolving the participating-media radiation transport. The solver error-tolerance parameters that control numerical resolution in the steady-state Fuego computations were also set at medium levels. (See [3] for details of discretization and solver resolution levels.) The medium levels used for spatial discretization, radiation transport, and solver computations are representative of what are routinely used for production calculation work by Sandia fire analysts. This is a pragmatic choice to achieve reasonable calculation run times (a few weeks on a thousand processors or less), but is also heuristically supported by comparisons against considerably finer resolution levels also tried on other projects (see, e.g., [17]).

In any case, if any discretization-related errors happen to be in the wrong (antagonistic) direction that would work against validation margins, the allowable margins for error are tens of degrees Kelvin, as already established. This is significant room for error, but experience indicates that discretization effects *could* be this large or larger. If a quantification of discretization-error uncertainty was available from resolution-refinement studies (e.g., by the methods presented in [18]), then in the validation analysis these would be handled like the simulation graphical processing uncertainties were. Finally, it is noted that significant code verification efforts for Fuego have been undertaken [31], which counts toward the coding correctness of the equations involved.

VII. Conclusions

In view of the arguments just made and the many conservatisms in the validation processing, the preponderance of the available evidence strongly suggests that the *Fuego fire-dynamics model for wind-driven fire conditions (including the uncertainties in Table 1 for intrinsic modeling uncertainty in the computational fluid dynamics and combustion subgrid scale models) captures the experimental steady-state temperature results here according to the pragmatic validation methodology applied.*

That is, the predicted calorimeter temperatures effectively bound the experimental temperatures at the eight diversely representative locations chosen for comparison on the calorimeter. Furthermore, there is considerable room for errors in the uncertainty estimates and processing of experimental and simulation results before this validation conclusion of a substantiated model would be overturned. The errors would have to be sufficiently large in magnitude and sufficiently conspiring (combining in antagonistic directions and/or sufficiently avoiding cancellation) to exceed the indicated considerable margins for error. The probability of these joint events occurring may be very low, and cannot reasonably be expected to be high, *but could occur.* In particular, the greatest concerns are 1) the error associated with using the small-sample (two sample) standard deviation as representative of the standard deviation from a large number of repeat fire tests and 2) error in the calculated steady-state results due to underresolution in the discretized model and computation.

Nevertheless, on balance the analysis supports the model-affirming conclusion arrived at for this experimental/physical setting. To establish model adequacy in this setting, however, another validation requirement must be met. Consider, for example, that infinite uncertainty bars on the simulation results (say due to infinite model-intrinsic uncertainties) would also contain the experimental uncertainty in this project, therefore leading to a substantiated model according to the validation paradigm employed. Yet with infinite intrinsic uncertainty, the model could not be claimed to be useful for prediction purposes, and therefore the model could not be considered to be adequate from the perspective of its intended use (per the validation definition statement in Sec. I).

Hence, substantiation that the model results bound reality does not automatically equate to model adequacy. Model usefulness must be established with respect to allowable uncertainty tolerances in design, analysis, and/or decision-making purposes for which the model will be used. It turns out that such assessment has been performed for the present model and its usefulness and therefore adequacy have been established. This is a significant result for the Fuego fire-dynamics model and for Sandia's fire modeling and weapon risk-assessment programs. The current work also represents a significant advancement in demonstration of a practical and versatile model validation paradigm and methodology.

Acknowledgments

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Appendix A: Estimation of Experimental Steady-State Temperatures by Time-Window Averaging

Here, a quantitative procedure is applied to estimate representative steady-state temperatures from the noisy temperature data of experiments 6 and 7. The results here support the uncertainty ranges stated in the body of the paper for TC5 steady-state temperatures in experiments 6 and 7.

Figure A1 shows TC5 temperature data from experiment 6. Also plotted are various smoothed representations of TC5 data. These are obtained by time-filtering the data with various time-averaging windows (of lengths 4, 6, and 8 min). For example, consider the smoothed data corresponding to the 6 min averaging window in Fig. A1. For each abscissa value of this curve, the ordinate value is obtained by averaging the experimental data over a 6 min time window centered on the abscissa value. Thus, the plotted temperature at x minutes is a smoothed value obtained by averaging over the $x \pm 3$ min . time window extending from $x - 3$ min to $x + 3$ min.

Figure A2 is a close-up of the region in Fig. A1, in which TC5 temperature reaches a plateau. The 32–38 min time window plotted in Fig. 4, with time-window mean of 1394 K cited in Table 3, corresponds to the point (35 min, 1394 K) on the curve with 6 min smoothing in Fig. A2. In the body of the paper it was said that the graphical processing uncertainty associated with the steady-state temperature for TC5 in experiment 6 was ± 5 K about the point estimate of 1394 K. The ± 5 K graphical uncertainty was attributed to the fact that different but still reasonable time-averaging windows than the one plotted in Fig. 4 (and used to get the point estimate of 1394 K) could yield different point estimates for TC5's steady-state temperature. Figure A2 shows that between the times of 29 to 36 min the curve of smoothed temperature using the 6 min filter remains within the stated 1394 ± 5 K uncertainty range. That is, the stated 1394 ± 5 K uncertainty range admits point estimates of steady-state temperature that come from 6 min time-averaging windows centered anywhere from 29 to 36 min into the experiment. The next paragraph

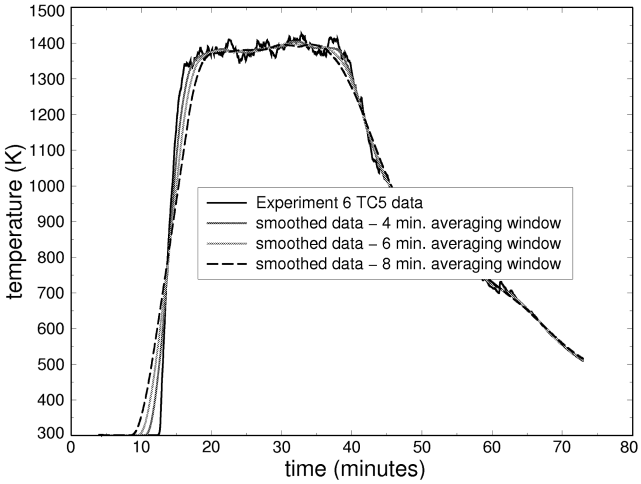

Fig. A1 Experiment 6 TC5 raw and smoothed temperature data (smoothing by data averaging over various time-window filter lengths as indicated).

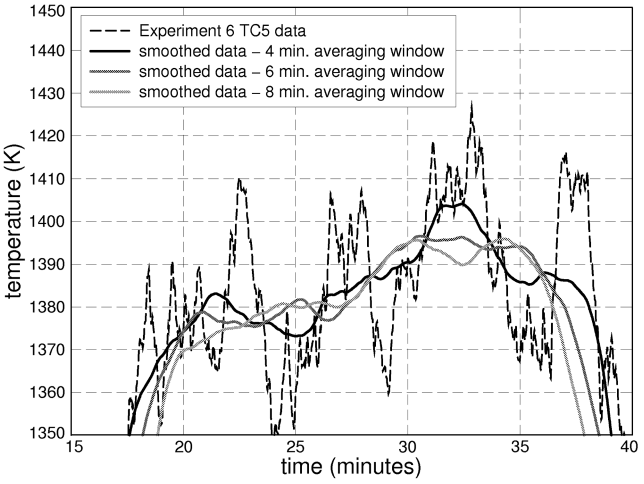

Fig. A2 Close-up of the region in Fig. A1, in which TC5 temperature reaches a plateau.

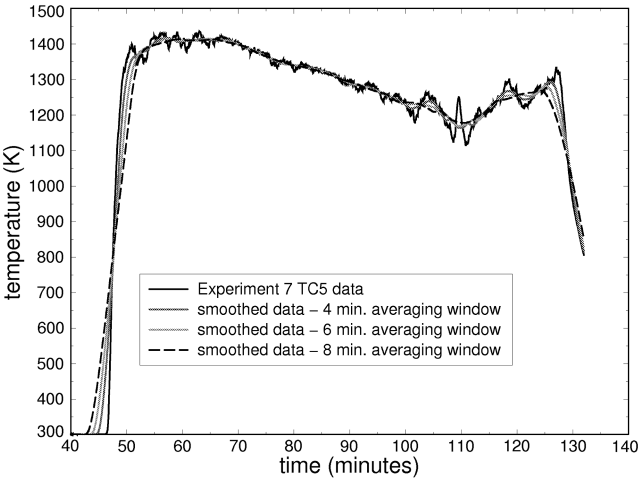

Fig. A3 Experiment 7 TC5 raw and smoothed temperature data (smoothing by data averaging over various time-window filter lengths as indicated).

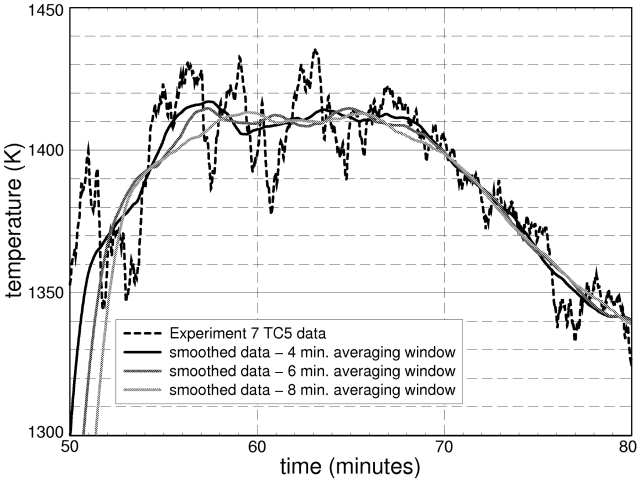

Fig. A4 Close-up of the region in Fig. A3, in which TC5 temperature reaches a plateau.

also establishes that the validation conclusions in the paper are not narrowly dependent on the particular time-window size used in the analysis and shown in Fig. 4.

The effect of filtering window size on point estimates of steady-state temperature will now be examined. Consider the curve in Fig. A2 corresponding to an 8 min filtering window. For this curve the 1394 ± 5 K uncertainty range admits point estimates from an 8 min time-averaging window also centered anywhere from 29 to 36 min in the data. Thus, the 1394 ± 5 K graphical uncertainty tolerance used in the validation analysis is reasonably representative and robust with respect to an 8 min filtering window as well. When using a 4 min filtering window, however, Fig. A2 shows that the same relatively large 29 to 36 min time span of robustness to translational perturbations of the filtering window would require a much larger graphical uncertainty tolerance of 1394 ± 10 K. But even if this larger graphical uncertainty were applied in the paper's validation analysis, the validation conclusions would not be affected. Moreover, a 4 min time window is considered to be at the small end of reasonable filter widths for deriving a representative steady-state temperature in these experiments.

Now consider steady-state temperature estimation for experiment 7 (TC5). Figures A3 and A4 for experiment 7 are the analogs of Figs. A1 and A2 for Experiment 6. It was stated in the body of the paper that the averaging time window drawn in Fig. 4 for experiment 7 yields a mean temperature of 1412 K (see Table 3). The 13 min filtering window is centered at about the 60.5 min mark in the data. This 13 min filtering window can be translated up and down in time over a range of 6.5 min (window centered as low as the 59 min mark to as high as the 65.6 min mark) while the temperature mean stays within the $1412 \text{ K} \pm 5$ K graphical uncertainty band cited in the paper for TC5's steady-state temperature in Experiment 7. Thus, the uncertainty tolerance employed in the paper for steady-state temperature allows a relatively large translational freedom where the 13 min filter is located. The next paragraph also establishes that the validation conclusions in the paper are not narrowly dependent on the particular time-window size used in the analysis and shown in Fig. 4.

The effect of filtering window size on point estimates of steady-state temperature will now be examined. Consider the curve in Fig. A4 corresponding to an 8 min filtering window. For this curve the employed 1412 ± 5 K uncertainty range admits point estimates from an 8 min filtering window centered anywhere from about 57 to 67 min in the data. Thus, the graphical uncertainty tolerance used in the validation analysis is even more robust with respect to an 8 min filtering window. A similar conclusion holds when a 6 min filter is used. The 1412 ± 5 K uncertainty range admits point estimates from a 6 min filtering window translated over an even larger range of times. The center of the window can be anywhere between the 56 and 68 min marks in the data. A 4 min filtering window can be translated over an even greater span of times, with the center allowed to range between 55 and 69 min in the data. Hence, the 1412 ± 5 K graphical uncertainty tolerance employed in the validation analysis is rather robust to various reasonable sizes and locations of the steady-state time-averaging window used.

Appendix B: Regression of Simulation Data to Estimate Asymptotic Temperatures

Here, a quantitative procedure is applied to estimate asymptotic stead-state temperatures from the incompletely stabilized simulation results. Figures B1 and B2 show regression curves fit to representative end ranges of simulation data for simulations 1–6. The regression curves are plotted out to where they reach a zero slope. The temperatures there are quantitative projections of the asymptotic steady-state temperatures of the simulations. The projected steady-state temperatures are listed in Table B1. The projected asymptotic temperatures all lie within the uncertainty bars drawn in Figs. 4 and 5 and defined in Table 5. These quantitative projections provide some corroboration for the visually estimated asymptotic steady-state temperatures and uncertainty ranges used in the paper's model validation analysis.

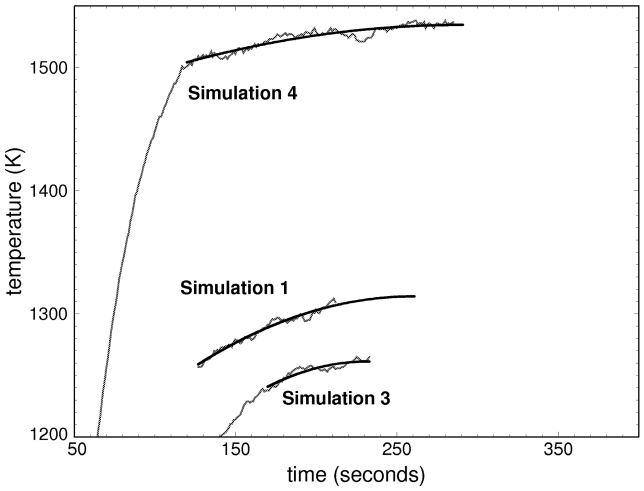

Fig. B1 Regression curves to simulations 1, 3, and 4 results.

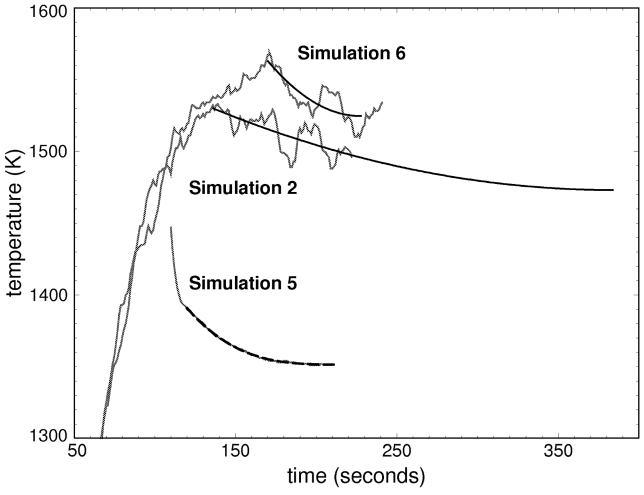

Fig. B2 Regression curves to simulations 2, 5, and 6 results.

Table B1 Projected asymptotic steady-state temperatures by regression of simulation results (TC5 location)

Simulation no.	Temperature
1	1314.3 K
2	1473.1 K
3	1261.0 K
4	1535.0 K
5	1351.3 K
6	1524.5 K

References

[1] Romero, V. J., "Some Issues and Needs in Quantification of Margins and Uncertainy (QMU) for Phenomenologically Complex Coupled Systems," 8th AIAA Non-Deterministic Methods Conference, AIAA Paper 2006-1989, Newport, RI, May 2006.

[2] Nakos, J. T., "JT4A-14B1 Abnormal Thermal Environment Qualification Activity: Test Data," Sandia National Labs., Rept. SAND2007-4286, Albuquerque, NM, July 2007.

[3] Luketa, A., Romero, V., Domino, S., Glaze, D., Sherman, M., and Figueroa, V., "Validation and Uncertainty Quantification of Fuego Simulations of Calorimeter Heating in a Wind-Driven Hydrocarbon Pool Fire," Sandia National Labs., Rept. SAND2009-7605, Albuquerque, NM, Nov. 2009.

[4] Domino, S. P., Moen, C. D., Burns, S. P., and Evans, G. H., "SIERRA/Fuego: A Multi-Mechanics Fire Environment Simulation Tool," 41st AIAA Aerospace Sciences Meeting, AIAA Paper 2003-0149, Reno, NV, Jan. 2003.

[5] *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*, AIAA Standards Series, Std. G-077-1998, AIAA, Reston, VA, 1998.
[6] "Accelerated Strategic Computing Initiative (ASCI) Program Plan," U.S. Dept. of Energy, Rept. DOE/DP-99-000010592, 2000.
[7] "DoD Modeling and Simulation Verification, Validation, and Accreditation," U.S. Dept. of Defense, Defense Modeling and Simulation Office, Instruction 5000.61, 13 May 2003.
[8] *Guide for Verification and Validation in Computational Solid Mechanics*, American Society of Mechanical Engineers, New York, 2006.
[9] *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer*, American Society of Mechanical Engineers, New York, 2009.
[10] *Technical Standard for Models and Simulations*, NASA, Std. 7009, July 2008.
[11] Romero, V. J., "On Model Validation and Extrapolation for Best-Estimate-Plus-Uncertainty Predictions," Sandia National Labs., Rept. SAND2005-7678C, Albuquerque, NM, Nov. 2005.
[12] Romero, V. J., "A Paradigm of Model Validation and Validated Models for Best-Estimate-Plus-Uncertainty Predictions in Systems Engineering," SAE International Paper 200-01-1746, 2007.
[13] Romero, V. J., "Validated Model? Not So Fast. The Need for Model 'Conditioning' as an Essential Addendum to the Model Validation Process," 9th AIAA Non-Deterministic Methods Conference, Honolulu, AIAA Paper 2007-1953, April 2007.
[14] Romero, V. J., "Data & Model Conditioning for Multivariate Systematic Uncertainty in Model Calibration, Validation, and Extrapolation," 12th AIAA Non-Deterministic Approaches Conference, Orlando, FL, AIAA Paper 2010-2511, 2010.
[15] Romero, V. J., "A New Interval-Based Real-Space Approach to Model Validation Involving Aleatory and Epistemic Uncertainty," Sandia National Labs., Rept. SAND2009-3160P, Albuquerque, NM, 30 April 2009.
[16] Romero, V. J., Luketa, A., and Sherman, M., "Application of a Pragmatic Interval-Based 'Real Space' Approach to Fire-Model Validation Involving Aleatory and Epistemic Uncertainty," 11th AIAA Non-Deterministic Approaches Conference, AIAA Paper 2009-2279, Palm Springs, CA, May 2009.
[17] Black, A. R., Hobbs, M. L., Dowding, K. J., and Blanchat, T. K., 2007, "Uncertainty Quantification and Model Validation of Fire/Thermal Response Predictions," 18th AIAA Computational Fluid Dynamics Conference, AIAA Paper 2007-4204, Miami, FL, June 2007.
[18] Roache, P. J., *Fundamentals of Verification and Validation*, Hermosa, Albuquerque, NM, 2009.
[19] Felske, J. D., and Tien, C. L., "Calculation of the Emissivity of Luminous Flames," *Combustion Science and Technology*, Vol. 7, No. 1, 1973, pp. 25–31.
doi:10.1080/00102207308952339
[20] Ertesvag, I. S., and Magnussen, B. F., "The Eddy-Dissipation Turbulence Energy Cascade Model," Dept. of Applied Mechanics, Thermodynamics and Fluid Dynamics, Norwegian Univ. of Science and Technology, Trondheim, Norway, 1997.
[21] Nicolette, V., Tieszen, S., Black, A., Domino, S., O'Hern, T., "A Turbulence Model for Buoyant Flows Based on Vorticity Generation," Sandia National Labs., Rept. SAND2005-6273, Albuquerque, NM, 2005.
[22] Rouson, D., Tieszen, S., and Evans, G., "Modeling Convection Heat Transfer and Turbulence with Fire Applications: A High Temperature Vertical Plate and a Methane Fire," *Proceedings of the Summer Program*, Stanford Center for Turbulence Research, Stanford, CA, 2002.
[23] Romero, V., Black, A., Brown, A., Tieszen, S., Domino, S., Blanchat, T., et al. "Verification, Validation, and Uncertainty/Sensitivity Studies of Fire Modeling at Sandia National Labs.," Sandia National Labs., Rept. SAND2008-2010C; also *Proceedings of the National Institute of Standards (NIST) Building and Fire Research Laboratory Annual Conference* [CD-ROM], Gaithersburg, MD, 2008.
[24] Nakos, J. T., Suo-Anttila, J. M., and Gill, W., "Shroud Boundary Condition Characterization Experiments at the Radiant Heat Facility," Sandia National Labs., Rept. SAND2004-5080, Oct. 2004.
[25] Mehling, H., Kuhn, J., Valentin, M., and Fricke, J., "Change of Infrared Emissivity of Metal Surfaces During Oxidation," *High Temperatures—High Pressures*, Vol. 30, 1998, pp. 333–341.
doi:10.1068/htec140
[26] Worner, B., and Neuer, G., "Emittance Measurements on Steel During Oxidation," *High Temperatures—High Pressures*, Vol. 15, 1983, pp. 455–462.
[27] Siegel, R., and Howell, J., *Thermal Radiation Heat Transfer*, Taylor and Francis, New York, 2002.
[28] Nakos, J. T., "Uncertainty Analysis of Thermocouple Measurements Used in Normal and Abnormal Thermal Environment Experiments at Sandia's Radiant Heat Facility and Lurance Canyon Burn Site," Sandia National Labs., Rept. SAND2004-1023, April 2004.
[29] Romero, V. J., Sherman, M. P., Dempsey, J. F., Johnson, J. D., Edwards, L. R., Chen, K. C., et al. "Development and Validation of a Component Failure Model," 45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, AIAA Paper 2005-2141, Austin, TX, April 2005.
[30] Figueroa, V., "Effects of Parameter Variations Associated with $1/16''$ Mineral Insulated Metal Sheathed Thermocouples Installation on the Surface Temperature Measurement of a Uniformly Heated Plate," M.S. Thesis, Univ. of New Mexico, School of Mechanical Engineering, Albuquerque, NM, April 2006.
[31] Domino, S. P., Wagner, G., Luketa, A., Black, A., and Sutherland, J., "Verification for Multi-Mechanics Applications," 9th AIAA Non-Deterministic Methods Conference, Honolulu, HI, AIAA Paper 2007-1933, April 2007.