Journal of Testing and Evaluation
Selected Technical Papers

STP 1541

Uncertainty in Fire Standards
and What to do About It

John R. Hall, Jr.
JTE Guest Editor

Journal of Testing and Evaluation Selected Technical Papers STP1541

Uncertainty in Fire Standards and What to Do About It

JTE Guest Editor:
John R. Hall, Jr.

ASTM International
100 Barr Harbor Drive
PO Box C700
West Conshohocken, PA 19428-2959

Printed in the U.S.A.

ASTM Stock #: STP1541

Library of Congress Cataloging-in-Publication Data

ISBN: 978-0-8031-7527-3

Journal of Testing and Evaluation (JTE) Scope

This flagship ASTM journal is a multi-disciplinary forum for the applied sciences and engineering. Published bimonthly, JOTE presents new technical information, derived from field and laboratory testing, on the performance, quantitative characterization, and evaluation of materials. Papers present new methods and data along with critical evaluations; report users' experience with test methods and results of interlaboratory testing and analysis; and stimulate new ideas in the fields of testing and evaluation.

Major topic areas are fatigue and fracture, mechanical testing, and fire testing. Also publishes review articles, technical notes, research briefs and commentary. All papers are peer-reviewed.

Photocopy Rights

Peer Review Policy

Each paper published in this volume was evaluated by two peer reviewers and at least one editor. The authors addressed all of the reviewers' comments to the satisfaction of both the technical editor(s) and the ASTM International Committee on Publications.

The quality of the papers in this publication reflects not only the obvious efforts of the authors and the technical editor(s), but also the work of the peer reviewers. In keeping with long-standing publication practices, ASTM International maintains the anonymity of the peer reviewers. The ASTM International Committee on Publications acknowledges with appreciation their dedication and contribution of time and effort on behalf of ASTM International.

Citation of Papers

When citing papers from this publication, the appropriate citation includes the paper authors, "paper title", J. ASTM Intl., volume and number, Paper doi, ASTM International, West Conshohocken, PA, Paper, year listed in the footnote of the paper. A citation is provided as a footnote on page one of each paper.

Printed in Bay Shore, NY
February, 2012

Foreword

THIS COMPILATION OF THE *Journal of Testing and Evaluation, (JTE)*, STP 1541, Uncertainty in Fire Standards and What to Do About It, contains only the papers published in JTE that were presented at a symposium in Anaheim, CA on June 16, 2011 and sponsored by ASTM Committee E05 on Fire Standards.

The Symposium Chairman and Guest Editor is John R. Hall, Jr., Division Director of Fire Analysis and Research, National Fire Protection Association, Quincy, MA, USA.

Contents

Overview

When it comes to measuring product fire performance in standard tests or standard calculations, uncertainty is the elephant in the room. We honor the need for uncertainty measurement and interpretation in the abstract and set uncertainty-related requirements for every standard (such as the precision and bias requirements). When it comes to actually calculating uncertainty and incorporating uncertainty into our use of test and calculation results, however, we often fall short. We act as if the elephant really isn't there, because we don't know what to do about it if it is there.

On July 16, 2011, ASTM Committee E05 on Fire Standards conducted an all-day symposium with 15 papers on the subject of uncertainty in fire standards and what to do about it. The objective of the symposium was to discuss different issues related to uncertainty in fire standards and to cover how different parties – testing laboratories, enforcement authorities, manufacturers, practicing engineers – incorporate uncertainty into their use of results from fire safety tests and calculations. The symposium was also designed to look at larger implications of different approaches and provide overviews of some of the newest methods and approaches for handling uncertainty.

An effort has been made to post all 15 presentations at the E05 website for a limited time at http://www.astm.org/COMMIT/e05_presentations.htm.

The first four presentations provided a basic familiarity with existing methods and procedures and with relevant ASTM and other standards. Because these presentations were not designed to provide new information– only to lay a solid foundation for the later presentations – they were not converted into published papers. This STP contains papers based on the other eleven presentations.

Because you, the reader, may not have access to those first four presentations, this Overview will provide a brief description of the contents of those presentations as well as places to go for more information.

William Guthrie of NIST led off with his presentation on "Assessing Uncertainty in Measurement Results: The Big Picture."

- He began by linking the need for uncertainty to situations where the threshold for acceptable product fire performance lies within the uncertainty range around the point estimate or single-value measurement of that performance.
- He identified the major factors that contribute to test uncertainty, including variations in the sample, the test method, the test environment, and the calibration of the instruments.

- The analyst needs to select a statistical approach that will accurately describe how uncertainty in each of these factors combines (or propagates) to produce combined uncertainty in the final measurement.
- He discussed both frequentist and Bayesian approaches and provided example calculations for both.
- He introduced the audience to ISO's "GUM" standard, which is the shorthand name for ISO JCGM (Joint Committee for Guides in Metrology) 100, *Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement*, an essential document for frequentist calculations of uncertainty, which is the approach used with nearly all uncertainty calculations. The GUM is accessible at http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf.
- For Bayesian analysis, he referred the audience to D.J. Lunn, A. Thomas, N. Best, and D. Spiegelhalter, "WinBUGS – A Bayesian Modelling Framework: Concepts, Structure and Extensibility," *Statistics and Computing*, volume 10 (2000), pp. 325-337.

Marc Janssens of Southwest Research Institute followed with two presentations – "Relevance of ASTM 2536 in Fire Safety Engineering Analysis" and "Precision and Uncertainty of Fire Tests – What's the Difference?"

- He described an example application using the cone calorimeter to develop input data for use in a fire dynamics model, either CFD or zone.
- ASTM E2536, *Standard Guide for Assessment of Measurement Uncertainty in Fire Tests*, is the principal ASTM reference for such an exercise.
- ASTM E2536 fully addresses measurement uncertainty but only partially (if at all) addresses uncertainty associated with the test specimen or the test procedure.
- Picking up on Guthrie's key step of selecting an appropriate statistical approach, Janssens illustrated the complex calculations required to estimate uncertainty more comprehensively in this example case.
- He noted that "seemingly small changes in the test conditions can have dramatic effects on the test results."
- In his second presentation, he explained the difference between uncertainty, which measures the magnitude of errors associated with a value, and precision, which focuses on variations between or within laboratories in repeated applications of a specified test method to a specified material.
- He then illustrated the calculation of both measures for a very simple example, which was the application of ASTM E691 to the total burning

time for 50 ml of 91% IPA alcohol in an empty tuna can taken from a specific brand and type of tuna.

Hershal Brewer of the International Accreditation Service provided the last of the presentations on basics in "Measurement Uncertainty for Fire Test Laboratories in the Accredited Environment Under ISO/IEC 17025:2005."

- ISO/IEC 17025, *General Requirements for the Competence of Testing and Calibration Laboratories*, is the standard that governs laboratory accreditation.
- Clause 5.4.6.2 of ISO/IEC 17025 requires laboratories to have and apply procedures to estimate uncertainty of measurement.
- ISO/IEC 17025 refers users to the GUM for methods to discharge its requirements. Brewer also cited ANSI/NCSL Z540-2-1997, which is the U.S. edition of the GUM, and NIST Technical Note 1297.
- He then walked through the application of these references to ASTM E84 tests.

The first paper in the STP is based on the fifth and final presentation in the introductory section of the symposium. John Hall's paper "Who Gets the Benefit of the Doubt from Uncertainty?" focuses less on the calculation and more on the framing and interpretation of uncertainty information, including the points in the decision-making process where imbalances in knowledge or in access can introduce biases in the decisions.

The next four papers in the STP were presented in the "Applications to Specific ASTM Fire Tests" section of the symposium.

- "Measurement Uncertainty in Fire Tests – A Fire Laboratory Point of View," by Javier Trevino and Rick Curkeet, provides a perspective on the way that measurement uncertainty rules are applied, simplified and sometimes declared non-applicable in practice.
- "Bench Tests for Characterizing the Thermophysical Properties of Type X Special Fire Resistant Gypsum Board Exposed to Fire," by Paul Shipp and Qiang Yu, is a detailed description of research conducted to address difficulties in estimating precision for ASTM E119, ASTM's most used standard fire test, relative to a specific product.
- "Measurement Uncertainty and Statistical Process Control for the Steiner Tunnel (UL 723, ASTM E84)," by John Resing and colleagues at Underwriters Laboratories, examines uncertainty measurement issues for ASTM E84, ASTM's second most used standard fire test.
- "Precision of the Cone Calorimeter and ICAL Test Methods," by Joe Urbas, examines uncertainty measurement issues for two of the relatively newer ASTM fire test methods, including the cone calorimeter,

which is probably the most used source for input data for fire modeling and calculation.

The next two papers in the STP were presented in the "Uncertainty in Engineering Calculations" section of the symposium.

- "Uncertainty in Fire Protection Engineering Design," by Morgan Hurley, walks through the uses of uncertainty in fire modeling and other fire protection engineering calculations. He references "Uncertainty," Chapter 5-4 in the 4th edition of the *SFPE Handbook of Fire Protection Engineering*. He also references the SFPE Engineering Guide G.06.2011, *Guidelines for Substantiating a Fire Model for a Given Application*, and ASTM E1355, *Predictive Capability of Fire Models*. Both guides are about verification, validation, and assessment of fire-related models.
- "Fire Pattern Repeatability: A Study in Uncertainty," by Daniel Madrzykowski and Charles Fleischman, begins with a review of uncertainties in calorimeter tests of various materials, then switches to a discussion of the change in uncertainties when test results are used instead as input data for computer model analysis and estimation.

The final four papers in the STP were presented in the "New Methods and Other Issues" section of the symposium.

- "In Search of Standard Reference Materials for ASTM E05 Fire Standards," by Norman Alvares and Harry Hasegawa, reviews the history of work in identifying, accessing and using standard reference materials to calibrate test instruments and test operators, providing better measurement and control of measurement uncertainty.
- "What Have We Learned About Uncertainty? Are We Still Playing with Fire?", by Ned Keltner, reviews a wide range of issues in thermal measurement, along with some candidate solutions and strategies to improve performance.
- "Heat Flux Measurements and Their Uncertainty in a Large-Scale Fire Test," by Cecilia Lam and Elizabeth Weckman, focuses on the contribution of heat flux gages to uncertainty in heat flux measurements.
- "Development of a Proposed ASTM Guide to Continued Applicability of Reports on Fire Test Standards," by Marcelo Hirschler and Timothy Earl, describes ongoing work on a proposed new ASTM guide that would translate uncertainty and variation considerations into guidance on the continued use of, or the need for new, fire test reports.

Also worthy of note is the one poster at the symposium, which was "Parametric Uncertainty of Moisture Content – Ignition Time Relationship for Cellulosic Materials," by Joseph Kahan and M. Ziefman of FM Global.

The intent of the symposium was to provide something useful for every member of the audience, regardless of their level of experience or their work responsibilities. For newer professionals, the symposium's front half was designed to give them a quick exposure to the basics and leads to major guidance documents and references that would allow them to be fully capable practitioners. For experienced professionals, the symposium's second half was designed to provide an overview of current activities designed to improve practice and identify the leaders of those efforts, so that interested audience members could become part of the change.

The intent of this STP is the same – helping the newer practitioners to build their skills quickly and helping the leading-edge researchers to attract additional minds and energy to their important work. If we are successful, we may soon be able to acknowledge the elephant in the room without flinching and set about bringing it under our control.

Many thanks are herein offered to all the speakers who made up this outstanding program; to the attendees, whose good questions and comments added value to the program; and to the several members of ASTM staff who made the symposium and this STP possible. It was a pleasure and a privilege to be part of this team.

John R. Hall, Jr.
Division Director
Fire Analysis and Research
National Fire Protection Association
Quincy, MA, USA

Reprinted from JTE, Vol. 40, No. 1
doi:10.1520/JTE103860
Available online at www.astm.org/JTE

John R. Hall, Jr.[1]

Who Gets the Benefit of the Doubt From Uncertainty?

ABSTRACT: Codes, standards and regulations require compliance with criteria stated as the results of tests or calculations. These results have associated uncertainty. This paper discusses different approaches to the use of uncertainty in the determination of compliance. In particular, most discussion addresses the many ways in which the decision-making protocol, intended to result in a determination of compliance, may shift the benefit of the doubt of the uncertainty between different interested parties.

KEYWORDS: uncertainty, fire test standard, fire risk assessment, statistical significance

Introduction

Nearly all of the technical guidance available on the subject of uncertainty has to do with uncertainty estimation techniques. Very little guidance or even discussion is provided on the subject of what to do with the uncertainty estimates. Once you begin to focus on the uses of uncertainty, it becomes more obvious whether and how the rules for using uncertainty are or are not favoring one interest over another. Therefore, before we can discuss who receives the benefit of the doubt, we need to spend some time considering context.

Suppose you have a basic decision tree, where the branching is done based on parameters that could be results of standard tests, estimates of relevant probabilities or likelihoods, or results of probabilistic and/or deterministic modeling. Every parameter has uncertainty attached to it.

For example, your decision problem might be how to prevent building structural collapse, principally through design of fire resistance protection for the structural elements. Your basic decision model might say that the fire challenge is a full floor burnout and the fire resistance is a set of specifications that

Manuscript received March 15, 2011; accepted for publication August 16, 2011; published online September 2011.
[1] National Fire Protection Association, 1 Batterymarch Park, Quincy, MA 02169-7471.

Cite as: Hall, Jr., J. R., "Who Gets the Benefit of the Doubt From Uncertainty?," *J. Test. Eval.*, Vol. 40, No. 1. doi:10.1520/JTE103860.

pass a fire resistance test on a target assembly. Traditional test-based decision-making would convert this single fire challenge to a set of test specifications and use that test to approve the fire resistance design. Traditional performance-based equivalency decision-making would convert this single fire challenge and the fire resistance design to a set of input parameters and other modeling specifications, and use the model to predict whether collapse would occur. In this traditional format, you may be able to address uncertainties internal to the test, but you have no obvious way of discussing, let alone addressing, the uncertainties in the steps that led up to the specifications you set for the test.

Unlike most of the papers in this symposium, the topic here is not how well we measure uncertainty, but how we incorporate uncertainty considerations into our decisions and what kinds of unanalyzed biases may be introduced in the process.

Types of Uncertainty

The first step in this paper is to separate types of uncertainty, in order to focus on those types where critical decisions occur outside the set of actions and calculations to which uncertainty assessment is normally applied. Start with the steps involved in using an ASTM E5 Standard test as the basis for evaluation; for example, to establish compliance with a code or regulation.

(1) The code must identify a *set of outcomes that define required safety*, such as a likelihood of fatal injury no higher than that associated with a reference or baseline condition.
(2) The code must translate the outcomes that define safety into a *set of measureable physical conditions such as measures of exposure by harmful conditions to protected targets* (people, property, the environment, etc.).
(3) The code or the user must *translate test outcomes into the same scales as those used in step 2*, which will probably require the specification of other conditions, such as the dimensions of the space into which fire will grow or the number, location and capabilities of exposed occupants.
(4) The code or the user must select one or more *fire challenges* that will collectively provide the basis for evaluating the test item.
(5) These fire challenges must be translated into *test specifications*.

Each of these steps involves choices and uncertainties, but none of them are primarily dependent on proper handling of variation in test results from test to test or from laboratory to laboratory, the kind of uncertainty addressed by precision and bias statements. Therefore, the choices and uncertainties at the center of this paper do not require or benefit from a detailed discussion of *aleatory uncertainty*, the uncertainty associated with randomness that is easiest to describe using known probability distributions and associated statistics, or systematic bias, which is more difficult to capture mathematically but has been extensively studied and addressed in standards.

Shown below are some of the types of uncertainty and variation that are central to the five steps in the decision-making process:

(1) *Uncertainties or mismatches in the test results as proxies for real-world outcomes*. The most familiar example of this is probably the correlation (or lack of correlation) between bench scale (or other scale of the test

procedure) and real scale. Scale-related variation can be and has been addressed mathematically, although it is not automatic for such variation to be explicitly addressed in the guidance on use of test results from ASTM E5 Standards. Much more complex is the variation associated with the use of observable, physical conditions and events as proxies for certain types of fire damage, which are the outcomes we really want to avoid. Temperatures and smoke concentrations in the neighborhood of a burning test specimen are only indirectly related to deaths, injuries and property damage. Deflections in a structural element sample do not automatically translate into building collapse and may not translate into an inability to continue using the building. In all such cases, many other uncertain and variable factors and events will come into play before the final outcomes do or do not occur. Our level of knowledge — the source of epistemic uncertainty — is usually quite limited as we attempt to link the test results to the outcomes of interest.

Furthermore, the selection of a level of harm that will constitute failure involves variation in judgments of acceptable risk. How much likelihood per year of how much harm will define failure? Opinions will differ. Therefore, this aspect of variation involves not only uncertainties in ability to predict outcomes but also in our agreement (or lack thereof) regarding the acceptability of outcomes once known.

(2) *Uncertainties and mismatches in the selected fire challenge(s) as proxies for the full range of fires that may expose or involve the product*. There is a wide range of possible fires that may challenge a product in a built environment. A standard test typically chooses one fire challenge or at best a controlled-growth fire challenge that is meant to represent a range of fire conditions. No matter how severe the test fire challenge is, a more severe fire challenge is always possible, with a certain likelihood.

We use standardized tests to represent the expected performance of a product against the range of fires that may occur where it is used. Our goal is to reduce risk to acceptable levels, and we cannot be sure we have accomplished that goal if our assessment procedures do not capture important parts of the total risk, associated with more severe fire challenges that we have not explicitly considered and do not understand.

Furthermore, in a systems design of a built environment, the likelihood of a fire challenge as severe as the test conditions or more severe than the test conditions may be significantly affected by a part of the design other than the tested product. For example, the likelihood of a full-floor burn out, the fire challenge used to assess structural elements, will be greatly affected by the use or non-use of sprinklers. By not explicitly tailoring our decision-making protocols to include all relevant design aspects and other conditions, we not only create the potential for unacceptable harm from very damaging fires that are more frequent or less well-handled than we thought, based on our limited tests. We also create the potential for needless expense through over-engineering our product to resist severe fires that are extremely unlikely to occur, thanks to other aspects of the systems design.

(3) *Uncertainties and mismatches in the use of a test specimen from a newly produced product as a proxy for the range of real products over the product's entire life cycle*. Test specimens are not subject to the performance degradations associated with age, poorly controlled production, poorly performed installation, poor or no maintenance, or any of the other events that predictably occur in the real-world application of the product. Some of these degradations can be addressed in testing. For example, when certain flame-resistance treatments for clothing were found to wash out after only a few washings, the test standards for flame resistance of clothing were modified to require washings of the test specimens before testing. However, for every one of these factors and conditions that are recognized and incorporated into the test standard, there are undoubtedly many more that are not recognized or not practical to incorporate into test modifications.

(4) *Uncertainties or mismatches in the selection of test results submitted as proxies for all the results of all the tests conducted*. It may be bad form to point out that fire testing can be subject to the same kind of "venue shopping" as other judicial or quasi-judicial forums, like courts, but in some circumstances, it is possible to collect a range of varied test results and submit only the ones — or the one — that fell into the compliant part of the range of variation. This practice is unethical and may be illegal, but that is not the same as saying that it does not occur or that our protocols for preventing or punishing such acts are well developed and effective.

Type I Versus Type II Error

Regardless of the type or source of uncertainty — whether it is routinely included in analysis of uncertainty or is normally overlooked, users of test results to assess compliance in the face of uncertainty are faced with the following pure strategies in interpreting test results:

(1) Compliance is assumed unless failure is by more than the uncertainty.

(2) Failure is assumed unless the margin of compliance is more than the uncertainty.

The first approach may routinely expose the public to unsafe conditions. The latter approach may impose unrealistic demands on manufacturers, because costs may rise exponentially as ever tighter tolerances are sought, and some margins of safety may be technically infeasible. And yet it is impossible to act so as to simultaneously assure that neither of these unacceptable conditions will occur.

The technical analysis of this dilemma can be set up as analysis of Type I versus Type II error for hypothesis testing of a null hypothesis that a measured quantity from the test results falls within the compliant range. (The alternative hypothesis is that the measured quantity does not fall within that range and in fact is different from it in a direction that implies a different conclusion about the tested product.) For purposes of this paper, it is worth spending some time on this issue because the first pure strategy above is tantamount to giving the

threatened occupants or property the benefit of the doubt of any uncertainties, while the second pure strategy is tantamount to giving the product manufacturers and sellers the benefit of the doubt.

If the null hypothesis specifies a single number, it may be fairly straightforward to construct a probability distribution (for Type I error), based on knowing the aleatory uncertainty, for the tested result around any particular number, such as the number that marks the transition from safety to failure. Then one can select the specified number for the null hypothesis to achieve any desired low probability that the product is really unsafe when its tested performance is graded as safe. The difference or ratio between the null hypothesis number (the tested result) and the number that marks safety is the safety margin or safety factor, and this approach is based on a rule that a product is to be judged unsafe unless it is proven safe.

It is, of course, possible to use the uncertainty the opposite way so that the product is judged safe unless it is proven unsafe.

None of this tells us anything about the size of the Type II error. We have designed our test criteria to set a low maximum on the likelihood that a product, tested as safe, is really unsafe. However, we have no idea how likely it is that a product, judged by test to be unsafe, is really safe. That likelihood could be quite high, depending on the size of the aleatory uncertainty, but also depending on whether the shape of the uncertainty probability distribution is the same around a true unsafe number as it is around a true safe number.

The best way to escape this dilemma is to introduce some quantification of the cost of different types and degrees of error, as is routinely done in Bayesian analysis. However, our state of knowledge regarding such cost functions is typically very limited, and the opinions of different parties, in the absence of a shared set of best data, will often vary widely. The end result is that explicit analysis using cost functions for error are extremely rare in fire testing and extremely controversial when proposed.

Is Any Test Result Truly Safe or Unsafe?

The discussion of Type I versus Type II error showed that even when the issues and the math are relatively simple and straightforward, there are non-obvious choices to be made that will have the effect of assigning more of the benefit of the doubt to one or the other group of interested parties. Now we can return to the five-step decision-making process and consider the many choices where the issues and the math are not at all simple or straightforward.

Any explicit treatment of the process of translating test results into physical conditions and then into safety outcomes will reveal so many choices and so many points of uncertainty that it will undercut any simple notion that *any* test result is perfectly safe or perfectly unsafe. It is far more likely that as test results slide from the less safe end of the spectrum to the safer end of the spectrum, the practical result is that the product will perform acceptably in a wider range of fires and under a wider range of other conditions and factors driving the outcomes. To the defender of safety or the defender of low-cost products, this may look like a family of slippery slopes, providing no obvious choice of "best"

threshold for grading test results. Those slippery slopes can easily become a reason for defenders of the testing and interpretation status quo to reject any suggestions for changes based on incorporation of previously unincorporated types of uncertainty.

For example, let us return to the testing of structural elements against a goal of preventing large losses of lives and property due to structural collapse. Structural collapse seems like the ultimate black-or-white event. With collapse, you have losses many orders of magnitude greater than without collapse. The damage if collapse occurs is so great that it can be difficult to impossible to accept any likelihood of collapse greater than zero. Unfortunately, that requires you to reject the evidence that such a goal is not achievable and that one can only choose between different low likelihoods of collapse.

Under those conditions, how would a supporter of the current test protocols respond to the suggestion that the current low level of likelihood of collapse will be preserved if we replace current requirements with a combination of high- (but not perfect-) reliability sprinklers and reduced fire resistance? Consideration of such a proposition requires a prior acceptance of the idea that current requirements do not completely eliminate the risk of collapse, because no strategy dependent on elements with less than perfect reliability can possibly provide performance equal to a system assumed to have perfect reliability.

At this point, there should be enough examples on the table to make the point that uncertainty issues arise at many points along the way of converting a goal to a standard test and rules for interpreting results of that test. It is not all, or even mostly, a matter of fully and properly addressing the aleatory uncertainty arising from variations in the conduct of the test itself. For this paper, all of that discussion was leading up to the real topic, which is how different parties may maneuver around this complicated landscape in order to pursue their differing interests — and what if anything can be done to improve the decisions that result.

Interests of Different Parties

In the discussion of Type I versus Type II error, there was a very brief reference to the idea that different interested parties have different interests and different concerns. That notion deserves to be examined in more detail. From an economics point of view, optimum decisions will result if decisions are made by a person, entity or process that fully reflects and fully responds to all the consequences of any decision. When a product is involved in a fire, the parties whose decisions shaped the product's fire performance do not automatically experience all of the consequences of harm arising from that fire. Legal arrangements intended to impose more of the consequences of harm on the parties whose choices contributed to the occurrence of the harm — starting with tort liability — have all kinds of gaps and rigidities that reduce their effectiveness, and there are huge costs associated with access to and use of these legal arrangements, costs which further distort the decision-making calculus of the decision-making parties.

This means that in any decision about choices and risk, the parties with the most influence on those choices are likely to be able to avoid exposure to many or most of the consequences of harm associated with the risks they have chosen.

Parties that are likely to avoid many or most of the consequences of harm have an incentive to ignore those consequences in making their decisions. They have an incentive to oppose regulations that seek to reduce harm by limiting the range of allowable product decisions. If regulations exist, they have an incentive to weaken those regulations in their implementation so that fewer choices are disallowed. They have an incentive to make compliance easier to achieve and to make non-compliance harder to detect, harder to prove, and harder to punish.

Opportunities to Influence the Process

Having established that different interested parties will have reason to want to influence the evaluation process so as to place more weight on their concerns and less weight on other concerns, it is useful to examine in more detail some of the points in the full five-step decision-making process where parties have an opportunity to exert that influence.

Setting Criteria for Acceptable Risk

The first opportunity is to argue that currently experienced risks must be acceptable because people have experienced them. The counter-argument is that people routinely accept risks only so long as they do not know a practical way to avoid them. This shifts the argument to the practicality of the approach embodied in the test standard.

The second opportunity is to argue that no risk should be regulated if it arises primarily or exclusively from choices and actions of the victim. In theory, this argument could be used to oppose any regulation. In practice, it tends to argue for excluding intentional fires from the argument (even though product fire performance may be very successful in mitigating the harm caused by an intentional fire and may even prevent a large share of attempts at fire-setting) or targeting only large-loss fires in public settings (even though fires that are caused by strangers to the victims are not limited to such large incidents).

The third opportunity is to argue that some risks are too small to worry about. The *de minimis* principle, stated simply, is that some level of risk is too low to justify our concern and attention [1]. This relatively self-evident general principle becomes problematic when you translate "too low" into specific criteria. If you chop up the fire problem associated with a product into enough distinct parts, you may well be able to argue that each of them is too small to worry about, even though the combination clearly is not that small.

The fourth opportunity is to argue that some risks are not the fault of the product. Tort liability works off formal legal principles setting the responsibilities of various parties when someone is harmed. The informal version is an argument that a product cannot be blamed for harm if anyone did anything stupid or wrong to contribute to the occurrence or severity of the fire. A typical severe fire starts and becomes severe as a result of many contributing factors, and it is a rare loss that does not involve some stupid or wrong act by someone — the victim or someone else — at a critical point. If you exclude any fires where the product's performance was not the whole story, you can exclude a great deal in the

calculation. This is how intentional fires and other fires that cannot be prevented by technology but can be mitigated if they occur come to be excluded altogether.

Linking Test Results to Outcomes

For a restrictive test, the opportunity is to identify cases where a product that fails the test — and in fact has inadequate performance as defined by the test (that is, fails not because of a large safety factor but because it does not perform up to the test) — is nevertheless safe in terms of the outcome criteria. For fire resistance to prevent building collapse, one could argue that some of the failure criteria for a test (such as deflections of structural elements) would not in practice result in harm to people, damage in need of repair, or the loss of the use of the building. For design of smoke alarms, one could argue that the smoke obscuration criteria would not in practice cause harm to people or lead to harm with high probability.

For a test that does little to restrict products, the opportunity is to minimize evidence that a product that passes the test — and in fact has acceptable performance as defined by the test — is nevertheless unsafe in terms of the outcome criteria. For spray-on fire resistance, one could argue that evidence of frequent loss of fire resistance due to poor application or routine stresses on the structural elements is not germane to an evaluation directed at test specimens.

The common element is that the interested parties are not arguing over the best decisions in the face of uncertainty but over the least disruptive and restrictive decisions that can be made to seem acceptable within the process.

Selecting the Fire Challenges

The opportunity here might be a multi-stage argument. First, narrow the focus of the discussion from attempts to address more than one design fire challenge or to at least explicitly examine the ability of the selected fire challenge(s) to adequately represent the entire range of possible fires to which a product might be exposed. The argument here would be an argument for less complex and less costly testing.

Second, argue for less severe fire test conditions based on difficulty of creating more severe conditions in the laboratory or even based on the dangers of more severe fires to lab personnel. The effect of this second set of arguments and related decisions on the representativeness of the test protocol will be less apparent if the participants have already abandoned any attempt to formally and explicitly examine representativeness.

Third, continue to nibble away at challenging fire conditions using arguments about the difficulty of reproducing such conditions in the lab (e.g., avoid the unusual challenges posed by fires in oddly configured concealed spaces by pleading an inability to set up the lab for routine testing under such conditions).

Setting Safety Factors or Safety Margins

If you have a well-defined probability distribution for measurement error, then there will be a natural tendency to build safety factors or safety margins around 95 % confidence intervals or some other well-established and widely used basis.

If you do not have a well-defined probability distribution to work with, then it is easier to fall back on rules of thumb or round numbers, but our knowledge in this area and our standard practices are pretty advanced, compared to our tools for dealing with other aspects of uncertainty.

However, there is considerable exposure to unacceptable and unintended risk if you do not look at Type II error. If you take a 95 % confidence interval at face value, you might wonder whether you are not accepting a process that permits one out of 20 products tested to be unsafe. In practice, it is not that simple. For one thing, the safety factor may well be calculated around a calculation that already includes other conservative or safety-factor-modified considerations.

There may be an unspoken assumption that true product fire performance has less variation than the fire test itself has. There may be an unspoken assumption, too, that manufacturers have introduced statistical quality control procedures that achieve far greater uniformity in product fire performance than was seen in the products tested and used to set the precision and bias characteristics of the test. In other words, there may be a confidence, warranted or unwarranted, probably not explicitly stated, that the true safety margins are better than one would believe from the tests.

Post-Purchase Factors

The opportunity here is to take all such factors off the table as not within the scope of a standard test protocol.

Is the Best Strategy Always to Reduce the Uncertainty?

In open debate on general principles where all interested parties are present, it is very difficult to sustain an argument that the benefit of the doubt for uncertainty should not be assigned to the people who will be harmed if fires occur. Behind closed doors in the detailed implementation of the decision-making protocols, it is very difficult to block all the opportunities for interested parties who will not experience the consequences to incrementally but systematically shift the benefit of the doubt back onto the people who will be harmed.

Both sides therefore have reason to favor efforts to reduce uncertainty so that the loser in the battle to assign benefit of the doubt does not lose that much. That seems like a simple motherhood and apple pie prescription, but like everything else in this paper, the reality is much more complicated. Consider two situations: The first is decision problems where the uncertainty cannot be substantially reduced by a modest investment in more tests or analysis or cannot be substantially reduced at all. The second is decision problems where interested parties who do not want to pay for more tests or analysis can make seemingly principled arguments in favor of ignoring large parts of the uncertainty.

What If the Uncertainty Cannot Be Reduced?

Some test procedures have a large uncertainty relative to the test output value. Acceptable precision can require far more replications than anyone is normally

willing to consider. For example, any test whose result has a binary form — such as ignition versus no ignition — will have a base variance that is solely a function of the underlying probability, p. The variance will be p times (1-p). If the underlying probability is 0.5, then the variance will be 0.25 and the standard deviation will be 0.5. If you would like a precision for the average as an estimate of the true underlying probability that is, say, 10 % of the estimated value (that is, two standard deviations are plus or minus 0.05), then you need 400 data points. If the underlying probability is 0.1, then the variance will be 0.09 and the standard deviation will be 0.3. A 10 % precision now requires 900 data points.

What If the Uncertainty Does Not Relate to the Number of Samples Tested?

Some calculations have to deal with a wide variation in target vulnerability. For example, toxic potency or hazard calculations routinely use a factor of a half or a whole order of magnitude to reflect variations in vulnerability of people to toxic insults. In this situation, you cannot compensate by running multiple calculations to get a better estimate of the average, because your safety goal is set in terms of the fraction of the target population you can protect. You can theoretically use analysis to try to balance the added cost of a better-performing product against the diminishing returns of safety delivered to ever more vulnerable people, but this Bayesian-style analysis requires more information, which is more expensive, and is likely to require more subjective estimates, which will hurt the credibility of the results with any interested parties who do not like the results.

Other calculations have to deal with significant issues of reliability. Deviations from design conditions or performance may mean unsatisfactory outcomes, but there may be technical limits or severe cost implications to attempts to improve reliability. The basic standard test is not normally configured to provide information on reliability but only on performance when it works. Therefore, running more tests is not an option here either; you need an entirely different protocol for gathering relevant information.

Summary of Results

(1) For anyone seeking to construct a decision-making protocol based on test results, there are many points — translation of general safety goals into specific acceptable outcomes, translation of specific acceptable outcomes into specific acceptable test results, specification of fire challenges, post-purchase factors, and setting safety factors or margins — where uncertainty arises and the rules of good practice do not lead everyone to the same choices.

(2) Different participants in the processes of designing tests and applying tests to decisions have different personal priorities, and the myriad types of uncertainty arising outside the traditional focus of random variation in testing provide myriad opportunities for participants to pursue those personal priorities.

(3) Better testing or repeated testing can reduce uncertainty to a more manageable size, but some types of uncertainty for some types of safety

goals and objectives cannot be reduced by better or repeated testing. Also, better or repeated testing costs more, and that increases the incentive to ignore large parts of the uncertainty rather than try to reduce them.

(4) There are asymmetries among the interested parties all over these decision-making processes. Simply put, one group of interested parties have the most direct control over the decisions made about the product or the building design, have considerable ability to avoid dealing with the consequences of failure of safety, and tend to be the ones paying the bills for the technical professionals who are supposed to sort all this out. Another group of interested parties have no direct control over the decisions made about the product or the building design, are at least as affected by the consequences of the failure of safety as by the costs of providing safety, and have more of the moral high ground in debates about general principles but lack the technical expertise to participate effectively in all the myriad implementation decisions and parameter specifications that will ultimately drive the risk actually experienced.

Conclusions: What is the Answer and What is the Strategy?

(1) The first step in solving a problem is acknowledging its existence.
(2) The best strategy is to be more explicit and more comprehensive in identifying and addressing the myriad choices and the myriad uncertainties at every step of the process.
(3) More information will not necessarily reduce the uncertainties, but it may bring the participants closer to a real consensus on their choices. In addition, the information necessary to improve the process may often be considerably less expensive — though possible no less complex — than would be more testing or more elaborate testing.
(4) Expanding the analysis and presentation of test results to cover Type II error and its implications for the conclusions should not be that difficult or that expensive; it draws on the same expertise and the same data currently being used to support fire test reports in the current format.
(5) Collecting and analyzing data on the range of fire challenges and the representativeness of specific design scenarios does not cost as much as conducting more lab tests.
(6) Summarizing available information on post-purchase factors and discussing whether and how to incorporate it into test procedures or the use of test results need not cost as much as conducting more lab tests.
(7) Developing a more explicit logic chain linking test results to outcomes to safety goals may be more expensive initially but can be done once in a general way and then used over and over on a wide range of fire tests.

Reference

[1] *De Minimis Risk*, C. Whipple, ed., Plenum Press, NY, 1987; cited in *SFPE Engineering Guide: Fire Risk Assessment*, Society of Fire Protection Engineers, Bethesda, MD, 2006.

Reprinted from JTE, Vol. 39, No. 6
doi:10.1520/JTE103911
Available online at www.astm.org/JTE

Javier O. Trevino[1] *and Rick Curkeet*[2]

Measurement Uncertainty in Fire Tests—A Fire Laboratory Point of View

ABSTRACT: Since the adoption of ISO/IEC 17025, testing laboratories have been required to perform Measurement Uncertainty analysis for the tests within their scope. Four points of recurring debate are discussed: (1) The variability in fire test results due to unforeseen/uncontrolled variables is generally far greater than the measurement uncertainty of the result. (2) It is important not to confuse "measurement uncertainty" (MU) with "precision" of results. MU has a very specific meaning as used in ISO/IEC 17025, ISO/IEC Guide 98-3 Guide to the Expression of Uncertainty in Measurement (GUM) and ISO Guide 99 International vocabulary of metrology—Basic and general concepts and associated terms (VIM). (3) An uncertainty result is not used to justify passing or failing a product with results very near the pass/fail limit. Where the measured result is subject to a measurement uncertainty evaluation and reporting, compliance limits may or may not require extending the test result by the MU value in making a compliance determination. (4) ISO/IEC 17025 specifically exempts standards that specify limits on sources of uncertainty and specify the form of reporting from a required MU statement. This makes uncertainty estimates inapplicable to those fire tests.

KEYWORDS: fire testing, fire calorimetry, fire resistance, flame spread, steiner tunnel, furnace, time-temperature curve, heat release rate, HRR, measurement uncertainty, variability

Introduction

In the past few years laboratories that conduct fire resistance tests of building assemblies have updated procedures and policies to conform to the requirements of ISO/IEC 17025 [1]. These procedures are reviewed and monitored by

Manuscript received April 13, 2011; accepted for publication June 30, 2011; published online August 2011.

[1] Priest and Associates Consulting, San Antonio, TX 78232, e-mail: javier.trevino@priestassociates.com
[2] Intertek Testing Services, Middleton, WI 53562, e-mail: rick.curkeet@intertek.com

Cite as: Trevino, J. O. and Curkeet, R., "Measurement Uncertainty in Fire Tests—A Fire Laboratory Point of View," *J. Test. Eval.*, Vol. 39, No. 6. doi:10.1520/JTE103911.

accrediting bodies on a global basis. An important requirement of ISO/IEC 17025 is for laboratories to report measurement uncertainty (MU) when necessary to allow for proper interpretation of test results.

First, one must understand what measurement uncertainty is and is not. The term "uncertainty" is defined in the *International Vocabulary of Basic and General Terms in Metrology* (VIM) as follows: uncertainty (of measurement)—parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

MU is only applicable to results of quantitative numerical measurement and is expressed as a plus/minus range for a specific confidence interval. This provides the user of the measurement result with a clear indication of the potential difference between the value of the measurement reported and what the *true* value of the measured property might reasonably be. For example if a reported result for the weight of a gold bar is 1.0532 kg ± 0.0001 kg at 99 % confidence, the user of the report will know that the true weight lies within a range from 1.0531 to 1.0533 with only a 1 % chance that its may be outside this range.

MU does not apply to test results that are not a quantitative property of the measurand. Since many test procedures are essentially qualitative in nature, MU is not required in reporting results of many types of tests. When a specimen is exposed to a certain condition, the product either passes or fails the test based on its condition after the required exposure. There is a measurement uncertainty associated with exposure conditions. However, this uncertainty cannot be directly related to the result of the test. In fact, such uncertainties in exposure conditions for most tests of this nature are negligible when compared to the effects of numerous uncontrolled variables that cannot be readily quantified. It is usually quite easy to determine and report the uncertainty of the actual exposure condition. While this information would provide the reader of the report with assurance that the test was conducted correctly, there is no direct relationship to the potential deviation disclosed and the end result of the test.

Misunderstandings

The requirement for providing Measurement Uncertainty (MU) in test reports started nearly a decade ago when ISO/IEC 17025 adopted the mandate to require testing laboratories to calculate or estimate MU for the standards within the laboratories Scope of Accreditation. At first, auditors required that at the very least, calibration records for equipment must contain a statement of uncertainty. Gradually, auditors required the laboratories to provide an uncertainty budget for each test method/standard within their scope. Ultimately, auditors began requiring full uncertainty analysis for individual test report files and if applicable in the test report itself. Buried beneath all of this was the fact that in many cases MU is not required to be reported unless the client (test sponsor) demanded the MU to be reported.

Misunderstandings arose when the auditor's understanding of MU differed from the lab's understanding. From the laboratory point of view, first order MU analysis was sufficient to fulfill the requirements. Specifically, if a test involved three instruments to calculate a result, and if the result had a clearly defined

mathematical relationship for the instrument values, then a first order MU calculation was clearly understood. However, as auditors gained confidence in the lab's ability to conduct MU analysis, the baby steps were sometimes increased to establish more precise MU calculations. In specific cases, some auditors required what some might consider second order variables to be considered in the MU analysis.

The Case of Calorimetry (Heat Release Rate)

For example, when measuring Heat Release Rate (HRR), the three variables are duct gas temperature, duct oxygen concentration, and differential pressure across a bi-directional probe. These are all first order variables which have a clear mathematical relationship when calculating HRR

$$\dot{q} = \dot{m}\, E 1.10 \frac{X^{\circ}_{O2} - X_{O2}}{1.105 - 1.5X_{O2}} \tag{1}$$

$$\dot{m} = C\sqrt{\frac{\bullet P}{T}} \tag{2}$$

The three variables are X (gas concentration), P (pressure), and T (temperature). E is the fuel value of the item under test. E is the heat produced per kg of O_2 consumed by the fire— not to be confused with heat of combustion which is heat produced per kg of material burned. It is well accepted that E has a value of 13.1 ±5% MJ/kg for most common materials to be tested or burned under a calorimetry hood. When burning propane, E has a specific value of 12.54 MJ/kg. The standards require that high purity propane be used to conduct the calibration burn. It has been calculated that commercial propane (with additives of various other fuels) has a heat of combustion (Hc) within 1% of high purity propane (internal Intertek document).

C is the calibration factor for the specific hood design which *can* be calculated (22.1 times Area of duct) but is actually measured during a calibration burn. C is measured by burning propane fuel at a fixed rate for a fixed duration so that one could calculate the theoretical heat produced (THRth = mass of fuel burned times Hc for propane which is 46.54 MJ/kg), and compare that value to the measured value (THRmsr) with a fixed initial value for C. One then adjusts C (no more than 20%) so that THRmsr matches THRth.

It can be argued that X, P, and T are first order variables. From the lab point of view these are the dynamic variables defined in the standard for calculating HRR. However, E and Hc are second order variables, i.e., E and Hc are fixed constants, each with a very small uncertainty value.

It is clear that the dynamic variables are constantly changing and each time step has a unique set of values for calculating HRR. What is not considered is that each first order measurement is time shifted (Fig. 1).

During a test or calibration burn, at any given moment, the pressure and temperature change almost instantaneously with respect to the fire, while the oxygen depleted air has to flow through tubing, into an analyzer. No matter

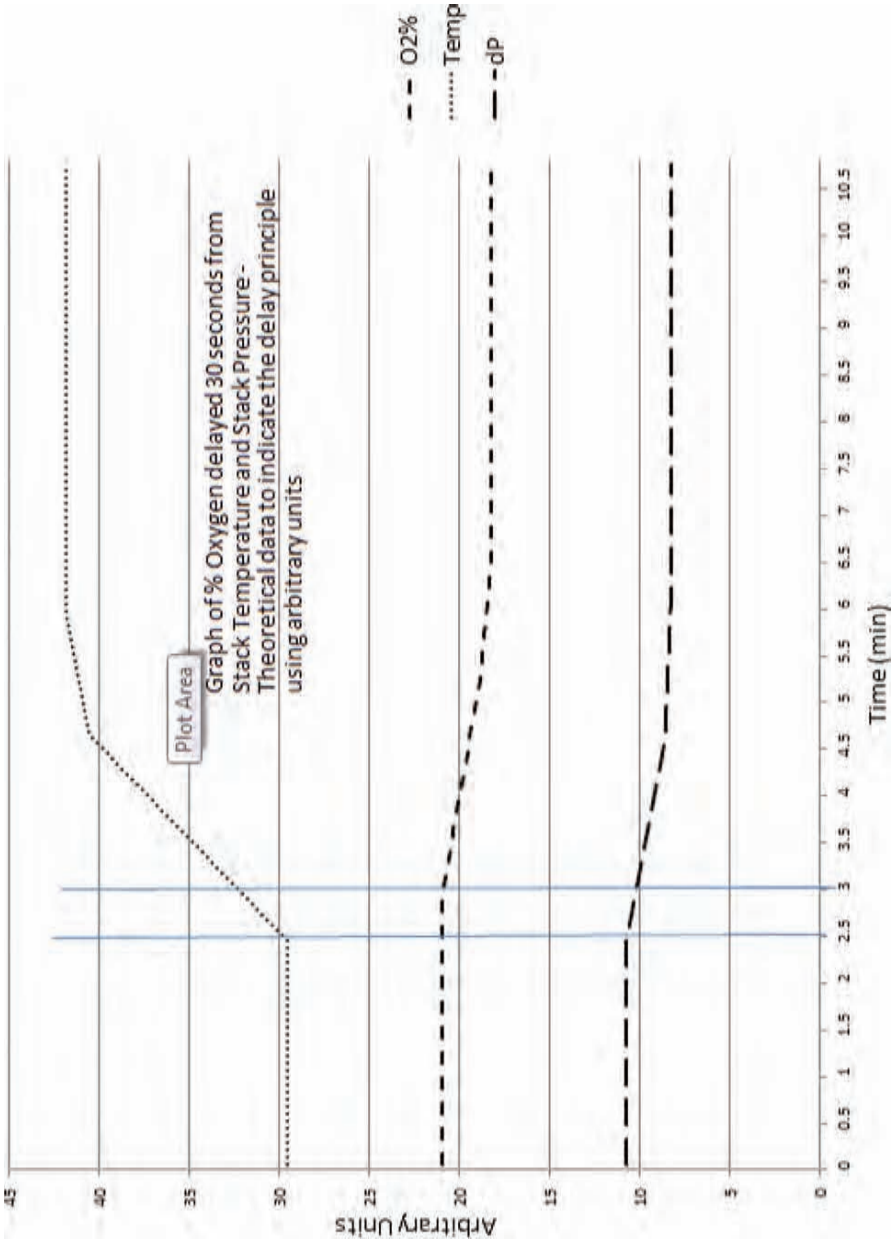


FIG. 1—*Time Shifted O_2 Signal versus Stack Pressure and Stack Temperature.*

how one adjusts the data to account for the time shifted data, it is always wrong. If the variables had a zero time constant (delay time), this would not be the case, i.e., one could adjust the data accurately. However, each dynamic variable has specific time constants, i.e., it takes a certain amount of time for a step change in value to be represented by the instrument. The accepted norm is to adjust the oxygen data temporally to account for 90 % of the time it takes to achieve a 100 % reading in a stepwise change. To complicate things further, each magnitude in step change has its own unique time constant. A step change of 1 % oxygen change may have a quicker response time than a 4 % step change in oxygen. So there is no "real" value for the time constant. It lies within a (fortunately) narrow band.

So, what is the first order MU for HRR? How does it compare to a second order estimate? How does it compare to actual data (Table 1)?.

When conducting Calorimetry (HRR) tests on interior finishes, furniture, or mattresses, one can actually "see" that the variability in fire testing of identical products depends more on the product and fire dynamics than the MU of the Calorimetry device.

During a HRR test, the reaction to fire depends on many variables. Some of the first order variables are: ignitability, burner position, random laboratory drafts, conditioning, fire induced drafts, lab geometry, specimen mounting, specimen chemistry, specimen component variables (many), lamination uniformity, and a myriad other specimen variables. Second order variables may be room material thermal conductivity, humidity, altitude, barometric pressure, and another myriad of other environmental factors which minimally affect the ignition or burning process.

The most obvious effect is how flames spread over a surface which ultimately affects the HRR curve. Consider an NFPA 286 room burn in which the 40 kW burner fire first heats the specimen surface. If ignition occurs the time to ignition can vary somewhat depending on burner position, flame induced drafts, chemistry, texture of specimen surface, etc. Sometimes, melting can move material away from the burner flames such that the flames do not directly impinge on the specimen and ignition never occurs. Once ignition occurs, the surface flame then begins to travel upwards in the test corner. This in itself has variability but is the most consistent parameter (based purely on observation of many hundreds of room corner fire tests). The amount of upward travel will result in an initial peak HRR if the surface flames begin to extinguish due to material no longer burning.

TABLE 1—*MU for 160 kW Fire is approximately 9.8 kW or 6% using modern calorimetry equipment.*

Variable	Value	MU Budget	MU Contribution	HRR MU
E (MJ/kg-O2)	13.1	0.68	0.0042	
P (Pa)	300	1.74	0.000234	
T (K)	472	0.55	0.000047	
O2 (fraction)	0.2058	0.0002	0.002513	
HRR (kW)	160	–	–	9.8 kW (6%)

However, if the flames reach the ceiling, a whole new onslaught of variables begin to take effect. Once the flames reach the ceiling, horizontal flame spread may or may not occur, but if it does, the effect depends on fire induced drafts, extinguishment of material in the corner, pieces falling off the corner or walls and melting or delamination of wall materials. Without going further into the test sequence (i.e., the 160 kW exposure), one can clearly see that the HRR curve can vary greatly depending on many specimen specific effects.

Similar arguments can be made for furniture or mattress burns. Consider the HRR graphs in Figs. 2, 3, and 4 of three identical mattresses burned sequentially in accordance with 16 CFR 1633. Sample A, Sample B, and Sample C were each burned in triplicate and represent different mattress designs.

Notice that in two cases, a large peak HRR for one burn while the other two burned at low HRR values. This was due to the flames spreading toward a vulnerable spot on the mattress and ultimately burning inside the mattress. Not only is the peak HRR different significantly, but the shape and time to peak HRR differs greatly. This is due to flames spreading differently in each case.

In the third set of mattresses (Sample C), the peak HRR varies and the time to peak HRR shifts. In this case, the flames spread to the vulnerable area from different directions and at different rates.

So, this is a clear example of immense variability in a test result due to the different ways a specimen reacts to an ignition source. However, the MU of the apparatus is small compared to the variability of the data.

So what is the MU for each test? Or what is the MU for that mattress based on statistics of three burns? The second question is invalid. One cannot confuse MU with test variability.

The Case for Fire Resistance Tests

Fire resistance testing of building assemblies and components is fundamentally a qualitative test where the assembly is exposed to a prescribed condition for a specific period and observations are made of its condition to determine whether failure occurs. These observations include breaching of the assembly, flaming of the unexposed face, temperature increase of the unexposed surface and the ability of the assembly to carry superimposed loads or high temperature of structural elements. The exposure conditions specify the furnace time and temperature relationship in the form of a standard curve. The accuracy of the exposure in meeting that specified is determined by integration of the area under the actual exposure Time-Temperature curve and comparing it to the integrated area under the standard curve. Thus, the only quantifiable uncertainties in the fire tests are related to the uncertainty in the measurement of the furnace temperature and the measurement of elapsed time from the start of the test as well as the uncertainty of load and unexposed face temperature measurements.

The furnace exposure uncertainty can be readily analyzed empirically by standard mathematical procedures as detailed in the ISO Guide to the Expression of Uncertainty in Measurement. Given the uncertainty of Special Limits of Error Type-K thermocouples of $\pm 2\ ^{\circ}F$ or 0.4 % of reading and an uncertainty of ± 0.1 min in elapsed time, the measurement uncertainty in the degree-minute

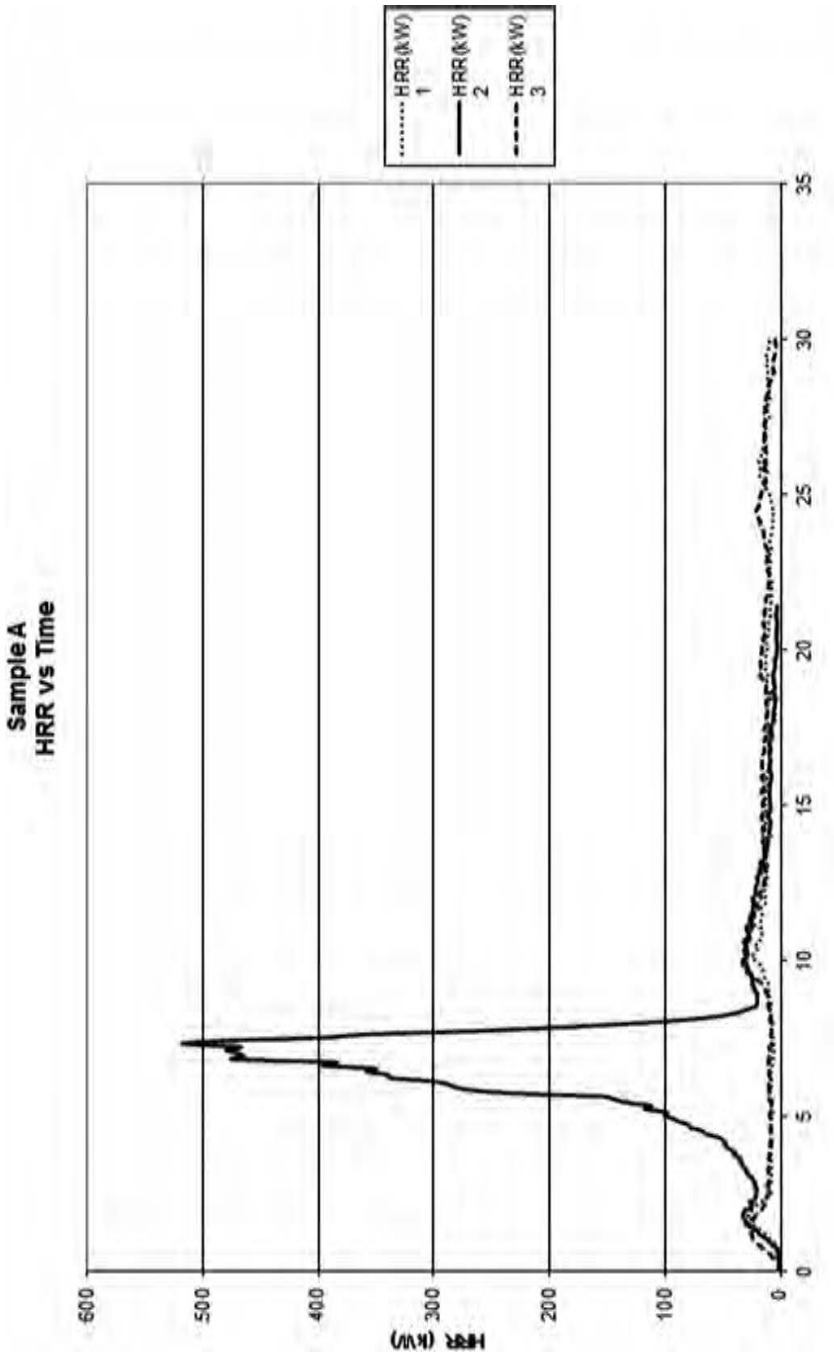


FIG. 2—*Mattress A Burned in Triplicate in accordance with 16 CFR 1633.*

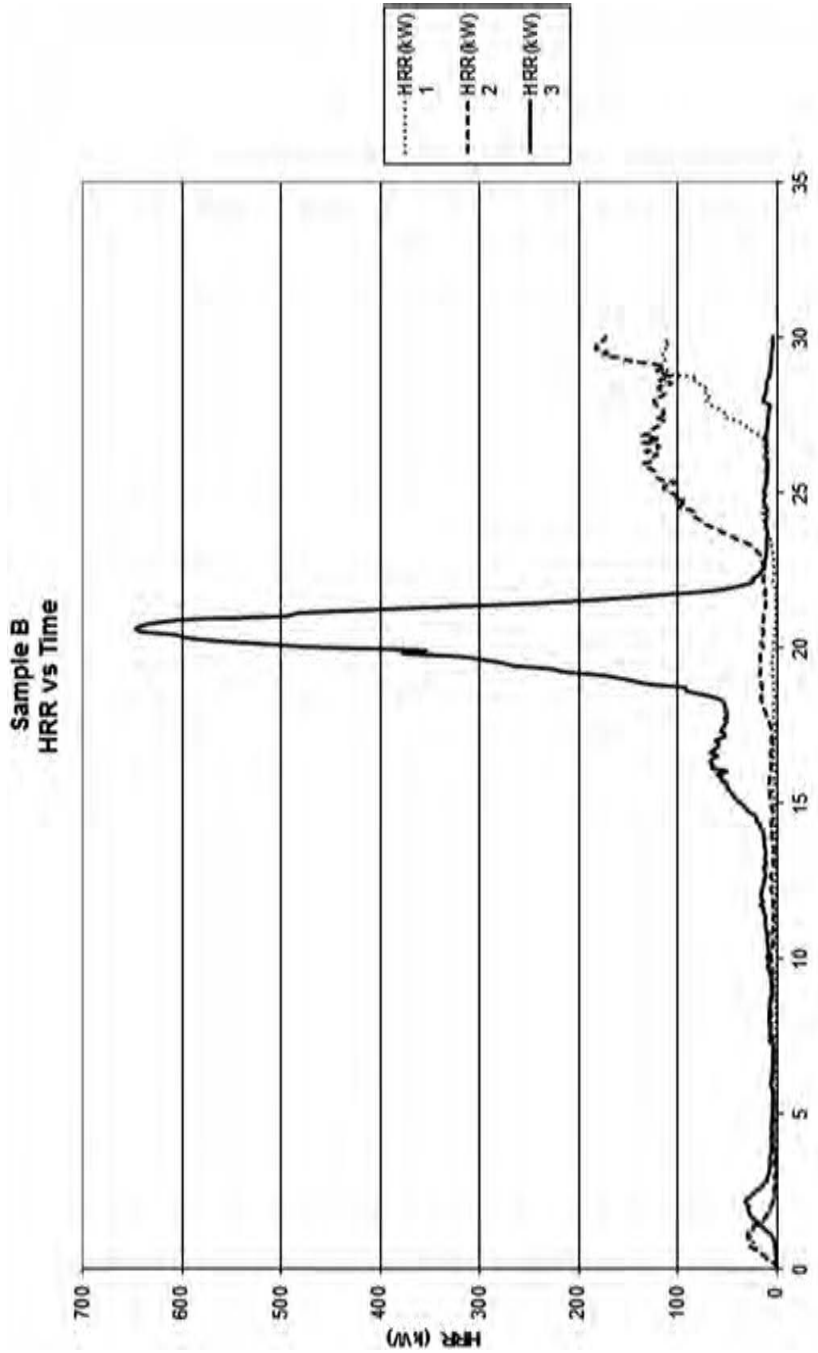


FIG. 3—*Mattress B Burned in Triplicate in accordance with 16 CFR 1633.*

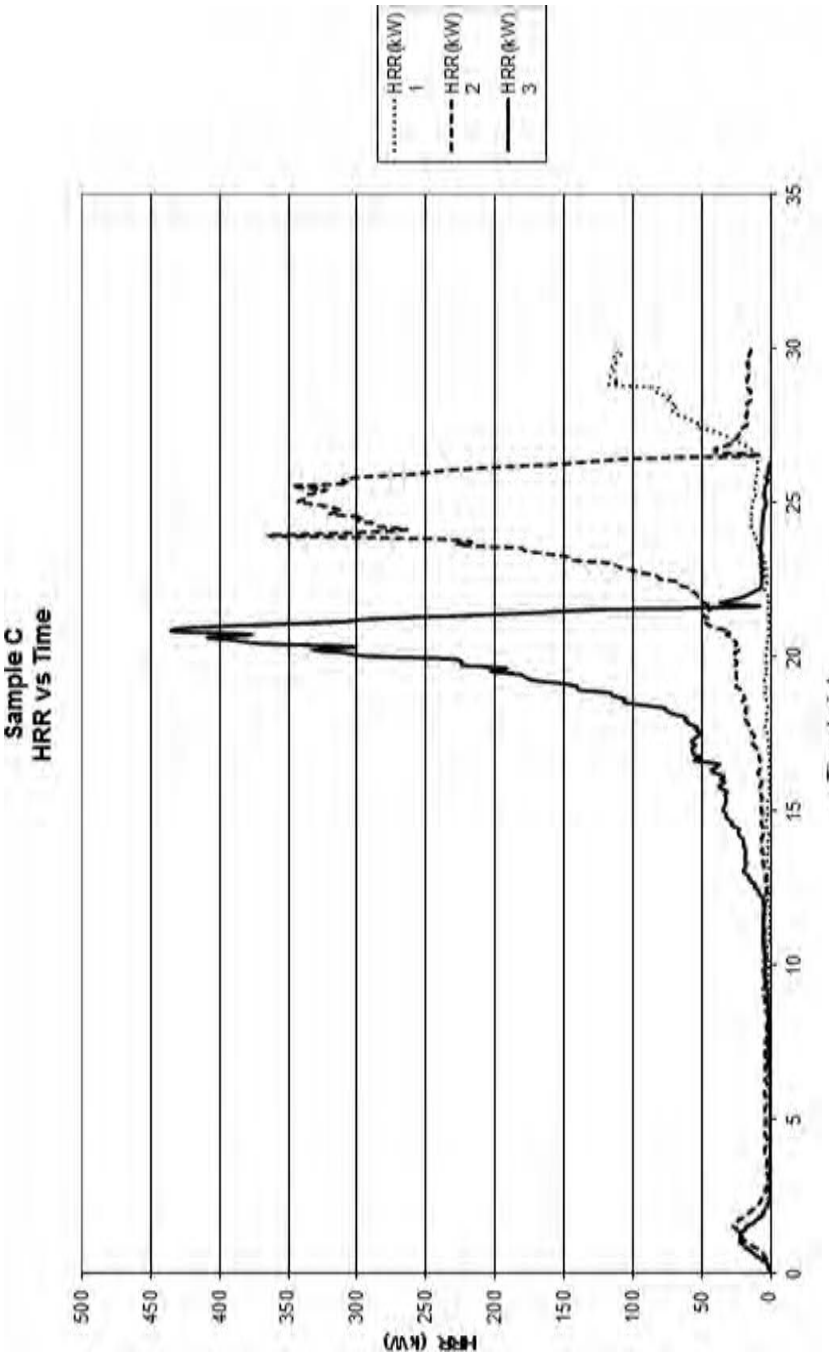


FIG. 4—*Mattress C Burned in Triplicate in accordance with 16 CFR 1633.*

exposure is typically in the range of ±3% for shorter tests and about ±1% for tests lasting several hours. This uncertainty is not significantly greater even if the temperature uncertainty is 5 fold larger (±2% of reading). Since the fire test standards allow deviations ranging from ±10% to ±5% depending on test duration, the MU associated with the temperature and time measurements is not particularly significant.

In reality, the area under the curve alone is not a good indicator of the equivalence of the exposure. For example, an assembly with a large combustible content can force the furnace temperature well above the standard curve early in a test requiring the operator to run below the curve later in the test to achieve the required total area. This can result in a very different actual exposure than a test that closely follows the standard curve from start to finish. In addition, the furnace design and type of burners used can substantially affect the oxygen concentrations within the furnace and this, in turn can affect the heat release from burning assembly materials.

The design and type of thermocouples used to measure furnace temperature also affect the apparent time-temperature exposure. For example the low mass "plate thermometers" specified in EN and ISO fire test standards have a faster response time than thermocouples in heavy metal or ceramic shields specified in ASTM E119. It has been recognized that the relatively long time constant of the E119 type thermocouple shield increases the severity of the test relative to tests run using faster response temperature measurement devices. However, the effect of this difference diminishes in longer duration tests due to the large amount of area under the curve accumulated during the portion of the test where the rate of temperature change is small. Still it is understood that different furnaces may produce significantly different time-temperature curve profiles even though the area under the curve can be virtually identical.

Measurement of superimposed loads and surface temperatures also have readily quantifiable MU, but these variations cannot be readily related to variation in the assembly's passing or failing a test. For example a ±2 °F uncertainty in surface temperature measurement thermocouple is likely to be insignificant compared to the potential variation in local surface temperatures of the assembly under test. Simply put, while we can have confidence that the uncertainty of the actual temperature of the thermocouple junction is ±2° F, the variability in what that temperature actually is that results from variation in the fire exposure, assembly materials, thermal conduction, effectiveness of the insulation pad, etc., is far larger.

Fire resistance test results are reported in terms of exposure duration, [3/4]-h, 1-h, 2-h, 3-h, etc., which to some appears to be a numerical performance measurement and thus subject to some uncertainty statement. In reality, fire resistance ratings are not numerical measurements but, rather, an identification of the exposure condition an assembly has passed. Because a standard fire test is terminated when the desired exposure duration is met, there is no specific information on how long the assembly might have continued to perform before a failure occurred. Thus, we do not know if an assembly rated at 2-h (120 min) might have failed at 122 min or 200 min. This does not mean that the test result has an uncertainty that can be quantified. Fire resistance test results are not numerical measures of performance and thus no uncertainty statement can be made.

There is a need to establish the repeatability and reproducibility (r&R) of fire resistance test results. However, this can only be done by means other than measurement uncertainty analysis. For this type of test method there are well established methods of evaluating repeatability (within a single lab) and reproducibility (between labs). This process requires carefully designed experiments and cooperation of a statistically significant number of laboratories. For example, data produced on a specific assembly design by 6 laboratories that each repeat the test 3 times would produce 18 data points that should be sufficient to make statistically valid conclusions about the test method's r&R at least as it relates to the assembly tested. Additional similar test programs on a wide range of assembly types could lead to a general statement of the test's r&R.

It should be noted that results of this type of experiment include evaluation of variability in results that is influenced by factors well beyond measurement uncertainty. For any given test specimen design there is some inherent variability that arises from numerous sources. One would expect results of a properly conducted evaluation process to produce results that are normally distributed and the standard deviation would thus be the primary descriptor of the variability. Measurement uncertainty in the test process would clearly be only one contributor to the overall variability and most likely its effect would be trivial compared to sources such as variation in materials, construction quality and differences in design of test furnaces. This issue is recognized in the Precision and Bias statement in the current edition of ASTM E119-10b [2] as follows:

11. Precision and bias

11.1 No comprehensive test program has been conducted to develop data on which to derive statistical measures of repeatability (within-laboratory variability) and reproducibility (among-laboratory variability). The limited data indicate that there is a degree of repeatability and reproducibility for some types of assemblies. Results depend on factors such as the type of assembly and materials being tested, the characteristics of the furnace, the type and level of applied load, the nature of the boundary conditions (restraint and end fixity), and details of workmanship during assembly.

It is clear that laboratories currently conducting fire resistance testing (and in fact many other types of tests similar in nature) do not have sufficient information upon which to base statements regarding the potential variability or uncertainty of their test results. It is also clear that the method required to establish a basis for such statements is to organize and carry out a statistically valid inter-laboratory study. Such a program could require 100's of full scale fire tests on a variety of product or assembly types. Given that these tests cost in the range of $5000 to $20 000 each, this is a very expensive proposition and it is the large cost that has been the primary roadblock to carrying out such programs. In considering whether or not funding for such a program can be raised and from where, the first question is probably "how badly does this need to be done?" In order to answer this question, one must ask many more questions. For example:

1. Is there an indication that variable fire test results have resulted in or contributed to performance failures in the field?

2. Are there indications that assemblies or products that pass in one lab routinely fail in others?
3. Do laboratories that run these tests have data or other information that indicates the degree of reproducibility of their tests? Can this data be shared within the fire testing community?
4. Can the industry organize proficiency test programs that will begin to build a database that could eventually provide a sound basis for determining r&R of these test methods?

The first round of such a program, organized by the North American Fire Test Laboratories Consortium (NAFTLC), was completed in 2008 and a report was presented by NIST at the Fifth International Conference on Structures in Fire (SiF'08) [4].

This proficiency test program ultimately included 10 North American labs and 5 Japanese labs. The results showed that in all labs the assembly achieved the 1-h design rating and the failure mode was similar, either exceeding the maximum average temperature rise or individual temperature limit. In all but one lab these two failure modes occurred within 1 min of each other. The average time to failure was 65.8 min with a 95 % confidence interval of ±5 min (Table 2).

We can conclude from this that, in spite of the variables associated with test furnace designs and all the variables involved in construction of assemblies,

TABLE 2—*North American Fire Test Laboratory Consortium ASTM E119 Proficiency Test Program Results for 10 NA labs and 5 Japanese Labs on 1 h steel stud and gypsum wall board non-bearing wall.*

Laboratory	Fire Resistance Rating	Time to First Failure, Minutes	Failed Thermocouple Reading	Other TC's Failing within 1 minute
NA-1	1 h	64	average	TC3, 5, 7
NA-2	1 h	62.8	TC3	average
NA-3	1 h	66.8	average	TC4, 6, 8
NA-4	1 h	67.5	average	TC3, 6, 8
NA-5	1 h	65.8	average	TC6
NA-6	1 h	70	TC3	TC4, 5, 6, 7, average
NA-7	1 h	60.6	TC7	TC3, 5, 6, average
NA-8	1 h	61.9	average	TC3, 4, 6
NA-9	1 h	65.8	TC5	none
NA-10	1 h	65	Average, TC5	TC4, 6, 7
J-1	1 h	67.7	TC7	TC6, average
J-2	1 h	67.3	TC7	TC6, average
J-3	1 h	66	TC6	average
J-4	1 h	65.5	average	TC3, 4, 5, 6, 7
J-5	1 h	68	Average, TC6	TC4, 5, 7
J-6	1 h	68	TC7	Average
average	1 h	65.8 ± 5 (95% C.L)	Average, TC3, 6, 7	–

laboratory conditions and conduct of tests, the reproducibility, at least for the specific design tested, is reasonably good. Repeatability, while not evaluated in this study, is virtually always better than reproducibility. Given all the potential sources of variability in large scale fire resistance tests, it is reasonable to conclude that actual measurement uncertainty is unlikely to be a significant contributor to the variability in test results.

The Case of the Steiner Tunnel

Consider the simplicity of the Steiner Tunnel Test. The instruments involved are a linear transducer for measuring flame spread; and a photocell system for measuring smoke opacity. Each instrument, by itself has a relatively small uncertainty, but things get complicated when one considers the operation of a Steiner Tunnel Test.

First, one must conduct an airflow "calibration". The tunnel blower is adjusted to produce a 240 ± 5 fpm average airflow velocity with all turbulence bricks removed. A vane airstream straightener is placed in the tunnel to produce even laminar flow at the measuring point near the end of the tunnel. This is typically done with a hot wire anemometer or a wind vane anemometer, each with a relatively small measurement uncertainty value. We must state that the damper which controls the air velocity is connected to either a pneumatic controller or a PID controller to maintain a specific differential pressure (vacuum) near the inlet of the tunnel. The tuning of these devices is critical for fast response when a fire is spreading along the tunnel. As the air expands due to heat, the pressure at the inlet changes and the system must respond to maintain the pressure as constant as possible. Most systems respond within 5 to 10 s to a stepwise change in tunnel pressure. During the adjustment period, the air velocity in the tunnel is no longer 240 fpm. Also consider the oscillation of the PID response after the adjustment is made. As stated before, the tuning is critical to minimize the oscillations, and respond quickly to changes in tunnel pressure.

Next, one conducts a flame spread "calibration" using red oak flooring. The flooring is conditioned to achieve a specific moisture content of 7 ± 0.5 %. After conditioning and preheating of the tunnel, the material is placed on the tunnel ledges. Note however that the turbulence bricks are placed in their specific locations. Without the bricks, the wood does not spread flame readily due to the oxygen starved upper layer zone. The bricks create turbulence to mix inlet air sufficient for ignition of downstream fuel. The turbulence is critical to achieve the required "calibration" parameters which state that the flames must spread to the end of the tunnel within a specific period of time – 5.5 min ± 15 s. One sets the fuel flow to an initial value of approximately 88 kW, and conducts the test. The fuel input is then adjusted on each run to achieve the calibration requirement. It may take 3 to 5 runs to finally achieve calibration.

Now, having established an initial calibration of the tunnel, one documents the flame spread and smoke developed (SD) areas for that test. Every 200 tests, the tunnel must be re-calibrated. The resulting SD areas are then averaged over the last 5 calibrations for use in calculating Smoke Developed index (SDI) values for specimen tests.

Flame Spread Index (FSI) and SDI results are then rounded to the nearest 5. If the SDI is greater than 200, then the SDI is rounded to the nearest 50.

Now, a first order MU can be made on the actual flame spread and smoke measurements based solely on the instrument uncertainties; however, since the standard requires rounding, the implied uncertainty of the FSI is 5 and SDI is 5 or 50 depending on the result. This comes directly from section 5.4.6 of ISO/IEC 17025

- 5.4.6 Estimation uncertainty of measurement
  - 5.4.6.2.
- NOTE 2: In those cases where a well-recognized test method specifies limits to the values of the major sources of uncertainty of measurement and specifies the form of presentation of calculated results, the laboratory is considered to have satisfied this clause by following the reporting instructions.

Considering the arguments made about variability in HRR in combustion calorimetry tests above, similar arguments can be made for variability in smoke and flame spread results of actual product tests.

A case in point are the reproducibility and repeatability (r&R) results published in ASTM E84 based on a Round Robin (RR) of 11 laboratories, 6 materials, and 4 replicates of each material. The results reveal that for most of the products tested, repeatability (within lab variability, standard deviation) of the FSI was within the rounding value (except for Douglas Fir plywood), and the reproducibility (among lab variability, standard deviation) was within the rounding value except for untreated and FR treated Douglas Fir Plywood.

The SDI values were not published due to the discovery that some labs had smoke/photocell systems different than what the standard described. It is assumed that the results were not very good. In fact, a new study is being conducted to use heptane as a smoke source for calibrating the tunnel SD area instead of red oak. This implies that it is well known that calibration using red oak is producing inconsistent reproducibility results among labs for a given material.

Conclusions

The fire laboratory community has responded to MU reporting in a relatively unorganized manner. Some labs are reporting uncertainty of instruments only, while others are presenting “attempted” uncertainties in fire test results. This is partially due to auditors being inconsistent in their understanding and interpretation of uncertainty compliance requirements and partially due to the lab’s interpretation of the requirements.

There has been debate on the need for uncertainty analysis for fire laboratories. Until the adoption of ISO/IEC 17025, fire testing laboratories did not typically perform uncertainty analysis for compliance work. However, for research testing, uncertainty analysis is typically required by the end user, especially when the data is used as inputs for fire modeling.

An understanding of the primary sources of uncertainty does little to reduce the overall variability in results. For fire models, the overall uncertainty analysis

of the outputs will lack the input uncertainty needed for a meaningful assessment of the result. For compliance assessment, it means that standard-compliant uncertainty analysis can do little to improve the precision of the estimates or the quality of the evaluation.

"Precision" is expressed in terms of repeatability and reproducibility and includes MU as well as many other sources of variability. There are problems with either a statistical approach or a calculated approach to estimating precision. In many cases, a fire test result lacks a clear mathematical relationship tying all measurements and other variables together needed to perform a traditional uncertainty calculation. Precision determination in terms of repeatability and reproducibility is possible through Interlaboratory Studies, but often requires an unaffordable large number of replications [3].

For variables used in models, a trained engineer may be able to perform an estimate of uncertainty with engineering judgment or experience, but for a compliance assessment, a meaningful measurement uncertainty estimate will be poorly understood, and will typically account for a minor fraction of the variability observed in actual test results.

Most standards are silent on the issue of using MU to determine if a product passes or fails a test. It is, however, quite common in product safety standards to recognize the variability, both due to MU and other sources of variability, of test results and set limits that are conservative to assure that MU and precision issues will not result in a compliant product exceeding a real critical limit. However, as a risk management tool, uncertainty estimates as well as r&R precision information can be used to aid in the assessment of a safety result. If variability of results is generally unavoidably large in fire tests, use of these variability measures in compliance assessment is likely to mean rejecting compliant products (if the result indicates a FAIL for a product that is compliant) or accepting non-compliant products (if the result indicates a PASS for a non-compliant product). Neither of these outcomes is acceptable, but mandatory use of variability estimates in compliance assessment would require one or the other, if reporting it implies using it. Suppose that it is established that the variability associated with a 1 h fire resistance test is ±5 min. Manufacturers would argue that a test that passes for at least 55 min should be considered a pass while the regulatory community would argue that a test should run a minimum of 65 min without a failure to be considered passing. The compromise position is, of course, to judge compliance on the actual test result. This is known as the "Shared Risk" principle and is applied broadly within the product certification system. But the question of using variability data in compliance work is still a debatable issue.

It is possible to quantify and report Measurement Uncertainty as it relates to specific aspects of fire test methods. An example is using ASTM E119's requirements for the fire test exposure in terms of the time-temperature curve. This example demonstrates the Measurement Uncertainty of the "degree-minutes" fire exposure which can be quantified and is actually quite small and thus is unlikely to be a significant factor in explaining variability of test results in furnace tests. For example in a 1-h fire wall test the uncertainty of the fire exposure in terms of measured degree-minutes is about ±2%, but the reproducibility (precision) obtained in inter-laboratory tests is about ±10%.

A second example shows that in the Steiner Tunnel, the uncertainty in the flame spread distance is typically reported as the uncertainty of the linear transducer which has a small uncertainty (less than 1%). Now consider the operator's judgment as to where the flame-front is. It is an accepted fact that this uncertainty is on the order of 6 inches. However, it is well known that a small change in the air flow velocity and fuel flow, coupled with the uncertainty of the products reaction to the fire (cracking, melting, etc.) can affect a flame spread result by more than 20%.

One implication of these arguments is that if uncertainty were really as large as current methods indicate and as unrelated to the variables controlled by good practice, then why are we not seeing a lot of product failures in the field, when manufacturers commonly engineer products to just meet pass/fail criteria?

What does this mean for the case of requiring MU analysis for fire tests?

Measurement uncertainty has a specific meaning and does not apply to test that do not result in a quantitative measurement.

Fire resistance tests are essentially qualitative procedures and therefore the result cannot be properly associated with a measurement uncertainty statement.

Measurement uncertainty should not be confused with "reproducibility" and "repeatability" determinations which can be developed for fire resistance test methods through interlaboratory experimental test programs.

The need for development of r&R data needs to be weighed against the cost and the potential for funding such a program. To date there is not a substantial record of problems or failures that are available for use in a cost/benefit analysis.

Proficiency test programs, while not sufficient to establish repeatability and reproducibility of fire test methods in the short term are still valuable as indicators of whether or not a serious problem may exist.

If MU is typically small compared to data variability due to unforeseen parameters (specimen quality control, preparation, fire exposure, random events, etc.), why is MU seen as a priority when a discussion of possible sources of variability may be more useful to the end user?

Acknowledgments

The authors would like to thank Dr. John Hall for allowing us to participate in the ASTM E5 Symposium on Uncertainty 2011. We also want to thank Intertek and Priest and Associates consulting for allowing the time to work on this paper.

References

[1] ASTM E119-10b, "Standard Test Methods for Fire Tests of Building Construction and Materials," ASTM International, West Conshohocken, PA.

[2] ISO/IEC 17025, *General Requirements for the Competence of Testing and Calibration Laboratories*.

[3] ASTM E691-09, "Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method," ASTM International, West Conshohocken, PA.

[4] Manzello, S. L., Grosshandler, W. L., and Mizukami, T., "Furnace Testing of Full-Scale Gypsum Steel Stud Non-Load Bearing Wall Assemblies: Results of Multi-

Laboratory Testing in Canada, Japan, and USA," *Proceedings of the Fifth International Conference on Structures in Fire (SiF'08), Publ. Organizing Committee - 5th International Conference - Structures in Fire*, K. H. Tan, V. Kudor, T. H. Tan, Eds., SiF'08 Nanyang Technological University, Singapore, May 28–30, 2008, pp. 687–698.

[5] ISO/IEC Guide 98-3:2008, *Guide to the Expression of Uncertainty in Measurement (GUM)*.

[6] ISO/IEC Guide 99:2007, *International Vocabulary of Metrology - Basic and General Concepts and Associated Terms (VIM)*.

Reprinted from JTE, Vol. 39, No. 6
doi:10.1520/JTE103914
Available online at www.astm.org/JTE

Paul H. Shipp[1] *and Qiang Yu*[2]

Bench Tests for Characterizing the Thermophysical Properties of Type X Special Fire Resistant Gypsum Board Exposed to Fire

ABSTRACT: ASTM C1396 Standard Specification for Gypsum Board defines type X special fire resistant gypsum board on the basis of the fire resistance of a load bearing wood framed gypsum partition tested in accordance with ASTM E119 Standard Test Methods for Fire Tests of Building Construction and Materials. Monitoring individual product performance in a wall system fire test presents serious challenges to the manufacturing facility. The expense and complexity of operating an ASTM E119 wall furnace makes it impossible for plants to run the defining test on site. In addition, many factors influence the outcome of a fire resistance test making it difficult to assess the performance of the gypsum board independently of other wall system and laboratory influences. In 2003 the Gypsum Association formed an ad hoc Product Standard Task Group in an attempt to develop an improved definition of type X gypsum board. The new definition had to be equivalent to the current ASTM C1396 specification but based on bench scale tests of the high temperature thermophysical properties of type X gypsum board alone. The tests must also be suitable for use as in-plant quality assurance procedures. A suite of three tests resulted. Their development and correlation to the ASTM E119 test has been documented by the authors in an earlier paper. An analysis of the precision of these three test methods is presented here.

KEYWORDS: fire resistance, gypsum board, type X gypsum board, high-temperature core cohesion high-temperature shrinkage, high-temperature thermal insulation

Manuscript received April 16, 2011; accepted for publication May 18, 2011; published online August 2011.
[1] Ph.D., PE, USG Corporation, Libertyville, IL 60048, e-mail: pshipp@usg.com
[2] Ph.D., USG Corporation, Libertyville, IL 60048.

Cite as: Shipp, P. H. and Yu, Q., "Bench Tests for Characterizing the Thermophysical Properties of Type X Special Fire Resistant Gypsum Board Exposed to Fire," *J. Test. Eval.*, Vol. 39, No. 6. doi:10.1520/JTE103914.

Introduction

The specification for type X (special fire resistant) gypsum board can be found in ASTM C1396 Standard Specification for Gypsum Board [1]. This stipulates that 5/8 in. (16 mm) thick type X gypsum board shall provide a fire resistance rating of at least 1 h when a single layer of gypsum boards is installed on each side of an uninsulated load-bearing wood stud partition and tested in accordance with test methods ASTM E119 [2]. Likewise, 1/2 in. (13 mm) thick type X gypsum board must achieve at least 3/4 h fire resistance rating in a single layer application on both sides of an uninsulated load-bearing wood stud wall. While this establishes a specified minimum performance level for a particular fire resistance test, the definition lacks specificity with regards to the properties of the gypsum board itself. The performance of the gypsum board during the fire test is masked by other key elements of the test that also influence its outcome, e.g., the wood studs and furnace. This presents challenges to gypsum board manufacturers in that the defining test is too expensive and difficult to conduct at manufacturing facilities and yet the specification provides no guidance as to what extensive or intensive material properties are required of type X gypsum board. Thus, the ASTM E119 specification is unsuitable for monitoring and gauging the individual effects of raw material or process variations that can occur in production settings.

In lieu of running ASTM E119 wall tests at the plants, gypsum board manufacturers rely on in-house quality assurance procedures to monitor whatever parameters each manufacturer deems sufficient to insure compliance with the ASTM C1396 type X definition. In 2003, the Gypsum Association (GA) formed an ad hoc Product Standard Task Group in an attempt to develop a standardized set of laboratory tests that could be used for an improved definition of type X gypsum board. The new definition must provide the same level of fire resistance performance as the current ASTM C1396 specification but be based on specific properties of the board relevant to its performance as a fire resistant membrane. This six year industry research program produced three bench scale test methods whose outcomes collectively predict fire resistance performance by means of simple linear regression correlation. The three tests are:

1. High temperature core cohesion.
2. High temperature shrinkage.
3. High temperature thermal insulation.

The engineering rationale for these tests and the regression analysis correlating the bench test outcomes to the ASTM E119 fire resistance test are documented by Shipp and Yu [3]. In the process of developing and evaluating the tests, a series of round robins, or interlaboratory studies (ILS), were completed. The results of the ILS statistical analyses used to determine the precision of these test methods are presented here.

Background

Gypsum plaster has been used to protect buildings from fire for centuries. With the development of type X special fire resistant gypsum board in the mid-20th century, its use in fire resistant construction has become a staple of modern

architectural design. A fire resistance membrane is usually not a structural element of the building assembly. Instead, its fire protection functions are twofold: (1) shield the assembly from corrosive attack by fire and, (2) retard heat transmission through the assembly. To fulfill the first function, the membrane must be able to deform with the movement of the underlying structural framing as it heats up. The membrane must therefore demonstrate a capacity for high temperature strain without developing gaps or openings through which flames can penetrate to attack the underlying structure. This places significant demands on membranes confronted with post-flashover conditions. Simultaneously, the membrane must also thermally insulate the unexposed elements of the assembly from heat transfer. Thus, the membrane must have low thermal diffusivity to retard heat conduction and provide a radiation barrier. The importance of thermal diffusivity is illustrated by a thin sheet of steel which, by itself, is not an effective fire protective membrane because of its high thermal diffusivity even though it is fully capable of moving with the structure and remaining intact without being breached by flames or hot gases during a fire.

When exposed to elevated temperatures, gypsum undergoes a series of endothermic phase changes that contribute to its low thermal diffusivity during post-flashover exposures. Compositional changes and molecular realignment in the crystal structure result in shrinkage during these changes. Numerous studies have been conducted on the thermal and mechanical properties of gypsum and gypsum board at elevated temperatures [4–8]. However, the test methods employed are better suited to a research laboratory than a gypsum plant. The three bench tests discussed herein were specifically developed for simplicity and robustness to be practical for use in plant quality assurance. A summary of each test method and its requirements is given by Shipp and Yu [3]. A detailed description of the three test methods can be found in ASTM Committee C11 letter ballot C11 (09-04) [9].

Interlaboratory Studies (ILS)

An engineering analysis of the requisite mechanical and thermodynamic properties of type X gypsum board resulted in the three tests listed above. Several interlaboratory studies (ILS), also called round robins, were conducted during their development. In alphabetical order, the following companies participated in one or more ILS:

- American Gypsum Co. LLC.
- CertainTeed Gypsum Inc.
- Georgia-Pacific Gypsum LLC.
- Lafarge North America Inc.
- National Gypsum Co.
- PABCO Gypsum.
- Temple-Inland.
- United States Gypsum Co.

As there is no standard reference material for type X gypsum board, no quantitative statement can be made regarding the accuracy or bias of these tests. Instead, the ILS analyses are restricted to repeatability and reproducibility statistics based on three common statistical measures of precision:

- Repeatability limit (r) —The repeatability limit "r" is the interval representing the critical difference between two test results for the same material, obtained by the same operator using the same equipment on the same day in the same laboratory. Two test results obtained within one laboratory shall be judged not equivalent if they differ by more than the "r" value for that material.
- Reproducibility limit (R) — The reproducibility limit "R" is the interval representing the critical difference between two test results for the same material, obtained by different operators using different equipment in different laboratories. Two test results shall be judged not equivalent if they differ by more than the "R" value for that material.
- Coefficient of variation (CV) —A normalized standard deviation used as a measure of variation within a probability distribution.

For each of the bench tests, at least one ILS was completed to evaluate the test procedure and its precision. Multiple round robins were conducted on the high temperature shrinkage and high temperature thermal insulation tests as they evolved. Initially, the round robins were restricted to 5/8 in. (16 mm) gypsum boards that met or exceeded the requirements for type X classification. Later round robins were extended to include samples of 1/2″ (13 mm) and 3/4″ (19 mm) special fire resistant gypsum boards. The statistical analyses of the ILS results were conducted in accordance with ASTM E691 Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method [10]. This standard practice entails determining whether the collected data are adequately consistent to form the basis for a test method precision statement, determining if there are inconsistencies in the data that require action, and calculating the statistical parameters for the precision statement. The definitions and equations used to calculate these parameters are given below for reference in the ensuing discussion.

p = number of laboratories, ($j = 1,\ldots p$)
n = number of test results per laboratory, ($i = 1,\ldots n$)
$x_{i,j}$ = individual test result i, laboratory j
t = Student's t value, (p-2) df, 0.5% two-tailed significance level
F = F-ratio at (n-1) and (p-1)(n-1) df, 0.5% significance level

$$\overline{x}_j = \sum_{i=1}^{n} \frac{x_{i,j}}{n} = \text{average for laboratory } j \tag{1}$$

$$s_j = \sqrt{\sum_{i=1}^{n} \frac{\left(x_{i,j} - \overline{x}_j\right)^2}{(n-1)}} = \text{standard deviation for laboratory } j \tag{2}$$

$$\overline{\overline{x}} = \sum_{j=1}^{p} \frac{\overline{x}_j}{p} = \text{average of all laboratory averages} \tag{3}$$

$$d_j = \left(\overline{x}_j - \overline{\overline{x}}\right) = \text{deviation for laboratory } j \tag{4}$$

$$s_{\bar{x}} = \sqrt{\sum_{j=1}^{p} \frac{d_j^2}{(p-1)}} = \text{standard deviation of laboratory averages} \tag{5}$$

$$s_r = \sqrt{\sum_{j=1}^{p} \frac{s_j^2}{p}} = \text{repeatability standard deviation} \tag{6}$$

$$s_R = \max\left\{ s_r, \sqrt{(s_{\bar{x}})^2 + \frac{(s_r)^2(n-1)}{n}} \right\} = \text{reproducibility standard deviation} \tag{7}$$

$$h_j = \frac{d_j}{s_{\bar{x}}} = \text{between-laboratory consistency statistic} \tag{8}$$

$$k_j = \frac{s_j}{s_r} = \text{within-laboratory consistency statistic} \tag{9}$$

$$h_{critical} = \frac{(p-1)t}{\sqrt{p(t^2+p-2)}} = \text{critical value of } h \tag{10}$$

$$k_{critical} = \sqrt{\frac{p}{1+\frac{(p-1)}{F}}} = \text{critical value of } k \tag{11}$$

$$CV_j = \frac{s_j}{\bar{x}_j} = \text{Coefficient of variation for laboratory } j \tag{12}$$

$$\overline{CV} = \sum_{j=1}^{p} \frac{CV_j}{p} = \text{average Coefficient of variation, all laboratories} \tag{13}$$

$$r = 2.8\ s_r = \text{Repeatability limit} \tag{14}$$

$$R = 2.8\ s_R = \text{Reproducibility limit} \tag{15}$$

High Temperature Core Cohesion Test

Type X gypsum board is distinguished from standard gypsum board primarily by the addition of chopped glass fibers to the gypsum core. The original ASTM definition of type X incorporated an early version of the core cohesion test. This was replaced in 1975 by the current type X definition based on the ASTM E119 fire test. The core cohesion test is a reliable indicator of the quantity and type of

glass fibers used although it alone is not sufficient as the defining test for type X. As such, it remained in use in Europe where it evolved into the current test method specified by the European Union in EN 520 [11]. The EN 520 high temperature core cohesion test was therefore selected for reintroduction into the type X definition. This test is a simple categorization, *i.e*. a binary pass/fail result, based on whether or not the test specimen breaks into two separate pieces at the specified strain while being heated on both faces by Meker burners. Seven gypsum manufacturers participated in a round robin that entailed four different type X products. In addition, one of the companies tested a mix of type X and regular gypsum boards. In all, over 400 gypsum board specimens were subjected to the core cohesion test. The test consistently segregated type X boards from regular gypsum board with no false positive or false negative results returned. Due to the nonquantitative nature of the test results; however, no precision and bias statement can be made.

High Temperature Shrinkage Test

While high temperature core cohesion is a necessary attribute of type X board, it is not sufficient to predict performance as a fire protection membrane. The amount of shrinkage exhibited by the gypsum core when exposed to fire also influences the degree and severity of cracking that will occur in the membrane. Traditionally, gypsum manufacturers have measured the high temperature shrinkage of type X gypsum boards at 1000 °F (538 °C). By the time it reaches that temperature the gypsum has converted to its anhydrite form ($CaSO_4$) with all of the crystalline water driven out, a process known as calcining. As the temperature rises further; however, anhydrite undergoes additional phase transitions in which the molecular packing changes and further shrinkage occurs near 1500 °F (815 °C). This latter shrinkage event can exceed the shrinkage from calcining. The temperature of the shrinkage test was therefore raised to 1562 °F (850 °C) to capture all of the shrinkage caused by post-flashover fire conditions. For personnel safety and to avoid specimen breakage due to thermal shock, an additional change is that rather than using a preheated oven the test begins with the oven at room temperature, heats the specimens to 1562 °F (850 °C), holds that temperature for 20 min and then shuts off the oven to allow the samples to cool before opening the oven door.

The measured quantity from the test method is the reduction in diameter of circular disk test specimens. This can be expressed as a decimal number or %. For the purpose of the statistical analysis, decimal values were chosen for the shrinkage data, e.g., 3.2% shrinkage is expressed as the value 0.032. A series of four round robins were conducted on the shrinkage test as the procedure was refined to its final form. For each ILS, the data were arranged in a spreadsheet and the statistical parameters calculated. The between-laboratory and within-laboratory consistency statistics were then examined graphically to flag inconsistencies and act upon them as prescribed in ASTM E691.

The three high temperature shrinkage round robins each consisted of four sample materials with six or more laboratories participating. Twelve test specimens from each sample material were tested. Data from a single study, Shrinkage Round Robin 2 (SRR2), are shown below to illustrate the analysis process

used for each ILS. Eight gypsum manufacturers (Labs A – H) participated in SRR2. The four type X material samples, each from different manufacturers, were labeled S, T, U, and V. The data and calculated statistical parameters for sample material T are shown in Table 1. Similar tables were constructed for the other three sample materials S, R, and U to complete the analysis.

Four tables such as Table 1 were generated and evaluated as a group for each ILS. A convenient aid in looking for inconsistencies in these data is to use a graphical representation of the between-laboratory and within-laboratory statistics as shown in Figs. 1(*a*)–1(*d*).

Examination of the consistency statistics, averages and standard deviations shows internal variations for Lab H comparable to those of the other labs, but the average values for Lab H stand out from the rest of the group for all four materials. Looking more closely at the within-lab variation, none of the calculated k_j statistics approach or exceed the critical k value. A different conclusion emerges from the between-lab variations where the $h_{j,j=H}$ values for all four materials approach or exceed the critical h value. The h_j values for all other laboratories are well below the critical h value. It is concluded; therefore, that the Lab H data are systematically inconsistent from the other laboratories in this ILS. The statistical parameters are then recalculated without laboratory H, Table 2, to verify that the data set has no other inconsistencies.

This procedure was applied to analyzing all of the high-temperature shrinkage test round robins. In the first two round robins, the test procedure was still being refined. SRR4 was conducted using the final version of the test method. The results for all of the round robins are summarized and presented below in Table 3 after the following section.

High-Temperature Thermal Insulation Test

As noted earlier, the primary function of the thermal barrier is to insulate the structure and unexposed surface from the fire. A new test method was developed to measure the thermal insulation effectiveness of the gypsum board at elevated temperatures and a total of five round robins were run. As with the shrinkage studies, each round robin examined four different sample boards with five to seven laboratories participating. For the thermal insulation studies, four test specimens ($n = 4$) were tested from each of the four sample materials. The most significant modification to the test method occurred between ILS no. 4 and no. 5 in which the oven temperature was increased from 400 to 500 °C and the test specimens were placed in a rack that held them at the geometric center of the oven volume.

Results and Observations

The precision analysis results for the high temperature shrinkage test and the high temperature thermal insulation test interlaboratory studies are shown in Table 3. Of key interest are the single laboratory repeatability limits (r), the between-laboratories reproducibility limits (R), and the average coefficients of

TABLE 1—*High-Temperature Shrinkage data and calculated parameters, High Temperature Shrinkage Round Robin 2 (SRR2), Material T, Laboratories A – H.*

Material T Laboratory	Test Results, xi, j															h-critical 2.15	k-critical 1.49	
	1	2	3	4	5	6	7	8	9	10	11	12	Avg	s	d	h	k	CV
A	0.043	0.040	0.052	0.053	0.044	0.046	0.041	0.040	0.050	0.053	0.044	0.044	0.046	0.005	0.006	0.669	1.381	10%
B	0.040	0.037	0.047	0.045	0.043	0.045	0.038	0.036	0.045	0.046	0.045	0.051	0.043	0.004	0.004	0.371	1.288	10%
C	0.039	0.043	0.042	0.037	0.044	0.044	0.039	0.046	0.044	0.041	0.044	0.043	0.042	0.003	0.003	0.284	0.782	6%
D	0.040	0.041	0.045	0.044	0.039	0.038	0.037	0.043	0.042	0.041	0.036	0.036	0.040	0.003	0.001	0.076	0.893	8%
E	0.043	0.047	0.041	0.043	0.044	0.040	0.039	0.048	0.041	0.043	0.047	0.042	0.043	0.003	0.004	0.374	0.812	7%
F	0.042	0.044	0.046	0.046	0.038	0.038	0.041	0.044	0.045	0.045	0.037	0.038	0.042	0.004	0.002	0.250	1.015	8%
G	0.042	0.044	0.046	0.046	0.042	0.040	0.043	0.045	0.046	0.044	0.044	0.042	0.044	0.002	0.004	0.416	0.557	4%
H	0.015	0.012	0.014	0.021	0.022	0.020	0.018	0.018	0.013	0.018	0.013	0.013	0.016	0.004	−0.023	−2.440	1.013	21%
												Avg of cell avgs, X:	0.040	0.003	= Avg std dev		Avg CV:	9%
												Std dev of cell avgs, Sx:	0.010					
												Repeatability std dev, Sr:	0.003				Repeatability limit, r:	0.010
												Reproducibility std dev, SR:	0.010				Reproducibility limit, R:	0.028

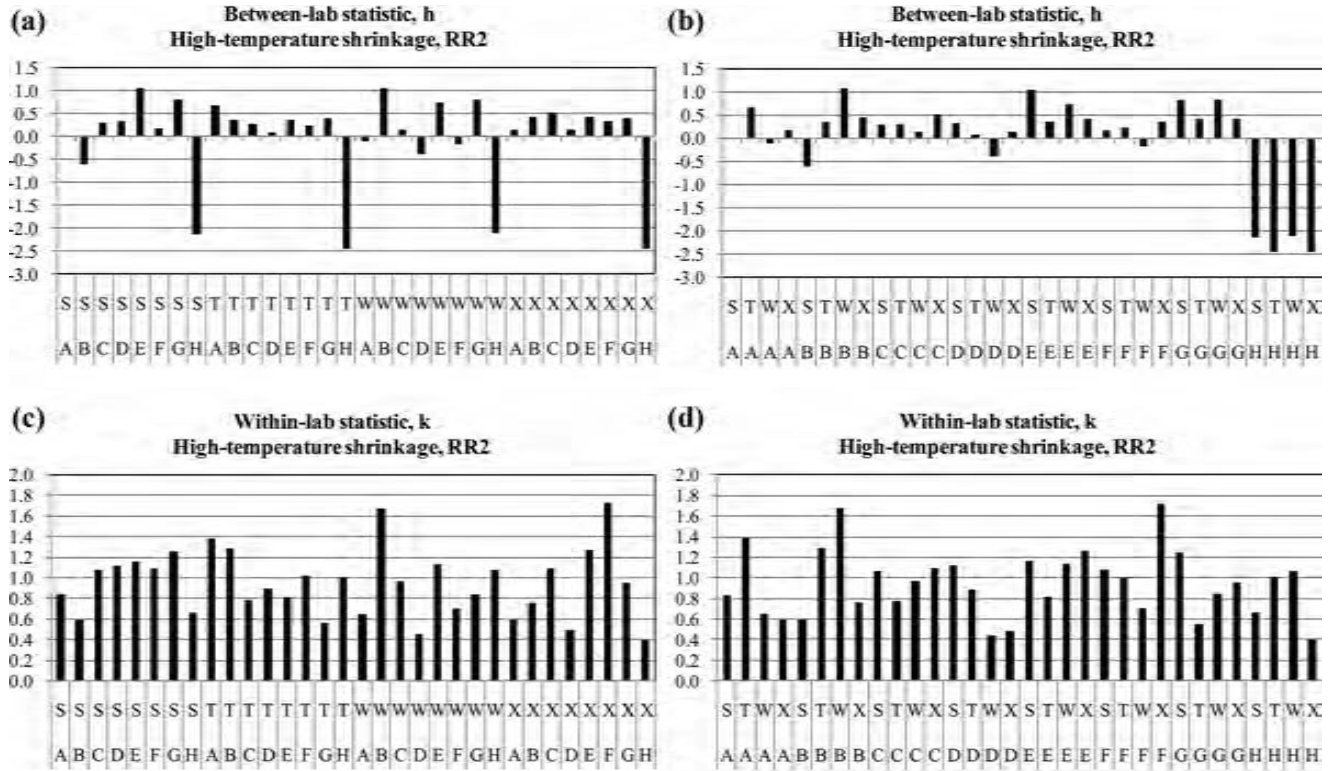


FIG. 1—*(a) Between-laboratories variability statistic, h. High-Temperature Shrinkage Round Robin 2 (SRR2), sorted by sample board (S, T, W, X), laboratories A – H. (b)– Between-laboratories variability statistic, h. High-Temperature Shrinkage Round Robin 2 (SRR2), sorted by laboratories (A – H), sample boards S, T, W, X. (c) Within-laboratory variability statistic, k. High-Temperature Shrinkage Round Robin 2 (SRR2), sorted by sample board (S, T, W, X), laboratories A – H. (d) Within- laboratory variability statistic, k. High-Temperature Shrinkage Round Robin 2 (SRR2), sorted by laboratories (A – H), sample boards S, T, W, X.*

SHIPP AND YU, doi:10.1520/JTE103914 37

TABLE 2—*High-Temperature Shrinkage data and calculated parameters, High Temperature Shrinkage Round Robin 2 (SRR2), Material T, Laboratories A – G.*

																h-critical	k-critical	
	Test Results, xi, j															2.05	1.48	
Material T Laboratory	1	2	3	4	5	6	7	8	9	10	11	12	Avg	s	d	h	k	CV
A	0.043	0.040	0.052	0.053	0.044	0.046	0.041	0.040	0.050	0.053	0.044	0.044	0.046	0.005	0.006	0.669	1.381	10%
B	0.040	0.037	0.047	0.045	0.043	0.045	0.038	0.036	0.045	0.046	0.045	0.051	0.043	0.004	0.004	0.371	1.288	10%
C	0.039	0.043	0.042	0.037	0.044	0.044	0.039	0.046	0.044	0.041	0.044	0.043	0.042	0.003	0.003	0.284	0.782	6%
D	0.040	0.041	0.045	0.044	0.039	0.038	0.037	0.043	0.042	0.041	0.036	0.036	0.040	0.003	0.001	0.076	0.893	8%
E	0.043	0.047	0.041	0.043	0.044	0.040	0.039	0.048	0.041	0.043	0.047	0.042	0.043	0.003	0.004	0.374	0.812	7%
F	0.042	0.044	0.046	0.046	0.038	0.038	0.041	0.044	0.045	0.045	0.037	0.038	0.042	0.004	0.002	0.250	1.015	8%
G	0.042	0.044	0.046	0.046	0.042	0.040	0.043	0.045	0.046	0.044	0.044	0.042	0.044	0.002	0.004	0.416	0.557	4%
													Avg of cell avgs, X:		0.043	0.003	= Avg std dev	Avg CV:	8%
													Std dev of cell avgs, Sx:		0.002				
													Repeatability std dev, Sr:		0.003			Repeatability limit, r:	0.010
													Reproducibility std dev, SR:		0.004			Reproducibility limit, R:	0.010

TABLE 3—*Precision statistics for the high-temperature shrinkage test and the high temperature thermal insulation test interlaboratory studies.*

ILS	No. of Labs, p	Materials	Between-labs $h_{critical}$	Between-labs $\lvert h_{max}\rvert$	Within-lab $k_{critical}$	Within-lab $\lvert k_{max}\rvert$	Average, x_{avg}	Repeatability Std Dev, s_r	Reproducibility Std Dev, s_R	Repeatability Limit, r	Reproducibility Limit, R	Coefficient of Variation, CV_{avg}
High-Temperature Shrinkage												
SRR1	7	S, T, W, X	2.05	1.57	1.48	1.75	0.050	0.0046	0.0062	0.0129	0.0172	9%
SRR2	7	S, T, W, X	2.05	1.07	1.48	1.73	0.047	0.0038	0.0050	0.0108	0.0141	8%
SRR4	6	R, U, V, Z	1.92	1.39	1.46	1.17	0.064	0.0044	0.0051	0.0124	0.0143	7%
High-Temperature Thermal Insulation												
TI1	7	S, T, W, X	2.05	1.99	1.87	2.11	51.4	1.62	3.75	4.55	10.49	3%
TI2	6	S, T, W, X	1.92	1.91	1.84	2.30	54.7	3.88	5.77	10.87	16.17	5%
TI4	5	R, U, V, Z	1.74	1.69	1.79	2.14	49.3	1.89	3.36	5.29	9.41	3%
TI5	6	P50	1.92	1.48	1.84	1.33	17.7	0.18	0.92	0.50	2.58	1%
TI5	5	P60	1.74	1.31	1.79	1.32	26.3	0.30	0.72	0.83	2.01	1%
TI5	6	P64	1.92	1.28	1.84	1.76	27.9	0.54	0.72	1.50	2.03	2%
TI5	6	P75	1.92	1.79	1.84	1.77	31.7	0.63	1.57	1.77	4.41	2%

Note 1: Study No. 3 not reported. Only the thermal insulation test was studied and problems were reported with the revised procedure.
Note 2: Studies Nos. 1 - 4 tested 5/8 in. (16 mm) gypsum boards exclusively. For simplicity, group averages reported for *x*avg, S_r, S_R, r, R.
Note 3: TI5 tested 1/2 in. (13 mm),5/8 in. (16 mm) and 3/4 in. (19 mm) gypsum boards. Individual samples statistics reported separately.
Note 4: TI5 is final configuration for thermal insulation test. Shrinkage not included in ILS No.5.

variation (*CV*) for the group of laboratories. The repeatability and reproducibility limits have an approximate 95% probability of being correct.

The high-temperature shrinkage test was found to provide excellent repeatability and reproducibility. The reduction in the coefficient of variation from 9% to 7% was achieved by introducing several key features to the test:

- The furnace must be unvented.
- The test specimens are placed on pedestals to expose both faces directly to oven conditions and prevent the specimens from curling or warping during the test.
- Care must be exercised not to thermally shock or mechanically jar the test specimens when they are removed from the oven.

Likewise, the high-temperature thermal insulation test also exhibits excellent repeatability and reproducibility. A key modification to the test proved to be the increase in the oven temperature which reduced the time required to run the test by roughly half (25–30 min versus 50–55 min for 5/8 in. (16 mm) gypsum board) with a coefficient of variation less than 3%.

The high temperature core cohesion test was found to consistently and accurately differentiate between type X gypsum board and regular gypsum board. As noted earlier, no precision statement is made for this test due to the binary test results that do not lend themselves to the statistical treatment employed here.

For both the high temperature shrinkage test and the high temperature thermal insulation test it can be observed that the coefficients of variation were already acceptable from the beginning. Improvements to the test methods evaluated in subsequent round robins resulted in further reducing the *CV* values but laboratory repeatability remained quite good throughout the program.

Likewise, between-laboratories reproducibility for the high temperature shrinkage test was quite acceptable (<2%) throughout the test development program with the final version demonstrating the best performance (1.4%).

Three iterations of the high temperature thermal insulation test were required to achieve the final precision levels. This is likely because this was the most novel test of the three and there was a learning curve involved in identifying and standardizing the key parameters. The order of magnitude improvement in *r* and *R* values as well as the reduction in *CV* from 5% to 1–2% demonstrate excellent agreement with the final configuration of the test.

Conclusions

It can be concluded that these tests are practical for use in a plant quality assurance department and provide satisfactory precision in measuring the thermophysical properties of type X special fire resistant gypsum board. The precision of the high temperature shrinkage and high temperature thermal insulation tests have been determined and are herewith documented in terms of their repeatability, reproducibility and coefficients of variation.

An additional observation can be drawn from the relative consistency of the precision statistics exhibited by the bench test methods throughout their development program. Basic design and development of the three test methods based

on the fundamental thermophysical properties of a fire protection membrane took less than seven months. The initial round robins were completed within 12 months of project inception. However, attempts to correlate the initial bench test results to the ASTM E119 tests produced less than satisfactory results. At the time, this was interpreted as a deficiency in the design of the new bench tests and/or lack of familiarity with the new tests by laboratory personnel. However, from the precision results reported above and the final correlation of the bench tests to the ASTM E119 fire resistance test following completion of more than 35 wall fire tests [3] it is now apparent that the difficulties encountered with the statistical correlation were rooted in the variability of the defining ASTM E119 fire resistance test itself rather than the variability of the bench tests. Achieving an acceptable statistical correlation simply required conducting a sufficient number of replications of the very expensive wall fire test rather than further refinement of the bench tests. This underscores the need for product specifications to be based on the material properties of the product rather than on a system application where the properties of the material can be masked or confounded by the performance of other key elements in that system.

References

[1] ASTM C1396/C1396M – 09a, 2009, "Standard Specification for Gypsum Board," *Annual Book of ASTM Standards*, Vol. 4.01, ASTM International, West Conshohocken, PA.

[2] ASTM E119 – 10b, 2010, "Standard Test Methods for Fire Tests of Building Construction and Materials," *Annual Book of ASTM Standards*, Vol. 4.07, ASTM International, West Conshohocken, PA.

[3] Shipp, P. H. and Yu, Q., "Thermophysical Characterization of Type X Special Fire Resistant Gypsum Board," *Proceedings of the Fire and Materials 2011 International Conference*, San Francisco, January 31–February 2, 2011, Interscience Communications Ltd., London, UK, pp. 417–426.

[4] Lawson, J. R., "An Evaluation of Fire Properties of Generic Gypsum Board Products," *NBSIR 77-1265*, U.S. Dept. of Commerce, National Bureau of Standards, Aug 1977.

[5] Manzello, S. L., Park, S-H, Mizukami, T., and Bentz, D. P., "Measurement of Thermal Properties of Gypsum Board at Elevated Temperatures," *Proceedings of the Fifth International Conference on Structures in Fire (SiF'08)*, May 28–30, 2008, Nanyang Technical Univ., Singapore, pp. 656–665.

[6] Bentz, D. P., "Fire Resistaive Materials: Thermal Barriers between Fires and Structures," *Thermal Conductivity 30/Thermal Expansion 18: Proceedings of the 30th International Thermal Conductivity Conference and the 18th International Thermal Expansion Symposium*, DesTech Publications, Inc., Lancaster, PA, pp. 108–119, 2010.

[7] Frangi, A., Schleifer, V., and Fontana, M., "Experimental and Numerical Analysis of Gypsum Plasterboards in Fire," *Fire Technol.*, Vol. 46, 2010, pp. 149–167.

[8] Cramer, S. M., Friday, O. M., White, R. H., and Sriprutkiat, G., "Mechanical properties of gypsum board at elevated temperatures," *Proceedings of the Fire and Materials 2003 International Conference*, San Francisco, Jan 2003, Interscience Communications Ltd., London, UK, pp. 33–42.

[9] *ASTM ballot C11 (09-04)*, Item 1, Aug 31, 2009.
[10] ASTM E691-09, 2010, "Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method," *Annual Book of ASTM Standards*, Vol. 14.02, ASTM International, West Conshohocken, PA.
[11] EN 520 Gypsum Plasterboards – Definitions, requirements and test methods, CEN European Committee for Standardization, Brussels, Nov 2004.

Reprinted from JTE, Vol. 39, No. 6
doi:10.1520/JTE103949
Available online at www.astm.org/JTE

John V. Resing,[1] *Pravinray D. Gandhi,*[2] *Dwayne E. Sloan,*[3] *and Randall K. Laymon*[4]

Measurement Uncertainty and Statistical Process Control for the Steiner Tunnel (UL 723, ASTM E84)

ABSTRACT: In the United States, the Steiner Tunnel (UL 723 [2008, "Test for Surface Burning Characteristics of Building Materials," Ninth Edition, Underwriters Laboratories Inc., Northbrook, IL], ASTM E84-10 [2010, "Standard Test Method for Surface Burning Characteristics of Building Materials," *Annual Book of ASTM Standards*, Vol. 4.07, ASTM International, West Conshohocken, PA.]) is an important fire test apparatus used by the building codes to assess the flammability and smoke generation characteristics of building products (e.g., insulation, sheathing materials, foamed plastics, wood-based products). This paper examines how various tools and methodologies can be used to quantify, improve, and control measurement uncertainty in fire tests, such as the ASTM E84 tunnel test. This paper also assesses uncertainty in the Steiner Tunnel using both a Gage Repeatability and Reproducibility (Gage r&R) study and historical results for reference materials. The improvements discussed were achieved by using a widely

Manuscript received May 19, 2011; accepted for publication August 2, 2011; published online September 2011.

[1] Lean Sigma Black Belt, Lean Sigma Office, Underwriters Laboratories, Inc., 333 Pfingsten Rd., Northbrook, IL 60062-2096 USA (Corresponding author), e-mail: John.V.Resing@us.ul.com

[2] Director of Research and Development, Underwriters Laboratories, Inc., 333 Pfingsten Rd., Northbrook, IL 60062-2096 USA.

[3] Principal Engineer - Reaction to Fire, Underwriters Laboratories, Inc., 333 Pfingsten Rd., Northbrook, IL 60062-2096 USA.

[4] Senior Conformity Assessment Staff/Chair ASTM E84, Underwriters Laboratories, Inc., 333 Pfingsten Rd., Northbrook, IL 60062-2096 USA.

Cite as: Resing, J. V., Gandhi, P. D., Sloan, D. E. and Laymon, R. K., "Measurement Uncertainty and Statistical Process Control for the Steiner Tunnel (UL 723, ASTM E84)," *J. Test. Eval.*, Vol. 39, No. 6. doi:10.1520/JTE103949.

known Lean/Sigma principle—the define, measure, analyze, improve, control (DMAIC) method. Benefits derived from the DMAIC method include: (i) an improved understanding and control of the sources of variation; (ii) the development of process control charts to monitor Steiner Tunnel performance; and (iii) a reduced need for calibration and re-verification of Steiner Tunnel performance. The Gage r&R results show that the major contributor to uncertainty in flame spread and smoke developed results is primarily due to the performance difference in the test samples, and not due to the measurement device or the person conducting the test. The uncertainty, in the form of the standard deviation, is also shown in the results of each of the materials tested in the study. Monthly data for the historical reference material, red oak, and a proposed reference material, heptane, are shown to demonstrate stability of the measurement system over time. The heptane test is presented to help measure and reduce uncertainty of the smoke measurement in the Steiner Tunnel.

KEYWORDS: ASTM E84-10, building materials, flame spread, smoke developed, statistical process control, Gage r&R, UL 723, uncertainty

Introduction

When ASTM E84 [1] (UL 723 [2]) was developed in the middle of the last century, it was in the context of dramatically improving the performance of interior finish materials in the wake of several historically disastrous fires [3]. The test played a key role in identifying interior finish materials that may cause rapid spread of flame and smoke in buildings. Now, the ASTM E84 test is an important fire test in North America for building materials that include interior finish, insulation, and foamed plastics [4]. Current U.S. building codes use the results of the E84 test to categorize interior finish materials into the following hierarchy of performance:

Class A—Flame Spread Index (FSI) of 0–25, Smoke Developed Index (SDI) of 0–450,

Class B—FSI of 26–75, SDI of 0–450,

Class C—FSI of 76–200, SDI of 0–450, and

Plenum Materials—FSI of 0–25, SDI of 0–50.

A photograph of the Steiner Tunnel at Underwriters Laboratory (UL) is shown in Fig. 1.

One challenge that faces Steiner Tunnel operators is ensuring consistency in the data generated from a complex system such as the Steiner Tunnel. The consistency of the FSI and SDI depend upon the control of a number of equipment processes, such as air and gas flow rates. The define, measure, analyze, improve, control (DMAIC) [5] methodology was used to provide rigor to the analysis along with other techniques that support DMAIC. These techniques included (i) Statistical Process Control; (ii) Poke-yoke; and (iii) Design of Experiment (DOE). The DMAIC method and use of these tools within DMAIC enabled identification of the system processes that may cause variation in FSI and SDI. The improved processes were then implemented, and a Gage

FIG. 1—*Photograph of Exterior of Steiner Tunnel.*

Repeatability and Reproducibility (Gage r&R) study was conducted in an effort to assess the measurement uncertainty of Steiner Tunnel results.

It was expected that the rigorous data gathering, documentation, and analysis demanded by DMAIC would also identify the natural operational limits of the Steiner Tunnel. This would then be used to identify systemic trends to ensure that corrective measures were taken, when necessary.

DMAIC Method

The DMAIC method is a general process improvement methodology composed of the following five steps:

1. **Define**—What is important to work on? Establish the problem statements.
2. **Measure**—How is the process doing in the current state?
3. **Analyze**—What is causing our problems?
4. **Improve**—How do we make it better?
5. **Control**—How do we maintain the gains?

Define

In the *define phase*, the processes and their control systems were identified as the potential sources of variation for the FSI and SDI calculations. The variations commonly came from system failures or degradation of a measurement device, such as process equipment, dampers, a photocell, a lamp, thermocouples, or a degrading pressure measurement unit. Because the inlet air influences the combustion of the test materials, the control of inlet air, and sub-systems that affect it, were considered potential sources contributing to variation in FSI and SDI results. The output problem statement from the define phase focused on reducing variation of these elements to improve the overall quality of test results.

Measure

In the *measure phase*, we gained a deeper understanding of our current state process through examining outputs from the Steiner Tunnel equipment, and reviewing the many standard calibration and check tests performed routinely. A tool called Statistical Process Control (SPC) charting was used to establish baseline performance for both the average results and the typical variation in the form of upper and lower control limits. Through SPC charting, over 50 test parameters from regular tests, check tests, and calibration tests were examined to better understand our current state process.

SPC charts are a powerful tool to determine if variation is due to either *chance* or some *assignable* cause. UL has deployed SPC concepts not only for the regular "Standard" monthly calibrations, but also for day-to-day and for test-to-test. In addition, this tool was employed for both "Standard" requirements as well as supplemental "Diagnostic Indicators" established for internal use by UL.

Figure 2 shows a typical SPC chart which indicates variation of one monitored parameter. The normal variation for this parameter is bound by upper (UCL) and lower (LCL) control limits determined by monitoring the system over a period of time. In the figure, one data point is "out of control" (i.e., exceeds the UCL). By investigating the specific causes for this event, and eliminating the identified sources of variation, the equipment and processes are thus improved continuously.

Analyze

Now, after using tools such as SPC and identifying special causes of variation, another Lean/Sigma tool called Design of Experiment (DOE) was used in the *analyze phase*. In DOE, an experiment is designed to alter the specific variables

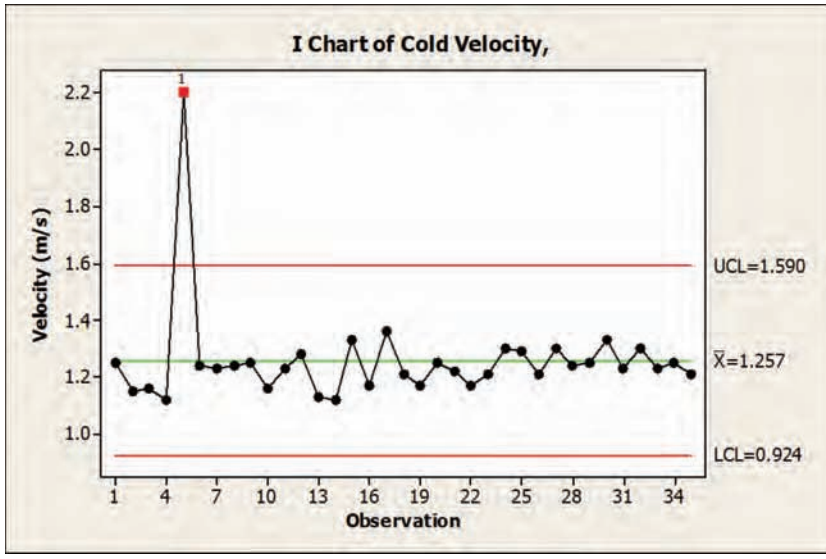


FIG. 2—*Control Chart of Cold Duct Velocity Showing Outlier.*

that may affect the process and results. By altering input variables and studying the corresponding output variables, we are able to gain a deeper analysis of the root cause of the problems. The application of the statistical design and analysis of experiments accelerates the learning process because it allows us to test multiple factors at the same time. DOE can be contrasted with the typical approach of One Factor at a Time (OFAT) where the experimenter "has a hunch" that some problem observed is due to a cause that he or she may already "know" how to fix. So, the experimenter simply goes about the business of fixing that cause. If that does not work, he or she fixes something else, and so on until maybe the situation gets better, or maybe it gets worse.

As an example, UL used DOE in the analyze phase to complete an air flow study. Steiner Tunnel air supply stability is important for testing consistency, repeatability, and reproducibility. The DOE took into account many factors including the inlet control damper; elements of the heating, ventilation, and air conditioning (HVAC) systems; process dampers; and the various pollution control equipment. The goal of the experiment was to find the air handling settings that optimize inlet air supply stability. By using DOE, UL was able to identify the equipment settings that would need to be addressed to improve the stability of the air flow.

Improve

In the *improve phase*, there was a deeper focus on the root causes of the factors causing the problems and efforts were put forth to either eliminate them or put countermeasures in place to mitigate them.

A common tool used by DMAIC practitioners in both improve and control phases is called "Poke-yoke" a Japanese term meaning "mistake proofing" or "fail-safing." This is loosely defined as any mechanism that allows an operator to avoid mistakes. UL established "Go" and "No Go" values based on SPC and included alarms in the automated data acquisition system to check for potential faults. If the selected Steiner Tunnel parameter is outside the control limits, the technician is required to take action before continuing. An example of this technique is shown in Fig. 3 for the pre-test procedures at the start of each day of tunnel operations. In this example, the automated data acquisition screen for UL's morning "Cold Velocity" shows a red cell and also indicates that the inlet temperature is out of control limits. The "Poke-yoke" method thus alerts the tunnel operator of potential problems.

Control

Once the upper and lower limits for the different processes were identified through the define-measure-analyze-improve phases, regular calibration and verification steps were included to ensure that the Steiner Tunnel system remained in control and was stable. The control charts and Poke-yoke techniques were made part of the improved Steiner Tunnel processes. An automated program was designed to walk through the calibration process each month, consistently, step by step. In addition, several simple checks were incorporated into the calibration process, identified as "Diagnostic Indicators." These parameters were identified as being critical to the consistent performance of the test. If the diagnostic tests fall within the control limits, the technician can be

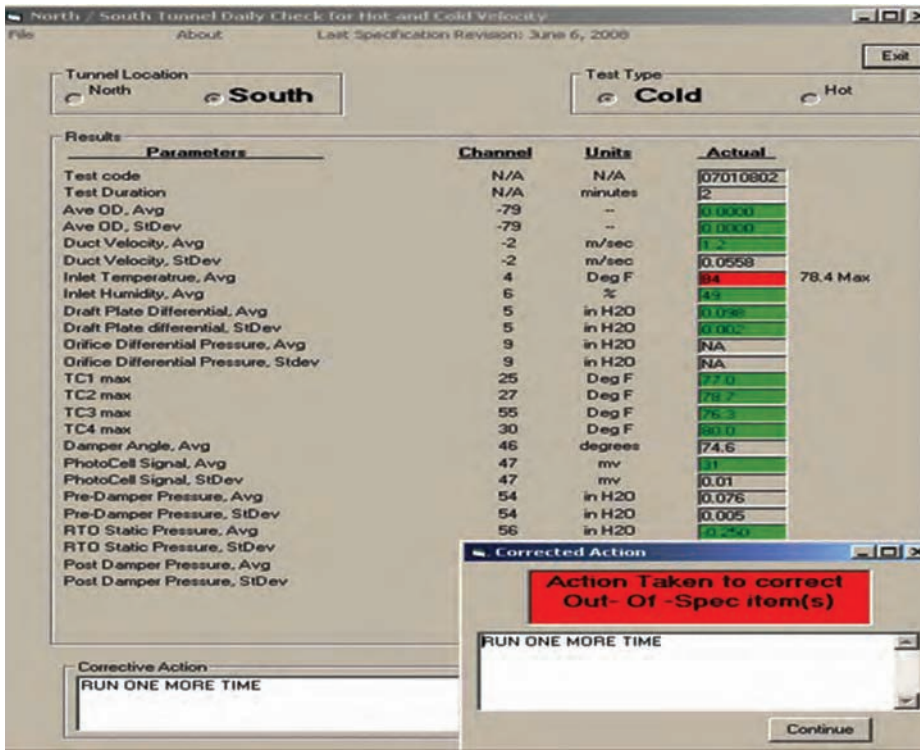


FIG. 3—*Data Automation Program for Daily Check Up.*

confident to continue to the next step in calibration. If the result is outside the limits, we can be fairly confident that there is some *assignable* cause.

Results

Process Control

Using the DMAIC method, UL has incorporated the use of SPC, DOE, and Poke-yoke to monitor, improve, and control Steiner Tunnel performance. The monthly calibration using reference materials including red oak, cement board, and heptane now provide a regular verification of the Steiner Tunnel performance. SPC assists in monitoring systems and subsystems daily between these monthly calibrations. One significant benefit of the DMAIC process was the ability to identify the natural variation in the Steiner Tunnel and to incorporate process control charts for the critical systems (e.g., air flow, gas flow, and so on). Thus, abnormal conditions and trends are readily identified so that corrective actions can be taken. This also prevents unnecessary changes to the system.

Figure 4 shows SPC charts for daily tests on air flow differential pressure across the inlet draft plate with the burner off (cold) and on (hot) and furnace

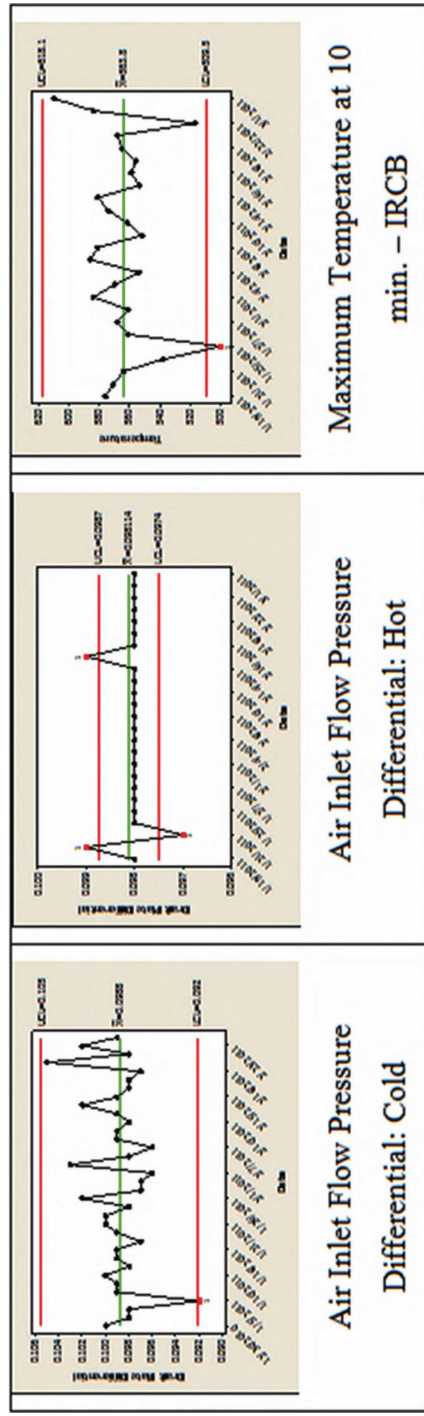


FIG. 4—*SPC Charts Showing Stability and Control of Steiner Tunnel.*

air temperature at 10 min for the inorganic reinforced cement board test. The process control charts in the figure demonstrate the identification of abnormal characteristics indicated by the red, out of control data points that may need operator attention. These include data outside of the control limits, as well as systemic trends that may be developing.

Another verification test used by UL is the heptane pan test to check the smoke measurement system. This test is not currently in the standard and is presented as a tool to help measure and reduce uncertainty of the smoke measurement. In this test, 310 g of laboratory grade heptane is used in a round metal pan. The pan is placed near the burner, and the heptane is ignited using an electronic igniter. The total smoke generated from the test is used to monitor the consistency of the smoke measurement system.

The measured percentage obscuration-minutes (% obs-min) results, before and after implementation of the process control improvements, are as shown in Fig. 5.

Once the sources of variation were identified (e.g., precise quantity of heptane, temperature of the tunnel prior to test, air flow velocity), better controls were implemented resulting in reduced standard deviation. The standard deviation was reduced from 13.9 to 4.6 for the example shown. It is noted that there is an increase in the mean for the second case. Some of this shift is accounted for in the elimination of the low outliers and some to the improved test method.

Fig. 6 shows control charts after implementing the process control monitoring for the calculated Flame Spread and Total Smoke (% obs-min) for red oak calibration material. The control charts demonstrate stability of the measurement system over time.

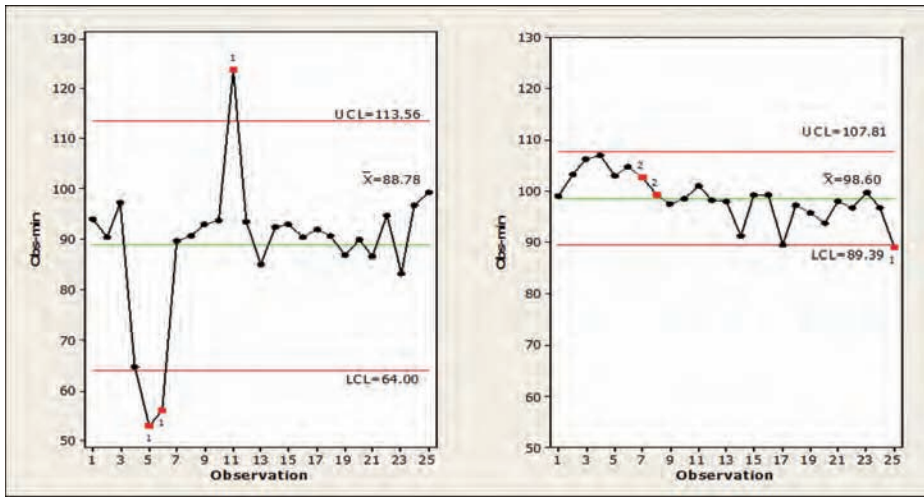


FIG. 5—*Control Charts Showing Improvement of Heptane Test— % Obscuration-Minutes.*

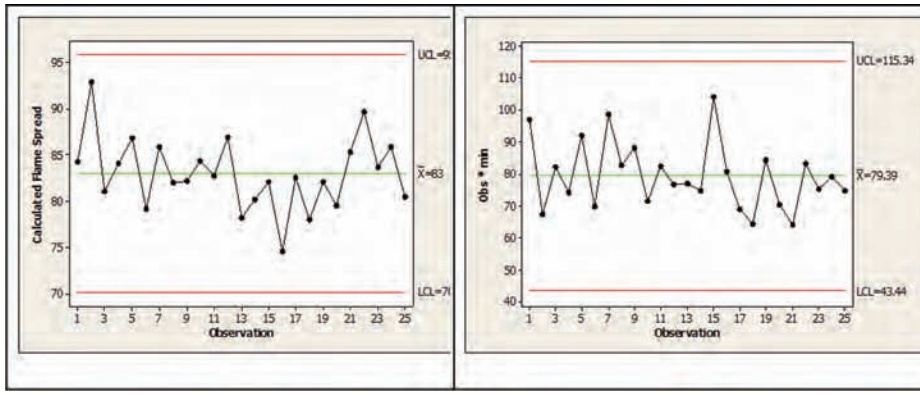


FIG. 6—*Control Charts for Red Oak.*

An important realized benefit from the DMAIC method is that it enabled a reduction in the number of verification tests required by approximately 40 %. This has resulted in less downtime and increased confidence of the Steiner Tunnel performance.

Gage r&R

To assess the results of DMAIC improvements, a Gage r&R study was conducted. Repeatability refers to the ability of an operator to repeat the same measurement and get the same result. Reproducibility refers to the ability of two operators to take the same measurement and get the same result. Two components of variation may be identified when measuring repeatability and reproducibility (commonly called Gage r&R): (i) the variation between materials; and (ii) the variation of the measurement system. The first type is also called part-to-part variability. The second type of variation in the measurement system can further be subdivided into that due to the operator and that which is intrinsic to measurement (e.g., measurement resolution).

A good measurement system will be able to provide discrimination between different materials (e.g., E84 Class A versus Class B materials); and should also be able to measure the same material time and time again with statistically similar results.

UL recently completed a Gage r&R study where six different building materials were selected representing a range of product types and fire performance. Each material was tested four times using two operators; each completing two tests for each material. The tests were conducted in a random order. The experimental design is shown in Table 1. The MINITAB statistical software [6] was used to set up the experimental design and then analyze the Gage r&R study. The test results for calculated flame spread (CFS) and calculated smoke developed (CSD) are also presented in Table 1.

To assess the results of the Gage study, the primary metric is termed percentage repeatability and reproducibility, % r&R. This metric results from a calculation of the standard deviations due to the measurement system divided by

TABLE 1—*Gage r&R Design.*

Material	Test Sequence	Operator	CFS	CSD
3.5″ Kraft-faced fiberglass insulation R13	2	X	1178	31.9
3.5″ Kraft-faced fiberglass insulation R13	6	Y	1217	35.1
3.5″ Kraft-faced fiberglass insulation R13	14	Y	1238	31.9
3.5″ Kraft-faced fiberglass insulation R13	18	X	1333	31.7
1/2″ Gypsum wallboard	4	Y	0	0
1/2″ Gypsum wallboard	13	X	0	0
1/2″ Gypsum wallboard	16	X	0	0
1/2″ Gypsum wallboard	19	Y	0	0
3/4″ Acoustic ceiling tiles	3	Y	12	18.8
3/4″ Acoustic ceiling tiles	15	X	14	19.4
3/4″ Acoustic ceiling tiles	22	Y	12	20.8
3/4″ Acoustic ceiling tiles	23	X	12	20.9
0.125″ Medium density fiberboard paneling	5	Y	142	138.4
0.125″ Medium density fiberboard paneling	11	X	128	154.7
0.125″ Medium density fiberboard paneling	12	X	125	132.3
0.125″ Medium density fiberboard paneling	21	Y	140	128.9
Reinforced plastic	1	Y	29	382.6
Reinforced plastic	8	X	28	388.2
Reinforced plastic	9	Y	26	394.9
Reinforced plastic	20	X	30	423.3
1″ Foil faced polyisocyanurate foamed plastic	7	X	26	91.4
1″ Foil faced polyisocyanurate foamed plastic	10	Y	29	86.4
1″ Foil faced polyisocyanurate foamed plastic	17	Y	42	83.5
1″ Foil faced polyisocyanurate foamed plastic	24	X	29	102.1

the total standard deviation (as a percentage) determined by the study. Because variability from the measurement system should be small relative to the total variability, a typical criterion for acceptability is 10 % or less [5]. The results from this study indicated a % r&R of 6.5 % for CFS and 6.3 % for CSD, indicating a good measurement system.

To better illustrate the results, Fig. 7 shows graphically the relative contribution of the useful variation (between parts, or samples) versus the error in the measurement system. This chart of components of variation is the standard graphical output of the MINITAB software. The chart shows that the Gage r&R error is extremely low, which means that there is low variability in the measurement system. The high part-to-part variation indicates that the measurement system is very capable of detecting variation between samples.

The results from the study were also used to investigate the Steiner Tunnel's ability to discriminate between materials with different characteristics (e.g., Class A, Class B, Class C). The average and standard deviation for CFS for the six materials in the Gage r&R are presented in Table 2. The average and standard deviation for CSD for the six materials in the Gage r&R are presented in Table 3.

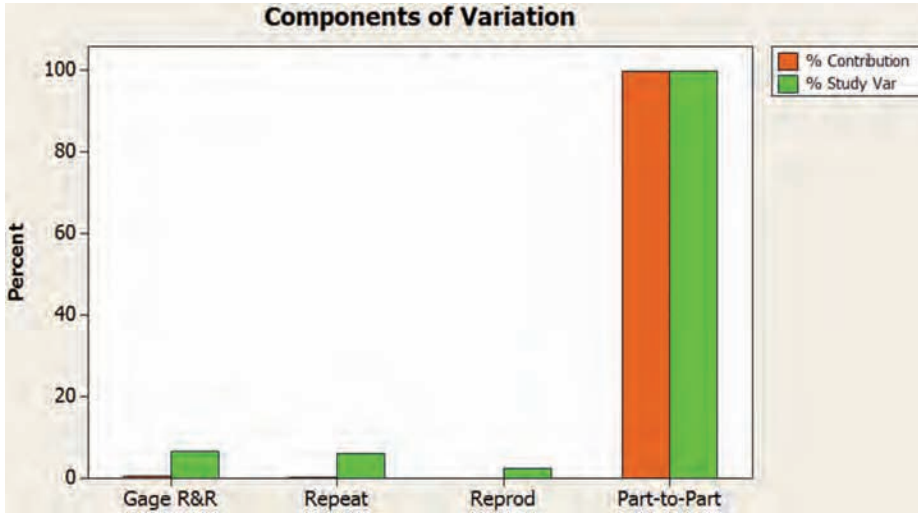


FIG. 7—*Gage r&R Results for Components of Variation.*

TABLE 2—*Comparison of CFS.*

Material	Average	Standard Deviation	% Standard Deviation
3.5″ Fiberglass insulation R13	1242	66	5.3
1/2″ Gypsum wallboard	0	0	0.0
3/4″ Acoustic ceiling tiles	13	1	8.9
0.12″ Medium density fiberboard paneling	134	16	11.8
Reinforced plastic	28	2	6.2
1″ Foil faced poly isocyanurate foamed plastic	32	7	22.5

TABLE 3—*Comparison of CSD.*

Material	Average	Standard Deviation	% Standard Deviation
3.5″ Fiberglass insulation R13	33	1.6	5.0
1/2″ Gypsum wallboard	0	0.0	0.0
3/4″ Acoustic ceiling tiles	20	1.0	5.2
0.12″ Medium density fiberboard paneling	139	11.4	8.3
Reinforced plastic	397	18.1	4.6
1″ Foil faced poly isocyanurate foamed plastic	91	8.2	9.0

The Gage r&R study results showed that the significant portion of variation was due to the samples themselves and not attributed to the measurement system or the person conducting the test. The study confirmed to UL that the Steiner Tunnel, as a measurement system, is very capable measuring within the range that we need for the products under evaluation, and that the tunnel can indeed differentiate between different products and materials that have different characteristics.

Conclusion

This paper was presented at the "Symposium on Uncertainty in Fire Standards and What to Do about It." The ASTM E84 test has a long history of providing fire professionals with useful surface flammability and smoke data on a variety of materials. It is understood that fire tests are relatively expensive, destructive, and time consuming. These factors limit our ability to collect the quantity of data that rigorous statistical analysis ideally requires. Still, test laboratories have the responsibility to provide and assess levels of test equipment performance. This paper described methods and tools that can be used for a variety of fire test methods and specifically outlined some of the work UL has conducted over the years to improve measurement uncertainty for UL 723/ASTM E84.

The paper discussed the following:

1. The use of the Lean/Sigma DMAIC method to better understand and control sources of variation and monitor, improve, and control Steiner Tunnel performance.
2. The use of tools and methodologies used within the DMAIC method such as SPC, DOE, and Poke-yoke.
3. The use of SPC and Gage r&R to confirm the results of DMAIC based improvements and gain confidence in the Steiner Tunnel as a measurement system.

For the Steiner Tunnel apparatus and its complexities, identifying problems, analyzing root causes, and implementing improvements and controls is difficult. The methods, tools, and results provided in this paper provide some examples of what can be done about the uncertainty of results using the Steiner Tunnel and UL 723/ASTM E84.

Acknowledgments

This paper was made possible through the support of Juan Amador of UL's Lean Sigma Office. The experiments performed in this study were conducted with the support of Phil Pastor, Albert Hislop, Sly Smith, Doug Zamisky, Joseph Mack, David Weinstein, and James Smith.

References

[1] ASTM E84-10, 2010, "Standard Test Method for Surface Burning Characteristics of Building Materials," Vol. 4.07, ASTM International, West Conshohocken, PA.

[2] UL 723, 2008, *Test for Surface Burning Characteristics of Building Materials*, 9th ed. Underwriters Laboratories Inc., Northbrook, IL.
[3] Laymon, R. K., "Assessing the Burning Characteristics of Interior Finish Materials—Standard Test Method for Surface Burning of Building Materials," *Fire Protection Engineering*, 4th Quarter, 2004, pp. 42–44.
[4] Hirscler, M. M., "Use of the Steiner Tunnel for Fire Testing in North America," *Proceedings of the Fire and Materials 2011 Conference*, San Francisco, CA, January 31–-February 2, 2011, Interscience Communications, London, UK.
[5] McCarty, T., Daniels, L., Bremer, M., and Gupta, P., *The Six Sigma Black Belt Handbook*; McGraw-Hill, New York, 2004.
[6] *Minitab Version 16.1.1*. (2010). Minitab Inc., State College, PA.

Reprinted from JTE, Vol. 39, No. 6
doi:10.1520/JTE103916
Available online at www.astm.org/JTE

Joe Urbas[1]

Precision of the Cone Calorimeter and ICAL Test Methods

ABSTRACT: Repeatability and reproducibility are typically determined for standard fire test methods on the basis of an interlaboratory test program (round-robin) to define their precision. Significant differences in precision were found in the past between the repeatability and reproducibility of small-scale and intermediate- and large-scale fire test methods that all utilize oxygen calorimetry to measure heat release rate. Repeatability of heat release rate related measurements with small-scale apparatus has been found to be significantly better than the repeatability of the intermediate and large-scale apparatuses. In this paper, the results of two round-robins are compared on the basis of relative repeatability standard deviations and relative reproducibility standard deviations calculated for individual materials for individual test parameters. The first one was a cone calorimeter round-robin and the second was an intermediate-scale calorimeter (ICAL) round-robin. The objective and subjective factors that might have contributed to the differences between the two test methods in both repeatability and reproducibility were analyzed. The most important factors that caused the differences in the ICAL round-robin were higher theoretical uncertainty of the ICAL, inadequate pre-round-robin calibrations, the small numbers of participating laboratories and samples tested, prevalence of fire-retardant-treated materials, and possible failure of some of the laboratories to fully comply with the standard. These factors, especially the fact that the testing procedures and apparatuses in the participating laboratories may not have been in complete accordance with the standards, indicate that the results of the round-robins did not reflect the real precision of the test methods. However, the terminology using the two

Manuscript received April 15, 2011; accepted for publication July 19, 2011; published online August 2011.
[1] Univ. of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223.

Cite as: Urbas, J., "Precision of the Cone Calorimeter and ICAL Test Methods," *J. Test. Eval.*, Vol. 39, No. 6. doi:10.1520/JTE103916.

components of precision, repeatability and reproducibility, is maintained in this paper except where a specific distinction is noted. Recommendations are made on how to improve the results of future fire test round-robins.

KEYWORDS: precision, repeatability, reproducibility, cone calorimeter, intermediate scale calorimeter

Introduction

The round-robin results of two relatively new fire test methods are compared in this paper. Both methods are designed to make heat release rate (HRR) and related measurements using oxygen consumption calorimetry. The two methods are the cone calorimeter [1] and the intermediate-scale calorimeter (ICAL) as described in a now outdated International Standardization Organization (ISO) Technical Report 14696 [2].

Precision of fire test methods has typically been determined for standard fire test methods developed within the ASTM and the ISO. Repeatability and reproducibility, or very similar repeatability standard deviation (s_r) and reproducibility standard deviation (s_R) are used as measures of precision. Sometimes, the results are expressed in terms of relative repeatability standard deviation (RSD_r) and relative reproducibility standard deviation (RSD_R), which are calculated as s_r or s_R multiplied by 100 and divided by the mean of the individual measurements, and also expressed in percentages. Interlaboratory test programs (round-robins) are conducted to determine the two precision parameters. Samples of the same material are tested by a number of laboratories according to a standard test procedure. A number of replicates of the same material are tested by each laboratory, and the results are compared across laboratories.

Procedures for conducting a round-robin and analyzing the results, to calculate the precision parameters, are described in ASTM E691 [3] and ISO 5725 [4]. According to these standards, repeatability is an indication of the variation in the results that is expected when multiple specimens of the same material are tested in a single laboratory under the same conditions within a short period of time. Reproducibility is a measure of the variation that can be expected when tests are performed in different laboratories using identical test procedures.

The results of the round-robins in the respective reports [5,6] are expressed as repeatability and reproducibility, the two components of precision as defined in ISO 5725 [4]. However, the modern fire test methods involve a large number of measurements performed by relatively complicated equipment. These measurements are used to calculate the desired test results, such as HRR using oxygen consumption calorimetry methodology. As will be discussed later in this paper, testing laboratories are, for various reasons, not always able to follow the standard procedures and maintain the equipment to follow the standards exactly. Therefore, the results of the two compared round-robins, and some other recent fire method round-robins, may not reflect the real precision of the test methods as the measurements used to calculate repeatability and reproducibility are not performed as required in ISO 5725 [4]. Also, the numbers of materials used, and the number of participating laboratories may be lower than required by that standard [4], which may directly affect the way the statistical

laboratory stragglers and outliers are determined. However, to maintain consistency with the round-robin reports used for this analysis [5,6] and some other recent round-robins described in Ref. [7], we will continue to use the terms precision, repeatability, and reproducibility to express the results of the round-robins in this paper except where specifically indicated.

Janssens analyzed the variability of the types of tests that utilize oxygen calorimetry for HRR determination in Ref. [7]. His analysis was focused on the comparison of overall RSD_r values for various oxygen consumption-based test methods as determined in round-robins. Theoretical uncertainty and the discrepancies between precision and uncertainty, which are also discussed in Janssens' paper, will not be addressed here.

In his paper, Janssens [7] analyzed the repeatability of results for six cone calorimeter and five intermediate- and large-scale calorimeter round-robins. He observed significant differences between the repeatability standard deviations determined for the cone calorimeter and those for the intermediate- and large-scale calorimeters. The results for the cone calorimeter were significantly better than for the intermediate- and large-scale calorimeter tests (with one exception). He attributed the poor results for the intermediate- and large-scale tests to the following:

1. It is much more difficult to perform large-scale fire tests in a consistent manner.
2. A relatively small number of laboratories participated in the intermediate- and large-scale test round-robins in comparison to the cone calorimeter ones.
3. Because of the high cost of large-scale round-robin testing, it is difficult in practice to meet the ASTM 691 [3] and ISO 5725 [4] minimum requirements for the number of participating laboratories.
4. A small number of participating laboratories tend to adversely affect the precision estimates, partly because it is more difficult to identify statistical outliers.
5. Too many fire-retardant materials in the material selection typically increases the variation of the results.
6. Some participating laboratories may not have followed the standard.

Janssens suggested that the cone calorimeter and one intermediate-scale test method round-robins were "…examples of carefully conducted round-robins with competent participating laboratories." To improve the situation with the intermediate- and large-scale test methods, he suggested that a proficiency program be introduced within the ASTM E5 Committee. The program would introduce the following pre-round-robin calibrations and measurements similar to those performed prior to some very successful round-robins:

- "Determine transport times, response characteristics, noise, and drift of individual instruments."
- "Perform multiple gas burner and/or liquid pool fire calibrations to reduce bias systematic errors and determine uncertainty."
- "Perform tests with standard reference materials, if available, to verify the uncertainty estimates."

The objective of this paper is to compare the complete precision results (of both repeatability expressed as RSD_r and reproducibility expressed as RSD_R) of

two round-robins, one conducted on a small-scale test method, and the other one conducted on an intermediate/large-scale test method. The first one was conducted on the cone calorimeter (included in Janssen's analysis [7]) [5]. The second was an ICAL round-robin conducted under the auspices of ISO TC92 in 2008 [6]. Janssens, however, compared the repeatability results of various round-robins only with theoretical uncertainty. The emphasis here is on identifying any differences in the results between the two different scale test methods, discussing the reasons for the differences, and, if required, making suggestions for improving the results of the two test methods as determined in the round-robins, to achieve results closer to real precision. Also, the intention is to determine whether Janssens' suggestions for the improvements of round-robin results of oxygen consumption-based fire test methods also apply to the two round-robins analyzed in this work, where reproducibility is included in addition to repeatability.

Description of Test Methods and Round-Robin Protocols

Cone Calorimeter

Representatives of seven project sponsors and four participating commercial laboratories together with a project coordinator formed the "Cone Calorimeter Round-Robin Consortium" (Consortium) to organize the project. The project sponsors funded the entire project including the cost of testing and the coordinator. A more detailed description of this project is provided in Ref. [5]. The Consortium defined the scope of the project, selected the materials to be tested, confirmed the participating laboratories, defined the calibration procedure, and confirmed the test protocol. All tests were conducted according to ASTM E1354 (1994 version) [1]. Each laboratory conducted three replicate tests on each of the round-robin materials at a single irradiance exposure. The following specifications were followed for the round-robin testing:

- Testing is in the horizontal orientation.
- Irradiance exposure is 75 kW/m^2.
- Retainer frame is used.
- A 2 s data collection interval is used.
- Data are collected for a period of 15 min plus the time delay of the oxygen analyzer.
- The heat release rate per unit area is calculated using $0.0088 m^2$ as the specimen surface area.

Sixteen materials were tested in the round-robin. All the materials were building materials. They are listed in Tables 1–5. The materials were carefully selected to include a wide range of material densities (20–2233 kg/m^3), a wide range of expected HRR measurements, and some fire retardant-treated materials.

The following test parameters were measured (following the standard test procedure described in Ref. [1]), and analyzed: time to ignition, 1-min average HRR, 3-min average HRR, 5 min average HRR, peak HRR, total heat released, maximum 60-s average HRR, and average heat of combustion.

TABLE 1—*Time to ignition RSD_r and RSD_R values.*

Material	Mean (s)	RSD_r (%)	RSD_R (%)
Untreated lumber—spruce pine fir (SPF)	8.3	25.3	25.3
Cellulosic insulation—FR treated	2.9	11.4	54.5
5/8 in. type X gypsum board	14.8	4.4	8.7
Paper-faced glass-wool (18 kg/m^3)	1.8	35.6	42.8
Sprayed fire resistant material (SFRM)	159.3	4.0	5.6
Polyvinyl (PVC) floor tile	22.0	4.4	5.8
Red oak (flame spread index = 100)	11.8	8.5	8.7
Paper covered particleboard	10.0	22.7	22.7
Imitation stucco	23.0	9.0	11.0
Textile wall covering on noncombustible board	9.8	7.2	10.1
OSB	8.5	18.6	21.4
FR plywood	8.2	11.1	13.9
Polyisocyanurate foam	1.3	0.8	44.6
High pressure decorative laminate bonded to particleboard	13.8	14.3	15.7
FR-treated lumber (flame spread index < 25)	24.4	70.5	75.7
Rubber foam insulation	1.4	22.8	38.9
Overall	20.1	16.9	25.3

TABLE 2—*Peak HRR RSD_r and RSD_R values.*

Material	Mean (kW/m^2)	RSD_r (%)	RSD_R (%)
Untreated lumber—SPF	226.4	7.0	12.8
Cellulosic insulation—FR treated	79.4	9.0	35.9
5/8 in. type X gypsum board	134.3	2.2	20.1
Paper-faced glass-wool (18 kg/m^3)	196.5	7.1	32.4
SFRM	21.7	11.8	15.1
PVC floor tile	181.1	2.9	9.2
Red oak (flame spread index = 100)	272.1	2.7	6.9
Paper covered particleboard	258.2	6.1	9.7
Imitation stucco	470.1	5.6	9.9
Textile wall covering on noncombustible board	225.0	4.2	20.3
OSB	330.9	6.4	9.0
FR plywood	169.1	8.5	10.0
Polyisocyanurate foam	138.8	22.2	91.8
High pressure decorative laminate bonded to particleboard	298.2	22.0	24.2
FR-treated lumber (flame spread index < 25)	82.8	15.3	18.1
Rubber foam insulation	121.2	3.4	19.5
Overall	200.4	8.5	21.6

TABLE 3—*One-minute average HRR RSD_r and RSD_R values.*

Material	Mean (kW/m^2)	RSD_r (%)	RSD_R (%)
Untreated lumber—SPF	169.1	7.1	7.1
Cellulosic insulation—FR treated	55.2	5.3	17.0
5/8 in. type X gypsum board	42.2	2.7	7.7
Paper-faced glass-wool (18 kg/m^3)	50.0	4.7	11.5
SFRM	14.8	11.4	38.6
PVC floor tile	118.5	5.8	13.4
Red oak (flame spread index = 100)	203.6	1.9	5.9
Paper covered particleboard	186.8	4.2	6.4
Imitation stucco	236.8	2.6	7.6
Textile wall covering on noncombustible board	70.3	6.7	11.4
OSB	239.8	2.5	9.9
FR plywood	104.8	14.4	20.7
Polyisocyanurate foam	76.8	10.3	13.9
High pressure decorative laminate bonded to particleboard	189.3	6.6	11.7
FR-treated lumber (flame spread index < 25)	66.9	14.7	20.8
Rubber foam insulation	79.7	3.3	9.7
Overall	119.0	6.5	13.3

TABLE 4—*Three-minute average RSD_r and RSD_R values.*

Material	Mean (kW/m^2)	RSD_r (%)	RSD_R (%)
Untreated lumber—SPF	125.7	8.9	8.9
Cellulosic insulation—FR treated	49.0	4.1	6.8
5/8 in. type X gypsum board	16.3	7.6	10.6
Paper-faced glass-wool (18 kg/m^3)	19.6	5.8	21.7
SFRM	8.6	14.8	23.5
PVC floor tile	94.5	2.0	11.0
Red oak (flame spread index = 100)	176.9	1.6	3.4
Paper covered particleboard	187.8	3.1	4.8
Imitation stucco	219.4	1.8	2.6
Textile wall covering on noncombustible board	30.9	7.3	10.1
OSB	216.0	1.6	3.5
FR plywood	88.8	9.4	12.1
Polyisocyanurate foam	56.3	1.7	4.4
High pressure decorative laminate bonded to particleboard	165.5	4.7	6.6
FR-treated lumber (flame spread index < 25)	60.3	11.4	14.1
Rubber foam insulation	58.1	2.1	4.5
Overall	98.4	5.5	9.3

TABLE 5—*Total heat release RSD_r and RSD_R values.*

Material	Mean (MJ)	RSD_r (%)	RSD_R (%)
Untreated lumber—SPF	108.8	6.4	6.6
Cellulosic insulation—FR treated	21.2	5.9	6.2
5/8 in. type X gypsum board	4.1	31.7	34.1
Paper-faced glass-wool (18 kg/m^3)	3.9	8.9	11.8
SFRM	2.5	14.4	28.8
PVC floor tile	22.4	5.8	12.3
Red oak (flame spread index = 100)	138.7	2.34	7.15
Paper covered particleboard	111.7	3.9	4.9
Imitation stucco	133.8	3.5	5.0
Textile wall covering on noncombustible board	7.4	15.9	15.9
OSB	106.8	4.0	4.9
FR plywood	53.4	5.7	5.7
Polyisocyanurate foam	17.8	9.4	9.9
High pressure decorative laminate bonded to particleboard	119.2	3.0	6.2
FR-treated lumber (flame spread index < 25)	47.7	15.3	18.6
Rubber foam insulation	26.3	2.0	4.3
Overall	57.9	8.6	11.4

The coordinator distributed the samples to the participating laboratories. Prior to distribution, instructions were sent to the laboratories as to which side of the material specimens should be tested. Four commercial laboratories (three located in the United States and one in Canada) participated in the project.

Before the round-robin tests were conducted, all the laboratories completed the "round-robin pre-testing procedure" consisting of calibration procedures and preliminary tests on two calibration materials (black PMMA with a relatively high heat release output and mineral ceiling board with a relatively low heat release output). These tests were conducted according to the ASTM E1354 [1] test method and the round-robin test protocol described previously. The calibration procedures were those described in the October 1996 version of ISO 5660-1 (document ISO TC92/SC1/WG5/N232) [8] and were distributed to the laboratories together with some additional requests for information in the form of a questionnaire. A new heat flux meter was circulated to the laboratories for heat flux verification. The initial results of the pretesting procedures revealed some inconsistencies among the laboratories. The main inconsistencies were: use of aluminum foil of significantly different thicknesses (the thickness of aluminum foil affects the time at which the foil melts, which influences the exposure areas of specimens and other aspects of specimen burning, and thus the test results), significantly varying specimen wrapping procedures, use of the shutter for specimen protection in two of the four laboratories (the use of the shutter is a mandatory part of the test procedure designed to protect the sample before it is exposed to heat radiation at the start of a test), use of different values

for specimen surface area in per square meter calculations, and significant deviation of retainer frame thickness and weight from the standard in one laboratory (slight deviations of retainer frame thickness result in the exposed surface area different from the one prescribed in the instructions). All the inconsistencies among the laboratories were corrected before the tests on the calibration materials were conducted. The tests conducted on black PMMA and mineral ceiling board revealed additional problems with the cone calorimeter hardware and software in most of the laboratories. Some of these problems were not identified in the pretest procedures and could not be expected under the existing ASTM standard. The problems included malfunctioning of the heater temperature control that resulted in large oscillations of heat output, inappropriate gas train composition that caused incorrect carbon dioxide concentration measurements, and scanning at irregular time intervals because of incorrect programming of data acquisition software. The irregular scanning led to an incorrect calculation of derived HRR values in one case and the inability to calculate mass loss rate and effective heat of combustion values as a function of time in two cases. (The data reduction software uses regular time intervals.) A major effort was made by the laboratories to correct most of the problems identified in the calibration tests. However, not all the problems could be corrected because of time constraints. The laboratories conducted additional PMMA tests after improvements were made on their cone calorimeters. These tests demonstrated that the problems that could not be corrected did not significantly impact the test results. Also, inappropriate results (stragglers, outliers) would be eliminated as part of the statistical analysis. Therefore, the Consortium approved the start of the round-robin testing in all the participating laboratories. The uncorrected problems probably contributed to somewhat higher repeatability and reproducibility than their true values. The coordinator visited and was present at the start of testing of the round-robin materials in all the laboratories. Heat flux meters in the laboratories were compared with the coordinator's meter, and problems discovered in the pretesting procedure were discussed and corrections suggested during the visits.

Intermediate-Scale Calorimeter

The ICAL round-robin was conducted under the auspices of the International Standardization Organization, Technical Committee 92, and Subcommittee 1 as a part of the effort to develop the ICAL test method from a Technical Report (ISO/TR 14696 [2]) into a full ISO standard. The round-robin was conducted strictly on a volunteer basis.

Test method repeatability and reproducibility information is required to be included in ISO standards. To develop this information, a round-robin was organized among the laboratories that had operational ICAL units built according to the ISO/TR 14696 specification [2]. This test method was technically identical to the ASTM E 1623 standard [9]. The test protocol strictly followed the standard specification. No additional requirements were made of the laboratories beyond those described in the standard.

Of seven laboratories that were invited, four agreed to participate voluntarily in the project. Three of these laboratories were commercial laboratories located in the

United States, and one was a research laboratory located in Japan. The coordinator was the ISO/TC92/SC1/WG7 ICAL project leader [6]. Of eight initially planned materials, four material samples were donated to the project and used for the testing. A high cost of a large amount of required samples and their distribution to the laboratories prevented some of the potential material suppliers from donating the samples. The ISO/TC92/SC1/WG7 decided to conduct the round-robin with the available materials [6]. The materials that were used for the round-robin are paper faced poly-iso-cyanurate foam (PIRF), oriented strandboard (OSB), fiberglass reinforced plastic (FRP), and fire retardant-treated plywood (FR plywood).

Before the round-robin was conducted, all the laboratories were asked to complete a diagnostic test of their equipment and a preliminary test on one specimen of oriented strandboard. The diagnostic test was designed to determine any problems with the heat release rate measurement equipment and instrumentation, and any deviations from the ISO/TR 14696 standard procedure [2]. Two out of four laboratories did not complete these tests. After some problems and inconsistencies were eliminated following the review of the results and recommendations by the project coordinator, three out of four laboratories conducted the preliminary test. The purpose of this test was to identify problems with the ICALs that were not detectable by the diagnostic test, such as potential problems with the ignition, smoke, and mass loss measurements. The results of this test were also reviewed by the coordinator, and recommendations were made to the laboratories of how to correct the problems. According to the laboratories, some of these problems were corrected. As indicated earlier, because of the voluntary basis of the project, the suggested diagnostic tests and the preliminary test were not completed in all the participating laboratories. Also, the testing equipment in the laboratories was not inspected for inconsistencies with the standard prior to the round-robin testing.

Three out of four materials/products that were included in the round-robin were fire retardant treated. Unfortunately, this condition could not be avoided because of great difficulties in obtaining the samples from the sponsors. More appropriate samples for the project could not be obtained because of a high cost associated with the purchase of and large amounts of 1 m^2 by 1 m^2 specimens required for the ICAL testing. The materials/products were tested in triplicate at the irradiances of 25 and 40 kW/m^2. The laboratories measured and reported time to ignition, peak HRR, 1- and 3-min average HRR, and total HRR, all determined as prescribed by the standard [2]. The effective heat of combustion and average smoke extinction area were intended to be measured. However, one laboratory was unable to perform the mass loss and smoke measurements. Consequently, the mass loss rate, effective heat of combustion, and smoke measurements were not included in the statistical analysis.

Results of Statistical Analysis

Cone Calorimeter Round-Robin

Mean values, RSD_r, and RSD_R values for the five measured parameters of the 16 tested materials are presented in Tables 1–5. All of these parameters were

TABLE 6—*Time to ignition RSD_r and RSD_R values.*

Material	Mean	RSD_r	RSD_R
PIRF 40 kW/m^2	6.8	11.2	32.2
OSB 40 kW/m^2	32.2	18.6	55.2
FRP 40 kW/m^2	152.5	8.6	29.1
FR plywood 40 kW/m^2	56.2	58.0	84.8
PIRF 25 kW/m^2	14.8	8.9	31.9
OSB 25 kW/m^2	159.6	9.0	33.0
FRP 25 kW/m^2	ND	ND	ND
FR plywood 25 kW/m^2	ND	ND	ND
Overall	70.4	19.0	44.4

determined at 75 kW/m^2 and as described in the description of the cone calorimeter test method. Overall values (averages of all the material results) for each of the parameters are given at the bottom of each table.

ISO ICAL Round-Robin

The same parameters as in the cone calorimeter were determined in the ICAL at two exposure fluxes, 40 and 25 kW/m^2. Mean values, RSD_r, and RSD_R values for the four tested materials are presented in Tables 6–10. Overall values (averages of all the material results) for each of the parameters are given at the bottom of each table.

Discussion of Results

The differences in relative RSD_r and relative RSD_R between the cone calorimeter and the ICAL are shown in Figs. 1 and 2, respectively.

The differences between the relative repeatability standard deviations and relative reproducibility standard deviations for the cone calorimeter are shown

TABLE 7—*Peak HRR RSD_r and RSD_R values.*

Material	Mean	RSD_r	RSD_R
PIRF 40 kW/m^2	243.9	14.3	41.1
OSB 40 kW/m^2	342.3	7.9	25.4
FRP 40 kW/m^2	233	22.5	34.5
FR plywood 40 kW/m^2	90.6	17.2	17.2
PIRF 25 kW/m^2	195.3	7.7	29.5
OSB 25 kW/m^2	239.5	25.3	30.5
FRP 25 kW/m^2	ND[a]	ND	ND
FR plywood 25 kW/m^2	ND	ND	ND
Overall	224.1	15.8	29.7

[a]ND—values were not determined because the material did not ignite.

TABLE 8—*One-minute average HRR RSD_r and RSD_R values.*

Material	Mean	RSD_r	RSD_R
PIRF 40 kW/m^2	131.0	14.7	17.5
OSB 40 kW/m^2	232.7	6.8	30.0
FRP 40 kW/m^2	153.1	17.9	20.0
FR plywood 40 kW/m^2	57.1	26.5	26.5
PIRF 25 kW/m^2	114.3	5.0	6.6
OSB 25 kW/m^2	168.9	14.3	37.8
FRP 25 kW/m^2	ND[a]	ND	ND
FR plywood 25 kW/m^2	ND	ND	ND
Overall	142.9	14.2	23.1

[a]ND—values were not determined because the material did not ignite.

TABLE 9—*Three-minute average HRR RSD_r and RSD_R values.*

Material	Mean	RSD_r	RSD_R
PIRF 40 kW/m^2	69.8	22.6	22.6
OSB 40 kW/m^2	172.3	5.7	20.4
FRP 40 kW/m^2	ND[a]	ND	ND
FR plywood 40 kW/m^2	49.0	29.2	31.8
PIRF 25 kW/m^2	65.0	9.1	20.8
OSB 25 kW/m^2	127.3	11.4	31.0
FRP 25 kW/m^2	ND	ND	ND
FR plywood 25 kW/m^2	ND	ND	ND
Overall	96.7	15.6	25.3

[a]ND—values were not determined because the material did not ignite.

TABLE 10—*Total heat release RSD_r and RSD_R values.*

Material	Mean	RSD_r	RSD_R
PIRF 40 kW/m^2	21.2	31.7	34.4
OSB 40 kW/m^2	56.3	13.1	30.7
FRP 40 kW/m^2	10.4	29.1	69.1
FR plywood 40 kW/m^2	42.2	22.3	39.6
PIRF 25 kW/m^2	31.5	17.3	54.8
OSB 25 kW/m^2	47.7	24.9	44.7
FRP 25 kW/m^2	ND[a]	ND	ND
FR plywood 25 kW/m^2	ND	ND	ND
Overall	34.9	23.1	45.6

[a]ND—values were not determined because the material did not ignite.

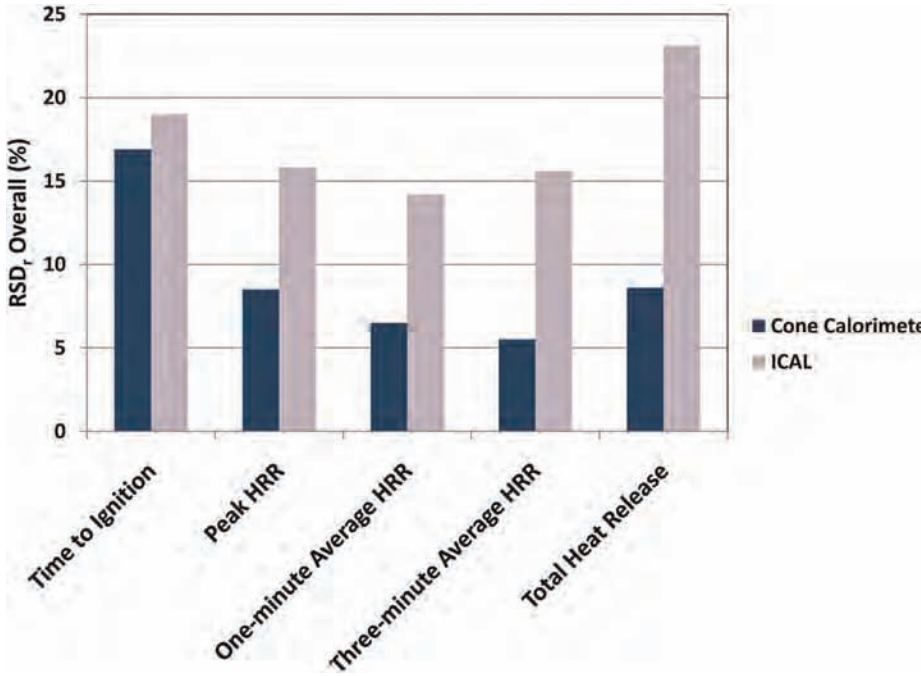


FIG. 1—*Overall relative RSD_r values for the cone calorimeter and ICAL round-robins.*

in Fig. 3; the differences between relative repeatability standard deviation and relative reproducibility standard deviations for the ICAL are shown in Fig. 4.

Overall RSD_r values for the ICAL are higher than for the cone calorimeter. The differences are small for the time to ignition (a factor of 1.1) and are somewhat higher for the HRR related parameters (a factor of 1.9 for peak HRR, 2.1 for 1-min average HRR, and 2.8 for 3-min average HRR). The ICAL total heat release values are higher than the cone calorimeter values by a factor of 2.7.

Overall RSD_R values for the ICAL are also higher than for the cone calorimeter. However, the difference in values is less than for the repeatability standard deviation values. The differences for the time to ignition are for a factor of 1.7, a factor of 1.4 for peak HRR, 1.7 for 1-min average HRR, and 2.7 for 3-min average HRR. The ICAL total heat release values are higher than the cone calorimeter values by a factor of 4.

The differences between the RSD_r and RSD_R for the cone calorimeter indicate higher values for the RSD_r for both the cone calorimeter and the ICAL. The values for the cone calorimeter are higher by a factor of 1.5 for the time to ignition, 2.5 for the peak HRR, 2.0 for 1-min average HRR, 1.7 for the 3-min average HRR, and 1.3 for the total heat release. The differences for the ICAL are a factor of 2.3 for the time to ignition, 1.9 for the peak HRR, 1.6 for the 1-min average HRR, 1.6 for the 3-min average HRR, and 2.0 for the total heat release. Obviously, the absolute RSD_R values are much higher in comparison to the RSD_R

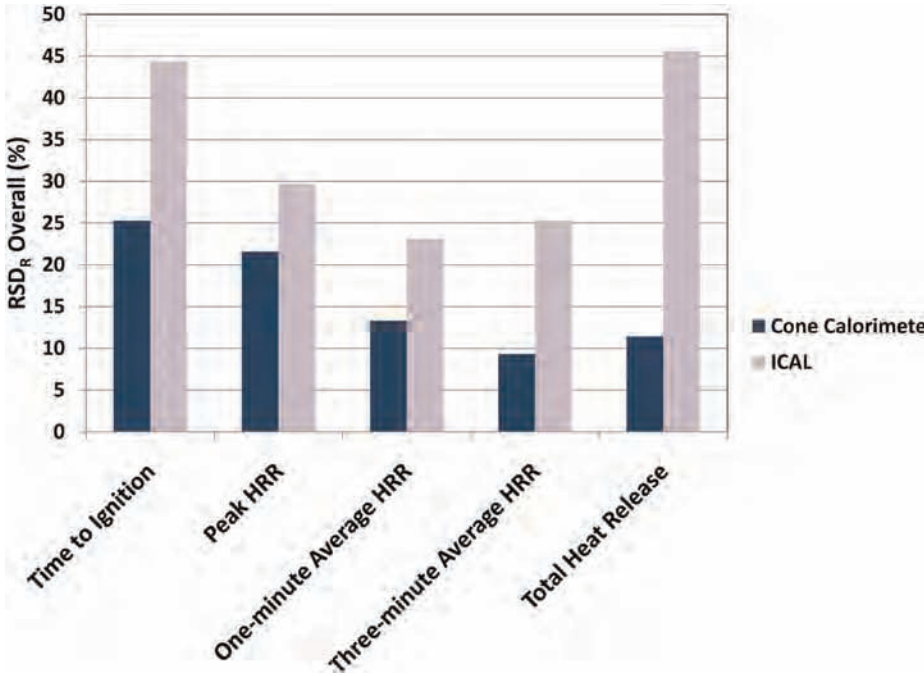


FIG. 2—*Overall relative RSD_R values for the cone calorimeter and ICAL round-robins.*

values for the ICAL than they are for the cone calorimeter as indicated previously.

The reasons for the differences can to some extent be deduced from the data shown in Tables 1–10 and are discussed further in the following:

The first obvious reason that can be determined from Tables 1 is the difference in the number of tested materials. Any very high numbers affect the average value more when the number of tested materials is low than when it is high.

The second reason is that a much higher percentage of FR-treated materials was used in the ICAL round-robin than in the cone calorimeter round-robin. These materials show higher relative RSD_r values than the non-FR-treated materials. In the cone calorimeter round-robin, the typical FR-treated materials were: FR plywood, polyisocyanurate foam, and FR-treated lumber, which amounts to 19% of the materials. Although cellulosic insulation was also an FR-treated material, this material is well homogenized, and does not show a high result variation. The typical FR containing materials used in the ICAL round-robin comprised 75% of the materials used and were: PIRF, FRP, and FR plywood. The differences for the measured parameters between the overall relative RSD_r values that include FR-treated materials and the overall relative RSD_r values that exclude the FR-treated materials (Diff. RSD_r) are shown in Fig. 5. The differences for the measured parameters between the overall relative RSD_R values that include FR-treated materials and the overall relative RSD_R values that exclude the FR-treated materials (Diff. RSD_R) are shown in Fig. 6.

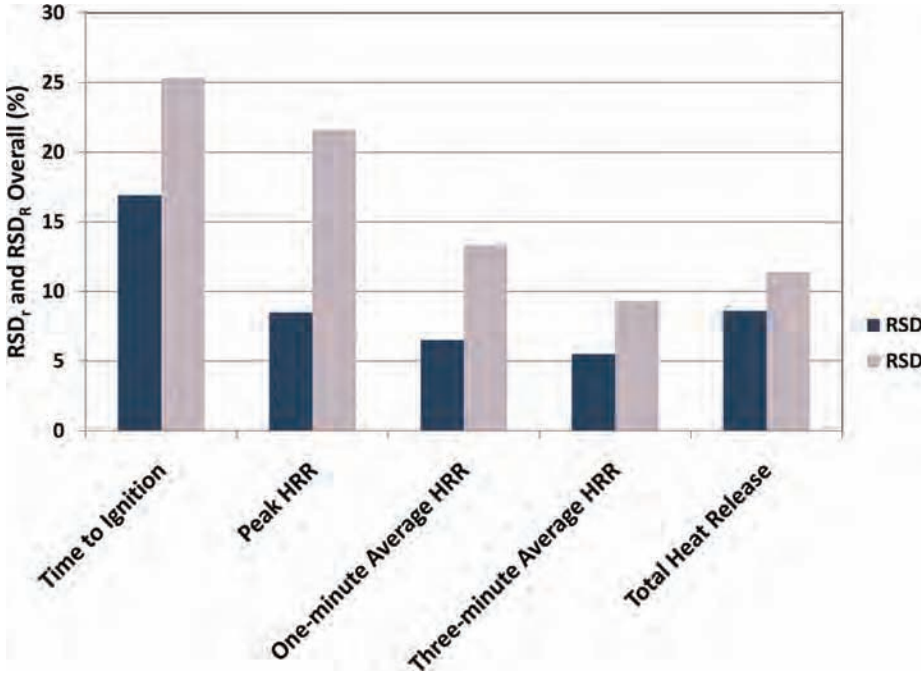


FIG. 3—*Differences between the overall relative RSD_r values and relative RSD_R values for the cone calorimeter.*

Figure 5 shows that the effect of the percentage of FR-treated materials on the overall RSD_r values in the ICAL round-robin is significantly higher than in the cone calorimeter round-robin. The exception is the peak HRR value, which can probably be attributed to the low number of materials tested in the ICAL round-robin. This result indicates that the high percentage of FR-treated materials in the ICAL round-robin in comparison to the cone calorimeter round-robin contributed to the higher RSD_r values determined in the ICAL round-robin. Figure 6 indicates that no clear effect of the percentage of the FR-treated materials on the RSD_R values can be identified based on the available data.

The third reason that can be inferred from the results of the round-robins shown in Tables 1–10 is related to the selection of the materials in the ICAL round-robin in association with the heat flux exposures at which the materials were tested. No results were obtained for two FR-treated materials (FRP and FR plywood) at 25 kW/m^2 exposure because the materials did not ignite and generated no heat. It is not known how this affected the results other than significantly reducing the number of measurements available for the statistical analysis and the calculation of the overall RSD_r and RSD_R values.

The reasons for additional differences that cannot be deduced from the statistical analysis results are based on the round-robin coordinator's observations, and include the following:

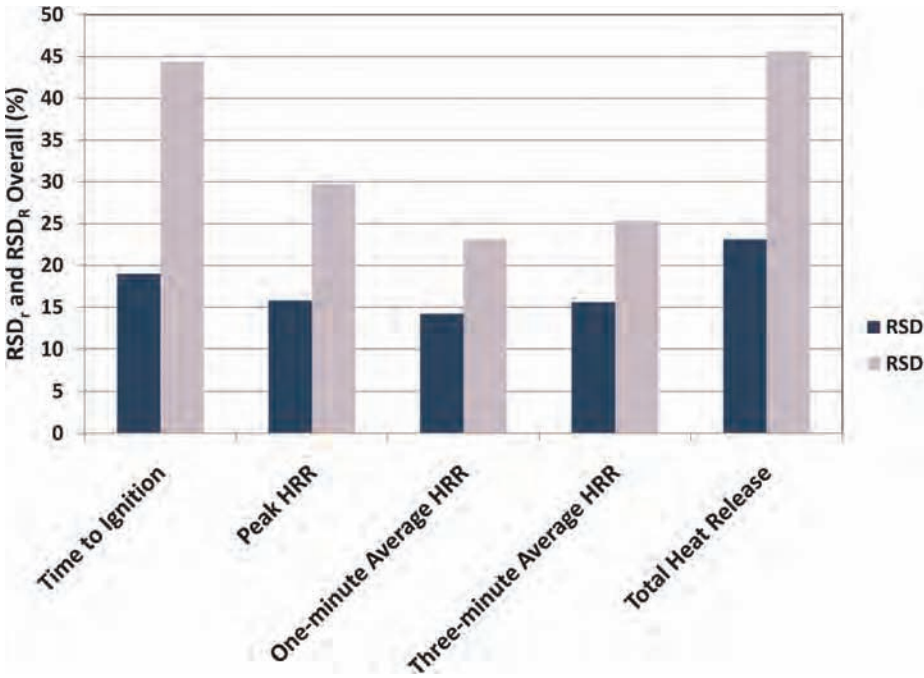


FIG. 4—*Differences between the overall RSD_r values and RSD_R values for the ICAL.*

1. The theoretical uncertainty is higher for the ICAL than for the cone calorimeter because of additional measurements involved in the ICAL test procedure for the determination of HRR and the derived parameters. The most significant one is the dynamic uncertainty related to the measurement of the natural gas flow to the radiant panel. This parameter needs to be subtracted from the total HRR measured by oxygen consumption.
2. The test procedure used in the ICAL round-robin was the one described in the Technical Report version of the ISO document (ISO/TR 14696) [2]. Significant improvements, now included in the ISO ICAL standard [10], were made to the apparatus. It is expected that the precision results of any future ICAL round-robins will be significantly better.
3. The pre-round-robin calibrations that were required to be performed prior to the ICAL round-robin were not performed in most of the laboratories because of time constraints and cost reasons. Any apparatus and testing procedure problems were neither identified nor corrected prior to the round-robin tests. These calibrations were, however, performed under the coordinator's supervision in the cone calorimeter round-robin.
4. The exposure fluxes in the cone calorimeter round-robin were 75 kW/m^2, whereas the fluxes used in the ICAL round-robin were 25 and 40 kW/m^2. It is known that lower flux exposures lead to larger variation measurements of test results [7].

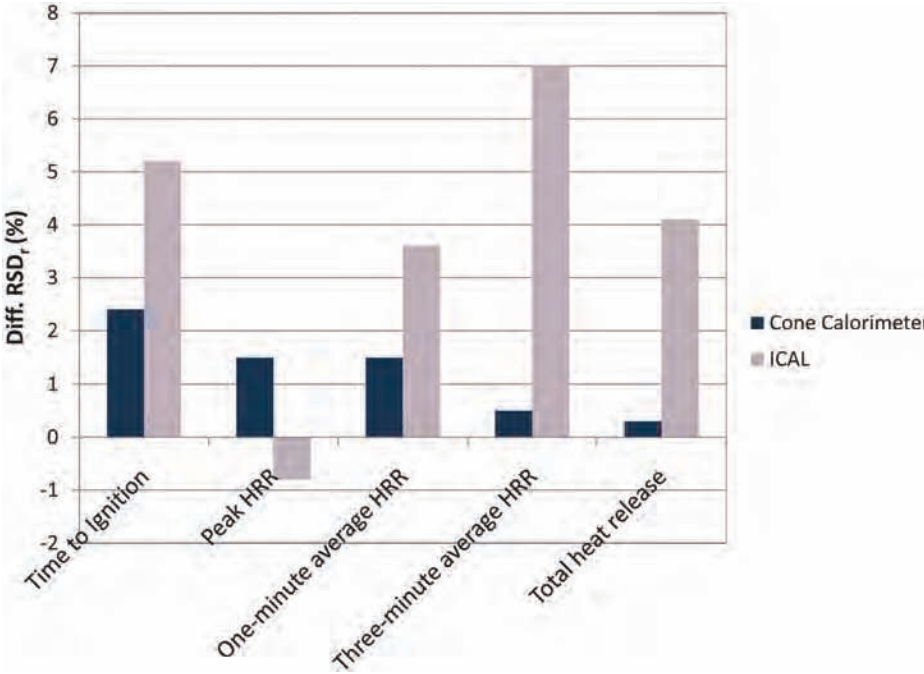


FIG. 5—*Differences for the measured parameters between the overall relative RSD$_r$ values that include FR-treated materials and the overall relative RSD$_r$ values that exclude the FR-treated materials.*

5. Some apparatuses used in the ICAL round-robin may not have fully complied with the standard.
6. Numerous testing parameters in the cone calorimeter round-robin were defined in a testing protocol developed by the project Consortium. These requirements were aimed at defining the testing methodology beyond the one described in the standard. This protocol was not followed in the ICAL round-robin. One of the essential requirements of the cone calorimeter round-robin was to limit the duration of the tests to 15 min. This limit significantly improved the total heat release precision results in that round-robin. The durations of the tests in the ICAL were determined based on the judgments of the laboratory personnel, as required in the standard, and resulted in a larger variation of the total heat release results.
7. Two of the participating laboratories had little prior experience with the ICAL apparatus.
8. The ICAL test utilizes a large-scale calorimeter, which performed relatively poorly in past round-robins. Janssens [7] reports RSD$_r$ values in the ranges of 7%–83% for the peak HRR and 5.2%–94% for the total heat release. Note: It is encouraging that the ICAL round-robin results presented in this paper are better than the results of most of the past

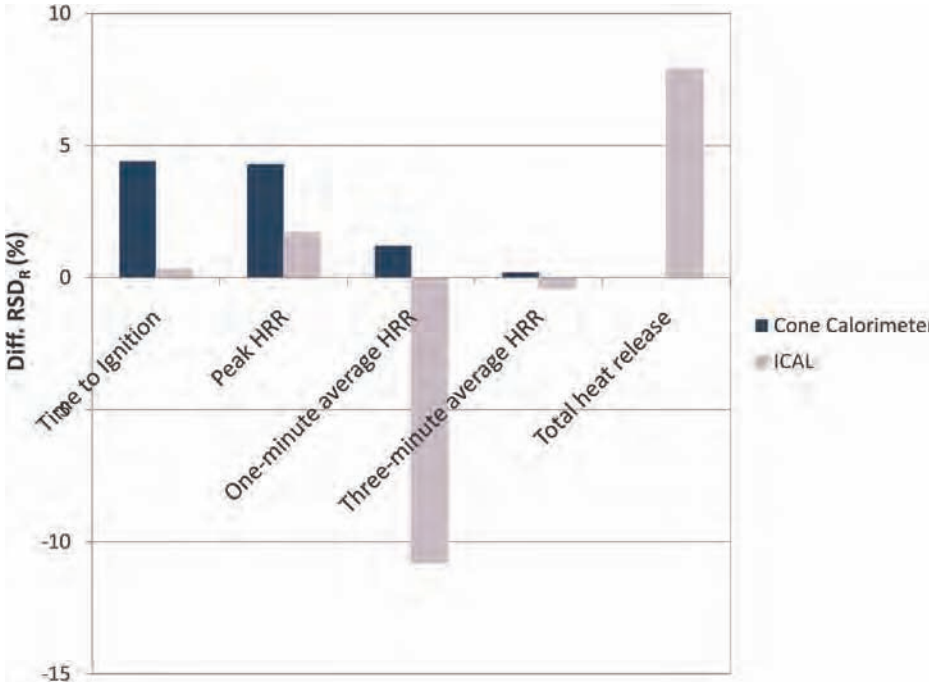


FIG. 6—*Differences for the measured parameters between the overall relative RSD_R values that include FR-treated materials and the overall relative RSD_R values that exclude the FR-treated materials.*

large-scale calorimeter round-robins. (The ranges of RSD_r values were 7.7%–25.3% for the peak HRR and 13.1%–31.7% for the total heat release.)

Suggestions for Improvements

Although the cone calorimeter precision results are better than those of the ICAL, there is still room for improvement. The following improvements are suggested in Ref. [5] to improve the cone calorimeter precision:

- The laboratories operating cone calorimeters should regularly check whether all the components of their apparatus comply with the latest version of the standard.
- A mechanism of control should be established to improve and maintain the high quality of the cone calorimeter operation in commercial and research laboratories.
- Improvements of the standards are necessary based on the findings of this round-robin and other cone calorimeter related projects. Future round-robins may reveal that even more improvements will be necessary.

- Although the cone calorimeter is already developed to a high degree of quality, further improvements based on systematic research and development might simplify its operation and make it less operator dependent. However, the improvements should be selected very carefully and not be overly based on a "black box" approach in which the operator loses control over individual parameters of the apparatus operation. Higher reliability and sufficient flexibility of operation are desired.
- One or more calibration material(s) should be developed and used for frequent evaluation of the cone calorimeter performance. A material bank should be maintained with a supply of the material(s) with known (not necessarily constant) heat release properties.

Large- and intermediate-scale fire test round-robins are expensive. The costs include the laboratory cost associated with large numbers of tests of large samples on relatively complicated test methods, the cost of purchasing or donating large numbers of specimens, and other costs. It has proven very difficult to conduct these large projects on a voluntary basis. The ICAL round-robin and other similar projects noted in Ref. [7] showed that it is very difficult to conduct intermediate- and large-scale fire test round-robins in the way required by ISO 5725 [4] or ASTM E691 [3] to determine precision. Therefore, the results obtained in these round-robins do not reflect the true precision of the test methods. The repeatability and reproducibility values obtained are higher than their true values for an unknown fraction. All efforts should be made in the future to determine true precision of the fire test methods.

The recommendation based on the ISO ICAL round-robin [6] is that future ISO round-robins, especially those involving large test methods, should be adequately funded. Adequate funding allows the projects to be completed within a reasonable time with more laboratories participating (three laboratories dropped out of this project), provides for purchasing and testing appropriate samples, and ensures project coordinator supervision of test apparatuses and test procedures in the participating laboratories.

Summary

The comparison of the cone calorimeter and the ICAL precision data shows higher ICAL overall RSD_r values than those of the cone calorimeter. The differences are small for the time to ignition (a factor of 1.1), and somewhat higher for the HRR related parameters (a factor of 1.9 for peak HRR, 2.1 for 1-min average HRR, and 2.8 for 3-min average HRR). The ICAL total heat release values are higher than the cone calorimeter values by a factor of 2.7.

Overall RSD_R values for the ICAL are also higher than for the cone calorimeter. However, the difference in values is less than for the RSD_r values. The times to ignition differ by a factor of 1.7, a factor of 1.4 for peak HRR, a factor of 1.7 for 1-min average HRR, and a factor of 2.7 for 3-min average HRR. The ICAL total heat release values are higher than the cone calorimeter values by a factor of 4.

Some of the reasons for the difference in precision as determined in the two round-robins are described in the paper. The reasons identified and suggestions for improvements are similar to those suggested by Janssens [7], although his

suggestions refer to repeatability only. The recommendations, based on the analysis performed in this paper, might improve the results of future oxygen consumption-based calorimetry fire test round-robins and their accuracy in everyday operation. The main goal of the following recommendations is to obtain round-robin results that would be as close as possible to true repeatability and reproducibility of the test methods.

1. Ensure funding for samples, round-robin testing, and a competent project coordinator prior to embarking onto a round-robin project.
2. Involve as many laboratories as possible, but not less than required by ISO 5725 [4] to allow proper elimination of outliers and stragglers.
3. Include as many materials as possible, but not less than required by ISO 5725 [4].
4. Select materials with a wide range of HRR and limit the number of fire retardant-treated materials to a small percentage (up to 25%) of all the materials used.
5. Define the testing protocol as precisely as possible and define the time of the test duration.
6. Ensure that the testing equipment, data reduction software, and test methodologies in all the laboratories strictly follow the standard. A project coordinator's visits to all laboratories are useful to ensure compliance.
7. Require pre-round-robin calibrations including: (a) determining transport times, response characteristics, noise, and drift of the individual instrument; (b) performing multiple gas burner or liquid pool fire HRR calibrations to reduce bias systematic errors; and (c) performing tests with standard reference materials to verify uncertainty estimates, or use homogenous materials with known consistent fire test results.
8. Carefully analyze the results of the pre-round-robin calibrations and communicate any problems to the laboratories.
9. Insist that all the problems be corrected before the laboratories start the round-robin testing.
10. Provide to the laboratories detailed instructions describing which side of the materials should be tested and how the samples should be oriented if they include joints or any other nonuniformities.
11. Provide detailed instruction to the laboratories on what/which measurements need to be recorded during the tests, what parameters should reported, and how the results should be reported.

References

[1] ASTM E1354-94, 1995, "Standard Test Method for Heat and Visible Smoke Release Rates for Materials and Products Using an Oxygen Consumption Calorimeter," American Society for Testing and Materials, West Conshohocken, PA.

[2] ISO/TR 14696, 1999, "Reaction to Fire Tests—Determination of Fire Parameters of Materials, Products and Assemblies Using an Intermediate-Scale Heat Release Calorimeter (ICAL)." International Organization for Standardization, Geneva, Switzerland.

[3] ASTM E691, "ASTM Practice for Conducting an Inter-Laboratory Study to Determine the Precision of a Test Method," American Society for Testing and Materials, West Conshohocken, PA.
[4] ISO 5725, 1986, "Precision of Test Methods—Determination of Repeatability and Reproducibility by Inter-Laboratory Tests," International Organization for Standardization, Geneva, Switzerland.
[5] Urbas, J., 2002 "BDMC Interlaboratory Cone Calorimeter Test Program," *J. Fire Mater.*, Vol. 26, pp. 29–35.
[6] Urbas, J., "ICAL Round Robin Report to ISO," Document No. 410, ISO/TC92/SC1/WG7 Large and Intermediate Scale Tests.
[7] Janssens, M. L., 2002, "Variability in Oxygen Consumption Calorimetry Tests," *Thermal Measurements: The Foundation of Fire Standards, ASTM STP 1427*, L. A. Gritzo and N. J. Alvares, Eds., ASTM International, West Conshohocken, PA.
[8] ISO 5660-1, 1996, "Reaction-to-Fire Tests—Heat Release, Smoke Production and Mass Loss Rate—Part 1: Heat Release Rate (Cone Calorimeter Method," International Organization for Standardization, Geneva, Switzerland.
[9] ASTM E1623 - 11, "Standard Test Method for Determination of Fire and Thermal Parameters of Materials, Products, and Systems Using an Intermediate Scale Calorimeter (ICAL)", ASTM International, West Conshohocken, PA.
[10] ISO 14696, 2009, "Reaction-to-Fire Tests—Determination of Fire and Thermal Parameters of Materials, Products and Assemblies Using an Intermediate-Scale Calorimeter (ICAL)," International Organization for Standardization, Geneva, Switzerland.

Reprinted from JTE, Vol. 40, No. 1
doi:10.1520/JTE103915
Available online at www.astm.org/JTE

Morgan J. Hurley[1]

Uncertainty in Fire Protection Engineering Design

ABSTRACT: The *SFPE Engineering Guide to Performance-Based Fire Protection* defines "uncertainty" as "the amount by which an observed or calculated value might differ from the true value." In engineering, there are two types of uncertainly: epistemic and aleatory. Epistemic uncertainty is uncertainty due to lack of (complete) knowledge. Aleatory uncertainty is uncertainty due to random variation. At present, no single, widely accepted methodology exists for dealing with uncertainty in fire protection engineering. This paper reviews the sources of uncertainty that are present in fire protection engineering and methods that are used to address them.

KEYWORDS: uncertainty, fire models, model validation

Introduction

Fire protection engineering frequently requires predictions of what might occur in a fire. Predictions might include the fire size and growth, resulting hazards, response of fire protection, or effects on people or property of the hazards created.

These predictions are generally developed through the use of fire models, which range in sophistication from simple algebraic correlations that can be solved with a hand calculator, to zone models or lumped parameter models that represent a space as a small number of elements (typically two or three), to computational fluid dynamics or field models which approximate a space as a large number of discrete volumes, potentially thousands or millions.

However, all fire models are approximations of reality. The difference between these approximations and reality can be called "uncertainty." While

Manuscript received April 15, 2011; accepted for publication August 24, 2011; published online October 2011.

[1] Society of Fire Protection Engineers, 7315 Wisconsin Ave., #620E, Bethesda, Maryland 20814.

Cite as: Hurley, M. J., "Uncertainty in Fire Protection Engineering Design," *J. Test. Eval.*, Vol. 40, No. 1. doi:10.1520/JTE103915.

uncertainty is omnipresent, it was not until the turn of the present century that it was identified as an important issue to address in fire protection engineering. Since then, there have been several suggestions for how to address this uncertainty. This paper reviews these proposals.

What is "Uncertainty"?

The *SFPE Engineering Guide to Performance-Based Fire Protection* [1] defines "uncertainty" as "the amount by which an observed or calculated value might differ from the true value." In engineering, there are two types of uncertainly: epistemic and aleatory.

Epistemic uncertainty is uncertainty due to lack of (complete) knowledge. For example, it may not be possible to calculate precisely a post-flashover fire temperature due to imprecision of models and input values. It is possible to develop estimates; however, these estimates will not be exact, even if the input variables are well-known. Epistemic uncertainty can be reduced by gaining additional information or knowledge.

Aleatory uncertainty is uncertainty due to random variation. For example, sprinklers may have slight variation in activation temperature and sensitivity, even if they come from a single assembly line. Listing standards may specify that these factors vary by no more than a certain amount, but it is impossible to reduce this uncertainty without changing the way the sprinklers are manufactured. Whether a coin toss will result in heads or tails is a form of aleatory uncertainty. No amount of additional research or information would reduce aleatory uncertainty.

In fire protection engineering, uncertainty primarily arises from the following sources [1].

Theory and Model Uncertainties

Fire models are frequently used in fire protection engineering analysis. While these models are essential tools in fire protection engineering, they also may introduce uncertainty. Calculation methods may be based upon an incomplete or less than perfect understanding of the underlying science.

For example, an empirically derived correlation may be based upon data from experiments that were conducted under a certain range of conditions. Without additional analysis, it is unknown how well the correlation would perform outside of this range. Similarly, in some cases, simplifications may be incorporated into a model or correlation based on an assumption that it will only be used within a limited set of conditions.

Data and Model Inputs

The data used as input into models or correlations is subject to uncertainty. For example, it may not be possible to precisely determine some material properties. The flashpoint of a flammable liquid provides an example—for any given flammable or combustible liquid, use of an open cup apparatus may yield a different measurement than use of a closed cup apparatus.

As input data becomes more complex, uncertainty can also become magnified. Burning rates used as input into a fire model illustrate this concept. If the

data used was gathered in a manner that is not similar to the scenario being modeled, uncertainty may result.

Calculation Limitations

Even in cases where the science underlying a model or correlation is well understood, simplifications may be incorporated to enable the model or correlation to be solved in a reasonable amount of time. For example, there are several techniques available for modeling room fires. Both zone models and field models are approaches for solving simultaneous equations of conservation of mass, momentum, and energy. The discretization technique employed in both types of models results in solutions that are approximate but never exact.

Fire Scenario Selection

Fire scenarios are generally used for the analysis of performance-based designs. Uncertainty may be introduced if the design fire scenarios selected do not represent the range of fires that might occur.

Uncertainty in Human Behaviors

Human behavior is a highly stochastic phenomenon. There are a number of actions that people may take in response to a fire. Many analyses do not address these actions. Where human behavior is considered, uncertainty can be introduced due to its stochastic nature.

Uncertainties in Risk Perceptions, Attitudes, and Values

This type of uncertainty pertains to the selection of objectives and performance criteria. In some cases, it may be difficult to accurately assess the level of safety that is desired by the stakeholders of a fire protection design. Determining the tolerable risk to life from fire in a building is a classical example.

All of these sources could be expected to contribute uncertainty to fire protection design. However, uncertainties in risk perceptions, attitudes and values would only be of concern in cases where a full risk assessment is performed. For comparative analysis, whether performed on a hazards (deterministic) basis or on a risk assessment basis, uncertainties in risk perceptions, attitudes and values are generally not of concern. This is because a proposed alternative design is generally compared to a code compliant design, and the underlying prescriptive code is generally accepted as providing a level of fire performance that is acceptable to society.

Uncertainty can arise from a number of sources. There are a number of techniques that can be used to compensate for uncertainty in fire protection engineering.

Notarianni

One of the first people in the fire protection engineering profession to recognize the importance of formally addressing uncertainty was Kathy Notarianni. Notarianni observed that decisions based on fire protection engineering calculations

generally impacted public health, safety and welfare, so it is important to address uncertainty to increase confidence in engineering decisions [2].

Notarianni focused on the issue of switchover. Switchover occurs when a predicted outcome changes from a determination that is acceptable to a determination that is not acceptable. Notarianni identified seven aspects that could cause switchover.

1. Uncertainty in the selection of performance criteria. Performance criteria are the metrics that are used to determine the acceptability of a performance-based design. Performance criteria are frequently developed on an ad hoc basis, and the values used may not adequately address all people or items that a design purports to protect.
2. Uncertainty in design fire selection. Design fires represent the "loads" that a performance-based design must withstand. Like performance criteria, design fires are typically selected by a design team. While designers generally wish to select bounding fire scenarios, the design fires that are used may not bound all possible fires that could occur.
3. Human behavior in fire. Many performance-based fire protection designs are intended to protect people from fire. However, human behavior in fire is highly stochastic. Mobility and cognitive disabilities can affect a person's actions leading to or following a fire. Failure to address this variability can lead to designs that are not safe for some fraction of the population.
4. Fire models have limitations. As simplifications of reality, all predictive fire models have some limitations. Failure to recognize and address these limitations can result in predictions that differ from what would physically occur.
5. Model output. Even when fire models are used within their limitations, there is uncertainty in their predictions. While fire models can produce output with several significant figures of precision, their output is an approximation of what might occur in a fire.
6. Working outside one's expertise. Performance-based design may involve a number of analyses, such as predictions of fire size, detector response, suppression effectiveness, human behavior, or response of other targets. Not all fire protection engineers are proficient in all of these areas. In some cases, an engineer may unwittingly work in an area outside of their expertise. This can lead to well-intended decisions that are actually unsafe.
7. System reliability. No method of fire protection is totally reliable. Published estimates of system reliability can vary widely [3]. How well this reliability is addressed, particularly for deterministic analyses, can lead to uncertainty.

Notarianni proposes a 10-step process for addressing uncertainty in performance-based designs. This process is placed in the context of the broader fire protection design process established in the *SFPE Engineering Guide to Performance-Based Fire Protection* [1].

Steps 1–3: Define Scope, Goals, and Objectives

The scope contains an identification of occupant and building characteristics. If changes in the occupant or building characteristics could cause switchover, the values at which this could occur should be identified.

The project goals address loss tolerance. The tolerability of loss can vary depending upon the perspective of the stakeholder. Because goals form the basis for performance criteria, changes in goals could change the metrics that are used to judge a performance-based design. If a value, attitude or risk perception could vary, and if this variation could cause switchover, this possibility should be identified. Similarly, objectives, which are a quantification of loss tolerance, could vary, and the ability of this variance to cause switchover should be identified.

Step 4: Statement of Performance

Notarianni recommends developing probabilistic performance criteria, which would be used to judge the adequacy of a design. Probabilistic performance criteria would contain four elements: probability, time, value, and threshold.

Step 5: Develop a Distribution of Design Fire Scenarios

Notarianni recommends developing a distribution of design fire scenarios that represent the types of fires that might occur and their frequencies. The distribution should be large enough that the tails of the distribution are included. Developing this distribution consists of five major elements: (1) selecting calculation procedures, (2) identifying uncertain input parameters, (3) generating a distribution of design fire scenarios, (4) defining distributions of other input parameters, and (5) selecting a sampling method.

Step 6: Develop Trial Designs

Notarianni does not recommend any treatment of uncertainty at this step.

Step 7: Evaluate Candidate Designs

Using a distribution of design fire scenarios, each weighted by its probability of occurrence, will result in a distribution of outcomes. This distribution will also be affected by variability in input values and reliability of fire safety systems. The distribution of outcomes would be compared to the probabilistic performance criteria.

Some conclusions will be more sensitive to uncertainty than others. Notarianni recommends considering the potential impact of uncertainty on each of the four elements of performance criteria: probability, time, value, and threshold. Those aspects that could cause switchover should be given attention.

Notarianni recommends comparing a minimum of two different design strategies. One could be a code-compliant design when an equivalency is sought, or both could be designs developed as proposed solutions. The distribution of outcomes for each solution would be compared to determine which solution would be preferred and under which conditions.

Many values that are used in the evaluation of design strategies will be subject to uncertainty. Notarianni recommends ranking these values in order of those that contribute the most uncertainty to the final outcome. The values that contribute the most uncertainty would then be given the most attention.

Step 8: Judge the Design Based on the Four Elements of Performance

A design would only be acceptable if all four elements of performance are satisfied: probability, time, value, and threshold. Notarianni further recommends that for analyses where the uncertainty is very large, societal, and policy input would be necessary.

Steps 9 and 10: Select a Final Design and Prepare Documentation

Notarianni does not recommend any treatment of uncertainty at these steps.

Lundin

Lundin [4] evaluated the predictive capability of smoke transport models. As part of this evaluation, he investigated the model error and the effect of model error on predictions. By developing an understanding of the model error, it would be possible to adjust model results to improve the quality of predictions.

Lundin also observed that model predictions are subject to uncertainty, and failure to consider this uncertainty can result in designs that are either unsafe or unnecessarily expensive to build. Lundin defined four sources of uncertainty: (1) resources, (2) assumptions and decisions, (3) models, and (4) input data. Resources are generally externally imposed: such as the time that is available to conduct an analysis or the quality of the tools that are available.

Lundin's approach is based on comparison of model prediction with test measurements, which others have called "validation." He suggests the following criteria for conducting this comparison:

- An appropriate model is selected for comparison with observed results.
- The model is used correctly and the output is correctly interpreted.
- The model output data are consistent with the form of the observed data.
- The experiments used are representative of the scenario being modeled.
- The experiments were not used in the development of the model.
- The spatial locations of the observations and prediction are coincident.

Model error is determined through the comparison of experimental data with model predictions. Additionally, by considering possible variation in input and output data (e.g., spatial and temporal variation), a probabilistic representation of model error can be developed. A confidence value (e.g., 95%) can be selected, and only predictions that exceed performance criteria at the level of confidence desired would be considered acceptable.

SFPE Engineering Guide to Substantiating a Fire Model for a Given Application

The most recent guidance on how to compensate for uncertainty in fire models is the *SFPE Engineering Guide to Substantiating a Fire Model for a Given Application* [5]. This guide was written to apply to fire models ranging in sophistication from algebraic correlations to computational fluid dynamics models.

The guide recommends a five-step process for determining the suitability of a fire model. These steps include defining the problem, selecting a candidate

model, verifying and validating the model, determining the impact of uncertainty and user effects on the model results, and finally, documenting the model evaluation. This process is illustrated in Fig. 1.

Definition of the Problem of Interest

The first step described in the guide is to define the problem of interest. This definition includes the relevant phenomena and key physics, collecting available information and identifying the analysis objectives.

Available information might include geometry, timeline, materials relevant to the problem of interest and relevant material properties, initial and boundary conditions, and analysis objectives.

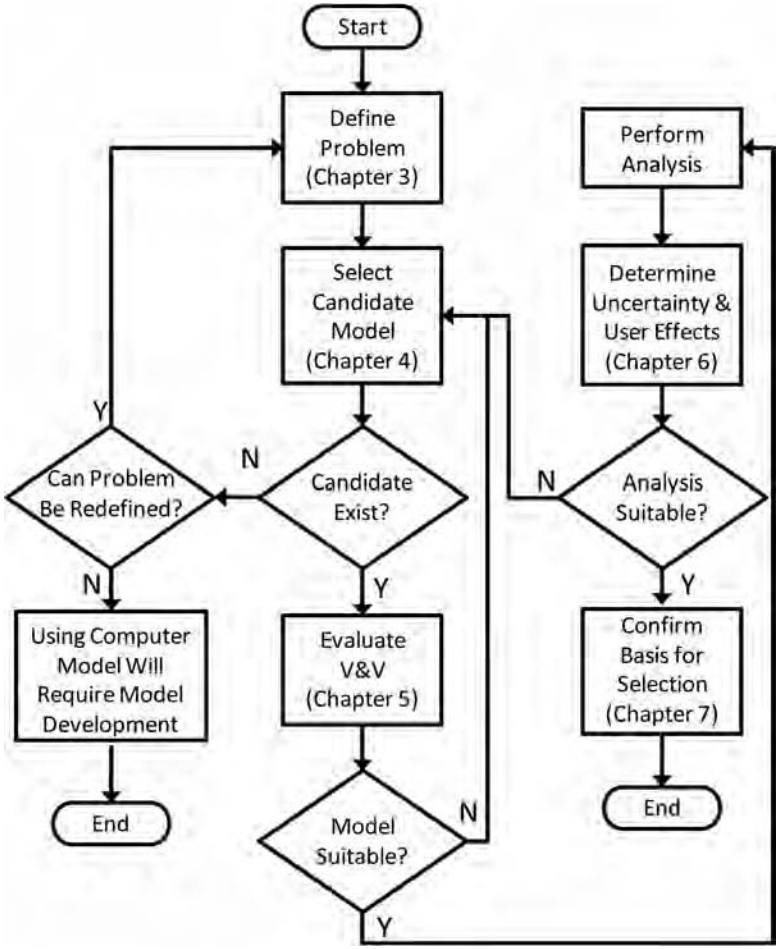


FIG. 1—*Fire Model Selection Flow Chart (see Ref. [5]).*

Select a Candidate Model

The next step identified in the guide is to select a candidate model. The guide recommends three major considerations for selecting a candidate model: determining the available model inputs, identifying the desired model outputs, and determining the available resources.

Determining the available model inputs requires the model user to identify the inputs that are available for a given problem and to identify the inputs that are not available but must be acquired before proceeding with an analysis. In cases where not all of the input data required by a model are available, the guide recommends three possible options:

- Perform preliminary calculations aimed at identifying the value of that specific variable.
- Make a reasonable assumption as to the range of values that the input could have and then perform a sensitivity analysis to determine the effect on the model results of changing that variable.
- Conduct experiments to obtain a value for the input.

There is often more than one model available that may provide a sufficiently accurate solution to a problem. In such cases, model selection can be based upon the resources that are available to run the model. While a CFD model may provide benefits, such as the ability to more exactly represent the geometry of a space and better visualization tools than a zone model, it may not always provide a more accurate solution to a problem. If time constraints and lack of computer resources prohibit a thorough sensitivity analysis using CFD, then for some problems it might be more appropriate to use a zone model or algebraic model in order to more thoroughly address uncertainty.

The guide suggests developing a resourcing plan that follows the following steps before starting large fire modeling projects:

- Determine the number of simulations needed to address any sources of uncertainty.
- Determine the amount of time required to run a simulation on the available computational resources.
- Determine whether or not several simulations can be run simultaneously.
- Determine the available time before the project must be completed.

Verification and Validation

Prior to using a model for a particular problem, the guide recommends determining whether the model is capable of generating a useful result. The formal process by which this is demonstrated is verification and validation (V&V).

Verification ensures that the model is working as designed, i.e., that the equations are being properly solved. It essentially is a check of the mathematics. Model verification serves two purposes. First, it ensures that the mathematical equations have been properly implemented. Second, it ensures that the model user understands the assumptions of the model.

The guide suggests that, at a minimum, model users should read the model documentation that describes efforts made by the developers to verify the model. Then, the user should supplement the work performed by the developers

to better address the specific application under consideration. The guide suggests a number of exercises that the model user can perform to supplement the verification efforts of the developers:

- Verify the basic functionality of the model—This typically involves creating simple test cases and comparing the model results to known analytical solutions.
- Verify consistency of input parameters—The user should address the appropriateness of input values, especially as they are used collectively.
- Verify that the input parameters are appropriately used—This generally involves studying the model documentation and diagnostic output.
- Verify the range of validity for input parameter values—Some values of the input parameters are only valid within a certain range. The model user should confirm that the input values are consistent with the underlying physical assumptions or experimental conditions.
- Verify consistency of results—In short, this involves demonstrating that the results make sense.

Verification ensures that the model is working as designed; that the equations are being properly solved. It essentially is a check of the mathematics. Validation is a check of the physics, i.e., whether the equations are an appropriate description of the fire scenario. Most often, validation takes the form of comparisons with experimental test data.

Typically, model validation involves a large amount of data—both in terms of model predictions and experimental measurements—and it can be difficult to succinctly display the results of the study. If all the experimental measurements can be quantified by the same total uncertainty, then a simple graph can be made to summarize the validation exercise. The graphs can indicate the uncertainty in the experimental data. If the model predictions lie within the band defined by the experimental uncertainty, then it cannot be said that the model predictions differ significantly from the measured data.

If the model predictions lie outside the uncertainty interval, this does not necessarily mean that the model is unsuitable. In such cases, the trend in the model's predictive ability needs to be evaluated in the context of the intended use.

User Effects

Once the verification and validation have been conducted, the next step is to focus on the uncertainty that arises in model predictions due to the use of a predictive model. The guide notes that possible sources of uncertainty include definition of the model space or computational domain, simplifying assumptions (in the application of the model), and the choice of input parameters. The result is a propagation of "error" or uncertainty through the model that should be understood, at minimum, at a qualitative level, but preferably, quantitatively.

In addition to uncertainty that exists within the model, the input data can introduce uncertainty into the model calculation. Input data, often based on assumed values or experimental data, is subject to many sources of uncertainty, including uncertainty in theory (for deriving the parameter) and measurement.

Such uncertainty in input imposes a limit on the confidence in model output. Variation in one or a combination of input parameters may substantially alter the model outcome. Treatment of uncertainty in the assumptions and input that define a problem is an important component of analysis that the modeler should address.

The guide suggests several methods of dealing with uncertainty introduced through the use of models.

- Performance Criteria. Fire models are often used as part of a fire protection design process in which the results are evaluated against threshold values of certain quantities or metrics, also known as performance criteria. The conclusions that may be drawn from an analysis are limited by the predictive accuracy of the model as well as the potential uncertainty in the input parameters. Performance criteria thresholds could account for limitations in the models and input.
- Safety Factors. Safety factors and margins of safety are used to provide a buffer to allow for uncertainty in the fire protection design process. A safety factor is a multiplier of a prediction for reference against a threshold or criterion. Safety margins are additive, not multiplicative.
- Sensitivity Analysis. A sensitivity analysis determines the relationships between the uncertainty in the input variables and the uncertainty in the resultant output. A sensitivity analysis provides information regarding how the uncertainty in the output of a model can be apportioned to different sources of variation in the input of a model. Sensitivity analysis allows the identification of those parameters that are most important to the outcome. It does not necessarily provide information regarding the value that should be used, but it can show the impact of using different values.
- Parametric Analysis. In a parametric analysis, a special form of a sensitivity analysis, detailed information of the effect of a certain input variable on model output is examined by systematically varying the input value of that variable, while holding others constant. A parametric analysis may be useful if detailed information regarding the potential variation of the input variable is unknown.
- Bounding. Bounding is a form of sensitivity study that evaluates the consequences of the extremes of possible values of an uncertain input quantity. If the outcome values at both extreme ends of the range of the uncertain input are acceptable relative to some criteria, further sensitivity analysis may be avoided. Bounding can be applied to not only input parameters but also selection of boundary conditions.
- Differential Analysis. For some models or systems, it is possible to solve directly for the partial derivative of the predicted values with respect to each of the input variables. The set of partial derivatives measures the sensitivity of the solution with respect to changes in the input parameters. A differential analysis has the advantages of being very quick and requiring very few resources to implement.
- Power Dependence. Less formal than differential analysis, power dependence assesses the proportionality or power-dependence of a model

target output to an input parameter. By examining the relationship of model output to input, the user will be able to identify the relative importance of the input. As a result, the user may be able to focus on refining the estimate for a "more" important input variable, while accepting perhaps a higher variability in a "less" important variable.

Once the modeler has determined that a model is appropriate for a given application, the modeler would likely use one or more of the methods decried above to address uncertainty that is introduced through use of the model. Since fire models are generally used in the context of a larger analysis—like a design or a forensic investigation—the *SFPE Engineering Guidelines for Substantiating a Fire Model for a Given Application* do not mandate any specific method for addressing uncertainty. The *SFPE Engineering Guidelines for Substantiating a Fire Model for a Given Application* do not specify a single method for addressing uncertainty since other standards or guides that are applicable to the larger analysis might contain requirements for addressing uncertainty.

Discussion

Because of the potential impact on public health, safety and welfare, it is important to consider uncertainty that is introduced when fire models are used to support fire protection engineering. Each of the methods that have been suggested has advantages and disadvantages.

Notarianni was the first person to develop a comprehensive method for handling uncertainty. However, the probabilistic nature of her methodology means that it is only useful for designs that are prepared on a probabilistic basis. At present, and for the foreseeable future, probabilistic designs are usually only performed in some specialized areas: e.g., the chemical process industry and nuclear power plant design. In other applications, probabilistic performance criteria are not available, as many stakeholders do not wish to acknowledge risk.

The method that Lundin developed is very labor intensive, and hence difficult to apply for use in an individual project. The approach could presumably be applied to a model by either the model developer or some third party. However, it is unlikely that this type of work would be undertaken without some financial incentive.

The approach published in the *SFPE Engineering Guide to Substantiating a Fire Model for a Given Application* is intended to be tractable as an adjunct to a larger project. The downside is that the model suitability would be narrowly established; however, for a given project this would be sufficient.

References

[1] *SFPE Engineering Guide to Performance-Based Fire Protection*, National Fire Protection Association, Quincy, MA, 2007.

[2] Notarianni, K., and Parry, G., "Uncertainty," *SFPE Handbook of Fire Protection Engineering*, National Fire Protection Association, Quincy, MA, 2008.

[3] Budnick, E., "Automatic Sprinkler System Reliability," *Fire Protect. Eng.*, No. 9, 2001, p. 7.
[4] Lundin, J., "Model Uncertainty in Fire Safety Engineering," *Report 1020*, Lund Univ., Lund, Sweden, 1999.
[5] *SFPE Engineering Guide to Substantiating a Fire Model for a Given Application*, Society of Fire Protection Engineers, Bethesda, MD, 2011.

Reprinted from JTE, Vol. 40, No. 1
doi:10.1520/JTE104261
Available online at www.astm.org/JTE

Daniel Madrzykowski[1] *and Charles Fleischmann*[2]

Fire Pattern Repeatability: A Study in Uncertainty

ABSTRACT: The National Institute of Standards and Technology (NIST) is conducting a multi-year study, with the support of the National Institute of Justice (NIJ) and the NIST Office of Law Enforcement Standards (OLES), to examine the repeatability of fire patterns. The primary objective of the study is assessing the repeatability of fire patterns on gypsum board exposed to a limited range of source fires. The focus of this paper is an overview of the uncertainties of the measurements. The study required the use of a variety of measurements to determine the repeatability of the source fires in terms of heat release rate, temperature, heat flux, and flame height. Replicate source fire experiments were conducted in an oxygen consumption calorimeter in order to examine the repeatability of the fires in terms of the heat release rate. The flame movement and height for each fire were recorded with photographs and videos. The fire pattern experiments were conducted in a three-walled structure with a full floor and partial ceiling constructed from wood framing and lined with painted gypsum board. The source fires were positioned against the rear wall, midway along its length. Replicate experiments were conducted with each fuel. The fire patterns were documented and analyzed for repeatability. The fire pattern height results are then compared to the mean flame height results to examine the level of agreement.

KEYWORDS: burn patterns, fire experiments, fire investigation, fire patterns, flame height, heat release rate

Manuscript received August 5, 2011; accepted for publication October 19, 2011; published online November 2011.

[1] National Institute of Standards and Technology, Gaithersburg, MD 20899.
[2] Ph.D., Dept. of Civil and Natural Resources Engineering, Univ. of Canterbury, Christchurch, New Zealand 8140.

Cite as: Madrzykowski, D. and Fleischmann, C., "Fire Pattern Repeatability: A Study in Uncertainty," *J. Test. Eval.*, Vol. 40, No. 1. doi:10.1520/JTE104261.

Introduction

Fires are investigated in order to determine the "origin and cause" of the fire and, in some cases, to determine who was responsible for setting the fire. The fire investigation process must follow the scientific method as documented in NFPA 921, Guide for Fire and Explosion Investigations [1]. There are numerous textbooks on the subject of conducting a fire investigation [2–7]. The NFPA guide and the texts provide information to a fire investigator collecting data from the scene, analyzing the data, and developing a hypothesis about a fire.

Patterns produced by a fire are, in many cases, a significant portion of the data collected at the scene. One of the basic methods of documenting the fire scene is to photograph fire patterns. As noted by Icove and DeHaan, "the ability to document and interpret fire patterns accurately is essential to investigators reconstructing fire scenes…" [5]. A fire pattern is defined in NFPA 921 as "the visible or measurable physical change, or identifiable shapes, formed by a fire effect or group of fire effects" [1]. Various types of fire patterns, such as "V-shaped," "hour-glass," and "inverted cone," have come from common observation of the pattern shape at actual fire scenes. Generally, the observations are qualitative in nature.

Previous fire pattern research by the National Institute of Justice (NIJ), the National Institute of Standards and Technology (NIST), and the United States Fire Administration (USFA) has shown that fire patterns provide data useful for the determination of fire origin. The reports noted the impact of ventilation on the development of the burn patterns [8,9]. A large number of other factors affect the formation of these patterns: burn time, heat release rate of the fire source, fire exposure, target fuel composition, adjacent fuel(s), compartmentation, and flashover, to name a few. Given the limited number of experiments in the literature and the large number of variables involved, it has been difficult to fully understand a cause and effect relationship between the fire scenarios and the resulting patterns.

In 2009, the National Research Council of the National Academies (U.S.) published a report identifying the research needs of the forensic science community. In the field of fire investigation, one of the specific needs identified was "…much more research is needed on the natural variability of burn patterns and damage characteristics and how they are affected by the presence of various accelerants" [10]. Previously, the Fire Protection Research Foundation convened a Research Advisory Council on Post-fire Analysis in 2002. Recommendations for research and development in their white paper included "advance the capabilities of computational fluid dynamics fire modeling, particularly as applied to fire ignition scenarios and fire pattern development and interpretation" [11].

A multi-year study to examine the repeatability of burn patterns has been started. In order to examine the repeatability of the burn patterns, the repeatability of the initial fires that cause the damage must be well characterized. This paper focuses on the uncertainties involved with the measurements of the fires and how these may be addressed when compared with results of the pre-flashover fire pattern repeatability experiments.

Technical Approach

A series of experiments were conducted to characterize a range of representative pre-flashover fires that can be expected from intentional and accidental residential fires. A fixed fire source footprint, 30.5 cm square, was chosen for this study and a range of common fuels was selected. The fires were positioned under an oxygen consumption calorimetry hood to determine the heat release rate. Heat release rate, flame/plume temperatures, and total heat flux from the flames were measured and flame movement and height were recorded with photographs and videos. Given the focus on residential fires, small source fires were selected so that the burn pattern would be limited to less than 2.4 m in height. Replicate experiments were conducted with each of the fuels, in order to examine the repeatability of the fires and to quantify the variability in the flames.

Fuels

Three different fuels were used in this study. Experiments were conducted with a natural gas fueled burner, a liquid fueled pan fire, and a solid fuel. Natural gas was chosen as a fuel because it is used for calibrating the oxygen depletion calorimeters in NIST's Large Fire Laboratory. Therefore, the heat content of the gas was monitored and mass flow controlled to provide a heat release rate based on fuel mass flow for comparison with the heat release rate values measured with the calorimeter. The gas burner arrangement provides the most idealized fire, in the sense that it can be ignited and brought up to a near steady state heat release rate within a few seconds and then, just as quickly, be shut down. This is important in future phases of the study for determining the amount of energy transferred to the exposed surface.

Gasoline was the fuel of choice for the liquid fueled pan fire. Gasoline is not an ideal fuel to use in experiments because it is a mixture of components and additives that are specific to manufacturers, changes from season to season, and varies based on requirements or restrictions of the locality. However, fire incident data shows that gasoline was the first item ignited in the majority of intentional fires where flammable or combustible liquids were used [12]. In addition, data from forensic laboratories, collected by Babrauskas and others, indicates that gasoline is the most prevalent accelerant found as part of the analysis of fire debris samples [13,14].

Polyurethane foam was chosen as the solid fuel. According to the Polyurethane Foam Association, flexible polyurethane foam is the most widely used cushioning material in upholstered furnishing and mattresses. More than 1.7×10^9 pounds are produced in the U.S. every year [15,16]. In Rohr's study, "Products First Ignited in U.S. Home Fires," upholstered furniture and mattresses and bedding were the first items ignited in 33% of the fatal fires in the U.S. based on averaged data from 1999 through 2002 [17].

The natural gas burner was 30.5 cm on a side and the top surface of the burner was 7.5 cm above the floor. The shell of the burner was made from steel and filled with "pea sized" gravel with an average diameter of 6 mm. A steady state heat release rate of approximately 80 kW was used for the experiments.

The gasoline pan fire used 500 mL of unleaded 87 octane gasoline. The pan was constructed from 6 mm thick steel. The inner dimension of the pan was 30.5 cm on a side and 1.9 cm high. The pan was elevated in order to bring the pan lip height up to 7.5 cm above the floor. The 500 mL of gasoline had a fuel depth of 5.5 mm and provided a burn time of approximately 300 s.

The solid fuel was composed of polyurethane foam, with a density of approximately 23 kg/m^3. Each block of foam was 30.5 cm on a side and was 7.5 cm thick. The bottom and sides of the foam block were wrapped in aluminum foil to prevent the fuel from moving or spreading as it burned to ensure that the fuel area remained the same the gasoline and natural gas.

Fire Source Characterization

The fire source heat release rate experiments were conducted in the NIST Large Fire Research Laboratory's 3 m × 3 m oxygen consumption calorimeter. This apparatus follows the methodology of ASTM E 2067-08, Standard Practice for Full-Scale Oxygen Consumption Calorimetry Fire Tests [18]. The fires were centered under the hood of the calorimeter and positioned in an open steel frame which provided support for the thermocouples and heat flux sensors. All of the data was recorded at intervals of 1 s on a computer based data acquisition system. A selection of the data is presented in the following sections.

Instrumentation and Measurement Uncertainty

The NIST Large Fire Research Laboratory's 3 m × 3 m oxygen depletion calorimeter has an estimated peak heat release rate capacity of approximately 700 kW. Details on the operation and uncertainty in measurements associated with NIST's 6 m × 6 m oxygen consumption calorimeter were documented by Bryant et al. [19]. Many of the same instruments are shared between the two calorimeters and calibration burns are conducted during each experimental series. As a result, the expanded uncertainty of this device has been estimated as ±11% on the measured heat release rate.

Water-cooled, Schmidt-Boelter total heat flux gauges were used to measure the heat flux. The sensing surface of the heat flux gauges were located 0.95 m horizontally from the center of the burner/fuel and 0.50 m above the top of the burner/pan lip/fuel surface. The manufacturer reports a ±3% calibration expanded uncertainty for these devices [20]. Results from an international study on total heat flux gauge calibration and response demonstrated that the uncertainty of a Schmidt-Boelter gauge is typically ±8% [21]. The gauges were mounted to uprights of the steel instrumentation support frame on opposite sides of the fire.

The fires were centered under the hood of the calorimeter on a noncombustible platform measuring 0.91 m × 1.2 m × 12 mm thick supported by the load cell. The load cell had a resolution of 1.0 g [22]. The expanded uncertainty is estimated to be ±5%.

Gas temperatures were measured with exposed-bead, Chromel-Alumel (type K) thermocouples, with a 1.0 mm nominal diameter. Starting from the exposed bead, the thermocouple wire was sheathed in a 3.2 mm diameter

Inconel shield, 0.76 m in length. The standard uncertainty in temperature of the thermocouple wire itself is ±2.2 °C at 277 °C and increases to ±9.5 °C at 871 °C as determined by the wire manufacturer [23]. The variation of the temperature in the environment surrounding the thermocouple is known to be much greater than that of the wire uncertainty [24,25]. Small diameter thermocouples were used to limit the impact of radiative heating and cooling. The estimated combined expanded uncertainty for temperature in these experiments is ±15%.

A schematic drawing of the instrumentation arrangement is shown in Fig. 1. The thermocouples were mounted on a steel frame. The frame enabled

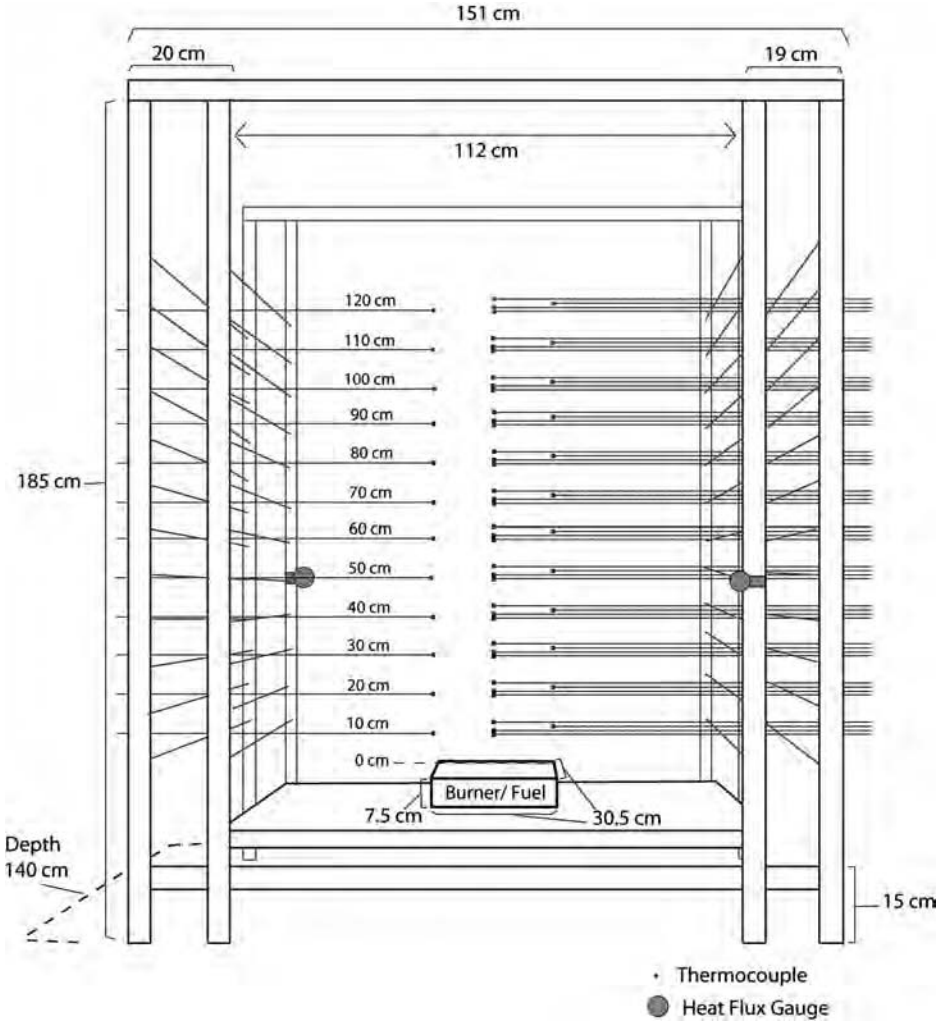


FIG. 1—*Source fire flame and plume instrumentation apparatus which was centered under the 3 m by 3 m oxygen consumption calorimeter. The heat flux gauges were facing toward the center of the burner/fuel area.*

the thermocouples to be installed at 10 cm intervals from 10 cm to 120 cm above the burner/pan lip/fuel surface. At each 10 cm interval, there were five thermocouples: one centered above the burner/fuel surface and one centered above each edge of the burner/fuel surface, for a total of 70 thermocouples. A plumb bob and steel tape measure, with a resolution of ±0.5 mm, were used to position and align the thermocouples. The expanded uncertainty for the location of each thermocouple tip is estimated to be ±2.5%. The same applies to the location of the heat flux sensor surfaces.

Heat Release Rate

Figures 2–4 are graphs showing measurement results of heat release rate versus time for the natural gas, gasoline, and polyurethane foam fueled fires. In each graph, the replicate heat release time histories are overlaid to give a sense of the experimental repeatability for each fuel.

On each fuel's averaged heat release rate curve, a bar representing the 95% confidence level (2σ), based on a Type A evaluation of the uncertainty is provided [26]. A Type A evaluation of the uncertainty is any statistical analysis of the data from a series of observations. In this case, the evaluation assumed a normal distribution of the data such that doubling the calculated standard deviation (2σ) would yield a confidence level of approximately 95%.

The natural gas burner provided the most repeatable results, with an average steady state heat release rate of approximately 80 kW, ±9 kW, with an average total heat release of approximately 23.3 MJ, ±1.6 MJ, over a 300 s period. The gasoline, based on the average over a 30 s interval bounding the peak heat release rate of each of the 17 experiments provided an average peak heat release

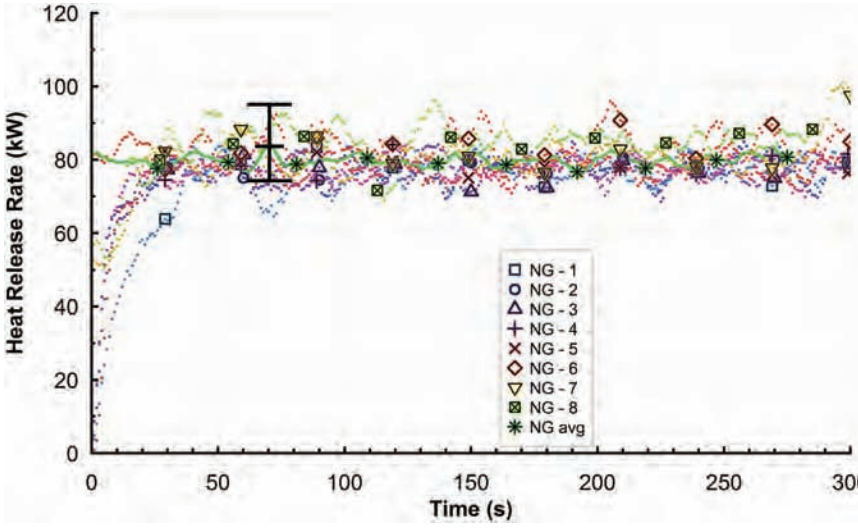


FIG. 2—*Heat release rates versus time for natural gas burner experiments.*

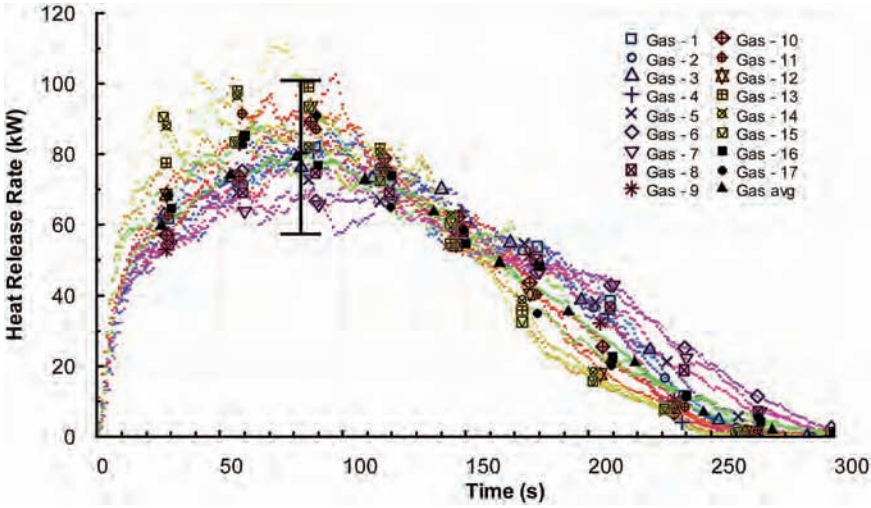


FIG. 3—*Heat release rates versus time for gasoline fueled experiments.*

rate of approximately 80 kW, ±24 kW. The total heat release of the gasoline over the 300 s period after ignition was approximately 13.6 MJ, ±0.6 MJ. The polyurethane foam demonstrated the largest variability, as shown in Fig. 4, of the three fuels in terms of peak heat release rate and time to peak heat release. Averaging over a 15 s interval bounding the peak heat release rate from each of the ten experiments yielded an average peak heat release rate of approximately

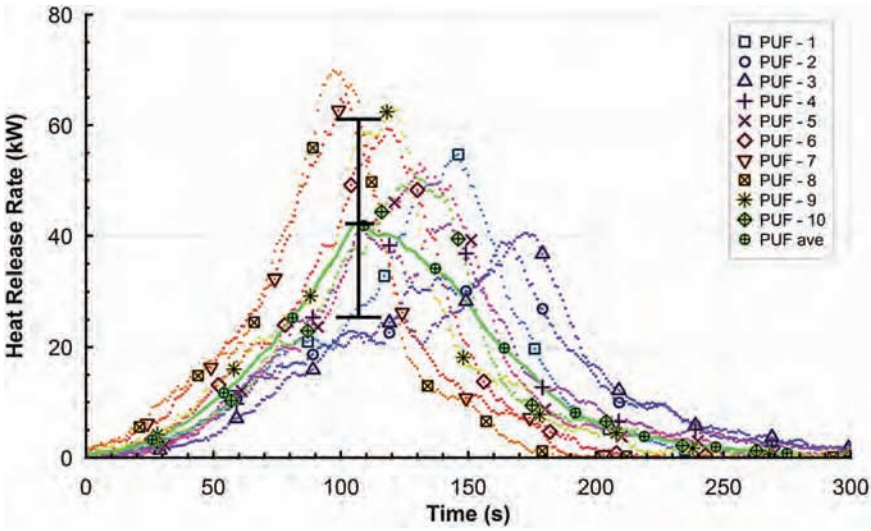


FIG. 4—*Heat release rates for polyurethane foam fueled experiments.*

45 kW, ±21 kW. The total heat release over the 300 s period after ignition was of 4.1 ± 0.2 MJ.

Temperatures

Plots of the centerline temperatures averaged over the replicate experiments for each of the fuels are presented in Figs. 5–7. The number following "TC" in the legend of the graphs represents the height of the thermocouple in increments of 10 cm (or dm) above the fuel surface. The natural gas temperatures were averaged over a "steady-state" burning period, as reflected by the data in Fig. 5. The temperature differences between the vertical thermocouple locations, for locations greater than 20 cm above the fire location, begin to decrease as the distances from the top of the burner increase. This trend is also similar for both the gasoline fires and the polyurethane foam fires. The most notable difference between the natural gas and the other fuels is the growth and decay of the fire during the test period.

Flame Heights

Each of the experiments was photographed with a digital single lens reflex camera, fixed on a tripod, with the capability of capturing at least five images per second. A video camera was located adjacent to the camera, recording images at approximately thirty 30 frames per second (fps). In this paper, the mean flame height for each fuel was estimated from a measurement of the uppermost continuous flame tip from the still photographic images and analysis of the video frames.

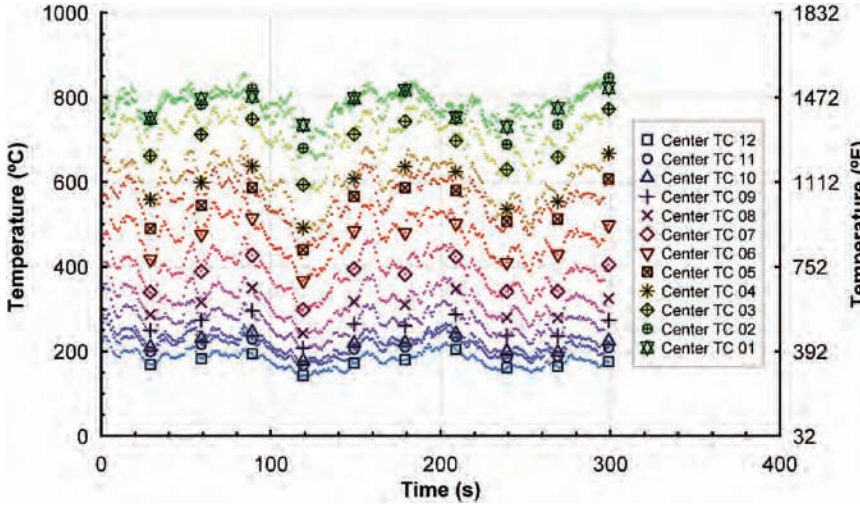


FIG. 5—*Averaged centerline temperature for the natural gas source fire characterization experiments. The number following the "TC" represents the height of the thermocouple in decimeters above the surface of the burner.*

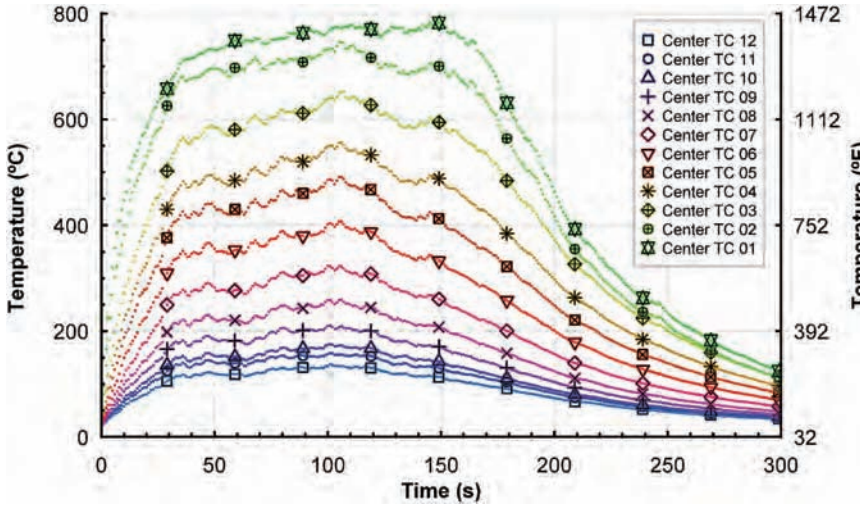


FIG. 6—*Averaged centerline temperature for the gasoline source fire characterization experiments. The number following the "TC" represents the height of the thermocouple in decimeters above the surface of the burner.*

Examples of the photographic images are presented in Figs. 8 and 9. Figure 8 presents a photograph of the natural gas flame taken during a period of "steady state" burning at approximately 80 kW and a photograph of the polyurethane foam fire taken near its peak heat release rate of approximately 60 kW.

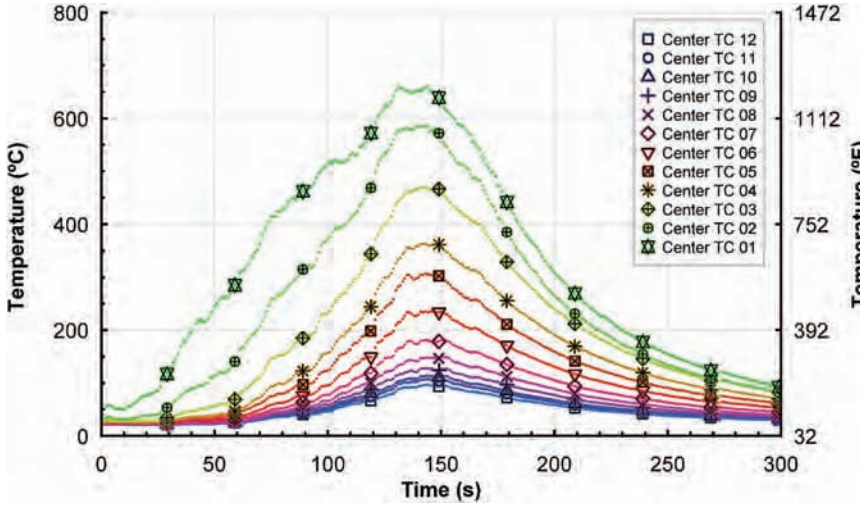


FIG. 7—*Averaged centerline temperature for the polyurethane foam source fire characterization experiments. The number following the "TC" represents the height of the thermocouple in decimeters above the surface of the burner.*

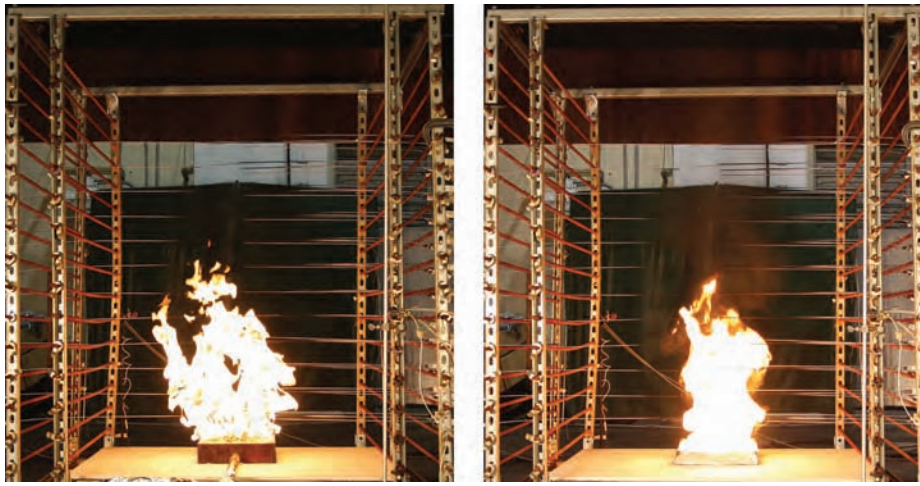


FIG. 8—*(a) and (b) Photographs showing a natural gas flame (left) and a polyurethane foam fueled experiment (right) under the oxygen consumption calorimeter.*

Figure 9 presents two photographs of the same gasoline fueled fire when it was burning near its peak heat release rate of 100 kW. The two photographs, taken 0.2 s apart, represent the range of the observed flame heights from one of the gasoline experiments from 0.65 and 1.0 m, respectively. Table 1 provides the mean visible flame height, along with the range of flame heights for each fuel, based on analysis of the photographs for three replicate experiments. In each case, the photographs were chosen for analysis while the fire was generating its

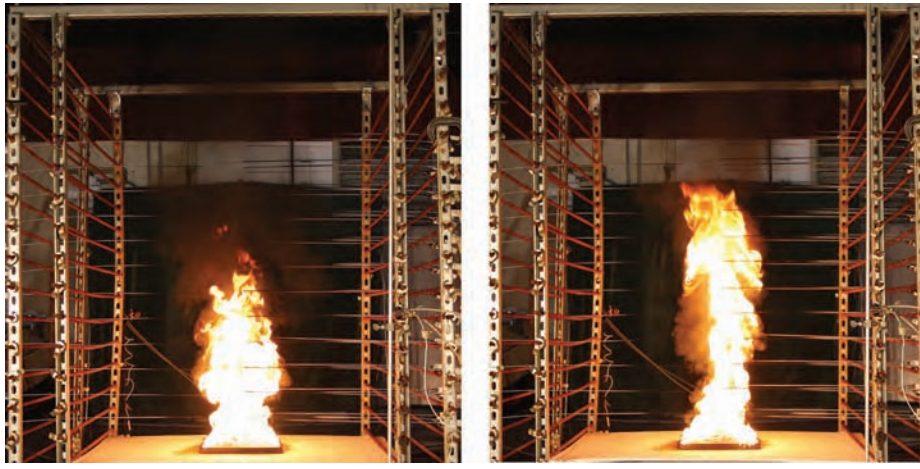


FIG. 9—*(a) and (b) Photograph showing a gasoline fueled experiment with a flame height of 0.65 m (left) and a photograph taken approximately 0.2 s later with a flame height of 1.0 m (right).*

TABLE 1—*Mean visible flame height and averaged minimums and maximums (range) of visible flame heights for each of the fuels during peak heat release rate.*

Fuel	Mean H_f Photographs, m	Range H_f Photographs, m
Natural gas	0.70	0.50–0.98
Gasoline	0.84	0.52–1.10
Polyurethane foam	0.46	0.30–0.78

peak heat release rate. The range of flame heights for each fuel is approximately twice the minimum flame height or the continuous flame height. Analyzing the photographs in this manner is time consuming and is based on making measurements on each photograph. Visible flame heights are manually measured with a scale from the base to the tip of connected flames.

Each experiment was also recorded with a digital video camera at a refresh rate frequency of 29.97 fps which is the National Television Standards Committee standard. The resulting image in each frame was 720 pixels wide × 480 pixels high. Segments of the video which corresponded to the peak heat release of each experiment were captured with a video editing program. For the natural gas experiments, approximately 30 s of "steady state" video were captured. In the cases of the gasoline and polyurethane fueled experiments, the video capture period bounded the peak heat release rate. In the case of the gasoline fueled fires, the peaks occurred over a long period of time, so that a 30 s video capture could be made which was representative of the peak burning rate. The peak duration for the polyurethane foam fires was shorter therefore, only a 15 s video capture was used.

Each frame from the video clip was saved as a full color tagged image file format (TIFF) file and the pixel aspect ratio was changed from 0.9 to 1, so that each pixel had the same dimension in both the horizontal and vertical directions. Once these images were rendered by the video editing program, each image was scaled and cropped to provide input for *Streams*, a suite of image processing programs developed by Nokes for analyzing fluid flow visualization experiments [27]. The images were scaled so that one pixel equals 5 mm in height and width.

The edited TIFF files were loaded into *Streams* from each of the peak heat release rate video clips. For the natural gas and gasoline fueled fires, approximately 900 images were used in each image sequence. The polyurethane foam experiments had approximately 450 images in each sequence for each experiment. Once the image sequence is created, a series of filters and amplifiers are applied to the images to eliminate all colors but red, based on a method developed by Goble [28]. Once all of the images were converted to a pure red "flame" which represented the area of visible fire, on a black background, the images are processed within *Streams* to develop an intensity field based on all of the images being "overlaid" on each other. The end result is a contour plot of the intermittency of the flame ranging from zero to one.

The calculations for the intensity field are based on a user defined computation grid. A grid sensitivity analysis was conducted with 5, 10, 20, 40, and

80 mm square grids. Since each pixel in the original images was equal to 5 mm, in the highest resolution case, the 5 mm analysis grid would match the best resolution of the original image. Four of the five cases resulted in similar results; only the 80 mm case generated significantly different results in flame height. As a result of the sensitivity analysis, a 10 mm grid was chosen as an optimum size for providing a sufficient level of resolution, enabling the efficient processing of 30 s worth of video frames. At the 5 mm grid resolution, limitations within the processing system would not allow a full 30 s of video analysis. Therefore, the intensity fields and the resulting plots and data generated for this study were developed with a 10 mm grid.

Examples of the time averaged contour plots are shown in Figs. 10–13 for natural gas, gasoline, and polyurethane fueled experiments respectively. Each contour line represents the percentage of time that the flame was at a given height. Figure 10 provides an example of a natural gas experiment. The region under examination, shown as a black bounding rectangle, is 1400 mm high and 800 mm in width. This example was averaged over approximately 29 s or approximately 870 images. The number above each contour line in the plot ranges from 0 to 1. The area bounded by the contour with the value 1 represents the area that the flame occupied 100% of the time period analyzed, or the persistent flame region as defined by McCaffery [29]. The remaining contours are in

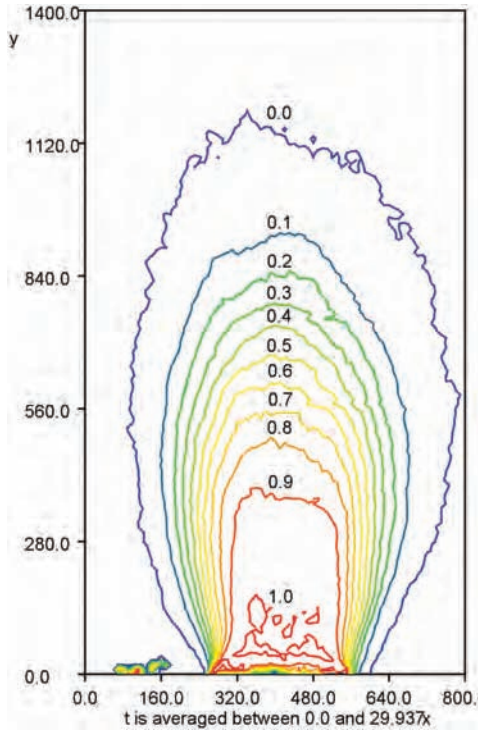


FIG. 10—*Natural gas contour plot.*

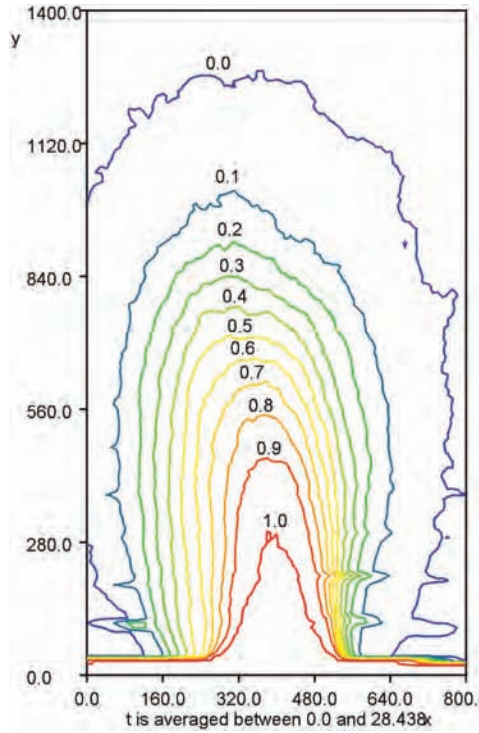


FIG. 11—*Gasoline contour plot.*

the intermittent flame region, with the 0.5 contour representing the mean flame height. Outside of the final contour, at value 0.0, no visible flame was recorded; hence, this contour marks the definitive start of the buoyant plume region.

In Table 2, each of the values represents the average flame heights, given in mm with 95% confidence limits based on a Type A statistical analysis of the data, for each of the source fuels shown at the percentage of time the flame was visible at that given height. For the natural gas fires, the averages are from ten 30 s segments, while the fire was being monitored at approximately 80 kW. In some cases more than one 30 s segment was sampled per heat release rate experiment. Sixteen 30 s video segments, which included the peak heat release rate, were successfully analyzed from the 17 gasoline heat release rate experiments. One of the videos was unusable for analysis due to reflected light which could not be edited out of the images. For the same reason, the polyurethane foam flame height averages result from eight of the ten heat release rate experiments.

For both the natural gas and the gasoline experiments, the 95% confidence limits show a less than 20% variation in visible flame height within the intermittent flame region. The percentage variation increases at the continuous flame region, although the magnitudes of the height variations are similar to other sections of the intermittent flame region. The percentage variation is higher in

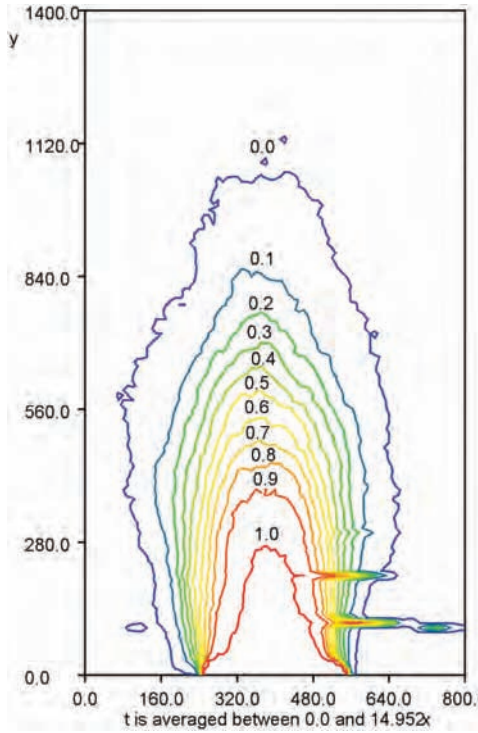


FIG. 12—*Polyurethane foam contour plot.*

the continuous flame region because the height of the continuous flame region is the smallest. The 95% confidence limits for the polyurethane foam are about twice those of the natural gas and the gasoline.

The results from the polyurethane fueled fire characterization experiments have an increased amount of variability due to the nature of the ignition source and the fuel itself. The polyurethane foam is readily ignited with a small flame similar to that used to ignite the natural gas and the gasoline. By comparison, a small flame ignited the gases flowing out of the top of the natural gas burner and the flames from the natural gas cover the entire burner surface within approximately 1 s. A similar ignition sequence occurs with the gasoline pan fires. The polyurethane foam has a more complicated flame spread sequence that results in significant changes to the peak heat release rate and the time to peak heat release. Once the polyurethane foam is ignited, it takes approximately 60 s for the flame to spread across the top surface of the fuel. As the flames spread horizontally across the fuel surface, the fire also burns down into the fuel. In some cases the fire burned completely through the fuel in the center of the foam sample, resulting in a hollow flame at the time of the peak heat release rate. An example of a contour plot, exhibiting the hollow or bifurcated flames, is shown in Fig. 13. Other methods of ignition for the polyurethane foam were considered in order to have a more uniform ignition of the top surface of the

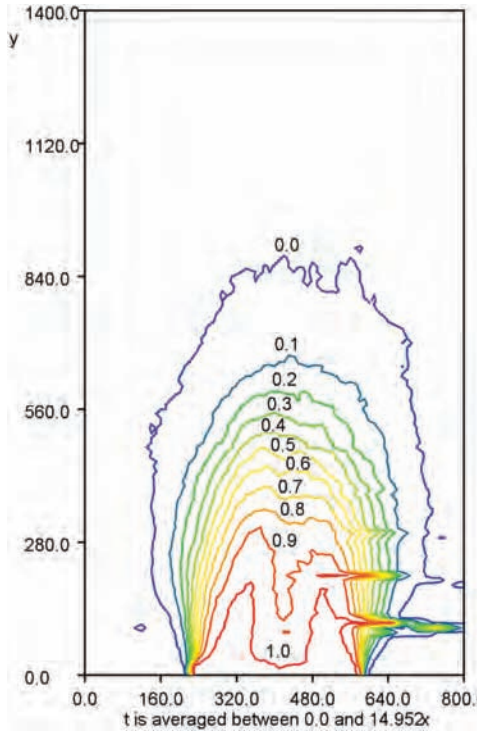


FIG. 13—*Polyurethane foam contour plot with bifurcated flames at the time of peak heat release rate.*

TABLE 2—*Average flame heights with 95% confidence limits for each of the source fuels shown as the percentage of time the flame was visible at that given height.*

Percentage of Time with Visible Flame, %	Visible Flame Heights, mm		
	Natural Gas	Gasoline	Polyurethane Foam
10	980 ± 155 (±16%)	930 ± 130 (±14%)	650 ± 220 (±34%)
20	890 ± 125 (±14%)	850 ± 120 (±14%)	590 ± 190 (±32%)
30	820 ± 125 (±15%)	790 ± 110 (±14%)	540 ± 180 (±33%)
40	760 ± 105 (±14%)	740 ± 95 (±13%)	500 ± 165 (±33%)
50	705 ± 106 (±15%)	695 ± 95 (±14%)	470 ± 160 (±34%)
60	650 ± 100 (±15%)	655 ± 85 (±13%)	435 ± 135 (±31%)
70	590 ± 90 (±15%)	610 ± 90 (±15%)	400 ± 125 (±31%)
80	530 ± 85 (±16%)	555 ± 90 (±16%)	360 ± 110 (±31%)
90	440 ± 75 (±17%)	490 ± 95 (±19%)	315 ± 110 (±35%)
100	220 ± 84 (±38%)	350 ± 110 (±32%)	210 ± 125 (±60%)

foam. Each method would have resulted in a larger release of energy from the ignition source or a starter fuel. Based on a few scoping experiments these methods did not improve the repeatability.

Fire Pattern Experiments

Each of the test compartments consisted of three 3.6 m long × 2.4 m high wood framed walls. The wood elements of the frame were composed of kiln-dried fir with a nominal cross section of 90 mm × 38 mm. The interior surface of the walls was composed of 12.7 mm thick regular core gypsum wallboard. As defined by ASTM, gypsum wallboard is designed for use on walls, ceilings, or partitions. It is composed of a non-combustible core, essentially gypsum, with paper bonded to the surface of the core [30]. A partial ceiling with a width of at least 1.2 m, measured from the back wall of the compartment, was installed across the width of the compartment. The back wall of the compartment consisted of three gypsum board panels, each 1.2 m wide and 2.4 m high. For the experiments discussed here, the center panel was exposed to the source fire. Each panel was painted with a primer coat and a cover coat of latex paint. At least 10 experiments were conducted with each fuel. The fuel or burner was centered on the middle panel of the rear wall. As in the fire source characterization experiments, the top of the fuel in each case was positioned 7.5 cm above the floor of the compartment. Each of the painted gypsum board panels that were used for a fire pattern had a 5 cm grid drawn on it to assist with the determination of the geometry of the fire pattern.

Each fuel or burner was positioned against the face of the painted gypsum board and ignited. The natural gas had a 300 s burn time, which was approximately the upper bound for the total burn time for the gasoline and the polyurethane foam. The gasoline and polyurethane foam fires were allowed to burn until all of the fuel was consumed and the fires self-extinguished. Once the fire was out, the fire pattern on the gypsum board was photographed for future analysis. Then each pattern was measured. The grids were used to assist with the determination of the area. In order to measure the dimensions of the fire pattern, a determination had to be made as to where the fire pattern stops.

DeHaan [4] categorizes the fire effects that form patterns as follows

1. Surface deposits.
2. Surface thermal effects.
3. Charring.
4. Penetration.
5. Consumption.

For this analysis, the outline of the fire pattern was defined where penetration and consumption of the paper covering of the gypsum wallboard stopped. Figure 14(*a*), 14(*b*) show a fire pattern caused by the natural gas burner fire. The most definitive line of demarcation for each pattern appeared to be the line where the paper covering of the gypsum board core was either burned completely away, exposing the gypsum (consumed), or burned and lifted with

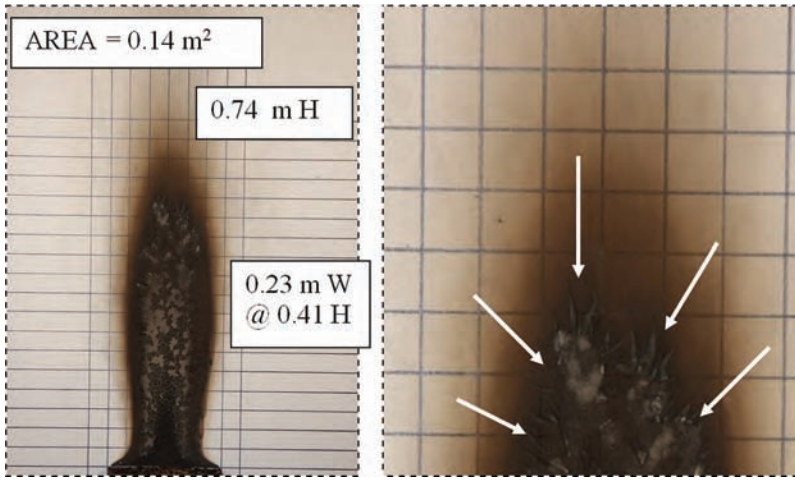


FIG. 14—*(a) and (b) A photograph (left) of a typical fire pattern from the natural gas fuel exposure with the measurements: area, height, and max width with the height at the max width labeled; and photograph (right) of the top potion of the same pattern with portions of the line of demarcation used for the analysis highlighted with arrows.*

portions still attached to the gypsum (penetrated). Penetration was the line of demarcation chosen for this analysis.

Figure 15 shows three different fire patterns that were caused by exposure to the gasoline fire source. The photographs in this figure provide a sense of the level of pattern repeatability under well-controlled conditions. Table 3 provides the average maximum height, average maximum width, average height at which the maximum width occurred and an estimate of the averaged area. The number of experiments included in each average is shown in the first column. Each dimensional value is presented with the 95% confidence level (2σ) based on Type A statistical analysis of the fire pattern measurements [26].

The patterns generated by the polyurethane foam fueled fires have a higher variability relative to the natural gas and gasoline generated patterns. The polyurethane foam exhibited a higher variability of fire growth, a lower peak heat release rate, and a lower total heat release than the other two fuels, which appears to have led to the larger fire pattern variability.

Comparison of Fire Patterns and Flame Heights

Each of the two sets of experiments, the fire source characterization and the fire pattern, had uncertainties associated with them. In the case of the fire source characterization, there were uncertainties associated with the fuel composition itself, the measurement of the heat release rate, temperature, heat flux, and flame height. The fire pattern experiments had additional uncertainties which included the composition and thickness of the gypsum wallboard, moisture content, and paint thickness. In both sets of experiments, efforts were made to limit

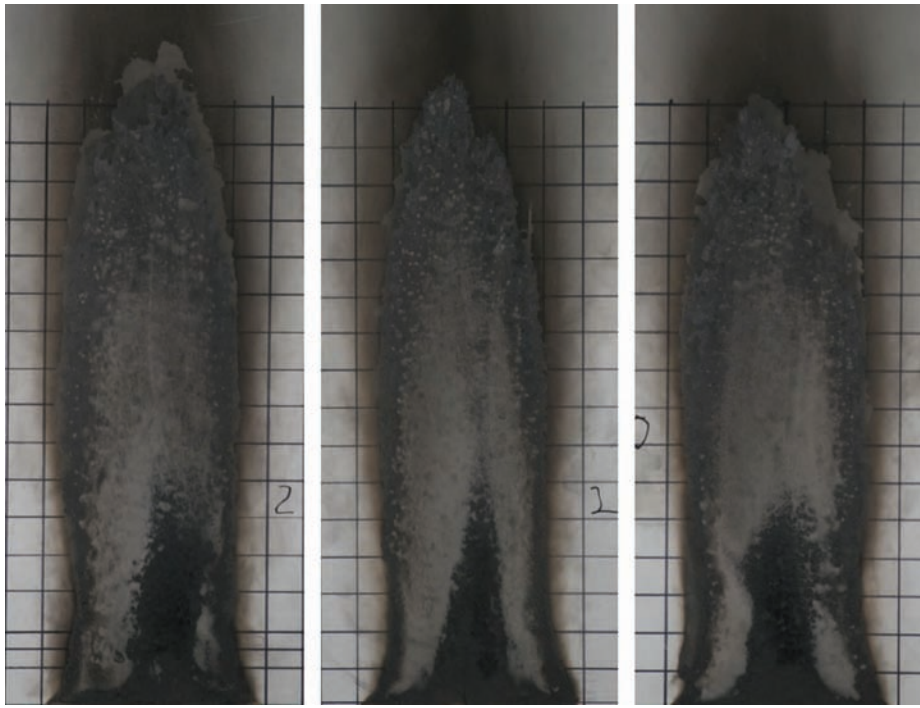


FIG. 15—*(a)–(c) Photographs from 3 different fire patterns generated with the gasoline fire.*

the uncertainties in the fuels, the measurement techniques, and the gypsum wallboard. In this laboratory-based study, the key question is—for a given fire pattern, is there a relationship that can be demonstrated to determine the source fire?

At this stage of the research, the answer to the above question is yes and no. Figure 16 shows the comparisons of the average maximum fire pattern heights with those of the mean flame height based on the analysis of the video images. The associated 95% confidence limits are also shown for each height.

TABLE 3—*Fuel comparison fire pattern dimensions with 95% confidence limits.*

Fuel, Number of Experiments	Height, m	Width, m	Height @ Max Width, m	Area, m^2
Natural gas, 10	0.74 ± 0.12 (±16%)	0.24 ± 0.06 (±25%)	0.41 ± 0.07 (±17%)	0.15 ± 0.05 (± 33%)
Gasoline, 12	0.83 ± 0.15 (±18%)	0.28 ± 0.09 (±32%)	0.44 ± 0.14 (±32%)	0.17 ± 0.04 (±24%)
Polyurethane foam, 10	0.24 ± 0.12 (±50%)	0.28 ± 0.08 (±29%)	0.04 0.24 (±60%)	0.05 ± 0.03 (±57%)

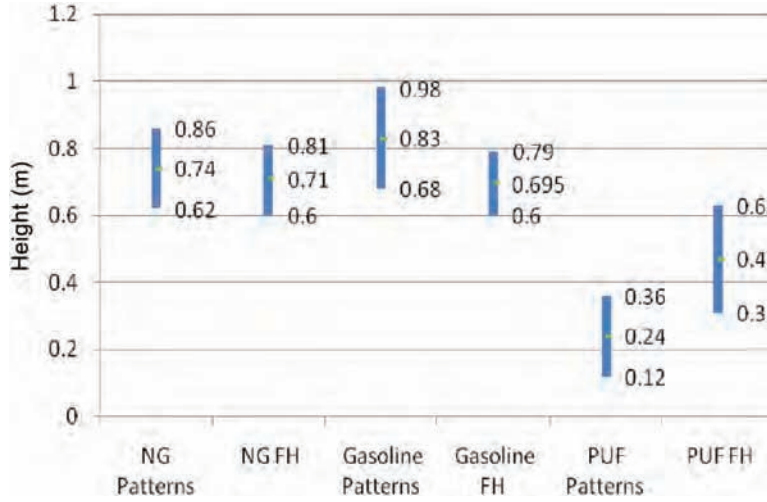


FIG. 16—*Comparison of fire pattern heights and flame heights. The mean height in m is bounded by the 95% confidence limits, also in m.*

The comparison for the fire pattern heights and flame heights from the natural gas fueled fires are very similar. This is expected, in part, because the natural gas fires are "steady state." The "steady" nature of the fire aids in the repeatability or the reduction of variability in the experiments. For example, the energy transfer to the gypsum wall board is also fairly "steady" for the duration of the experiment. Based on the average total heat released, approximately 23 MJ, more energy was transferred to the wall board when compared to the other two fuels. This energy is needed to "dry" out the gypsum wallboard, both of water absorbed from the environment and the chemically bound water, to allow the paper cover and paint to burn.

The gasoline fueled fires also had similar fire pattern heights and flame heights; however, the similarity was not as strong as the natural gas fires. The transient nature of the fire played a strong role here, so that there was less energy transferred to the gypsum wall board and the exposure time for portions of the wall board more distant from the fuel surface was less. Another issue that comes into the discussion is the length of time, 30 s, that was used for flame height analysis. This period of time was used to ensure that the flame height at the time of peak heat release rate was included in the analysis. However, this ultimately may have had the effect of reducing the mean flame height. The mean flame height based on the photo analysis was 0.84 m. The photo analysis was also performed over a 30 s period bounding the peak heat release rate, however, the number of images analyzed was about one sixth of the images analyzed with the video technique.

It should also be noted that the range of the flame heights increased, especially on the lower bound, with the video analysis compared to the photo analysis.

The polyurethane foam fueled fire heights and the associated fire pattern heights had the least similarity when compared to the other two fuels. This

follows with the other findings regarding the polyurethane fueled fires in this study. The most prominent finding being the complications added by the fuel being solid and the uncertainties of the flame spread and localized burning rates, leading to poor repeatability of the fire growth and the heat release rate relative to the natural gas and the gasoline fueled fires.

Mean flame height has been identified by Heskestad and others as having a strong relation to a given heat release rate from fires with a horizontal fuel surface [31,32]. However, the ability to relate the mean flame height to the fire pattern height, based on the limited tests here, is best suited to a "steady state" case. In cases where transient fire growth must be considered, flame height alone may not be enough to determine the heat release rate of the fire without further information to enable an energy balance calculation on the gypsum wall board or other target fuel even for "simple" pre-flashover fire patterns.

Summary

One objective of this paper was to highlight the many uncertainties involved in what appear to be simple fire experiments. This paper showed a variety of uncertainties related to the composition of the fuels, measurement methods, and analysis techniques, just to list a few of the areas considered. The objective of the research was to examine the repeatability of pre-flashover fire patterns generated from exposure to short duration (300 s maximum) well-characterized fires. Three different fuels were used: natural gas, gasoline, and polyurethane foam. Each fuel had a similar top surface area. The heat release rate data showed that the variability was greater for more complex fuels. While the variation in the peak heat release rate with the natural gas was similar to the expanded uncertainty of the measurement system, ±11%, the variation in the peak heat release rates of the gasoline and the polyurethane foam increased due to increased uncertainties in the burning behavior of the fuels. The patterns generated by the polyurethane foam fire had more variability than the natural gas and the gasoline fire patterns. This is likely due to the greater variability of the burning of the solid fuel and the lower heat release rate. The maximum fire pattern heights generated from the natural gas and gasoline fires were shown to have uncertainties of ±18% or less, based on a Type A statistical analysis with 95% confidence limits. The comparison of the fire pattern heights and the mean flame height demonstrated that the "steady state" natural gas fires exhibited the highest level of agreement and the polyurethane foam fueled fire exhibited the least agreement, with the gasoline fueled fires in between. This trend is the same as exhibited by the repeatability of the heat release rate measurements and the flame height measurements.

Future Research

The final step in this phase of the research program will be to use the NIST fire dynamics simulator (FDS) and Smokeview programs to examine the ability to reproduce the source fires and the resulting damage patterns to the gypsum wallboard based on the amount of energy transferred from the fire to the

gypsum wallboard. While the results from this project will provide some insight into the repeatability of pre-flashover fire patterns, a significant amount of research examining other fuels, wall coverings, the impacts of compartmentation, ventilation, and post-flashover conditions are just a few of the parameters that need to be examined to develop a more complete understanding of fire patterns.

Acknowledgments

This work is partially funded by the National Institute of Justice (NIJ) through an interagency agreement with the NIST Office of Law Enforcement Standards (OLES). The writers would like to thank Susan Ballou of OLES for her support of this project. This work was also partially supported by the U.S. Fire Administration. The writers also extend their appreciation to Roger Nokes for his development and continued assistance with Streams and to the New Zealand Fire Commission for their financial support to the University of Canterbury.

References

[1] NFPA 921, 2011, "Guide for Fire and Explosion Investigations," National Fire Protection Association, Quincy, MA.
[2] *Analysis and Interpretation of Fire Scene Evidence*, J. R. Almirall and K. G. Furton, Eds., CRC Press, Boca Raton, FL, 2004.
[3] *Fire Investigation*, N. N. Daeid, Ed., CRC Press, Boca Raton, FL, 2004.
[4] DeHaan, J. D., *Kirk's Fire Investigation*, 6th ed., Pearson Prentice Hall, Upper Saddle River, NJ, 2007.
[5] Icove, D. J. and DeHaan, J. D., *Forensic Fire Scene Reconstruction*, 2nd ed., Pearson Prentice Hall, Upper Saddle River, NJ, 2009.
[6] Lentini, J. J., *Scientific Protocols for Fire Investigation*, CRC Press, Taylor & Francis Group, Boca Raton, FL, 2006.
[7] Noon, R., *Engineering Analysis of Fires and Explosions*, CRC Press, Boca Raton, FL, 1995.
[8] Shanley, J. H., 1998, "USFA Study of Fire Patterns." U.S. Fire Administration, Emittsburg, MD.
[9] Putorti, A. D., "Full Scale Room Burn Pattern Study," *NIJ Report No.* 601-97, National Institute of Justice, Washington, D.C., Dec 1997.
[10] *Strengthening Forensic Science in the United States, A Path Forward*, National Research Council of the National Academies, The National Academies Press, Washington, D.C., 2009.
[11] Fire Protection Research Foundation, "Recommendations of The Research Advisory Council on Post-fire Analysis," Quincy, MA, Feb 2002.
[12] Rohr, K. D., "The U.S. Home Product Report, Forms and Types of Materials First Ignited in Fires: Flammable or Combustible Liquids," National Fire Protection Association, Quincy, MA, Dec 2001.
[13] Babrauskas, V., *Ignition Handbook*, Fire Science Publishers, Issaquah, WA, 2003. p. 683.
[14] Chasteen, C., Division of the State Fire Marshal, State of Florida, personal communication, April 2010.
[15] *Flexible Polyurethane Foam: Industry at a Glance*, Polyurethane Foam Assn., Inc., Knoxville, TN, 2007.

[16] *In Touch, Information on Flexible Polyurethane Foam*, Polyurethane Foam Assn., Inc., Wayne, NJ, Vol. 2, No. 3, 1992.
[17] Rohr, K. D., "Products First Ignited in U.S. Home Fires," National Fire Protection Association, Quincy, MA, April 2005.
[18] ASTM E 2067-08, 2008, "Standard Practice for Full-Scale Oxygen Consumption Calorimetry Fire Tests," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA, pp. 1192–1215.
[19] Bryant, R. A., Ohlemiller, T. J., Johnsson, E. L., Hamins, A. H., Grove, B. S., Guthrie, W. F., Maranghides, A., and Mulholland, G.W., "The NIST 3 MW Quantitative Heat Release Rate Facility: Description and Procedures," *NISTIR Report No.* 7052, National Institute of Standards and Technology, Gaithersburg, MD, Sept 2004.
[20] Medtherm Corporation Bulletin 118, "64 Series Heat Flux Transducers," Medtherm Corporation, Huntsville, AL, Aug 2003.
[21] Pitts, W. M., Murthy, A. V., de Ris, J. L., Filtz, J.-R., Nygard, K., Smith, D., and Wetterlund, I., "Round Robin Study of Total Flux Gauge Calibration at Fire Laboratories," *Fire Saf. J.*, Vol. 41, 2006, pp. 459–475.
[22] National Type Evaluation Program, *Certificate Number 00-075A1, Model K-series*, Mettler-Toledo, Worthington, OH, Dec 2002.
[23] *The Temperature Handbook*, Vol. MM, Omega Engineering Inc., Stamford, CT, 2004, pp. Z-39–40.
[24] Blevins, L. G., "Behavior of Bare and Aspirated Thermocouples in Compartment Fires," *Proceedings of the 33rd National Heat Transfer Conference*, HTD99-280. Albuquerque, NM, Aug 15–19, 1999, ASME, New York.
[25] Pitts, W. M., Braun, E., Peacock, R. D., Mitler, H. E., Johnsson, E. L., Reneke, P. A., and Blevins, L. G., "Temperature Uncertainties for Bare-Bead and Aspirated Thermocouple Measurements in Fire Environments," *Proceedings of the ASTM*, Dallsa, TX, Dec 3, 2001, ASTM Spec. Tech. Publ., Vol. 1427, pp. 3–15.
[26] Taylor, B. N., and Kuyatt, C. E., "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results," *NIST Report No.* TN 1297, National Institute of Standards and Technology, Gaithersburg, MD, Sept 1994.
[27] Nokes, R., *Streams*, Version 1.07, Dept. of Civil and Natural Resources Engineering, Univ. of Canterbury, Christchurch, NZ, Feb 2011.
[28] Goble, K., 2007, "Height of Flames Projecting from Compartment Openings," Master's thesis, Dept. of Civil and Natural Resources Engineering, Univ. of Canterbury, Christchurch, NZ.
[29] McCaffery, B. J., "Purely Buoyant Diffusion Flames: Some Experimental Results," *Report No.* NBSIR 79-1910, National Bureau of Standards (currently NIST), Gaithersburg, MD, Oct 1979.
[30] ASTM C1396/1396M-11, 2011, Standard Specification for Gypsum Board," *Annual Book of ASTM Standards*, Vol. 04–01, ASTM International, West Conshohocken, PA.
[31] Heskestad, G, "Fire Plumes, Flame Height and Air Entrainment," *SFPE Handbook of Fire Protection Engineering*, 4th ed., P. J. DiNenno, D. Drysdale, C. L. Beyler, W. D. Walton, R. L Custer, J. R. Hall, and J. M. Watts, J.M., Eds., Society of Fire Protection Engineers, Bethesda, MD., 2008.
[32] Beyler, C., "Fire Plumes and Ceiling Jets," *Fire Saf. J.*, Vol. 11, 1986, pp. 53–75.

Reprinted from JTE, Vol. 39, No. 6
doi:10.1520/JTE103927
Available online at www.astm.org/JTE

Norman J. Alvares[1] *and Harry K. Hasegawa*[2]

In Search of Standard Reference Materials for ASTM E05 Fire Standards

ABSTRACT: The first paragraph of E-691, Standard Practice for Conducting an Inter-laboratory Study to Determine the Precision of a Test Method [ASTM E691-05, 2005, "Standard Practice for Conducting an Inter-laboratory Study to Determine the Precision of a Test Method," *ASTM Book of Standards,* ASTM International, West Conshohocken, PA.] states: "Tests performed on presumably identical materials in presumably identical circumstances do not, in general, yield identical results. This is attributed to unavoidable random errors inherent in every test procedure; the factors that may influence the outcome of a test cannot all be completely controlled." Further in the same paragraph the factors are identified: "Many different factors (apart from random variations between supposedly identical specimens) may contribute to the variability in application of a test method, including: a. the operator, b. equipment used, c. calibration of equipment, and d. environment." The primary subjects of both the first and second paragraphs are: "presumably (supposedly) identical materials". If, in fact, "identical materials" were available, one of the variables of testing would be eliminated because the performance of the material would be a known, which could be used for calibration procedures. Thus, any variations in the test results would be caused by a, b, or d. The Holy Grail of ASTM fire standards is the precision and bias section. ASTM regulations require precision statements in all test methods in terms of repeatability and reproducibility. However, most E05 standards do not provide precision and bias data. In fact, only 33% of E05 standards have conducted an inter-laboratory study or round robin at some time. A path forward to promote more systematic calibration procedures for E05 fire test methods and to facilitate more frequent round robin studies, is to explore the identification or development of Standard Reference Materials (SRM's) with defined properties for calibration of ASTM E05 fire tests

Manuscript received April 20, 2011; accepted for publication May 18, 2011; published online July 2011.

[1] Fire Science Applications, San Carlos, CA, 94070, e-mail: nalvares@sbcglobal.net
[2] FireQuest, Oakland, CA 94618.

Cite as: Alvares, N. J. and Hasegawa, H. K., "In Search of Standard Reference Materials for ASTM E05 Fire Standards," *J. Test. Eval.*, Vol. 39, No. 6. doi:10.1520/JTE103927.

and to validate operational performance for the same test method conducted at different facilities. Surveys of a wide range of ASTM test standards show that some of them use well-characterized materials to provide specific output data as part of their calibration procedures. In fact, SRM's have been developed for test standards; E162, E648 and E662. This paper summarizes the properties of the materials used for test calibration and their potential use as reference standards or for identifying properties important to formulating a standard material(s) for ASTM E05 fire test methods.

KEYWORDS: round robins, standard reference materials, fire tests

Introduction

This review is motivated by the goal of Subcommittee E5.32, TG 5 on "Standard Test Material Feasibility." The TG's primary goal is to identify or develop materials that could become standard reference materials (SRM)s for selected E05 test methods. The ideal combustible material(s) or material combinations must be stable, have known physical properties, be environmentally inert, remain available, and not too expensive. A range of materials will be required, but the desirable characteristics that define a material's performance should be similar. Their use would become part of the calibration procedure for a specific test apparatus and, more importantly, to compare operational performance for the same test method conducted at different facilities.

If SRM's can be developed and accepted by E05 test operators, more round robin studies may be encouraged for E05 tests, resulting in a much higher percentage of test procedures with acceptable precision and bias sections.

An example of the utility of a standard reference material is the specially formulated PMMA specified as a calibration standard for ASTM E1354 in the Cone Calorimeter user's manual [1]. The PMMA test specimen is tested according to the standard and the results are compared to previous, well-characterized calibration runs. If there are discrepancies between the calibration data and the test specimen, troubleshooting and repairs are done to harmonize the data.

Although these calibration runs are for the purpose of diagnostic testing within 1 laboratory only, they probably account for the nominal variability in the repeatability of E-1354 results in that lab. This well characterized PMMA could well be a candidate SRM for small-scale tests methods where its melting behavior is not critical. Its use may not be feasible for many large-scale tests and for different orientations.

Background

Survey of ASTM fire test standards

Appendix contains all "ACTIVE" standards categorized by each E05 Subcommittee responsible for the test standards. Table 1 is a distillation of Appendix in which we summarized the number of test standards for each Subcommittee, the primary fire performance characteristics measured in each, and a brief description of the type of testing apparatus used to conduct the tests. The test

TABLE 1—*Fire test characteristics.*

Sub-Committee	Sub-Committee Name	Active standards	Measured Parameters	Type of Tests
E05.11	Fire Resistance	7	Fire Resistance	Furnace Tests
E05.14	External Fire Exposure	2	FSR, penetration resistance	Fire Exposure
E05.15	Furnishings & Contents	6	HRR, Ignition	Calorimeter Tests
E05.21	Smoke & Combustion Products	11	Obscuration, HRR, Ignition, mass loss, FSR	Smoke Chamber
E05.22	Surface Burning	8	FSR, ignition, critical flux, smoke density	Flamespread
E05.23	Combustibility	2	Mass loss, temperature rise	Tube furnace

methods require a variety of measurements in order to assess the tested material's performance. Some of these include measurements of temperature, mass loss, fuel gas flow rate, airflow, light transmission, heat flux, particulate collection, and gas analysis. Many of the tests provide data that yield more than one fire performance characteristic of the tested materials. The following tabulation compares the number of E05 test standards to measured performance parameter:

Heat release rate (HRR):	10 test standard
Ignition:	10 test standards
Flame spread rate (FSR):	9 test standards
Fire resistance:	8 test standards (includes 2 penetration resistance)
Mass loss:	4 test standards
Critical flux:	3 test standards
Smoke density:	2 test standards

Because one test standard can be used to determine more than one performance parameter, the apparent hierarchy is debatable; however it appears that the dominate performance characteristics are heat release rate, ignition, flame spread rate, and fire resistance. This hierarchy does; however provide some guidance with regard to what parameters we should consider when we attempt to develop/identify standard materials.

Excluding SRMs pertaining to standard cigarettes and cigarette ignition strength, there are two previously available SRMs for flame spread tests and one each for flaming and nonflaming smoke density tests. However, there is no documentation in the SRMs that describes the process used to select these materials. References contained in the SRMs indicate that "Engineering testing and analysis leading to the certification of the SRM was performed by personnel of the Fire Research Division or its predecessors at NIST. These SRMs pertain to a single fire performance characteristic, ideally, a single material might be identified or formulated that could be used to evaluate more than one of the fire performance characteristics. This may not be possible with currently available materials, but it may be conceivable to fabricate a material that could serve as an "all-around" SRM for ASTM flammability tests.

Survey of inter-laboratory studies (round robins)

ASTM regulations require precision statements in all test methods in terms of repeatability and reproducibility. One of the primary purposes of inter-laboratory studies is to develop information needed for a precision statement. In addition, these studies are useful in identifying problems or weaknesses in the operation and calibration of test standards. Two standards: E691 Practice for Conducting an Inter-laboratory Study to Determine the Precision of a Test Method and E2653 Standard Practice for Conducting an Inter-laboratory Study to Determine Precision Estimates for a Fire Test Method with Fewer than Six Participating Laboratories, describe techniques for planning, conducting, analyzing, and treating results of round robins. E691 requires 6 or more testing laboratories to provide statistical credence for the calculations of data precision

and bias. Recognizing that there may not be six or more laboratories available to participate, standard E2653 was established where a minimum of three laboratories are required to participate to provide valid data. Participation of less than three laboratories does not qualify to provide validated precision statements.

E2653 requires that no less than three test materials shall be selected for the round robin, and that a minimum of three replicates shall be done for each test material selected. Also, if a SRM is available for the test method, the material shall be included in the round robin. If a standard reference material is not available, a test specimen that consistently produces low variability test results shall be selected as a reference material for the round robin. Regrettably, we found no information in our review of the available round robins that described the "test specimen that consistently produced low variability results" selected as a reference material.

We surveyed the 2007 book of "ASTM Fire Standards and Related Technical Material" [2] to identify which standards had conducted round robins and published precision and bias statements as a part of the test standard. The number of test standards according to ASTM designation is listed below:

"C"	Sealants and Gaskets	7 standards,
"D"	Textiles	80 standards,
"E"	Fire Standards	67 standards,
"F"	Protective clothing	15 standards,
"G"	Oxygen Index	5 standards.

There are a total of 174 fire test standards, excluding standard guides, practices, and specifications.

Table 2 lists the E05 standards and Table 3 lists all those that are not E05 standards. None of the "C" designated tests were included in Table 3 because they did not conduct any inter-laboratory studies. Note that many of the test methods use the same test apparatus and we have not identified them because the focus of our analysis is the E05 tests listed in Table 2.

All of the fire test standards summarized in both Tables 2 are listed because they have participated in inter-laboratory testing at some time. In addition, they list whether or not the standard contains a precision bias statement, the number of labs involved in the round robins, dates when the round robins were done, and if they conducted the round robins using a reference and/or calibration standard. It should be noted that five standards were included in the tables because they contained a precision and bias section, or they identified a calibration or reference standard in their procedure, even though they did not conduct round robins.

From the tables, we determined that:

28 (35%) of the "D" standards conducted round robins and 15 (19%) contained precision and bias statements,

7 (47%) of the "F" standards conducted round robins and 6 (40%) had precision and bias statements,

1(20%) of the "G" standards conducted a round robin and included a precision and bias statement.

Out of 67 "E" standards, 22 (33%) conducted round robins and 14 (21%) contained precision and bias statements.

TABLE 2—*ASTM E05 Fire Standards That Conducted Inter-laboratory Studies.*

Standard No.	Prebias: Y/N	Reference Standard	Calibration Standard	Round Robin(s)	No. of Labs	Dates
E84-07	N	No	Cement asbestos, Red oak	1	11	?
E-162	N	Fibrous hardboard	Reinforced cement board	1	?	?
E648-06a	Y	Craft paper board	No	2	13	?
E662-06	Y	Alpha cellulose, plastic sheet	Inorganic mill bd.	1	20	?
E698-05	Y	No	No	1	8	2000
E711-87-04	Y	No	No	1	?	?
E906-06	N	AC composite	No	FAA/Boeing	27	2006-7
E1232-02	N	No	No	1	1	?
E1317-97a(02)	N	No	No	2	?	?
E1352-02	Y	No	No	1	5	?
E1354-04a	Y	No	PMMA	E05 & ISO	17	1990
E1461-01	Y	No	No	A number	?	?
E1474-04e1	Y	No	No	2	7	?
E1537-02a	N	No	No	1	4	?
E1590-02	Y	No	No	1	6	1997
E1623-04	Y	No	propane or methane sand burner	1	3	?
E1740-02	N	No	No	1	1	?
E1822-02b	Y	No	No	1	1	?
E1995-04a	Y	No	No	2	8,1	?
E2058-06	N	No	Acetone for heat of combustion	1	2	?
E2257-03	Y	No	No	1	5	1989-90
E2652-09a	Y	No	No	1	?	?

TABLE 3—*ASTM Fire Standards That Conducted Inter-laboratory Studies.*

Standard No.	Pre-Bias: Y/N	Ref. Standard	Cal. Standard	Round Robin(s)	No. of Labs	Dates
D635	N	No	No	2	?	1986, 1987
D1322-97	N	No	Hydrocarbon blends	No		
D1929-96	N	No	No	1	?	1994
D2584-02	N	No	No	1	7	?
D2671-00	N	No	No	1	7	?
D2843-99	N	No	No	2	6	1982, 1998
D2863-06a	Y	No	PMMA	3	12,18,29	1999,?,?
D2939-03	Y	No	No	No	–	
D3014-04a	N	No	No	1	6	1988
D3278-96	N	No	No	1	5	?
D3574-05	Y	No	No	multiple	6,10	1998-2000
D3675-05	Y	No	Cement board	1	11	1970's
D3801-06	N	No	No	1	18, 7	1994, 1995
D3806-98	Y	No	Asbestos-cement bd	No	–	
D3828-05	Y	CRM's	CRM's	1	6	1986
D4809-06	Y	No	Benzoic acid	?	?	?
D4986-03	Y	No	No	1	?	1990
D5048-03	N	No	No	1	13	1988
D5132-04	Y	No	No	1	11	2002
D5238-98	Y	No	No	1	7	1987-88
D5306-92	Y	No	No	1	8	?
D5865-04	N	No	benzoic acid in O2	No		
D6113-03	N	No	No	Ref. E1354		
D6413-99	Y	No	No	1	1	?
D6545-00	Y	No	No	1	1	1999
D6801-02a	Y	No	No	1	1	?
D7016-04	Y	No	No	1	4	?
D7140-07	N	No	No	1	1	?
F955-07	Y	No	Referenced fabrics	See ASTM	1	
F1060-05	Y	No	No	1	1	?
F1358-00	Y	No	No	1	1	?
F1868-02	Y	No	Fabric in Part C	1	6	2001
F1930-00	N	No	No	1	3	1998
F1939-99a	Y	No	No	1	1	1997
F1958/F-99	Y	No	No	1	1	?
G125-00	Y	No	No	1	18	?

A disturbing result of our survey was that the most recent inter-laboratory round robins were conducted between 9 to 40 years ago. The tables also show that only the "D" and "E" standards had reference standards, 1 for the "D" standards and 5 for the "E" standards. 7 of the "D" and 6 of the "E" standards

list calibration standards. 2 "F" standards contained calibration standards. It is equally disturbing to note that each of the "D", "E", and "F" standards had 4 or 5 test methods for which "round robins" were performed by a single laboratory. According to E691 and E2653, these "round robin" results cannot be used to make valid precision and bias calculations.

Extensive inter-laboratory studies were performed by the Boeing Aircraft Company for the aerospace industries using the E906 Standard Test Method for Heat and Visible Smoke Release Rates and by ASTM Institute for Standards Research (ISR) and International Organization for Standards (ISO) using the E1354 Standard Test Method for Heat and Visible Smoke Release Rates for Materials and Products Using an Oxygen Consumption Calorimeter.

The Boeing round robins were convened by both OSU and NBS (now NIST) along with 27 participating labs. Three materials were tested along with the FAA standard reference panel (BAC 5524 Glass Epoxy Laminate) manufactured by Schneller [4]. Results of these round robin tests showed a wide variance in results amongst laboratories. One reason for these large uncertainties (a factor of 2) is attributed to the misuse of the heat flux sensors used to specify the incident heat flux to the measured material [5].

At least two rounds of inter-laboratory tests were conducted for ASTM standard E1354. Participation involved both the ASTM ISR (4 labs) and the ISO (13 labs). Six materials were tested. Results of the 1990 series indicated a dramatic difference of repeatability and reproducibility from the previous series. It was concluded that this difference occurred because most ISR materials were treated with fire retardants. It is well known that the variance of fire test data is much greater for treated materials. Moisture content also played a role. These results demonstrate the importance of test material selection. This fire test standard is one of the few that contains a precision bias statement. E906 does not include a precision bias statement.

Existing standard reference materials (SRMs)

Many of the E05 tests and commercial agencies (e. g., the aircraft materials manufacturers) already have a requirement to use a defined material as a method of ensuring the proper operation of their testing apparatus as well as to compare results with other testing labs. Examples of these test methods are:

E-84 uses inorganic reinforced cement board and select grade red oak.
E-162 uses inorganic reinforced cement board.
E-906 uses methane to calibrate for heat release rate.
E-1354 uses methane for electronic mass flow meter calibration, specified PMMA for checking system calibration, ethanol for orifice coefficient checks.
E-1623 uses either propane or methane with a sand burner.
E-2058 uses acetone to determine effective heat of combustion.

As indicated above, PMMA is specified as a calibration standard for weekly calibrations of the E1354 test apparatus. The PMMA specification has changed once during development of the standard. The mounting protocol is rigid and after mounting in the apparatus, the standard test procedure is followed. The

output of test results are compared to previously determined ranges and "if any of these values do not conform to the expected ones, further troubleshooting and repair must be done."

The NIST division on Standard Reference Materials (SRM) certifies 5 subcommittee test materials, which are:

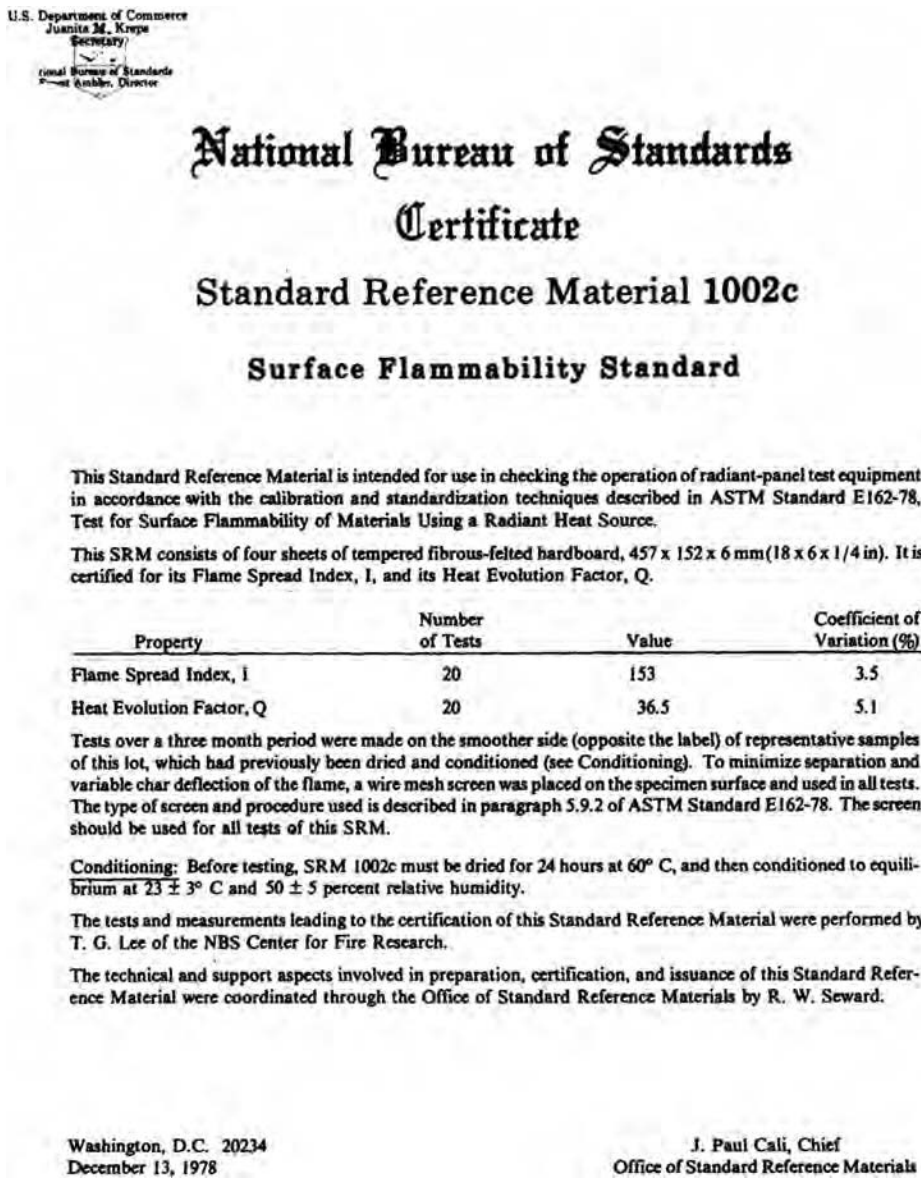
- E162-78: Surface Flammability, SRM 1002c.
- E648-78 (NFPA253-780): Flooring Radiant Panel, SRM 1012.
- E2187-04 and E2187-02b: Cigarette ignition strength and Standard cigarette for ignition resistant testing. SRM No's 1082 and 1096.
- E662-05 (NFPA 258-1998; Smoke Density Chamber. SRM 1006d, Nonflaming exposure, (paper). SRM 1007b; flaming exposure (plastic).
- Non ASTM Smoke toxicity Cup Furnace Standard; SRM 1048, Univ. of Pittsburgh Method, SRM 1049.

The SRM for E162 was preconditioned "fibrous-felted" hardboard, exposed on the "smooth side." To prepare the E162 apparatus for measurement, the operator must construct or buy a "black body" furnace to calibrate the pyrometer that sets the temperature of the gas fired radiant panel. The operator must also construct a line burner to calibrate the stack thermopile used to measure the HRR. Specification of the burner geometry is not detailed and the gas used for calibration can be: "manufactured methane, or natural gas, or combinations of these gases." Because of these compounding sources of uncertainty, E162 is a standard where a SRM would be very useful, and there is one. The very last paragraph in the Annex describing the Procedure for calibration describes how to obtain the Surface Flammability Standard. Unfortunately, the SRM is listed as "out of stock" in the SRM NIST catalog. Figure 1 is a reproduction of the last published SRM certificate (SRM 1002c) for E 162 published on Dec. 13, 1978. The published flame spread index (I) is a product of the Heat Evolution Factor and the average upward flame spread-rate. I for SRM 1002 appears to vary in time. i.e., I = 153 in 1978, 190 in 1971, and 131 in 1964. This is not the range of variation we would expect to see in an "ideal" SRM.

The equipment used in E648-78 is similar to E162, but the orientation and operational parameters are not. Moreover, there is much more detail in the description of the calibration procedures. The composition of the pilot burner gas is precise and, while the pyrometer calibration process is the same, there is a further recommendation to have the pyrometer compared to a standard pyrometer that has a calibration traceable to NIST. Note that unless we heed the recommendations of reference [Op cit 4] the use of Gardon-type pyrometers in this test method will have uncertainty problems similar to those of FAA's use of E906.

SRM 1012 is available for E648 and it consists of 3 sheets of conditioned kraft paper board; 104.1 cm long, by 25.4 cm wide and 3.05 mm thick. The certified value for average critical radiant flux is: CRF = 0.36 +/− 0.04 $W.cm^{-2}$. The last available certificate was published on Sept. 14, 1984. There is no mention of this SRM in Standard E648-06a in the current book of standards, although it is listed as available in the SRM NIST catalog.

Aside from the cigarette standards, the only other SRM's listed in the NIST catalog are for E662, the smoke density chamber. SRMs 1006d and 1007a are to



U.S. Department of Commerce
Juanita M. Kreps
Secretary

National Bureau of Standards
Ernest Ambler, Director

National Bureau of Standards
Certificate
Standard Reference Material 1002c
Surface Flammability Standard

This Standard Reference Material is intended for use in checking the operation of radiant-panel test equipment in accordance with the calibration and standardization techniques described in ASTM Standard E162-78, Test for Surface Flammability of Materials Using a Radiant Heat Source.

This SRM consists of four sheets of tempered fibrous-felted hardboard, 457 x 152 x 6 mm (18 x 6 x 1/4 in). It is certified for its Flame Spread Index, I, and its Heat Evolution Factor, Q.

Property	Number of Tests	Value	Coefficient of Variation (%)
Flame Spread Index, I	20	153	3.5
Heat Evolution Factor, Q	20	36.5	5.1

Tests over a three month period were made on the smoother side (opposite the label) of representative samples of this lot, which had previously been dried and conditioned (see Conditioning). To minimize separation and variable char deflection of the flame, a wire mesh screen was placed on the specimen surface and used in all tests. The type of screen and procedure used is described in paragraph 5.9.2 of ASTM Standard E162-78. The screen should be used for all tests of this SRM.

Conditioning: Before testing, SRM 1002c must be dried for 24 hours at 60° C, and then conditioned to equilibrium at 23 ± 3° C and 50 ± 5 percent relative humidity.

The tests and measurements leading to the certification of this Standard Reference Material were performed by T. G. Lee of the NBS Center for Fire Research.

The technical and support aspects involved in preparation, certification, and issuance of this Standard Reference Material were coordinated through the Office of Standard Reference Materials by R. W. Seward.

Washington, D.C. 20234
December 13, 1978

J. Paul Cali, Chief
Office of Standard Reference Materials

FIG. 1—*SRM 1002c.*

calibrate the chamber for nonflaming — and flaming—exposure conditions, respectively. According to Mr. Jonathan Jackson of Commercial Testing Company, SRM 1007a is no longer available from NIST.

Perhaps the most familiar SRM (actually, not a SRM) is the "select-grade" red oak used to make the reference deck for E84, the key test listed in Chapter 8

of the International Building Code (IBC). The annual calibration of the E84 apparatus to establish the Flame Spread Index ratings for construction materials has become a basic tool of the prescriptive fire ratings industry. Confidence in the test results from this standard are unshakable. If all E05 standards were validated with a well characterized SRM, they to would have enhanced credibility. Unfortunately, some E84 operators are concerned that good quality red oak is becoming difficult to obtain. Consequently, the periodic tunnel calibration is becoming problematic and a replacement for red oak is potentially necessary.

Current basis for SRM development

The NIST SRM web page outlines the criteria and bases for SRM's as follows.

"The certificates for NIST Standard Reference Materials (SRMs) traditionally have been in conformance with guidance criteria issued by the ISO Advisory Committee on Reference Materials. The criteria for the contents of certified reference material (CRM) certificates are contained in ISO Guide 31."

A survey of the ISO references led us to a power point presentation prepared by the Laboratory of the Government Chemist (LGC) on the practical implementation of ISO guide 34. Guide 34 contains the general requirements for the competence of reference material producers. They establish that "the committee on reference materials (REMCO) of ISO was established in 1973 and the aim of the committee is to carry out and encourage a broad international effort for the harmonization and promotion of certified reference materials. They define a reference material (RM) as Materials, sufficiently homogeneous and stable with respect to one or more specified properties, which have been established to be fit for its intended use in a measurement process."

A more formal characterization is for Certified Reference Materials (CRM). A CRM is:

"An RM characterized by a metrologically valid procedure for one or more specified properties, accompanied by a certificate that states the value of the specified property, its associated uncertainty, and a statement of metrological traceability."

This document also gives guidance regarding a logical approach for producing a RM:

Production Steps

- Material specification
- Confirmation of identity
- Sourcing, preparation and subdivision
- Homogeneity assessment
- Stability assessment
- Characterization of the assigned value(s)
- Calculation of the assigned value and its uncertainty
- Documentation and storage of the material

While these guidelines were formulated for the chemical industry, they are sufficiently general and can provide us with a template for development of contemporary SRMs for selected ASTM EO5 test methods.

Discussion

Are SRMs for ASTM E05 fire tests needed?

SRMs were developed for E05 standards E162, E648, E662, and E2187, where the earliest dated certificate was for E162 in 1964. Except for SRM 1082; the Cigarette Ignition Strength Standard for E2187, we do not know the criteria for their selection. However, they were certified by NBS and the purpose of the SRMs was to "check the operation of testing apparatus." There is very little information regarding the frequency with which the SRMs for E162, E648, and E662 were used, either for in-house calibration or for round robin campaigns. The last SRM issued for these tests was nonflaming α cellulose paper sheets, certified for E662 in 1999. This SRM is still available from NIST. The SRM for flaming exposure conditions in E662 was last issued in 1976 and is no longer available. Similarly the SRMs for E162 and E648 are no longer available. Obviously, no one has requested them and the certifications have lapsed. Perhaps this is the answer to the question poised in the heading for this paragraph, which is SRMs for E05 fire tests are no longer needed because no operators use them.

A better question may be Are precision and bias statements required as an integral part of all E05 fire test standards and should SRMs be developed to help facilitate the precision and bias calculations? Our survey of all E05 test standards show that, only 21% contain precision and bias statements. Since all ASTM standards are required to have them, most of the E05 standards are noncompliant, and apparently no one is very concerned.

What are the essential characteristics of SRMs for E05 fire tests?

Our survey of E05 fire tests show that the fire characteristics most frequently measured are heat release rate, flame spread rate, fire resistance, critical flux, mass loss, ignition and optical density/smoke release rate. The ideal SRM for fire resistance tests would be a material where the thermal conductivity, heat capacity and density are known and that these properties have either no temperature dependence or the temperature dependence is well defined. The most important characteristic of fire resistance tests is a defined thermal exposure scenario. Because the heat release rate and the smoke production of a material are generally related to mass loss rate, a possible SRM for these parameters may incorporate these inter-dependent properties. The SRM used for the Boeing/FAA E906 testing may be a possible candidate for tests that use small sample size. Another possibility is a tunable gas phase SRM, which may be an ideal media for this purpose, depending on the apparatus configuration. Indeed, methane sand burners are routinely used to calibrate oxygen depletion sensors for room scale tests. Similarly, a SRM could be developed that includes the related characteristics between critical flux, flame spread rate and ease of ignition. The composite FRP material used in the Boeing/FAA tests may also be a potential candidate SRM material for these tests. This would depend on the composite's scalability and physical behavior when large size samples are exposed to heat flux gradients common to these types of tests.

Although commercially available PMMA does consistently provide statistically reliable ranges of output data for an E1354, it is unlikely that PMMA could be used for tests where the sample sizes and orientation are substantially different. The obvious problem of thermoplastics, such as PMMA, is that they soften and melt at elevated temperatures. This problem is not bad at small scale unless the sample orientation is vertical or upside down. However, in large-scale tests, the physical characteristics of unsupported thermoplastics are problematic.

Other materials and composites that could be considered as candidate SRMs are:

- Thermosetting composites (Phenolic – Polyester – Polyester – Resins)
- Elastomeric materials
- Inert porous materials preloaded with thermoplastics or flammable gels.
- Honeycomb materials preloaded with thermoplastic or flammable gels.
- Reconstituted cellulosic's (particle board, melamine)
- Pre-loaded – NASA developed reentry materials.

All considered material would be required to be rated by the published production steps of ISO 34. Moreover, the recent criteria for NIST SRM certification will also be required. As stated in the 2008 SRM 1082, the Cigarette Ignition Strength Standard:

"A NIST certified value is a value for which NIST has the highest confidence in its accuracy and that all known or suspected sources of bias have been investigated or accounted for by NIST."

Can we afford SRMs?

In the NIST catalog of SRMs for engineering material, the price for the only in stock E05 SRM for E648 was $330.00 per unit, which is 3 sheets of the Kraft paper. Since the calibration requirement is for at least 3 replications, the total cost per calibration is ~ $1000.00. Similarly, the cost to construct a red oak deck for E84 calibrations at Commercial Testing Company in Dalton, GA is $225.00, including labor. Three replications are required and the total cost for annual calibration is ~ $700.00.

Once the SRM is available, the cost is not prohibitive. The cost to identify and validate the material to ISO 34 requirements would be very expensive. The last published NIST certification of the SRM for E162, required 20 samples plus a report to the NIST Measurement Services Division. At current cost and overhead charges, the price to validate a SRM is, without doubt, quite costly.

Conclusion

In an ideal world of flammability testing, all test standards would come with precision and bias statements. These statements would have been developed from a series of inter-laboratory studies where SRMs of specifically known properties, validated by the "production steps" listed in ISO 31 and certified by the Measurement Services Division of NIST, would provide each testing agency with a calibration standard of universally accepted accuracy. Developers of new

test methods would be required to identify candidate SRM materials as an integral part of the calibration process for the new test.

Paragraph 5.16 of ASTM Standard E535: Standard Practice for Preparation of Fire-Test-Response Standards details the requirements for ASTM Precision and Bias statements. Paragraph 5.16.2 states "inter-laboratory round robin evaluation is the best means for developing a satisfactory precision and bias statement." There is no requirement for SRMs in E535. However, the mandatory requirements for test calibration procedures are such that the availability of a SRM would be significantly useful as a calibration aid and to facilitate the required development of precision and bias statements.

APPENDIX

Committee E05.11 Fire Resistance

E119-10b Standard Test Methods for Fire Tests of Building Construction and Materials

Fire resistance. Time to reach unexposed face temperatures, flame through, are measured.

E814-11 Standard Test Method for Fire Tests of Penetration Firestop Systems

This method of testing through-penetration fire stops exposes fire stops to a standard temperature-time fire, and to a subsequent application of a hose stream. Fire resistance using E-119. Time to reach unexposed surface temperatures are measured.

E1529-10 Standard Test Methods for Determining Effects of Large Hydrocarbon Pool Fires on Structural Members and Assemblies

These test methods are intended to provide a basis for evaluating the time period during which a beam, girder, column, or similar structural assembly, or a nonbearing wall, will continue to perform its intended function when subjected to a controlled, standardized fire exposure. Fire resistance. For furnace type facilities calibration runs using instrumented noncombustible materials shall be conducted. Measurements are times to reach specified temperatures.

E1725-08 Standard Test Methods for Fire Tests of Fire-Resistive Barrier Systems for Electrical System Components

Fire resistance using either E-119 or E1529 time temperature curves. Acceptance temperature rise.

E1966-07 Standard Test Method for Fire-Resistive Joint Systems

Fire resistance using either E119 or E1529 time temperature curves.

E2032-09 Standard Guide for Extension of Data From Fire Resistance Tests Conducted in Accordance with ASTM E119

Extension of data.

E2226-10e1 Standard Practice for Application of Hose Stream

E2307-10 Standard Test Method for Determining Fire Resistance of Perimeter Fire Barrier Systems Using Intermediate-Scale, Multi-story Test Apparatus

Fire resistance using E119.

E2336-04(2009) Standard Test Methods for Fire Resistive Grease Duct Enclosure Systems

Fire resistance using E119 and E136.

Committee E05.11 Fire Resistance

E2748-10 Standard Guide for Fire-Resistance Experiments

E2749-10 Standard Practice for Measuring the Uniformity of Furnace Exposure on Test Specimens

This practice describes a procedure to gather data intended to measure the uniformity of exposure conditions upon test specimens for the fire test methods described in Test Methods E119, E814, E1529, E1725, E1966 and E2336

Committee E05.14 External Fire Exposures

E108-10a Standard Test Methods for Fire Tests of Roof Coverings

Surface spread of flame, and ability of the roof covering material or system to resist fire penetration. Also flying brands.

E2707-09 Standard Test Method for Determining Fire Penetration of Exterior Wall Assemblies Using a Direct Flame Impingement Exposure

The test method described herein measures the ability of the exterior wall covering material or system to resist fire penetration from the exterior to the unexposed side of the wall assembly under the specified conditions of exposure.

Committee E05.15 on Furnishings and Contents

E1352-08a Standard Test Method for Cigarette Ignition Resistance of Mock-Up Upholstered Furniture Assemblies

Ignition resistance of ulpholstered furniture to lighted cigarettes.

E1353-08ae1 Standard Test Methods for Cigarette Ignition Resistance of Components of Upholstered Furniture

Ignition resistance of ulpholstered furniture components to lighted cigarettes.

E1537-07 Standard Test Method for Fire Testing of Upholstered Furniture

The upholstered furniture specimen is allowed to burn freely under well-ventilated conditions after ignition using a propane gas burner. The most important fire-test-response characteristic measured is the rate of heat release.

E1590-07 Standard Test Method for Fire Testing of Mattresses

Heat release rate of mattresses.

E1822-09 Standard Test Method for Fire Testing of Stacked Chairs

Heat release rate of stacked chairs.

E2187-09 Standard Test Method for Measuring the Ignition Strength of Cigarettes

This test method enables comparison of the relative ignition strength of different cigarette designs.

E2280-09 Standard Guide for Fire Hazard Assessment of the Effect of Upholstered Seating Furniture Within Patient Rooms of Health Care Facilities

E2335-08 Standard Guide for Laboratory Monitors

E2653-09 Standard Practice for Conducting an Inter-laboratory Study to Determine the Precision of a Fire Test Method with Fewer Than Six Participating Laboratories

This practice describes the techniques for planning, conducting, analyzing, and treating results of an inter-laboratory study (ILS) for estimating the precision of a fire test method when fewer than six laboratories are available to meet the recommended minimum requirements of Practice E691.

Committee E05.17 on Transportation

E2061-09a Standard Guide for Fire Hazard Assessment of Rail Transportation Vehicles

Committee E05.11 Fire Resistance

E2230-08 Standard Practice for Thermal Qualification of Type B Packages for Radioactive Material

A package is considered qualified if material temperatures are within acceptable limits, temperature gradients lead to acceptable thermal stresses, the cavity gas pressure is within design limits, and safety features continue to function over the entire temperature range

Committee 5.21 Smoke and Combustion Products

E603-07 Standard Guide for Room Fire Experiments

E662-09 Standard Test Method for Specific Optical Density of Smoke Generated by Solid Materials

Determination of the specific optical density of smoke generated by solid materials and assemblies exposed to radiant and flaming exposures.

E800-07 Standard Guide for Measurement of Gases Present or Generated During Fires

E906/E906M-10 Standard Test Method for Heat and Visible Smoke Release Rates for Materials and Products Using a Thermopile Method

Determination of the heat release release rate and visible smoke from materials, products, or assemblies exposed to radiant heat.

E1354-11 Standard Test Method for Heat and Visible Smoke Release Rates for Materials and Products Using an Oxygen Consumption Calorimeter

Determination of the ignitability, heat release rates, mass loss rates, effective heat of combustion, and visible smoke development of materials and products.

E1474-10 Standard Test Method for Determining the Heat Release Rate of Upholstered Furniture and Mattress Components or Composites Using a Bench Scale Oxygen Consumption Calorimeter

Determination of the ignitability and heat release from the composites of contract, institutional, or high-risk occupancy upholstered furniture or mattresses.

E1623-11 Standard Test Method for Determination of Fire and Thermal Parameters of Materials, Products, and Systems Using an Intermediate Scale Calorimeter (ICAL)

Heat release rate of specimens exposed to a radiant panel.

E1678-10 Standard Test Method for Measuring Smoke Toxicity for Use in Fire Hazard Analysis

Determination of lethal toxic potency of smoke produced from a material or product ignited while exposed to a radiant heat flux.

E1740-10 Standard Test Method for Determining the Heat Release Rate and Other Fire-Test-Response Characteristics of Wall Covering or Ceiling Covering Composites Using a Cone Calorimeter

Determination of the ignitability and heat release rate of composites consisting of a wall-covering, a substrate, and all laminating adhesives, coatings, and finishes.

E1995-08 Standard Test Method for Measurement of Smoke Obscuration Using a Conical Radiant Source in a Single Closed Chamber, With the Test Specimen Oriented Horizontally

Measuring smoke obscuration of essentially flat materials, products, or assemblies (including surface finishes), in a horizontal orientation, exposed to thermal irradiance.

E2067-08 Standard Practice for Full-Scale Oxygen Consumption Calorimetry Fire Tests

Committee E05.11 Fire Resistance

E2102-09 Standard Test Method for Measurement of Mass Loss and Ignitability for Screening Purposes Using a Conical Radiant Heater

Measuring mass loss and ignitability, for screening purposes, from essentially planar materials, products, or assemblies (including surface finishes), exposed to radiant heating, with or without an external igniter.

E2257-08 Standard Test Method for Room Fire Test of Wall and Ceiling Materials and Assemblies

Test indicates the maximum extent of fire growth in a room, the rate of heat release, and if they occur, the time to flashover, and the time to flame extension beyond the doorway following flashover. It determines the extent to which the wall and ceiling materials or assemblies contribute to fire growth.

E2405-05 Standard Test Method for Determination of Fire and Thermal Parameters of Materials Using an Intermediate Scale Test with Vertically Oriented Specimen

Test method is primarily to determine time to- ignition, vertical flame spread rate, and lateral flame spread rate of materials, products and assemblies in a vertical orientation when exposed to radiant heat flux.

Committee E05.22 on Surface Burning

D2859-06 Standard Test Method for Ignition Characteristics of Finished Textile Floor Covering Materials

Methenamine ignition test on horizontal specimen. Char shall not extend to within 25.4 mm edge of the hole.

E84-10b Standard Test Method for Surface Burning Characteristics of Building Materials

Measurements of surface flame spread and smoke density.

E162-09 Standard Test Method for Surface Flammability of Materials Using a Radiant Heat Energy Source

Meaures surface flamespread rate.

E648-10e1 Standard Test Method for Critical Radiant Flux of Floor-Covering Systems Using a Radiant Heat Energy Source

Determines critical radiant flux which is one measure of the sensitivity to flame spread of floor-covering systems in a building corridor

E970-10 Standard Test Method for Critical Radiant Flux of Exposed Attic Floor Insulation Using a Radiant Heat Energy Source

Determines critical radiant flux which is one measure of the surface burning characteristics of exposed insulation floors or attics.

E1317-08b Standard Test Method for Flammability of Marine Surface Finishes

Measures ignitability, heat exposure for continued burning, critical flux at extinguishment, and heat-release behavior under varying flux exposure conditions applied.

E1321-09 Standard Test Method for Determining Material Ignition and Flame Spread Properties

Determines material properties related to piloted ignition of a vertically oriented sample under a constant and uniform heat flux and to lateral flame spread on a vertical surface due to an externally applied radiant-heat flux.

E2058-09 Standard Test Methods for Measurement of Synthetic Polymer Material Flammability Using a Fire Propagation Apparatus (FPA)

Committee E05.11 Fire Resistance

The Ignition and Combustion test methods involve the use of horizontal specimens subjected to a controlled, external radiant heat flux. The Fire Propagation test method involves the use of vertical specimens subjected to ignition near the base of the specimen from an external radiant heat flux and a pilot flame.

E2231-09 Standard Practice for Specimen Preparation and Mounting of Pipe and Duct Insulation Materials to Assess Surface Burning Characteristics

E2404-10 Standard Practice for Specimen Preparation and Mounting of Textile, Paper or Vinyl Wall or Ceiling Coverings to Assess Surface Burning Characteristics

E2573-07a Standard Practice for Specimen Preparation and Mounting of Site-Fabricated Stretch Systems to Assess Surface Burning Characteristics

E2579-07 Standard Practice for Specimen Preparation and Mounting of Wood Products to Assess Surface Burning Characteristics

E2599-11 Standard Practice for Specimen Preparation and Mounting of Reflective Insulation, Radiant Barrier and Vinyl Stretch Ceiling Materials for Building Applications to Assess Surface Burning Characteristics

E2688-10 Standard Practice for Specimen Preparation and Mounting of Tapes to Assess Surface Burning Characteristics

E2690-10e1 Standard Practice for Specimen Preparation and Mounting of Caulks and Sealants to Assess Surface Burning Characteristics

Committee E05.23 on Combustibility

E136-11 Standard Test Method for Behavior of Materials in a Vertical Tube Furnace at 750°C

Test method assists in indicating those materials which do not act to aid combustion or add appreciable heat to an ambient fire. Measurements include mass loss, surface and interior temperature rise.

E2652-09a Standard Test Method for Behavior of Materials in a Tube Furnace with a Cone-shaped Airflow Stabilizer, at 750°C

Test method assists in indicating those materials which do not act to aid combustion or add appreciable heat to an ambient fire. Measurements include mass loss and temperature rise.

Committee E05.31 on Terminology and Services/Functions

E176-10ae1 Standard Terminology of Fire Standards

E535-09 Standard Practice for Preparation of Fire-Test-Response Standards

E2536-09 Standard Guide for Assessment of Measurement Uncertainty in Fire Tests

Committee E05.33 on Fire Safety Engineering

E1355-11 Standard Guide for Evaluating the Predictive Capability of Deterministic Fire Models

E1472-07 Standard Guide for Documenting Computer Software for Fire Models

E1546-09a Standard Guide for Development of Fire-Hazard-Assessment Standards

E1591-07 Standard Guide for Obtaining Data for Deterministic Fire Models

E1776-07 Standard Guide for Development of Fire-Risk-Assessment Standards

E1895-07 Standard Guide for Determining Uses and Limitations of Deterministic Fire Models

References

[1] ASTM E691-05, 2005, "Standard Practice for Conducting an Inter-Laboratory Study to Determine the Precision of a Test Method" *ASTM book of Standards,* ASTM International, West Conshohocken, PA.

[2] Twilley, W. H. and Babrauskas, V., 2007, ASTM Fire Standards and Related Technical Materials, ISBN: 978-0-8031-5685-2, ASTM Stock#: FIRE07CD, ASTM International, W. Conshohocken, PA, 2007.

[3] Standard Core Panel, Schneller, 6019 Powdermill Rd., Kent, OH 44240.

[4] Keltner, N., *Discussion on E5.11*, rewrite, Jan 2, 2006, private communication.

Reprinted from JTE, Vol. 40, No. 1
doi:10.1520/JTE103919
Available online at www.astm.org/JTE

Ned Keltner[1]

What Have We Learned About Uncertainty? Are We Still Playing with Fire?

ABSTRACT: Accurate characterization of the thermal environment in fires and fire safety tests is crucial to understanding the impact of fire on materials and assemblies and the work on modeling those impacts. Standardized fire resistance test methods have been used to help evaluate building materials and assemblies since the early 1900s and have been part of a successful effort to improve fire safety. However, current methods use predominately temperature measurements and do not provide sufficient heat transfer data to support the development and validation of fire protection engineering models. This paper looks at measurement techniques that were used successfully in pool fire research and how they could be applied to modeling structure fires.

KEYWORDS: temperature, heat flux, uncertainty analysis, directional flame thermometers, fire resistance, pool fires

Introduction

A 1995 paper titled "Are We Playing with Fire?" considered nontraditional temperature and heat flux measurement errors in fire safety testing [1]. Nontraditional errors include such things as the effects of the sensor design and installation techniques along with dynamic measurement uncertainties. Studying nontraditional errors is important because these errors generally produce much larger uncertainties than the traditional uncertainty analyses would suggest. Large errors in temperature and heat flux measurements affect the quality of risk assessments for fires in buildings, manufacturing facilities, and transportation systems.

Standardized fire resistance test methods, such as ASTM E119 [2] and ISO834 [3], have been used to help evaluate building materials and assemblies

Manuscript received April 16, 2011; accepted for publication November 11, 2011; published online December 2011.

[1] Fire Instrumentation Research and Engineering Services, Albuquerque, NM 87112.

Cite as: Keltner, N., "What Have We Learned About Uncertainty? Are We Still Playing with Fire?," *J. Test. Eval.*, Vol. 40, No. 1. doi:10.1520/JTE103919.

since the early 1900s. Standardized test methods have been part of a successful effort to improve fire safety.

However, our use of data from fire resistance and other tests is changing. While we still need ratings and answers to the "did it pass or fail" question, there is a growing need to use this large scale test data to help develop engineering models of the thermal and structural performance of materials and assemblies during varied fire exposures. For these applications, we are on very shaky ground.

One of the recommendations that came out of the investigations by NIST of the World Trade Center was the need for quantitative heat flux measurements during fire safety tests. Unfortunately, the large scale standardized test methods from both ASTM and ISO rely heavily on temperature measurements to control thermal exposure and determine response. The result is that we do not have sufficient information to accurately model the 'fire environment' and the 'response to fire' of materials and assemblies exposed to it.

The need for high-quality experimental data to support model development and validation requires an increased effort to perform well-designed and well-characterized experiments of different sizes. It has placed additional emphasis on the appropriate application of existing measurement techniques and the development of new techniques. Uncertainty analysis is necessary for all of this work.

For fire protection engineering, an attractive alternative to conducting numerous large scale tests is a program coupling numerical simulations with well-designed experiments. In such a program, the development of numerical simulation tools must be supported by well-characterized measurements. Code validation using high fidelity experimental data is an essential component of the model development process. Comparing model predictions with experimental results also helps to refine future experiments and direct research efforts towards the most appropriate areas. This paper will look at the history of pool fire research as a possible guide for fire resistance research.

Temperature and Heat Flux Measurement Uncertainty

Measurements are often considered to be the standard against which we should compare and validate our analyses. The feeling is that what happens (at least, as it is described by our measurements) is what actually happens. If we ever hope to approach this measurement utopia, we must understand our measurements as well as the modelers understand their models.

Application measurements are different from laboratory or calibration measurements made under well controlled conditions. As a result, there is a need to understand how the sensors work and how the conditions in the application affect the accuracy of the measurement.

Making accurate measurements in any type of fire test is challenging. Understanding the total uncertainty is even harder. The problems are especially difficult because there are so many nonlinear phenomena in fires that can affect the measurements. In addition, there are dynamic uncertainties because the thermal conditions generally vary with time.

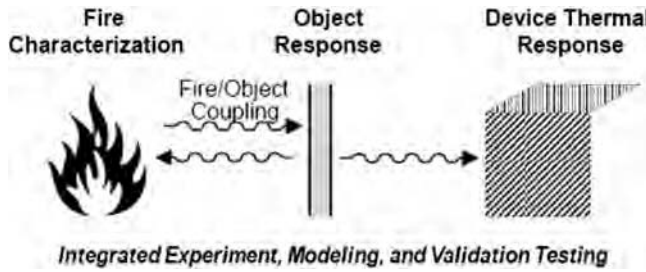


FIG. 1—*Schematic of how coupling affects thermal measurements.*

Unfortunately, it is easy to be lied to by (or to lie with) our measurements, especially in fire testing. The basic effect is called *Measurement Magic* [1]. In this type of illusion, the sensor appears to be giving a correct reading or at least one that agrees with your intuition. Measurement Magic is not an artificial effect, such as lying with statistics; and it is not as obvious. Detailed uncertainty analysis helps unlock the secrets of Measurement Magic. (Note that there is a corresponding form of illusion called *Model Magic* [1]. However, that's another story.)

What causes *Measurement Magic*? The basic process is shown in Fig. 1 [4]. For example, stick a bare wire thermocouple into a flame and it indicates an elevated temperature, as expected. However, insert a sheathed thermocouple and it will indicate a different temperature. Is either one correct? The different shapes and possibly sizes change the coupling between the flame (test) environment and the sensors. Each design has different stem conduction and dynamic response characteristics. The result is two different "measured flame temperatures." Detailed studies of thermocouple measurement uncertainties in fire environments are given by Nakos [5,6].

In the first "Playing with Fire" paper, one measurement problem example involved temperature (thermopile) measurements in the stack of the Ohio State University Heat Release Rate fixture [7]. The analysis showed that the energy balance for the stack thermopile was different between the "clean" calibration run, which used methane, and the "dirty" test runs where smoke and soot in the stack gases affected the radiation heat transfer and changed the coupling. The effect on the uncertainty of the predicted heat release rate is unknown. Heat release rate tests are now focused on using oxygen consumption measurements, as in the cone calorimeter [8], instead of relying on stack temperature measurements.

An example of a heat flux measurement problem involves the use of water cooled heat flux gauges in fire resistance tests. When inserted through the insulated furnace walls, the cold gauge surface trips the boundary layer on the hot surface. The turbulent boundary layer dramatically increases the convective heat transfer coefficient, which changes the coupling between the gauge and the furnace. This problem is further compounded by the fact that the radiant–-convective sensitivities of Gardon gauges [9] and Schmidt-Boelter gauges [10] are different.

Another heat flux measurement error comes from measuring hemispherical radiative heating with a water-cooled heat flux sensor that has been calibrated

with a narrow angle heat source or vice versa. This error stems from the fact that the hemispherical absorptivity is generally 3–5 % lower than the near normal absorptivity [11].

These problems are but a small part of a long list. Temperature and heat transfer measurements are not always easy!

Uncertainty Analysis in Fires and Fire Safety Tests

Of all the observable thermal characteristics of a fire, temperature is often considered to be the most important. Thermal transport within any fire involves a combination of radiation, including that from participating media, and convection in a turbulent, chemically-reacting, buoyancy-driven flow. Understanding the effect of fire on objects and materials requires an understanding of the temperature field in the nearby fire environment. Temperature is a convenient parameter for delineating fire boundaries and zones, thus helping to establish the fire structure.

Knowledge of the material and media temperatures allows other characteristics such as ignition, failure limits, buoyancy, reaction kinetics, and chemical species production to be inferred. Because temperature is important in understanding fire and characterizing fires, the apparent ease with which temperature data can be acquired adds to its desirability. However, material, and especially media, temperatures in fire tests can be difficult to measure and interpret.

Thermocouples are by far the most widespread instruments used for obtaining "fire temperature" measurements because they are rugged, economical, and relatively easy to install. Thermocouples are also the most common thermal instruments used in fire response tests. However, thermocouples are deceptively simple. Because they consist of just a pair of wires, they are perceived as being easy to use and thus easy to understand sensors. This is a false assumption! For details, see ASTM's Manual on the Use of Thermocouples in Temperature Measurement [12].

In Jakob's classic heat transfer text [13], the title of one section is "Heat Transfer, a Prerequisite of Thermometry." For thermocouple based temperature sensors, the "measured temperature" is the result of a continuous analog energy balance involving many different heat transfer paths. In fire resistance tests, these include radiation exchange with the bounding surfaces, radiative and convective heat transfer with the gases in the furnace volume, storage in the sensor due to the continually changing temperatures, plus any heat conduction down the thermocouple stem.

Due to this energy balance process, 'flame or furnace temperatures' measured by thermocouples are, at best, equal to the temperature of the thermocouple hot junction (i.e., a different physical object than the flames). Because thermocouples are physical objects in the fire environment, thermocouple temperature measurements cannot be directly compared to the calculated nodal temperatures which are readily available from fire models. However, virtual sensors offer the potential to relate the experimental and analytical results.

One classic paper concerning thermal measurement uncertainty is "Using Uncertainty Analysis in Planning of an Experiment," by Moffat [14]. Moffat also

developed a Temperature Measurement Error Analysis Hierarchy that is very useful [15]. It also applies to many heat flux measurements because most depend on measurements of temperature or temperature differences:

- Undisturbed value of measurand—this is the value we want.
- Available value—this is the best we can do with the sensor in place.
- Achieved value—this is what we can measure.
- Observed value—measurement returned by data acquisition system in engineering units.
- Corrected value—application of any corrections.

The difference between the values of the undisturbed, the available, and the achieved temperatures is due to the thermal coupling between the sensor and the application environment, as shown in Fig. 1. The coupling is affected by a number of items, including: the sensor geometry and thermal properties, its dynamic response characteristics and, by the installation techniques. All of these affect the heat transfer and thus change the measured temperature.

A report by Blanchet et al. used a combination of analysis and experiments to examine the differences between the heat fluxes calculated with a simplified response model and a detailed one for a sheathed thermocouple attached to a thin metal plate [16]. The report concludes that "missing physics" between the two models was the main contributor to the large errors (+60 %/ −20 %) in estimates of thermal exposure. This report provides an approach for estimating the uncertainty associated with using either the ASTM E119 shielded thermocouple [2] or the ISO834 plate thermometer [3] for estimating the thermal exposure in standardized fire resistance tests.

These differences reported in Blanchet et al. are a result of undiagnosed measurement errors or uncertainties, between the first and third levels of Moffat's Hierarchy [15]. This part of temperature measurement uncertainty analysis is often ignored. Common uncertainty analyses generally look at the difference between the achieved and observed values of temperature.

A common perception is that the fire environment can be specified by specifying the "fire temperature." As a result of such undiagnosed errors, data presented later will show that this assumption is only good to within a factor of 2! The data will also show that determining the actual thermal environment in a fire safety test requires independent measurements of the temperature and heat flux.

One important benefit of an uncertainty analysis is that it provides a road map of how to best reduce the overall uncertainty in the measured temperatures or heat fluxes. The uncertainty analysis is not intended to produce precise values, but rather to estimate the uncertainty. A quotation by Dowdell [17] summarizes these sentiments: "There is a lot of uncertainty in uncertainty analysis but some analysis is better than no analysis at all."

Thermal Measurements in Pool Fires

One of the recommendations that came out of NIST's World Trade Center investigations was the need for quantitative heat flux measurements in fire safety tests. The NIST report stated such data are needed to support the development

and validation of fire protection engineering models. A report developed for The Fire Protection Research Foundation expanded on this recommendation [18].

ASTM Committee E5 has been working to incorporate heat flux measurements into its two fire resistance test methods: ASTM E119 and ASTM E1529. Some perspective on the problems facing E5 can be gained by looking into the history from similar work on large pool or spill fires. In fact, these results were used to help develop E5's hydrocarbon-curve, fire-resistance test: E1529 [19].

For over 50 years, pool fires tests have been part of the testing protocol required for qualifying hazardous and radioactive material shipping containers. In the 1960s, the perception was that the pool fire environment could be specified by specifying the "fire temperature."

A landmark report published by Sandia National Laboratories in 1965 stated that the heat flux exposure in large, JP-4 pool fires could be accurately described as blackbody radiation from a 1010 °C (1283 K) source [20]. This simplified model was based on measurements made near the centerline of an 18 ft × 18 ft pool fire using thermocouples and specially designed calorimeters mounted within a 5 ft square in the center. This work helped form the basis for hazardous material transport regulations issued by the U.S. DOT, DOE, and NRC and the International Atomic Energy Agency.

Twenty years later, an extensive database of pool fire measurements had been developed. These helped provide a much better understanding of the fire environment. These data sets showed that the heat flux was not constant in the central region [21–24] and not uniform around an object. The data in Table 1 and Fig. 2 shows that the measured heat flux depends on the properties of the object being used to measure it. As a result of changes in the fire-object coupling shown in Fig. 1, the heat flux measured with a "physically large and thermally massive" object was lower than that measured by a smaller object [25,26].

Most references indicate radiation accounts for over 90 % of the total heat transfer in pool fires. To evaluate this radiation-dominated hypothesis, the measured heat flux versus the measured local temperature was compared to the heat flux calculated from the Stefan-Boltzmann Equation.

Figure 2 shows data from a number of large pool fire tests where the heat release rate is approximately 500 MW. The relationship between the measured temperature and measured heat flux is anything but clear.

Estimating the heat flux by calculating the blackbody radiative flux corresponding to the measured 'fire temperature' provides a value that is only good to within a factor of two of the measured heat flux. This factor of 2 uncertainty might suggest a temperature measurement uncertainty of +19 %/−16 %. Alternatively, it might be due to the different calorimeter designs.

TABLE 1—*Heat flux measurements in 500 MW JP4 pool fires [23].*

Calorimeter	Average Peak Cold Wall Flux, kw/m^2	Convection, %
10 cm, dia	165	20–25
20 cm, dia	165	20–25
1.4 m, dia	130	8

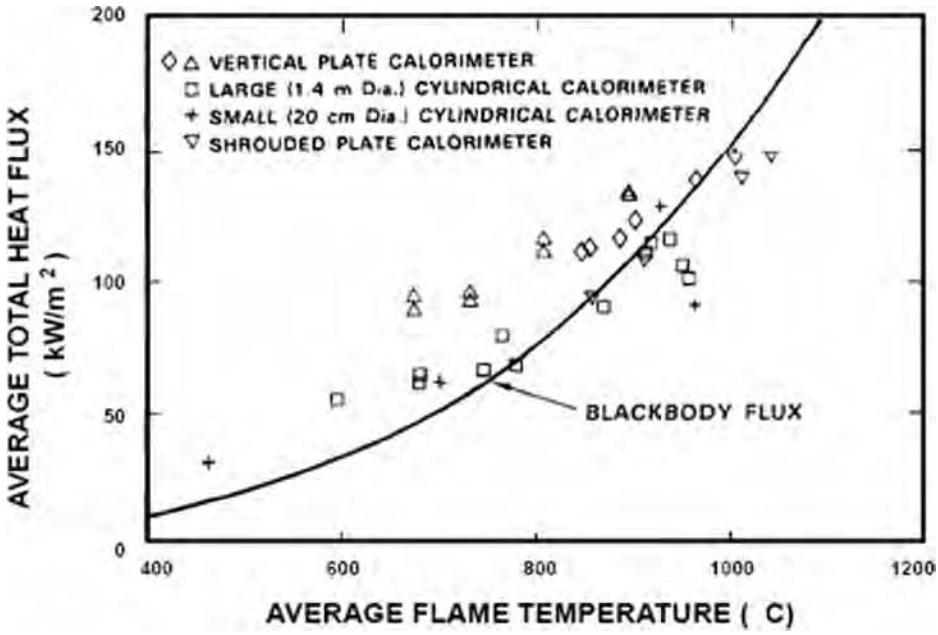


FIG. 2—*Average cold wall heat flux versus average measured flame temperature [21].*

Figure 2 also shows that the 150 kW/m^2 heat flux level {blackbody radiation from a 1010 °C (1283 K) source}, as described by Bader, is more of an upper limit than a specification [20]. This body of work demonstrated that independent measurements of heat flux and temperature are necessary for characterizing large fires.

As part of a probabilistic risk assessment program sponsored by the Defense Nuclear Agency, numerical simulations were needed to predict the thermal exposure in large aircraft crash fires in many different accident scenarios. This combination of models and experiments is similar to NIST's goal.

Advanced fire physics models require state-of-the-art submodels (combustion, multidimensional participating radiation, convection, etc.) which are coupled with the governing equations for the flow-field. Simplified models apply first principles only to the dominating physical phenomena and rely on empirical factors to represent the remaining physics. Such partially empirical models have been developed into predictive tools to reduce computer run times. Parametric analyses were developed to aid in the understanding of what affects the thermal coupling phenomena in fires and provide quick estimates [26].

To support the development of these various models, tests were conducted to simulate fuel spill fires that might occur under the wing of a C-141 transport aircraft. The thermal response of such a system in the fire strongly depends on the magnitude of the heat flux and the partitioning between convective and radiative heat fluxes. Measurements were made of the fire temperatures and heat fluxes to the underside of the wing and the under-wing fuselage surfaces. Again,

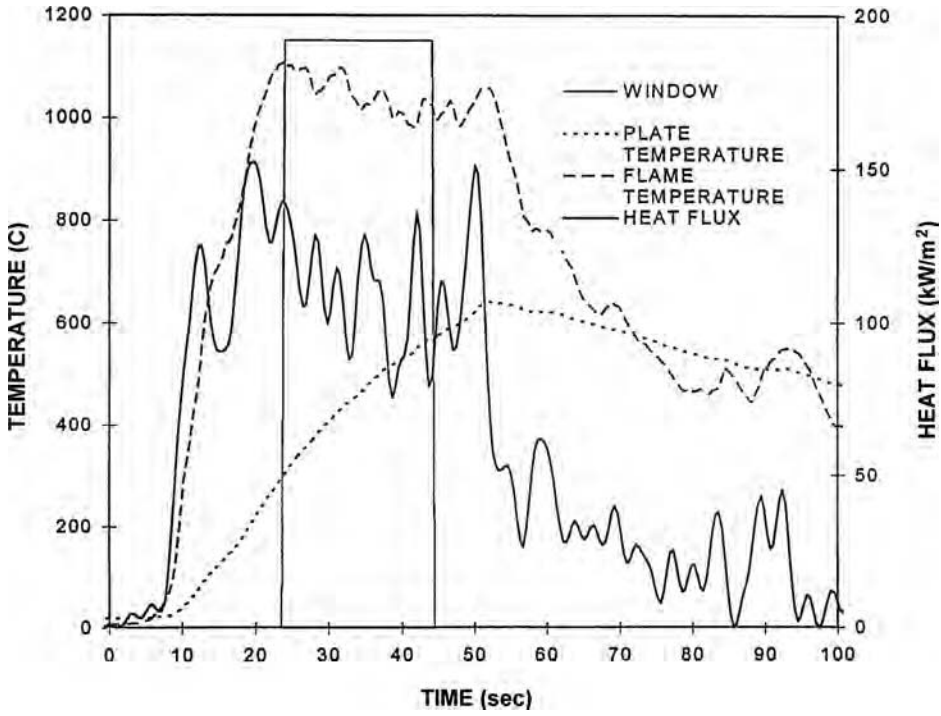


FIG. 3—*Heat flux measurements on a simulated aircraft wing.*

temperature and heat flux measurements were needed to adequately specify the fire exposure.

Figure 3 shows the data for one location in a spill fire under a simulated large airplane [27]. The wing was simulated by a 1.52 mm thick, 304 stainless steel plate. A sheathed thermocouple was attached to the unexposed surface. The "flame temperatures" were measured with sheathed and bare bead thermocouples mounted 10.2 cm below the wing. In this case, the estimated time constant of the plate was 20–22 s. For changes in the plate temperature, the dynamic response of the thermocouple and attachment was shown to be $1 - \exp(-t/\tau)$, where the time constant $\tau = 1.9$ s [28].

Differential compensation of the dynamic errors in the measured plate temperature history was used to provide an effectively sub-second response [20]; the compensated data were used with thin-skin calorimeter methods reported in ASTM E459 to calculate the heat flux history [29]. As can be seen in Fig. 3, the heat flux oscillates rapidly during the 20 s measurement window. In the measurement window, the total heat transfer coefficient was approximately 250 W/m^2 K.

Using heat flux versus local temperature data from the simulated aircraft fires along with earlier pool fire data sets, statistical analysis demonstrated that the measured *heat flux depends on the shape, size, location in the fire, and orientation of the object*. Polynomial regressions were used to provide mathematical expressions for each data set. Based on a forward F-test procedure, second

order polynomials provide the best representations of these two data sets [19]. For each regression curve (Q_{avg}), the 95 % confidence interval for the regression curve (Cl_{line}) and the standard deviation for the next observation (S_{point}) were calculated for each of the data sets.

Vertical Plate Calorimeters—Open Pools

$$Q_{avg} = 103.98 - 1.335E-1\,T + 1.696E-4\,T^2 \quad (1a)$$

$$CI_{line} = (59315 - 279.1\,T + 0.492\,T^2 - 3.84E-4\,T^3 + 1.13\,E-7\,T^4)^{0.5} \quad (1b)$$

$$S_{point} = (12957 - 60.66\,T + 0.107\,T^2 - 8.35E-5\,T^3 + 2.46E-8\,T^4)^{0.5} \quad (1c)$$

Underwing Fuselage Surface (Plates)

$$Q_{avg} = 10.42 - 1.154E-2T + 1.068E-4\,T^2 \quad (2a)$$

$$CI_{line} = (186.96 - 1.20T + 3.04E-3\,T^2 - 3.37E-6\,T^3 + 1.36E-9\,T^4)^{0.5} \quad (2b)$$

$$S_{point} = (126.6 - 0.312T + 7.91E-4\,T^2 - 8.77E-7\,T^3 + 3.54E-10\,T^4)^{0.5} \quad (2c)$$

The confidence limits are related to the standard deviations by the expression

$$S_i = CI_i / T_{n-2,1-\alpha/2} \quad (3)$$

where $t_{n-2,\,1-\alpha/2}$ is the $1-\alpha/2$ percentage point of the t-distribution with $n-2$ df. The standard deviation for the next point is related to the standard deviation of the line by the relation

$$S^2_{point} = \sigma^2 + S^2_{line} \quad (4)$$

where σ is the standard deviation of the residuals. Due to the significantly larger number of data points, the confidence intervals for the current test data are much smaller than those for the earlier tests.

Comparisions of these correlations with the data are shown in Fig. 4 [Note: Ref [7] includes two other correlations]. It is clear that the heat fluxes from pool/spill fires under the wing area of an aircraft are lower than the heat fluxes measured with the vertical plate and horizontal cylinder calorimeters in open pool fires. It is unknown if this is due to changes in air entrainment or something else. This difference demonstrates the importance of conducting fire tests in as realistic a configuration as possible in order to obtain the best data.

To use the data in Fig. 4, a technique for estimating the fire temperature was needed for the statistical analyses of the temperature and heat flux data to develop empirical temperature-to-heat flux mappings. Direct comparisons

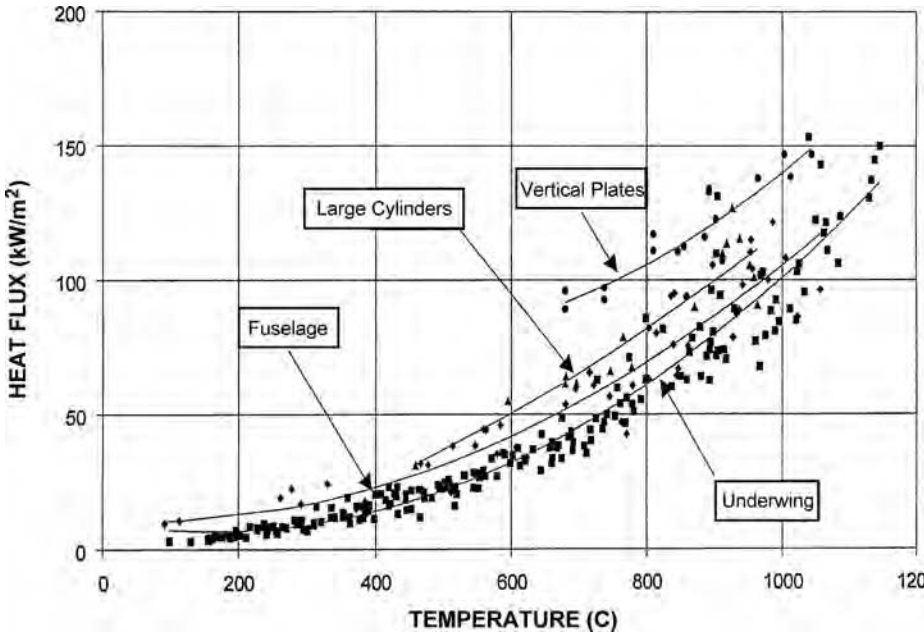


FIG. 4—*Temperature versus heat flux mappings for four calorimeter configurations.*

between the average centerline temperatures measured at a given elevation in different sizes of fires showed poor agreement. Figure 5 shows that specifying the location and the fire size by scaling the elevation with the total heat release rate gave good results [27].

All of these studies have shown that the heat flux cannot be accurately estimated from a measured fire temperature. Independent measurements of temperature and heat flux are necessary to characterize a "fire environment." The studies have shown that the heat flux is affected by the size, shape, orientation, and physical properties of the sensor or test specimen. This is the coupling process sketched in Fig. 1. The work has also shown that the best option for comparing temperature measurements with models is to build a thermal response model of the sensor, i.e., a virtual sensor, into the numerical simulation [30,31].

Thermal Measurements In Fire Resistance Tests

In standardized fire resistance tests, such as ASTM E119 or ISO 834 or IMO A754, the furnace temperature is controlled to a standard time-temperature curve. Historically, implicit assumptions were made that the thermal exposure could be described solely by the measured furnace temperature history and that the exposure would be repeatable from time to time and place to place. Historical variations in the qualitative fire protection ratings (e.g., 1 h) of up to 50 % or more between different furnaces or laboratories indicates these assumptions were not well founded.

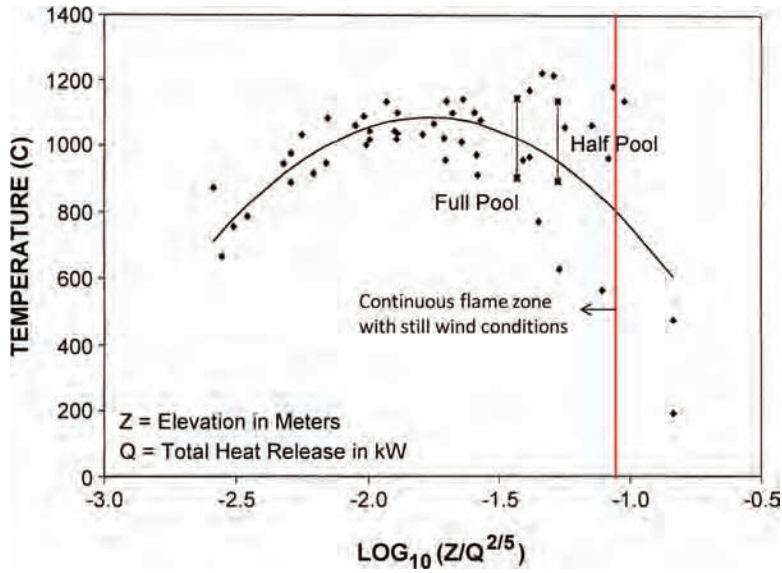


FIG. 5—*correlating local temperatures in pool/spill fires to accident conditions [27].*

In the mid-1990s, the U. S. Coast Guard authorized a study of fire resistance tests in furnaces. The goals were to better understand what factors produced the large uncertainties in current fire safety test methods and to help address any differences between national and international standards [30,31]. As in the pool/spill fire studies, one important conclusion was that heat transfer in furnaces could not be predicted solely from furnace temperature measurements without large static and dynamic uncertainties [32].

This conclusion was supported by subsequent characterization tests conducted in fire resistance test furnaces. A large actively-cooled heat flux measurement system, furnace characterization unit, was designed to simulate a marine fire barrier [33]. Heat flux data from these tests showed that the thermal exposure in vertical furnaces was different from that in horizontal furnaces. It also showed that the exposure varied with location in the furnaces, e.g., walls and ceilings and floors.

In fire safety tests, temperature and heat flux measurement uncertainty (quality) are intimately related to the ability to correctly understand and model the interaction of the sensor with its environment. Despite the apparent ease of acquiring thermocouple data in fire resistance tests, accurate interpretation of these measurements can be difficult. A good understanding of the heat transfer measurements is critical to a good understanding of the temperature measurements; the reverse is also true. Heat flux is a critical measurement in fire response tests. The measurement analysis techniques developed in pool fires can aid in the understanding of the temperature measurements and control of fire resistance testing along with the uncertainty of these measurements [26].

As with pool fires, the problem in fire resistance tests is the *furnace thermocouples can only measure of their own temperature*! There is no such thing as "a

furnace temperature" because the wall, ceiling, floor, and gas temperatures are all generally different. Therefore, we are left with only "furnace temperature measurements." As the pool fire research demonstrated, independent measurements of the heat flux are required because furnace temperature measurements cannot be used to accurately estimate the heat flux and its uncertainty.

In one of the initial studies of the temperature measurements in fire resistance tests, Babrauskas and Williamson compared the response of the ASTM E119 shielded thermocouple [2] and a 0.81 mm bare wire thermocouple in an E-119 furnace [34]. During the first five minutes of the test, the difference between the ASTM control thermocouple and the bare wire was as much as 550 °C.

In furnaces that use relatively clean-burning fuels, a large fraction of the radiative heat flux comes from the walls. The furnaces are often controlled by thermocouples mounted close to the test specimen. As a result of the energy balance previously described, the thermocouple readings are typically at an intermediate temperature between the temperatures of the walls, the furnace gases, and the specimen. As shown in the following text, these biased temperature measurements bias the thermal exposure estimates as well!

Even though the ASTM E119 [2], ISO 834 [3], and IMO A754 fire resistance tests use basically the same time-temperature curve, the thermal exposure (incident heat flux) in each is different because the thermocouples used for furnace control are different. The ISO and IMO currently use plate thermometers for furnace control; however, both test methods previously used bare bead thermocouples. The IMO method used to allow five different thermocouple designs.

Recent studies have shown that changing the sensor design changes both the measured furnace temperature and the thermal exposure in fire resistance tests [35–37]. Figure 6 shows the results of measurements in an ASTM E119 test from arrays of six different temperature sensor designs including:

- *ASTM Method E-119 Shielded Thermocouple*—beaded thermocouple mounted in a 0.84 in OD (21.3 mm) thermowell with a 0.11 in (2.8 mm) thick wall.
- *ISO 834 (and IMO A754) Plate Thermometers (PTs)*—100 mm × 100 mm Inconel 600 plate, 0.7 mm thick, with a surface emissivity greater than 0.7 [3]. A Type K, 1 mm OD, sheathed thermocouple is mechanically attached to the unexposed surface of the plate by a strap. The thermocouple is pressed against the surface by a10 mm thick ceramic fibreboard.
- *Directional Flame Thermometer (DFT)*—DFTs are thermocouple based heat flux sensors where are two Inconel 600 plates with a layer of insulation material in between [38,39]. Inconel sheathed thermocouples (1.6 mm OD) are thermo-mechanically attached to the unexposed surfaces. The plates are heavily oxidized—the nominal absorptivity is 0.85. The measured front plate temperatures are shown.
- *Bare Bead Thermocouple*—20 gauge, Type K wire twisted together and fused to form a 2 mm bead.
- *Inconel Sheathed Thermocouples*—6.5 mm OD, grounded and ungrounded junction designs [12].

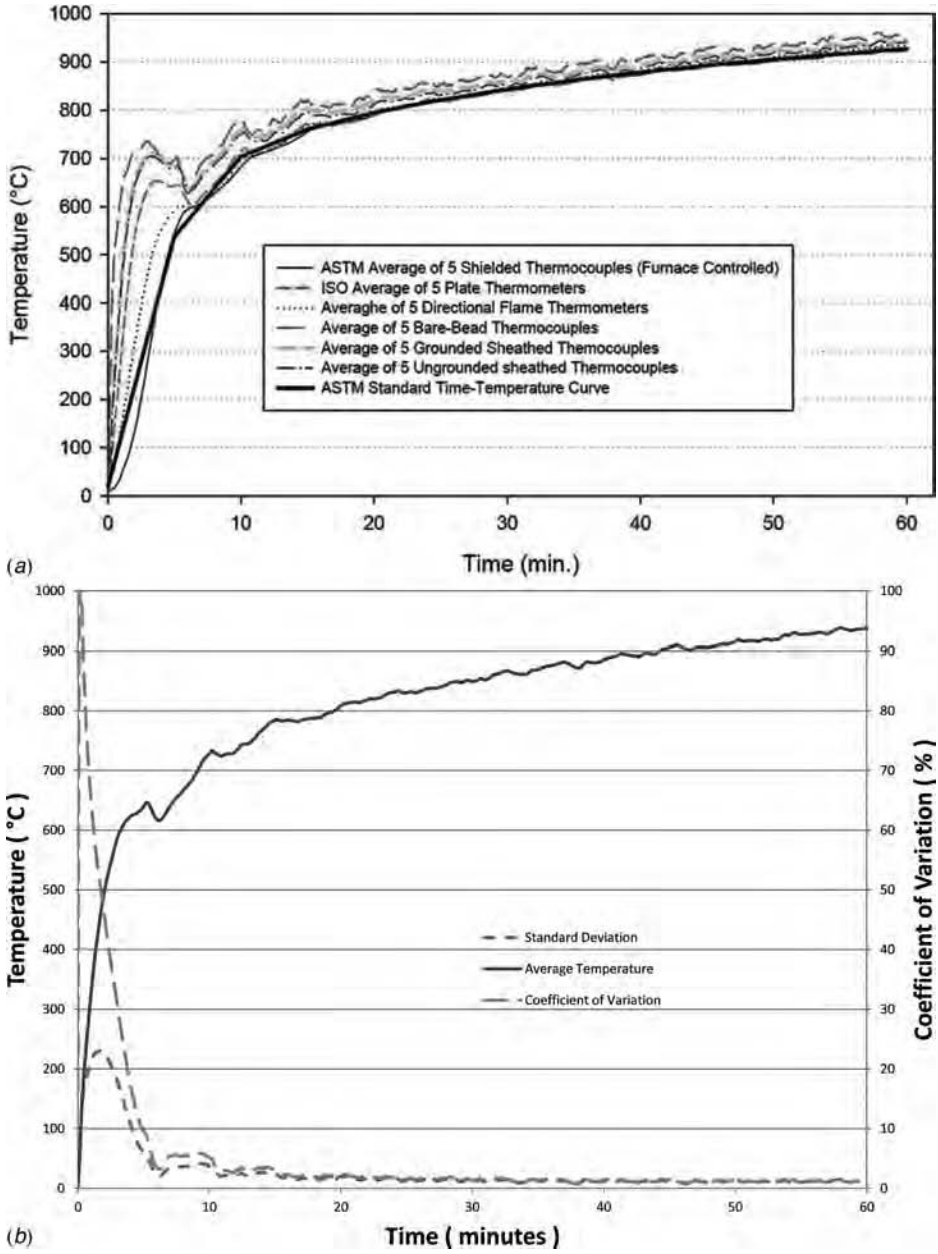


FIG. 6—*(a) Measured temperatures in a floor furnace test controlled by an E119 thermocouple. (b) Analysis of the measured temperatures in a floor furnace test.*

Figure 6(*a*) shows that each temperature sensor design provides a different furnace temperature measurement history. Figure 6(*b*) shows the average value, the sample standard deviation and the coefficient of variation (CoV) of the data. The CoV is defined as sample standard deviation/average value. The CoV peaks out just over 100 % during the first minute after ignition. At 2 min, the average measured temperature is about 500 °C; the CoV is about 45 % or 225 °C.

Figure 1 indicates that coupling between the sensors and the environment affects the actual temperature measurements. For example, consider the object in the center of the Fig. 1 to be the exposed surface of the temperature sensors. For all of the sensors except the bare bead thermocouple, the thermocouple hot junction is then the device on the right hand side.

The different sensor geometries and properties change the heat transfer from the fire to the exposed surface element. As compared with the bare bead thermocouple, the larger size of the other five temperature sensor designs reduces their sensitivity to convective heat transfer; i.e., it changes the radiative/convective partitioning. The surface absorptivity values for the bare wire and sheathed thermocouples are likely lower than those of oxidized surfaces of the directional flame thermometer [37,38], plate thermometer [3] and the E119 thermocouple [2]. The differences in absorptivity change the absorbed radiative heat transfer.

The convective and radiative effects may be more important at earlier times in the test when temperatures are rapidly changing.

The DFTs and PTs are the largest temperature sensors. The insulation layer in them increases the directional sensitivity by reducing heat transfer to or from the unexposed surface. The others are omni-directional.

Temperature sensors are often considered to be lumped parameter systems, where the temperature of the entire sensor assembly rises and falls almost uniformly. The bare bead thermocouples are likely the only sensor of the six designs that qualifies. The characteristic dynamic response of a lumped parameter temperature sensor is {1 − exp (time/time constant)}.

The time constant equation for such lumped parameter systems is

$$\tau = \rho \mathrm{VCp/hA} \tag{5}$$

where τ is the time constant, ρ is the density, V is the volume, C_p is the specific heat, h is the total heat transfer (convection + linearized radiation) coefficient, and A is the exposed surface area.

Note: the time constant is directly related to the thickness of the exposed surface element and inversely related to the total heat transfer coefficient h which varies continuously during a fire resistance test. As a result, the time constant of the exposed surface element also varies continuously.

The other five sensors are more likely to exhibit a double exponential dynamic response characteristic

$$1 - \mathrm{C1}^{*}\exp(\mathrm{time}/\tau 1) - \mathrm{C2}^{*}\exp(\mathrm{time}/\tau 2) \tag{6}$$

In a system with a double exponential dynamic response, the first exponential term describes the response of the exposed surface element (see Fig. 1), which

is likely to behave as a lumped parameter system, as previously described. The second exponential term describes the thermocouple hot junction's response to heating of the exposed surface element of the sensor. This term was experimentally determined to be 1.9 s ±5 % for DFTs [20]. Based on measurements made during cool-down, an overall (lumped-parameter) time constant of the E119 shielded thermocouple is designed to be in the range of 5.0 to 7.2 min. Values for the second time constant are unknown for the ISO 834 plate thermometers [3] and the ASTM E119 shielded thermocouples [2].

Figure 7(*a*) and 7(*b*) show the results of heat flux measurements made in a horizontal furnace with Gardon gauge heat flux sensors [37]. Figure 7(*a*) shows how the heat flux history varies when different temperature sensors are used for control. The spread between the minimum and maximum values is significant over the entire test. The value slowly grows from 10 min to 60 min, when it is approximately 20 kW/m^2. When normalized, the spread around the average heat flux falls steadily from ±18 % at 10 min to ±7 % at 60 min. The integrated heat flux over one hour is 306 kJ/m^2 ± 11 % for the six sets of data shown. Over the first five minutes, when using the E119 thermocouple for control, the integrated heat flux is over twice that when the plate thermometer is used. In the early times, E119 provides a much more severe exposure than ISO 834.

Figure 7(*b*) shows some statistical analysis of this data. The average value and the sample standard deviation are comparable for the first two minutes. The coefficient of variation, defined as sample standard deviation/average value, ranges from 80 to 160 % during the same time period. The coefficient of variation is 16.5 % at 10 min and falls to 5.2 % at 60 min.

Quantitative Heat Flux Measurements in Fire Safety Tests

Heat flux feedback determines the burning rate of pool fires. Heat flux is at the heart of ignition and fire growth in a compartment. As shown in Fig. 7, dramatic changes in the thermal exposure (heat flux) can result from changing the design of the furnace control thermocouples [35–37]. Although the focus in this paper is on fire resistance tests, the discussion is applicable to other large scale standardized tests such as room burns, room-corner tests, multi-story or external fire spread, etc.

Several sensor choices are available for measuring heat flux. Two are water-cooled sensor designs, circular foil heat flux sensors, also known as Gardon gauges [9], and thermopile sensors, also known as Schmidt-Boelter gauges [10]. Both sensors are commonly used for making cold-wall heat flux measurements. Directional flame thermometers (DFTs) are uncooled hot-wall heat flux sensors [38,39]. The DFTs are specified during the test in ASTM E1529 [19], which is the hydrocarbon-curve or high-rise curve, fire resistance test.

As noted earlier, heat flux measurements in an application are different from a laboratory/calibration measurement because the conditions in the application affect the accuracy or uncertainty. Even a recalibration will vary if the calibration fixture is different than the one used by the manufacturer. The data sheets for Gardon and Schmidt-Boelter gauges often state the accuracy as 3 %; this value is actually the manufacturer's calibration uncertainty. A round-robin

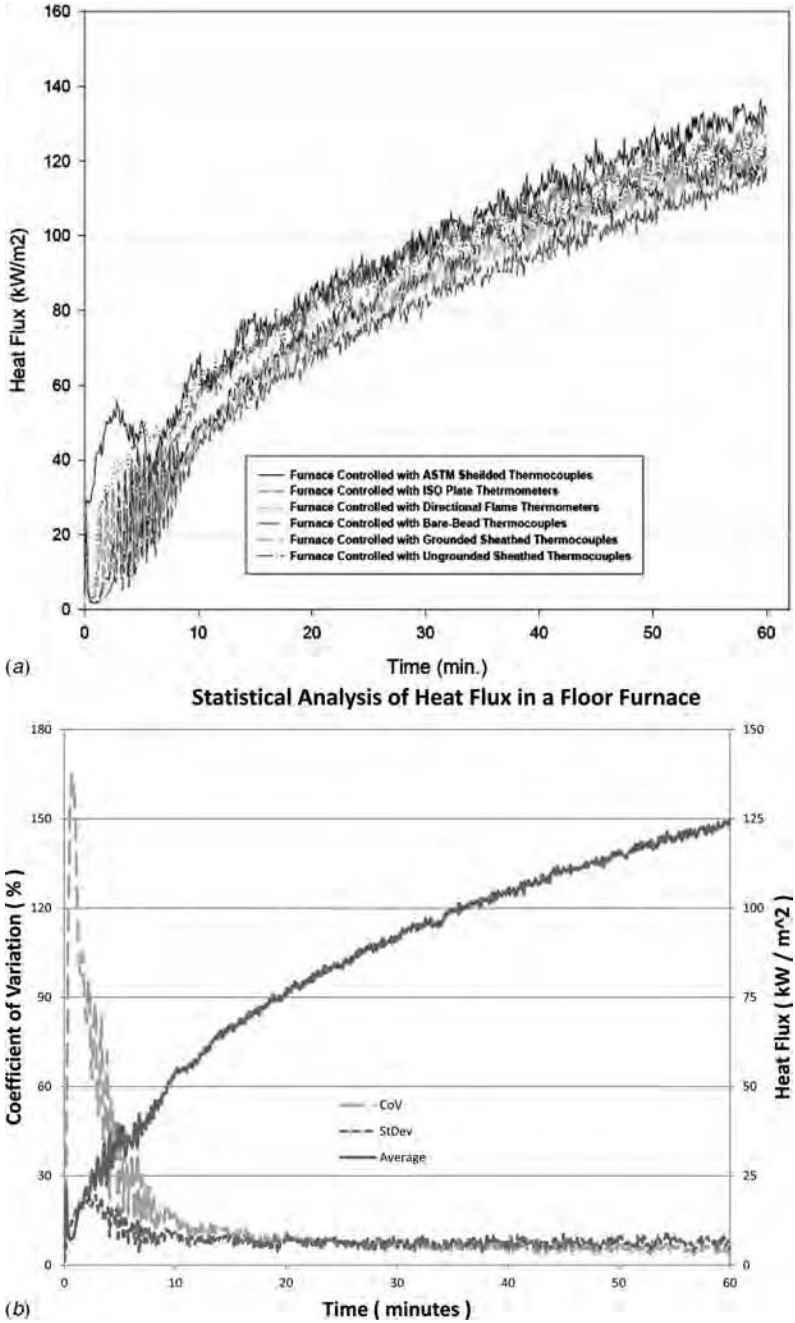


FIG. 7—*(a) Incident heat flux measured with ceiling mounted Gardon gauges in a floor furnace. (b) Average heat flux along with the standard deviation and coefficient of variation.*

project showed the inter-laboratory calibration uncertainty at the 95% confidence level was ±8.7% for Schmidt-Boelter gauges and for Gardon gauges it was ±9.2% [40].

Estimating heat flux measurement uncertainty is especially difficult in fires and fire safety testing because there are so many different nonlinear physical phenomena in fires that can affect heat flux measurements. In a Heat Flux Measurement Workshop sponsored by NSF/ASME/NIST in 1995 (search: NIST), an informal survey of the participants (manufacturers and researchers) estimated a total heat flux measurement uncertainty of 15–20% in applications.

One of the important problems is that the sensitivity of Gardon gauges to convective heating is up to 25% lower than for radiative heating [9,41]. It appears that this problem also affects Schmidt-Boelter gauges [10,42,43]. The ASTM Method E511 gives techniques for estimating uncertainty of Gardon gauge measurements due to this problem [9]. The ASTM Method 2683-09 covers Schmidt-Boelter gauges [10]. Bryant et al., provide total uncertainty estimates ranging from 7–25%, depending on conditions, at the 1σ (e.g., 67%) confidence level for Schmidt-Boelter gauges used in the ISO 9705 Room Corner Test [44].

Directional flame thermometers were developed for making total heat flux and two-flux measurements in pool fires. The first part of the process for using DFTs to measure heat flux involves measuring the temperature history of two well-characterized calorimetric objects (i.e., the DFT plates). The net heat flux is then calculated using a thermal model of the sensor; these models require accurate knowledge of the temperature dependent thermophysical properties. There are three analyses used for calculating the heat flux from the DFT temperature histories in ASTM E1529-10, Annex II [19]. These are:

(a) Analyzing at early times as thin-skin calorimeters [29]—cold-wall flux.
(b) Performing an energy balance on the front plate at later times—effective furnace radiation temperature (EFRT) (irradiance).
(c) Using inverse heat conduction analysis to calculate the hot-wall heat flux data for the entire test.

The nonlinear inverse heat conduction code, IHCP1D, calculates the net heat flux and provides detailed information on the DFT analysis. A new inverse filter functions technique provides real-time heat flux readouts or a quick-look analysis capability for large data sets [39].

The net heat flux is what crosses the exposed surface of the DFT; it is the total heat flux minus the re-radiated flux. At early times, re-radiation is low and the net heat flux approximates the total flux (convection + absorptivity * incident radiation). The nonlinear inverse heat conduction code calculation also returns the flux conducted into the insulation layer and the net heat flux at the back surface (i.e., facing the test unit) [39].

The DFTs have been used to obtain quantitative heat flux measurements in fire resistance tests [33,36–39]. Heat flux measurements made with DFTs have been used to support the development of large computational fluid dynamics (CFD) models for pool fires [21,22,24,25,27,45]. Preliminary uncertainty estimates for the DFTs range from 12–15% [39]. Final estimates are expected to be comparable to the uncertainties for Schmidt-Boelter and Gardon gauges, which are approximately ±20%.

Providing Heat Flux Measurements for Performance Based Models

This paper and others have demonstrated that heat flux measurements are needed for the development and validation of fire protection engineering models, for accurate structural calculations, for studying ignition of materials, and for fire propagation analysis. The question is how would they be applied?

Just as simultaneous temperature and heat flux boundary conditions cannot be specified for a surface of a solid, multiple studies show it is difficult to measure just one parameter in a fire resistance test (such as the furnace temperature) and calculate the other (thermal exposure). This is especially true during the first 5–10 min if the dynamic response of the temperature sensor has not been well characterized.

Part of the difficulty involves deciding how many simplifying assumptions are allowed. The quality of the estimate depends on the assumptions made. For example, estimates in fire resistance test furnaces where there is a good understanding of the overall conditions should, in general, be more accurate than those in an uncontrolled environment, such as a room burn or pool fire. A better approach is to measure the heat flux in every experiment

Janssens [46] showed excellent agreement between the DFT and Schmidt-Boelter measurements in an ASTM E119 fire resistance test when bare bead thermocouples were used to provide gas temperatures for making the comparison. Heat flux estimates made directly from furnace temperatures measured with either the E119 shielded thermocouples or plate thermometers were consistently low.

Janssens proposed the method below for applying heat flux measurements [46]. For the exposed-surface (x = 0), the boundary condition in numerical models is

$$-k\frac{\partial T}{\partial x} = \varepsilon_s \dot{q}_{irr} + h(T_g - T_s) - \varepsilon_s \sigma T_s^4 \tag{7}$$

Where k is the thermal conductivity of the exposed object, ε is the emissivity of the exposed surface, σ is the Stefan-Boltzmann constant while the subscript 'g' is for gas and 's' is for surface DFTs in conjunction with an inverse heat transfer code, can be used to determine furnace irradiance. A supplemental bare-bead thermocouple is needed to provide a measure of the gas temperature T_g. An estimate of the convective heat transfer coefficient h is also needed; the uncertainty of the coefficients is estimated to be ±20 to 25 % [43].

A second approach is to use the total heat flux exposure [38,39]. The total heat flux (fire exposure) is calculated by adding the net flux and re-radiation from the DFT hot-face. Results from fire resistance tests demonstrated that the total heat flux versus time curves obtained from a large furnace characterization unit and the relatively small DFTs agreed closely, except at very early times, when condensation might be important [33]. The total heat flux reached 50 kW/m^2 in the first 5 min and was approximately 100 kW/m^2 at one hour. The integrated heat flux values over the test period also agreed within a few percent in this furnace and all others tested. This approach requires knowledge of the temperature and absorptivity of the exposed surface.

After the first 7–10 min of the test, an Effective Furnace Radiation Temperature (EFRT) can be calculated from the DFT hot-face temperature with the net heat flux to the DFT and the heat loss from the unexposed surface. Since heat flux is a vector quantity, a DFT used for this purpose should be mounted 10 cm in front of the surface and point in the same direction as the surface [33,38,39].

The EFRT is used to calculate an equivalent blackbody heat flux. The EFRT approach is similar to the adiabatic surface temperature calculations based on plate thermometer data [47]. In some conditions, the effective furnace radiation temperature provides a more accurate measure of the thermal exposure because it includes transmission through the DFT, heat loss to the test unit, and any energy storage in the DFT plates [38,39].

Summary

Independent temperature and heat flux measurements with uncertainty estimates are needed to support the development and validation of engineering models of the performance of fire protection materials and structural element tests. This has placed additional emphasis on the appropriate application of existing and new measurement techniques. In this paper, the primary attention is focused on potential problems that can dramatically increase the overall uncertainty of these measurements in fire safety testing.

ASTM E5 has been working on the temperature and heat flux measurement problems and uncertainties discussed in this paper. Some of the accomplishments are:

(1) Standardized a new E119 shielded thermocouple design to aid harmonization.
(2) Incorporated the furnace uniformity measurement procedure into E119 (Part 2?). This measurement fixture adds ISO type furnace temperature measurements using plate thermometers and heat flux measurements using directional flame thermometers to the furnace performance evaluation process.
(3) ASTM E1529, the Hydrocarbon Curve (High-Rise) Fire Resistance Test Method, uses both temperature and heat flux measurements to demonstrate furnace control. In support of the NIST-BFRL request, E1529 is the first large-scale test to require continuous heat flux measurements during the test.

Sandia National Laboratories has incorporated heat flux measurements using Gardon and Schmidt-Boelter gauges along with DFTs and other calorimetric sensors into the development and validation CFD models of pool and jet fires. These are used in studies of transportation accidents including the thermo-structural safety analysis of hazardous material shipping containers. Most recently, such data were incorporated into a NASA-JPL/Sandia effort on risk assessments for radio-isotopic heat sources in rocket launch accidents [48].

Hopefully, ASTM E5 can develop fire safety test standards that provide data of the same quality to meet the request from NIST-BFRL and support the development of codes to accurately predict response to fire.

Acknowledgments

The writer gratefully acknowledges the data and support provided by Dr. Mohamed Sultan from the National Research Council of Canada.

References

[1] Keltner, N. R., "Thermal Measurements in Fire Safety Testing - Are We Playing With Fire?" *Special Symposium of Fire Calorimetry*, NIST, Gaithersburg, MD, July, 1995, *Fire Calorimetry Proceedings, DOT=FAA=CT-95=46*, FAA Tech Center, Atlantic City Int'l Airport, NJ.

[2] ASTM E119–11a, "Standard Test Methods for Fire Tests of Building Construction and Materials," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[3] ISO 834–1:1999, 1999, "Fire-Resistance Tests – Elements of Building Construction–Part 1: General Requirements," ISO, Geneva, Switzerland.

[4] Gritzo, L. A., Sandia National Laboratories, private communication, 2001.

[5] Nakos, J. T., Gill, W., and Keltner, N. R., "An Analysis of Flame Temperature Measurements Using Sheathed Thermocouples in JP-4 Pool Fires," *Proceedings of the ASME/JSME Engineering Joint Conference*, J. R. Lloyd and Y. Kurosaki, Eds., ASME Book No. I0309E, ASME, New York, 1991, pp. 283–289.

[6] 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," *SAND2004–1023*, Sandia National Laboratories, Albuquerque, NM, April, 2004.

[7] ASTM E906–10 / E906M-10, 2010, "Standard Test Method for Heat and Visible Smoke Release Rates for Materials and Products Using a Thermopile Method," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[8] ASTM E1354–11a, 2011, "Standard Test Method for Heat and Visible Smoke Release Rates for Materials and Products Using an Oxygen Consumption Calorimeter," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[9] ASTM 511–07, 2007, "Standard Test Method for Measuring Heat Flux Using a Copper-Constantan Circular Foil, Heat-Flux Transducer," *Annual Book of ASTM Standards*, ASTM International, West Conshohocken, PA.

[10] ASTM E2683–09, 2009, "Standard Test Method for Measuring Heat Flux Using Flush-Mounted Insert Temperature-Gradient Gages," *Annual Book of ASTM Standards*, Vol. 15.03, ASTM International, West Conshohocken, PA.

[11] Alpert, R. L., Orloff, L., and De Ris, J. L., "Angular Sensitivity of Heat Flux Gauges," *Thermal Measurements: The Foundation of Fire Standards, ASTM Spec. Tech. Publ.*, Vol. 1427, L. A. Gritzo and N. J. Alvares, Eds., ASTM International, West Conshohocken, PA, 2002.

[12] *Manual on the Use of Thermocouples in Temperature Measurement*, MNL12-Fourth Edition, ASTM Committee E20 on Temperature Measurement, ASTM International, West Conshohocken, PA, 1993.

[13] Jakob, M., *Heat Transfer*, Vol. II, John Wiley & Sons, NY, 1957.

[14] Moffat, R. J., "Using Uncertainty Analysis in the Planning of an Experiment," *J. Fluids Eng.*, Vol. 107, No. 2, June 1985, pp. 173–181.

[15] Moffat, R. J., "Describing the Uncertainties in Experimental Results," *Exp. Therm. Fluid Sci.*, Vol. 1, No. 1, 1988, pp. 3–17.

[16] Blanchet, T. K., Humpries, L. L., and Gill, W., "Sandia Heat Flux Gauge Thermal Response and Uncertainty Models," *SAND 2000–1111*, Sandia National Laboratories, Albuquerque, NM, May, 2000.

[17] Dowdell, R. B., Discussion: "Contributions to the Theory of Single-Sample Uncertainty Analysis," (Moffat, R. J., *ASME J. Fluids Eng.*, Vol. 104, 1982, pp. 250–258).

[18] Beyler, C., Beitel, J., Iwankiw, N., and Lattimer, B., "Fire Resistance Testing for Performance-Based Fire Design of Buildings," The Fire Protection Research Foundation, Quincy, MA, June, 2007.

[19] ASTM E1529–10, 2010, "Standard Test Methods for Determining Effects of Large Hydrocarbon Pool Fires on Structural Members and Assemblies," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[20] Bader, B. E., "Heat Transfer in Liquid Hydrocarbon Fuel Fires," *Report SC-R-64–1366A*, Sandia Corp., Albuquerque, NM, June 1965.

[21] Gregory, J. J., Mata, R., Jr., and Keltner, N. R., "Thermal Measurements in a Series of Large Pool Fires," *SAND 85–0196*, Sandia National Laboratories, Albuquerque, NM, Aug 1987.

[22] Gregory, J. J., Mata, R., Jr., and Keltner, N. R., "Thermal Measurements in Large Pool Fires," *J. Heat Transfer*, Vol. 111, May 1989, pp. 446–454.

[23] M. E. Schneider, Keltner, N. R., and Kent, L. A., "Thermal Measurements in The Nuclear Winter Fire Test," *SAND 88–2839*, Sandia National Laboratories, Albuquerque, NM, Jan 1989.

[24] Bainbridge, B. L. and Keltner, N. R., "Heat Transfer to Large Objects In Large Pool Fires," *J. Hazard Mater.*, Vol. 20, Dec 1988, pp. 21–40.

[25] Keltner, N. R., and Moya, J. L., "Defining the Thermal Environment in Fire Tests," *Fire Mater.*, Vol. 14, 1989, pp. 133–138.

[26] Gritzo, L. A. and Nicolette, V. F., "Coupled Thermal. Response of Objects and Participating Media in Fires and Large Combustion Systems," *Numer. Heat Transfer, Part A*, Vol. 28, 1995, pp. 531–545.

[27] Keltner, N. R., Gill, W., and Kent, L. A., "Simulating Fuel Spill Fires Under the Wing of an Aircraft," *Proceedings of the 4th International Symposium on Fire Safety Science*, T. Kashiwagi, Ed., Ottawa, Ontario, Canada, July 1994.

[28] Heinonen, E. W., McCarson, T. D., Jr., Stepetic, T. J., Kent, L. A., Gill, W., and Keltner, N. R., "Inverted Deluge System (IDS) Development Tests, Volume I: Fire Suppression Tests," *CEL-TR-92–71*, NM Engineering Research Institute, Albuquerque, NM, June 1992.

[29] ASTM E459–05, 2005, "Standard Test Method for Measuring Heat Transfer Rate Using a Thin-Skin Calorimeter," *Annual Book of ASTM Standards*, Vol. 15.03, ASTM International, West Conshohocken, PA.

[30] Brundage, A. L., Donladson, A. B., Walt Gill, W., Kearney, S. P., Nicolette, V. F., and Yilmaz, N., "Thermocouple Response In Fires: Part 1: Considerations In Flame Temperature Measurements by a Thermocouple," *J. Fire Sci.*, Vol. 29, May 2011, pp. 195–211.

[31] Brundage, A. L., Donladson, A. B., Walt Gill, W., Kearney, S. P., Nicolette, V. F., and Yilmaz, N., "Thermocouple Response In Fires: Part 2: Validation of Virtual Thermocouple Model For Fire Codes," *J. Fire Sci.*, Vol. 29, May 2011, pp. 213–226.

[32] Wittasek, N. A. "Analysis and Comparison of Marine Fire Testing Regulations and Procedures," °Master Of Science Thesis, Worcester Polytechnic Institute, Worcester, MA, 1996.

[33] Keltner, N. R., Nash, L., Beitel, J., Parker, A., Welsh, S., and Gilda, B., "Fire Safety Test Furnace Characterization Unit," *Thermal Measurements: The Foundation of Fire Standards, ASTM Spec. Tech. Publ.* Vol. 1427, L. A. Gritzo and N. J. Alvares, Eds., ASTM International, West Conshohocken, PA, 2002.

[34] Babrauskas, V. and Williamson, R. B., "Temperature Measurement in Fire Test Furnaces," *Fire Technol.*, Vol. 14, 1978, pp. 226–238.
[35] Sultan, M., "Fire Resistance Furnace Temperature Measurements: Plate Thermometers Vs Shielded Thermocouples," *Fire Technol.*, Vol. 42, No. 3, July 2006, pp. 253–267.
[36] Sultan, M., "Comparisons of Measured Temperature and Heat Flux in Fire Resistance Test Furnaces Controlled by Either ISO Plate Thermometers or by ASTM Shielded Thermocouples or by Directional Flame Thermometers," *ASTM E5 Symposium on Advances in The State of The Art of Fire Testing*, Miami Beach, FL, Dec 2008.
[37] Sultan, M. A., "Comparisons of Furnace Temperature and Incident Heat Flux in Wall and Floor Furnaces Controlled by Six Different Temperature Sensors," *NRCC-52644*, National Research Council of Canada, Ottawa, Ontario, Canada, July 2010.
[38] Keltner, N. R., "Directional Flame Thermometers: A Tool for Measuring Thermal Exposure in Furnaces and Improving Control," *Proceedings of the Interflam 2007 Conference* (CD-ROM), Royal Holloway College, Univ. of London, Interscience, London, Sept 2007.
[39] Keltner, N. R., Beck, J. V., and Nakos, J. T., "Using Directional Flame Thermometers for Measuring Thermal Exposure," *J. ASTM Int.*, Vol. 7, No. 2, 2010, JAI102280.
[40] Pitts, W. M., Murthy, A. V., deRis, J. L., Filtz, J. R., Nygard, K., Smith, D., and Wetterlund, I., "Round Robin Study of Total Heat Flux Gauge Calibration at Fire Laboratories," *NIST Spec. Publ.*, Vol. 1031, Oct 2004.
[41] Keltner, N. R. and Wildin, M. W., "Transient Response of Circular Foil Heat-Flux Gauges to Radiative Fluxes," *Rev. Sci. Instrum.*, Vol. 46, No. 9, Sept 1975, pp. 1161–1166.
[42] Kidd, C. T. and Nelson, C. G., "How the Schmidt-Boelter Gage Really Works," *Proceedings of the 41st International Instrumentation Symposium*, Aurora, CO, May 7–11, 1995, Instrument Society of America, Research Triangle Park, NC.
[43] Nakos, J. T. and Brown, A. L., "Schmidt-Boelter Heat Flux Gage Sensitivity Coefficients In Radiative And Convective Environments," *ASME/JSME 2011 8th Thermal Engineering Joint Conference (AJTEC2011)*, Honolulu, HI, March 13–17 2011 Interscience, London.
[44] Bryant, R., Johnsson, E., Ohlemiller, T., and Womeldorf, C., "Estimates of the Uncertainty of Radiative Heat Flux Calculated from Total Heat Flux Measurements," *Interflam'01. 9th International Fire Science and Engineering Conference*, Edinburgh, Scotland. Sept 17–19, 2001 Interscience, London.
[45] Gritzo, L. A., Gill, W., and Keltner, N. R., "Thermal Measurements to Characterize Large Fires," *Proceedings of the 41st International Instrumentation Symposium*, Aurora, CO, May 7–11, 1995, Instrument Society of America, Research Triangle Park, NC, pp. 337–346.
[46] Janssens, M., "Comparison Of Different Methods To Characterize The Thermal Environment In Fire Resistance Furnaces," *ISO/TC92/SC2 Fire Safety - Fire Containment, Workshop on Heat Transfer Calculations and Measurements in Fire Safety Engineering–Furnace Characterization and Control*, Sept 10, 2007, SP Technical Research Institute Of Sweden, Borås, Sweden.
[47] Wickstrom, U., "Adiabatic Surface Temperature for Calculating Heat Transfer to Fire Exposed Structures," *Proceedings of the Interflam 2007 Conference*, Royal Holloway College, Univ. Of London, London, Sept, 2007.
[48] Noravian, H., Nicolette, V. F., Woodbury, K. A., Reinhart, L. E., and Keltner, N., "VULCAN/SINDA Loosely Coupled Analysis Methodology for the NASA/JPL Rod Calorimeter," Joint Army-Navy-NASA-Air Force (JANNAF) Interagency Propulsion Committee, Boston, May, 2008.

Reprinted from JTE, Vol. 39, No. 6
doi:10.1520/JTE103922
Available online at www.astm.org/JTE

Cecilia S. Lam[1] *and Elizabeth J. Weckman*[2]

Heat Flux Measurements and Their Uncertainty in a Large-Scale Fire Test

ABSTRACT: Heat flux data from a series of controlled experiments involving a 2 m diameter, wind-blown pool fire are examined to highlight the difficulties involved in conducting heat flux measurements in a realistic, large-scale, hydrocarbon-fueled fire. Data were taken at several locations along the ground near the fire. At each location, three different heat flux sensors were positioned together: a Gardon gage, a directional flame thermometer (DFT) and a Sandia heat flux gage (HFG). Methods were first developed to correct measured values of heat flux for the slight differences in gage location relative to the fire. The remaining discrepancies between the values of heat flux measured by the different gages were then used to highlight uncertainties in heat flux measurements due to differences in gage surface temperature, in gage thermal response to the inherent modes of heating involved in the large hydrocarbon fire environment, and in conduction losses from the gage sensor plates. The importance of these sources of discrepancy varied depending on the magnitude of the measured heat flux and on whether the gages were located in a radiation-dominated or mixed radiative-convective environment within the fire.

KEYWORDS: heat flux gage, radiation, convection, fire testing, measurement uncertainty

Manuscript received April 16, 2011; accepted for publication July 11, 2011; published online August 2011.

[1] CanmetENERGY, Natural Resources Canada, 1 Haanel Dr., Ottawa, Ontario, K1A 1M1, Canada. (Corresponding author), e-mail: cecilial@alumni.uwaterloo.ca. (Formerly at Univ. Waterloo.)
[2] Dept. of Mechanical and Mechatronics Engineering, Univ. of Waterloo, 200 Univ. Ave. W., Waterloo, Ontario, N2L 3G1, Canada.

Cite as: Lam, C. S. and Weckman, E. J., "Heat Flux Measurements and Their Uncertainty in a Large-Scale Fire Test," *J. Test. Eval.*, Vol. 39, No. 6. doi:10.1520/JTE103922.

Nomenclature

h = Convective heat transfer coefficient, W/m^2 K
Q_{rad} = Radiative heat flux to gage, see Eq 1, W/m^2
Q_{tot} = Total heat flux to gage, see Eq 3, W/m^2
T_f = Flame temperature, K
T_g = Surface temperature of heat flux gage, K
x = Distance from fuel pan center in direction of wind, see Fig. 3, m
y = Distance from fuel pan center perpendicular to direction of wind, see Fig. 3, m
z = Height above floor, see Fig. 2, m
α_g = Absorptivity of heat flux gage
ΔQ_{conv} = Change in convective heat flux due to change in position, see Eq 4, W/m^2
ΔQ_{rad} = Change in radiative heat flux due to change in position, see Eq 2, W/m^2
ΔT = Change in flame temperature due to change in position, K
ε_f = Flame emissivity
σ = Stefan-Boltzmann constant, W/m^2/K^4

Introduction

In large-scale fire testing, knowledge of heat flux is typically needed to determine levels of fire exposure and thermal hazard, as indicated in standards such as ASTM E1529-06 [1]. Therefore, proper understanding of heat flux measurement techniques and their sources of uncertainty is critical to reliable fire testing and hazard assessment.

Although much work has been done on characterizing the responses of commercially-available heat flux gages in radiative environments [2,3], many fires involve potentially significant levels of convection [4–7], which affect the performance of these gages. Differences between radiative calibration environments and mixed-mode measurement environments can cause errors in collected data and thus increase difficulty in interpreting practical heat flux measurements. A few researchers have considered methods of calibrating commonly-used heat flux gages in convective conditions [8–10]; however, purely convective conditions are rarely typical in fires, which involve both radiation from hot combustion products and convection from buoyancy- or wind-driven flows [4–7]. A small amount of research has been performed on heat flux measurement in mixed radiative-convective environments [11,12]. Kuo and Kulkarni [11] proposed a method for correcting measurements made in mixed-mode environments for a Gardon gage calibrated in a radiative environment, but this method requires estimation of a convective heat transfer coefficient, which is itself associated with large uncertainty [3,4,12]. Despite such difficulties, measurement of heat flux in fires is necessary since it is critical to modeling fires and to assessing the severity of fire exposure [1,13].

The focus of this paper is to outline potential sources of uncertainty involved in measuring heat flux with different types of heat flux gages in a

FIG. 1—*Univ. of Waterloo Live Fire Research Facility.*

wind-blown fire environment. Three types of heat flux gages are placed near a large, wind-blown fire to expose the gages to different levels of radiation and convection. Although difficulties are encountered in quantifying the measurement uncertainty of the different gages, key issues that should be considered when measuring heat flux in large-scale fires with each type of gage are highlighted.

Experimental Setup

Tests were conducted in the large-scale test enclosure (19.5 m × 15.4 m floor area) of the Univ. of Waterloo Live Fire Research Facility (Fig. 1). Crosswinds were generated using a bank of six 2.0 m diameter fans, located in a three-by-two array at one end of the enclosure. Average wind speeds of up to 13 m/s could be generated. The winds issued through a 5.9 m high × 8.2 m wide × 8.3 m long fan plenum into the main test area and exited the test enclosure through a 7.9 m × 7.9 m door. The present test series involved average wind speeds of 3, 5, 7, and 10 m/s. A detailed characterization of the flow fields is provided by Best [14].

Fires were established in a 2.0 m diameter pan using Jet-A fuel floating on a water substrate. The pan center was located on the central longitudinal mid-plane of the test enclosure, 5.9 m downwind of the exit plane of the fan plenum. A raised floor surround, made of insulating fire bricks raised on concrete cinder blocks, was built around the fuel pan in a 2.7 m × 2.7 m area immediately surrounding the pan. As shown in Fig. 2, the top surface of the bricks was raised to be at the same level as the top rim of the pan. The upwind extent of the floor surround was lengthened using fiber-reinforced cement boards joined to cover a 2.4 m × 4.9 m area and raised to the same level as the top of the fire bricks. This extension was required to prevent the fire from becoming attached to the upwind edge of the floor surround. The central region of the remaining floor area of the test enclosure was protected by either a single or double layer of fire bricks, as illustrated in Fig. 2, because it was expected to be exposed to significant heat flux from the hot fire plume.

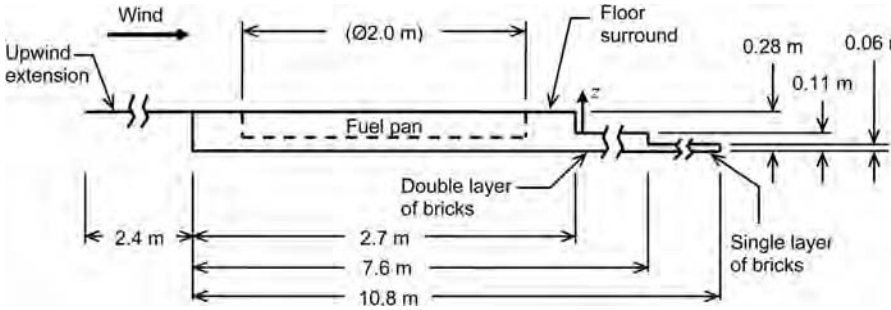


FIG. 2—*Geometry of the raised floor surround and surrounding brick layout.*

A network of 24-gage (0.51 mm diameter), Type-K thermocouples, insulated with ceramic-fiber sheathing and covered with protective metal braiding, was used to measure temperatures downwind of the fire. The thermocouples were formed with exposed junctions and mounted pointing upwind on vertical chains. As illustrated in Fig. 3, these chains were aligned along seven measurement planes that were oriented normal to the x direction and located between 1.5 and 9.2 m downwind of the fuel pan center. The thermocouples were distributed along each measurement plane to capture the cross-sectional extent of the fire plume. The total number of thermocouples was 396, distributed over 53 rakes.

Three types of gages were used to measure heat flux: the Gardon gage, the directional flame thermometer (DFT) and the hemispherical heat flux gage (HFG). A description of these three gages was previously given in Lam and Weckman [12]. For completeness, a summarized version of that description is provided here.

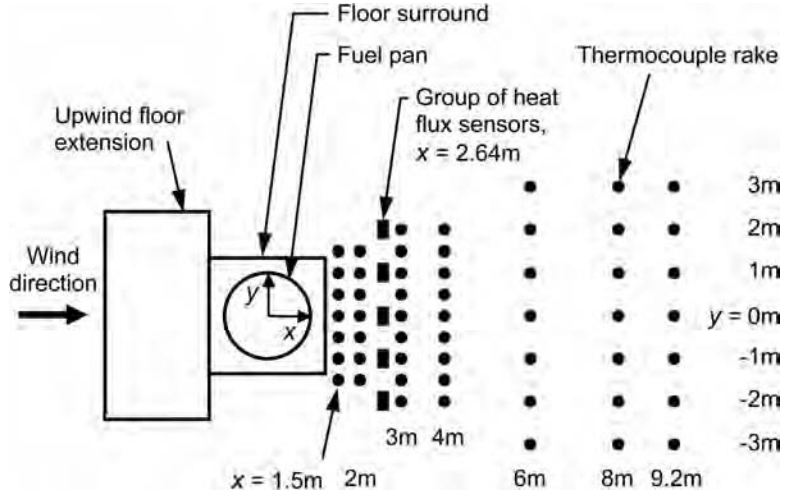


FIG. 3—*Plan view of experimental setup.*

The Gardon gage was a windowless, water-cooled sensor from the Thermogage[3] 1000-1 series made by Vatell Corporation of Christiansburg, VA. With an outer diameter of 25 mm, it produced a voltage corresponding to the difference in temperature between the center and water-cooled edge of a 4.7 mm diameter sensing foil [15,16]. The voltage signal was linearly related to the incident heat flux (rated up to 155 kW/m^2) via a manufacturer-provided calibration constant, which was based on irradiation over 2π sr and had a stated accuracy of ± 3 %. The sensing surface emissivity was specified by the manufacturer to be 0.94. The temperature of the cooling water (approximately 16°C) was above the dew point in all tests to prevent condensation on the gage surface [17].

Unlike the Gardon gage, the DFT and HFG were not water-cooled. For each, total incident heat flux was estimated using an inverse heat conduction analysis. The DFT, manufactured by Ktech Corporation of Albuquerque, NM, consisted of two 3.2 mm thick Inconel[4] sensor plates, each 120 mm $\times$ 120 mm and separated by a 12 mm thick layer of metal-felt insulation [18]. The outward, exposed face of each plate was coated with flat black paint with an emissivity of 0.85 to achieve a diffuse, gray surface [19]. A 1.6 mm diameter, Inconel-sheathed, Type-K thermocouple with an ungrounded junction was attached to the center of the unexposed face of each plate using retainer straps spot-welded to the surface [18]. The one-dimensional inverse heat conduction program IHCP1D (of Beck Engineering Consultants Company, Okemos, MI) was used to calculate the net heat flux from the measured DFT temperature data [20]. In this program, the DFT was modelled as a three-layer system in which data from both thermocouples were input to the program to prescribe a temperature-time history on each side of the insulation. Total heat flux was obtained by adding the radiation emitted from the gage surface (estimated using the emissivity and surface temperature of the gage as calculated by IHCP1D) to the net heat flux determined using the IHCP1D program.

The HFG, manufactured by Sandia National Laboratories of Albuquerque, NM, contained a 0.25 mm thick, stainless-steel sensor plate with a 50 mm diameter exposed sensing area in the center [19]. A 1.6 mm diameter, Inconel-sheathed, Type-K thermocouple with an ungrounded junction was attached to the center of the unexposed side of the sensor plate using retainer straps spot-welded to the surface. The sensor plate was sandwiched between two layers of 3.2 mm thick ceramic-fiber insulation and two 3.2 mm thick stainless-steel plates (each with a 50 mm diameter hole aligned with the exposed sensing area). This assembly was placed against one end of a 102 mm diameter $\times$ 102 mm long steel cylinder. The inside of the cylinder was filled with ceramic-fiber insulation to minimize heat losses from the unexposed face of the sensor plate. Similar to the DFT, the exposed area of the sensor plate was coated with flat black paint with an emissivity of 0.85 to achieve a diffuse, gray surface. Temperature data from the HFG were analysed using the data reduction routine described in Blanchat *et al.* [19] to estimate the total incident heat flux. This program, which was developed by Sandia National Laboratories specifically for

[3]Trademark of Vatell Corporation, Christiansburg, VA.
[4]Registered trademark of Special Metals Corporation, New Hartford, NY.

TABLE 1—*Y locations of heat flux gages and associated thermocouples ($x = 2.64$ m; $z = 0$ m for heat flux gages, $z = 0.03$ m for thermocouples).*

Nominal Location, m	−2	−1	0	1	2
Measured location of Gardon gage, m	−2.11	−1.12	−0.11	0.90	1.91
Measured location of DFT, m	−2.01	−1.03	−0.03	0.99	2.00
Measured location of HFG, m	−1.88	−0.90	0.10	1.10	2.12
Measured location of thermocouple, m	−1.94	−0.96	0.04	1.05	2.06

their HFG, was based on heat conduction equations, derived from first principles, that modelled the gage response to an applied heat flux as the one-dimensional response of a heated composite wall composed of the sensor plate and the insulation.

All three gages were used to measure heat flux to the ground at several locations downwind of the fuel pan. All of the gages were centered along a plane situated at a distance of $x = 2.64$ m downwind of the fuel pan center (Fig. 3). The gages (15 in total) were divided into five groups of three (with a Gardon gage, DFT, and HFG in each group), located nominally on the longitudinal midplane, ± 1 m from the midplane, and ± 2 m from the midplane. Within each group of gages, the three sensors were placed adjacent to each other at the same x location, while the y locations varied by no more than ± 0.12 m from the nominal positions listed in the preceding text. The measured y coordinate of each heat flux gage (measured to the center of the gage surface with an uncertainty of ± 0.005 m) is listed in Table 1. In addition, a 24-gage (0.51 mm diameter), Type-K thermocouple with an exposed junction was positioned (pointing upwind) between the HFG and DFT at a height of approximately 25 mm above the ground plane ($z = 0.03$ m) to measure the local gas temperature.

The heat flux gages were recessed into the brick floor such that the gage surfaces were approximately flush (within 1 cm) with the top surface of the bricks, as shown in Fig. 4. To minimize conduction between the non-water-

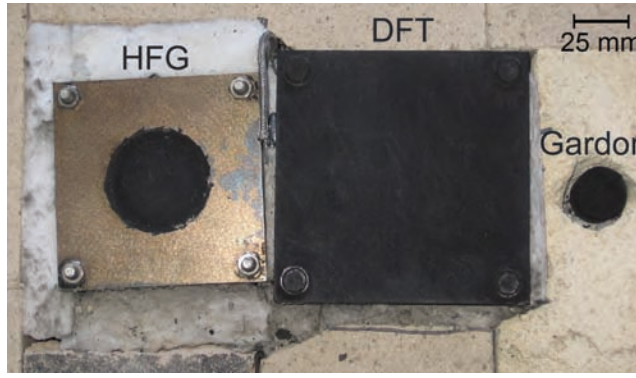


FIG. 4—*Setup of heat flux gages in brick floor.*

cooled gages and the surrounding floor, the DFT was placed on top of two layers of 25 mm thick ceramic-fiber insulation, while the cylindrical housing of the HFG was surrounded by one layer of the same insulation.

All wires leading to the thermocouples and heat flux gages were run underneath the protective layer of fire bricks on the test enclosure floor in order to minimize heat exposure to the instrumentation cables. Data acquisition was conducted using a PC-based, distributed system consisting of a modular set of backplanes linked through Ethernet cables to a computer. Data were sampled at a rate of approximately 0.4 Hz; this was limited by the presence of a significant quantity of other instrumentation in the experiment (described elsewhere [21]).

Results and Discussion

Temperature Contour Plots

Figures 5 and 6 show contour plots of the time-averaged increase in temperature (steady-state fire gas temperature referenced to ambient temperature) as measured by the thermocouple arrays for the lowest and highest wind speeds. The plots were produced using the Kriging method, which is a robust method

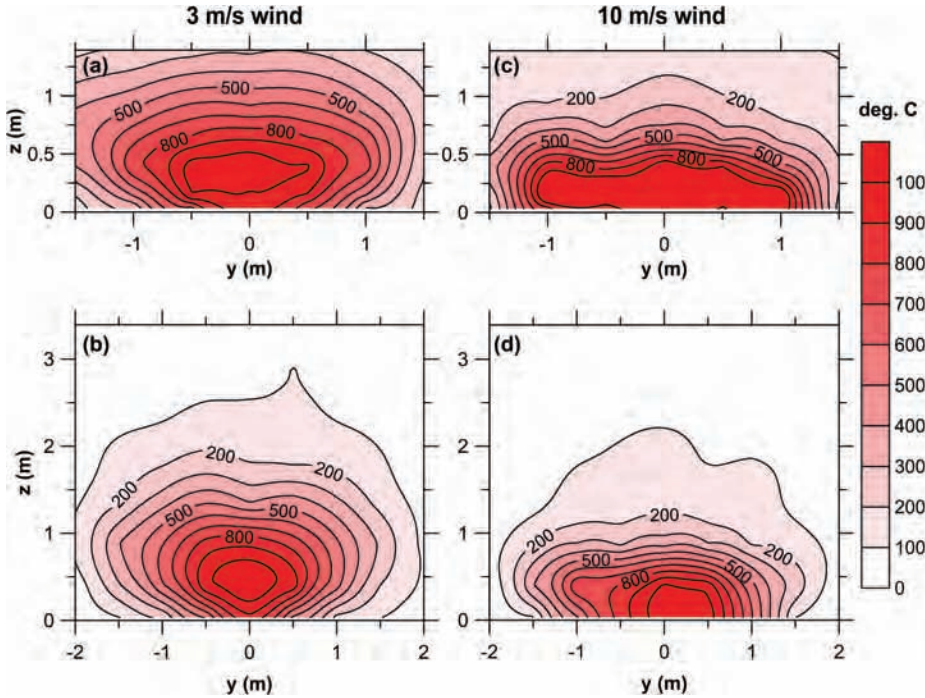


FIG. 5—*Example temperature contour plots, (a) x = 2 m, 3 m/s wind, (b) x = 3 m, 3 m/s wind, (c) x = 2 m, 10 m/s wind, and (d) x = 3 m, 10 m/s wind.*

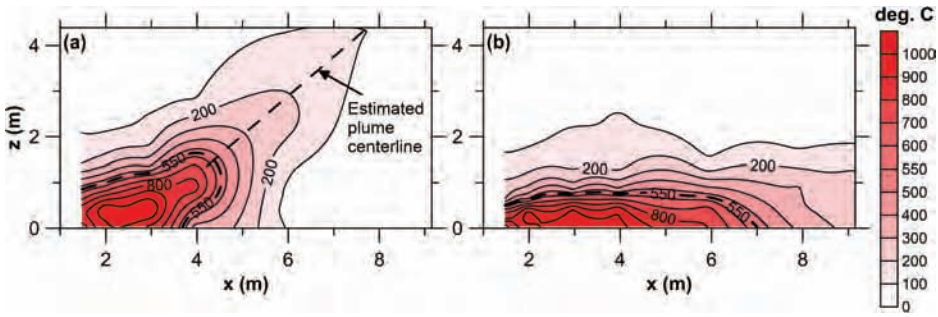


FIG. 6—*Example temperature contour plots along y = 0 m, (a) 3 m/s wind, and (b) 10 m/s wind.*

that is suitable for many types of data and provides a visually plausible representation of the data contours [22].

These contour plots are presented to show the general shape and direction of the fire plume, which will be of importance in later discussions of the necessary corrections to the heat flux data. As such, a detailed discussion of thermocouple measurement uncertainty is not provided and the remainder of the paper focuses on heat flux. While noting that heat flux cannot be directly measured, but is instead deduced from thermocouple temperature measurements, this paper will focus not on the uncertainty in the use of thermocouples, but instead on the uncertainty in the overall heat flux measured using three different designs of heat flux gage.

Heat Flux Measurements

Typical time traces of incident total heat flux to the Gardon gage, DFT, and HFG nominally located at $y = 0$ m are shown in Fig. 7(*a*). A trace of net heat flux absorbed by the DFT (as calculated by the IHCP1D program) with time is also included. The corresponding temperatures measured by the thermocouples attached to the DFT and HFG sensor plates, and by the exposed thermocouple located between the two gages, are shown in Fig. 7(*b*).

At the beginning of the test, the thin sensor plate in the HFG heated up more quickly than the thicker top plate of the DFT [Fig. 7(*b*)], however, by 223 s, the temperatures of the exposed sensor plates in the two gages were within 40°C of each other. Both temperatures were similar to those measured by the exposed thermocouple. The HFG and DFT data did not exhibit the same level of temporal fluctuation as the thermocouple data due to the larger thermal mass (slower time response) of the sensor plates in the heat flux gages. The heat flux time traces of the HFG and DFT [Fig. 7(*a*)] similarly exhibited less fluctuation than that of the Gardon gage, due to the greater mass of their sensor plates in comparison to the small circular sensing foil of the Gardon gage. With time, the measured total heat flux from all three gages increased to quasi-steady values, permitting the calculation and comparison of time-averaged values. For example, in Fig. 7, the time-averaging period lasted from 223 to 402 s.

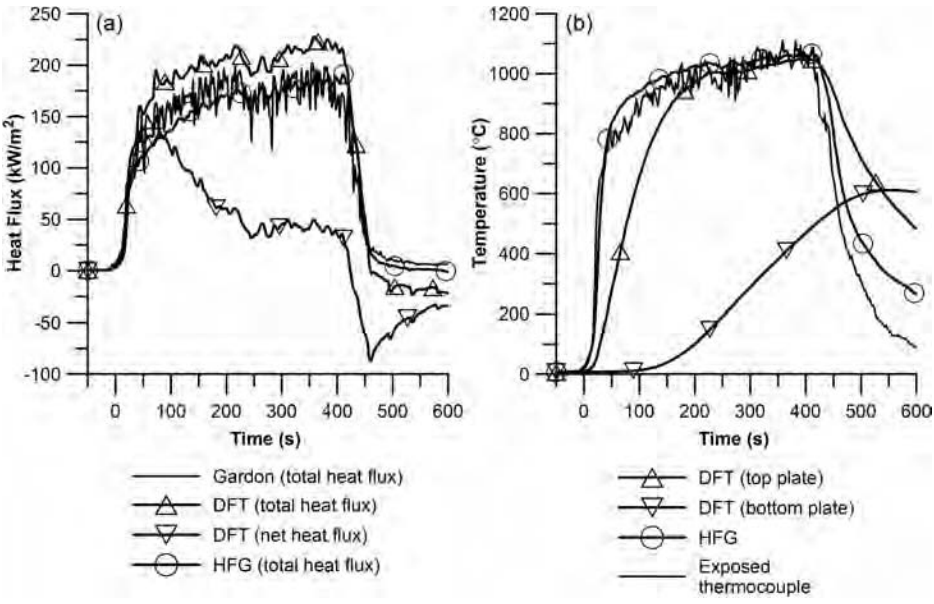


FIG. 7—*Example time traces, 7 m/s wind, x = 2.64 m, y = 0 m, (a) heat flux, and (b) temperature.*

Table 2 lists time-averaged values of the measured total incident heat flux for each gage, along with corresponding values of standard deviation for all measurement locations and wind conditions tested. The sensor plate temperatures from the DFT and HFG, together with temperatures from the neighbouring exposed thermocouple, are contained in Table 3. Reported values of heat flux and temperature were calculated by subtracting the initial value (measured immediately prior to the test) from the value averaged over the steady burning period. In this way, the values were corrected for background heat transfer between the gage and the surroundings along with any offset errors in the gage readings. No results were available for the HFG located at $y = 1$ m due to failure of the thermocouple on the sensor plate early in the test series. Similarly, no results were available from the Gardon gage at $y = 0$ m during the test with the 5 m/s wind.

Table 2 indicates that measured values of heat flux were highest on the fire axis, $y = 0$ m, and decreased towards the edges of the fire plume. Gages located at the same nominal distance to either side of the centerline (e.g., $y = \pm 1$ or ± 2 m) measured similar levels of heat flux, attesting to reasonable symmetry within the fire plume. For example, values from the two Gardon gages located at $y = \pm 1$ m differed by up to 19 kW/m^2 (35 % of the average of values measured by both gages), while those from the corresponding DFTs differed by up to 6 kW/m^2 (11 % of the average). At $y = \pm 2$ m, data from the Gardon gages, DFTs, and HFGs differed by up to 2, 4, and 3 kW/m^2, respectively. Due to the low levels of heat flux measured at this distance, these differences corresponded to larger percentage differences (e.g., up to 113 % for the HFG) than at $y = \pm 1$ m.

TABLE 2—*Increase in incident total heat flux to the ground, kW/m^2.*

Nominal *y* Location	– 2 m	– 1 m	0 m[a]	1 m	2 m
3 m/s Wind:					
Gardon	14 ± 3	40 ± 8	176 ± 14	51 ± 9	16 ± 3
DFT	7.9 ± 0.6	42 ± 3	212 ± 7	38 ± 3	5.1 ± 0.6
HFG	4.1 ± 0.6	41 ± 3	157 ± 10	…	1.2 ± 0.2
5 m/s Wind:					
Gardon	13 ± 2	45 ± 9	…	64 ± 10	13 ± 2
DFT	7.5 ± 0.9	48 ± 3	220 ± 4	54 ± 5	4.8 ± 0.7
HFG	3.8 ± 0.5	46 ± 4	176 ± 4	…	1.0 ± 0.1
7 m/s Wind:					
Gardon	8.9 ± 1.5	64 ± 11	175 ± 17	79 ± 10	8.1 ± 0.8
DFT	7.0 ± 0.8	64 ± 4	210 ± 10	69 ± 5	3.4 ± 0.5
HFG	2.7 ± 0.6	59 ± 4	177 ± 7	…	0.5 ± 0.1
10 m/s Wind:					
Gardon	9.7 ± 1.6	83 ± 11	199 ± 14	101 ± 10	7.9 ± 0.9
DFT	5.6 ± 1.0	93 ± 4	241 ± 8	99 ± 11	2.8 ± 0.5
HFG	2.4 ± 0.6	94 ± 4	199 ± 7	…	0.4 ± 0.1

[a]At $y = 0$ m, measurements from the Gardon gage were above the manufacturer-provided range of calibration (maximum 155 kW/m^2), but below the damage threshold for the gage (207 kW/m^2).

TABLE 3—*Temperature increases measured by heat flux gages and exposed thermocouples, °C.*

Nominal *y* Location	– 2 m	– 1 m	0 m	1 m	2 m
3 m/s Wind:					
DFT	187 ± 15	578 ± 7	1032 ± 12	537 ± 17	140 ± 10
HFG	225 ± 9	630 ± 8	1001 ± 17	…	105 ± 4
Thermocouple	54 ± 2	235 ± 9	977 ± 50	184 ± 10	40 ± 2
5 m/s Wind:					
DFT	157 ± 16	576 ± 18	1029 ± 9	604 ± 25	114 ± 12
HFG	208 ± 8	646 ± 13	1022 ± 6	…	89 ± 6
Thermocouple	44 ± 2	274 ± 14	1057 ± 29	251 ± 12	35 ± 2
7 m/s Wind:					
DFT	120 ± 15	645 ± 24	1015 ± 17	627 ± 45	62 ± 8
HFG	173 ± 9	720 ± 9	1037 ± 12	…	48 ± 3
Thermocouple	39 ± 2	375 ± 17	1024 ± 46	303 ± 13	27 ± 1
10 m/s Wind:					
DFT	97 ± 13	750 ± 22	1061 ± 13	750 ± 42	58 ± 7
HFG	154 ± 8	840 ± 8	1073 ± 13	…	41 ± 3
Thermocouple	36 ± 2	565 ± 19	1088 ± 32	477 ± 38	25 ± 1

Furthermore, Gardon data for $y = 1$ and 2 m were generally higher than those for $y = -1$ and -2 m, consistent with the fact that the Gardon gages on the $+y$ side of the test enclosure were approximately 0.2 m closer to the fire than those on the $-y$ side (Table 1). The reverse was true for the HFGs at $y = \pm 2$ m. No consistent trend was exhibited in the DFT data because the DFTs were centered on the nominal y locations and thus equidistant from the centerline of the fire plume (Table 1). The preceding results indicate that the plume was symmetrical across the $y = 0$ m plane, which is in agreement with the temperature contour plots of Fig. 5. This is further confirmed by the thermocouple data in Table 3; thermocouples at $y = \pm 1$ m measured temperatures within 88°C of each other (17 % of the average), while those at $y = \pm 2$ m were within 14°C of each other (29 % of the average).

While the overall symmetry of the fire plume could be confirmed by the consistency in measured heat flux from each gage type (Table 2), it is clear that magnitudes of measured heat flux differed from one gage type to another. For large, optically thick, hydrocarbon fires under low wind conditions, uncertainties in steady-state measurements of incident radiative heat flux have been found to be approximately 39 % for commercially available gages like the Gardon gage, 27 to 40 % for DFT-type gages based on inverse conduction methods, and 24 to 42 % for thin-skin sensors like the HFG [23]. For similar measurements under high wind conditions, the corresponding uncertainties were estimated to be 23 % for the Gardon gage, 25 to 27 % for the DFT and 21 to 31 % for the HFG [23]. Since these levels of measurement uncertainty still do not fully explain the differences in heat flux at all measurement locations in Table 2, other factors affecting gage response must be examined and, where possible, appropriate correction methods developed.

In the remainder of this paper, it is shown that differences in the location of the gages relative to the fire plume greatly contributed to differences in measured heat flux from each group of gages. This is affected by the overall response of each heat flux gage, combined with the sensitivity of any correction method, to the local conditions encountered in the experiment. As such, corrections for differences in the y location of the gages relative to the hot central core of the fire plume (the major source of radiation being measured) form the framework for much of the discussion, laying the foundation for the examination of other sources of uncertainty such as gage surface temperature, gage thermal response to the inherent modes of heating involved in the large hydrocarbon fire environment, and conduction losses from the gage sensor plates. While the latter effects are better known, the former type of correction is rarely discussed in the literature and can quite often be missed during large-scale experiments of any kind, yet it needs to be considered since it is a key factor in assessing the overall uncertainty in heat flux measurements in large fires.

Correction for Differences in Gage Position

As indicated in Table 1, the heat flux gages at each nominal measurement position varied in the y location by up to ± 0.12 m. In each case, since the DFT was closest to the nominal measurement position, HFG and Gardon data were

corrected to the location of the DFT. As discussed in this section, different correction methods were developed for each measurement location because at each point the gages were exposed to differing contributions of radiative and convective heat flux from the fire.

At $y = 0$ m, the heat flux gages were fully immersed in the fire as indicated by the temperature contour plots in Fig. 5. This is further substantiated by the temperature data in Table 3 where the sensor plate temperatures of the DFT and HFG agree closely (within 6 %) with the data from the exposed thermocouple.[5] Although some convective heating of the gages may have occurred, particularly at higher wind speeds, the thermal conditions at this location were radiation-dominated and the following equation was used to model the gage response to the fire

$$Q_{rad} = \alpha_g \varepsilon_f \sigma T_f^4 \tag{1}$$

For a given heat flux gage, a change in position was associated with a change in T_f, which is the temperature of the flames radiating to the gage. (Changes in ε_f were assumed to be negligible.) The resulting change in Q_{rad} was then expressed as

$$\Delta Q_{rad} = \frac{\alpha_g \varepsilon_f \sigma (T_f + \Delta T)^4 - \alpha_g \varepsilon_f \sigma T_f^4}{\alpha_g \varepsilon_f \sigma T_f^4} = \frac{(T_f + \Delta T)^4}{T_f^4} - 1 \tag{2}$$

Here, T_f was taken as the temperature measured by the exposed thermocouple in each group of sensors (Table 3). The change in flame temperature due to the change in position, ΔT, was estimated based on the temperature gradient along the lowest measurement height ($z = 0.03$ m) in the two measurement planes closest to the heat flux gages, $x = 2$ and 3 m (example plots are in Fig. 8). Data from these two measurement planes were averaged before being used in Eq 2. For simplicity, ΔT was always estimated based on the maximum offset distance of 0.12 m from the nominal y location (corresponding to the DFT), rather than on the measured y locations listed in Table 1. In Figs. 8(*a*), 8(*b*), for example, ΔT was -40°C on the $-y$ side and -45°C on the $+y$ side. Values of ΔQ_{rad}, which corresponded to the expected difference in measured heat flux caused by

[5]Due to differences in geometry, bare-bead thermocouples respond differently to a given set of thermal conditions than do the sensor plates of the DFT and HFG, often resulting in different temperature measurements. As seen in Table 3, temperatures measured by the DFT and HFG at $y = \pm 1$ and ± 2 m were significantly higher (up to 395C, or 168 %) than those measured by the exposed thermocouples. This was likely due to differences in convective effects and in the radiative response of the different sensors [24,25]. (Conduction losses might also be considered, although these are likely less important than the other two modes.) The temperatures measured by the thermocouple, DFT, and HFG at $y = 0$ m indicate minimal differences due to complete immersion of the sensing surfaces in the fire, thereby implying that radiation from optically thick flames was the dominant heating source.

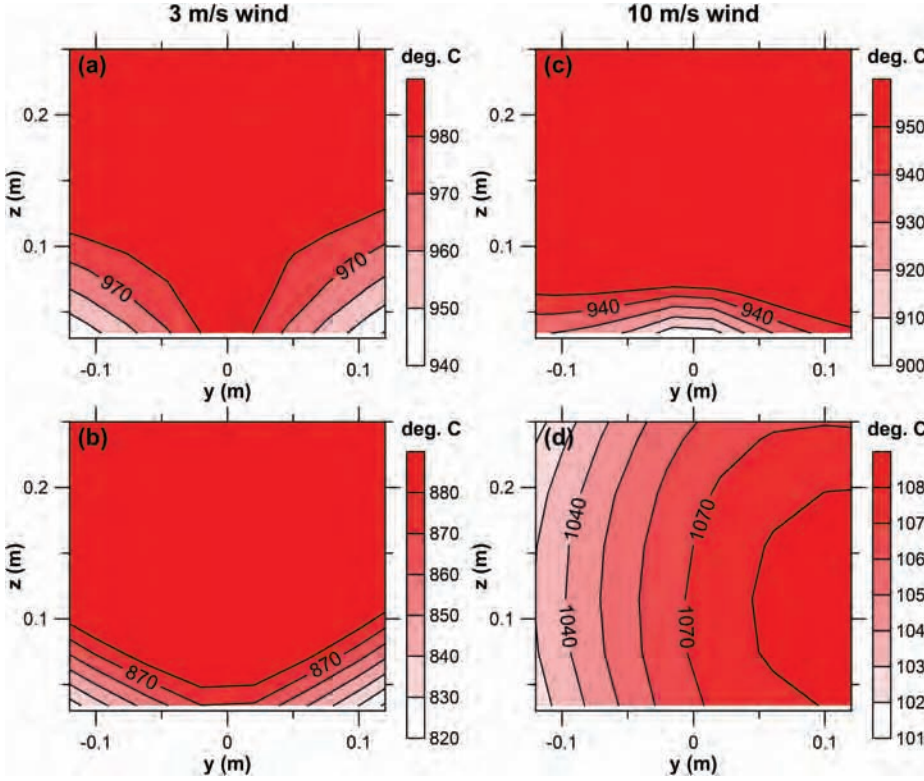


FIG. 8—*Example temperature contour plots, zoomed in near the ground at y = 0 m, (a) x = 2 m, 3 m/s wind, (b) x = 3 m, 3 m/s wind, (c) x = 2 m, 10 m/s wind, and (d) x = 3 m, 10 m/s wind.*

differences in the y position, were then determined for the Gardon gage and HFG. Results for the gages at $y = 0$ m are given in the middle column of Table 4.

In the preceding calculation, uncertainty in the values of T_f would be expected to affect ΔQ_{rad} much less than uncertainty in the values of ΔT. Uncertainty in the values of ΔT arose because they were determined using the temperature contour plots, which were themselves produced via interpolation of thermocouple measurements taken at 0.5 m intervals in the y direction (Fig. 3). As an example of the effect, ΔT ranged from -40 to 20°C in Figs. 8(*c*), 8(*d*); a variation in ΔT of $\pm$ 30°C corresponded to a change in ΔQ_{rad} of approximately ± 9 %.

For measurements made at $y = \pm 1$ m, the heat flux gages were situated near the edges of the fire plume, as indicated by the temperature contour plots in Fig. 5. Since these gages were not immersed in optically thick flames, both radiation and convection contributed to the total heat flux incident on each gage. To estimate the total heat flux, the following equation was considered

$$Q_{tot} = \alpha_g \varepsilon_f \sigma T_f^4 + h(T_f - T_g) \tag{3}$$

TABLE 4—*Predicted percent differences in heat flux relative to DFT (based on differences in y location of the gages).*

	−2 m	−1 m		0 m	1 m		2 m
Gage	ΔQ_{rad}, %	ΔQ_{rad}, %	ΔQ_{conv}, %	ΔQ_{rad}, %	ΔQ_{rad}, %	ΔQ_{conv}, %	ΔQ_{rad}, %
3 m/s Wind:							
Gardon	−9	−15	−15	−12	154	143	11
HFG	11	134	−30	−14	…	…	−9
5 m/s Wind:							
Gardon	−9	−12	−10	…	174	99	12
HFG	12	135	−35	−14	…	…	−9
7 m/s Wind:							
Gardon	−9	−17	−11	−17	126	64	11
HFG	11	97	−35	10	…	…	−9
10 m/s Wind:							
Gardon	−9	−29	−15	−10	65	27	11
HFG	11	57	−36	−3	…	…	−9

The first term in this equation describes the radiative contribution, while the second describes the convective contribution. For the radiative term, changes in position of the heat flux gage would be associated with changes in T_f and affect Q_{rad}, according to Eq 2. For the convective term, changes in gage position would also be associated with changes in T_f and, if the gage was not water-cooled, T_g. For non-cooled gages, the value of T_g depends on the value of T_f due to radiative and convective coupling between the gage surface and the surrounding gases. For the purposes of this analysis, changes in T_g were assumed to be much smaller than the corresponding changes in T_f and changes in h were assumed negligible, resulting in the following simple expression for ΔQ_{conv}

$$\Delta Q_{conv} = \frac{h(T_f + \Delta T - T_g) - h(T_f - T_g)}{h(T_f - T_g)} = \frac{\Delta T}{T_f - T_g} \tag{4}$$

Equation 4 describes the change in convective heat flux due to a change in position for a given sensor. For the HFG, T_g was taken as the temperature listed in Table 3, while for the Gardon gage, it was assumed to be 100°C. Results for both ΔQ_{rad} and ΔQ_{conv} for the gages located at $y = \pm 1$ m are listed in separate columns in Table 4.

The greatest sources of uncertainty in Eq 4 lay in the estimates of ΔT and T_g, particularly for the Gardon gage. Because the sensing foil in the Gardon gage is attached along its edges to a heat sink, the foil temperature is not uniform across its surface [16,17]. A variation in T_g (taken here as an average foil temperature) of 50°C caused ΔQ_{conv} to change by up to 9 % for the data at $y = -1$ m. For the data at $y = 1$ m at the lowest wind speed, a 50°C change in T_g

caused ΔQ_{conv} to change by up to 210 % because the value of T_f was close to that of T_g. As the wind speed increased, the value of T_f increased and corresponding changes in ΔQ_{conv} decreased to 4 %. Similarly, a change in ΔT of 30°C corresponded to a 6 to 36 % change in ΔQ_{conv}, with the largest difference occurring at the lowest wind speed at $y = 1$ m. This brief analysis points to the sensitivity of these corrections to details of the local conditions encountered by a heat flux gage in a given experiment.

At $y = \pm 2$ m, the gages were well outside the main fire plume based on the temperature contour plots in Fig. 5. Radiation from the flame to the gage would be expected to be the dominant mode of heat transfer, although there may be some convective effects due to differences between the local gas temperature and the surface temperature of the gage (Table 3). Nevertheless, because the gages were not situated in a region of large temperature gradients, changes in gage position would not be expected to result in significant differences in convection to the gage. Thus, only differences in radiation were considered.

The expected impact on measured heat flux due to changes in gage position at $y = \pm 2$ m was estimated via the change in view factor between the fire and each gage. To calculate this, the fire was represented as a tilted cylinder with a diameter equal to the fuel pan diameter and a length equal to the distance from the fuel pan center to a flame "tip" determined from the temperature contours along the midplane, $y = 0$ m (Fig. 6). The location of the flame "tip" was taken as the intersection of the 550°C temperature contour, which corresponds to 50 % flame intermittency [26], with the plume centerline. The angle of the tilted cylinder was taken as the angle between the vertical and the line used to determine the cylinder length. In Fig. 6(*a*), the flame "tip" is located at $x = 4.4$ m and $z = 1.5$ m, giving a cylinder length of 4.6 m and a tilt angle of 72°. In Fig. 6(*b*), the location of the plume centerline is not obvious, so the flame "tip" was estimated to be at $x = 6.0$ m and $z = 0.7$ m, where the 550°C contour just starts to slope down towards the ground.

View factors were estimated using two closed-form expressions developed by Mudan [27] for a horizontal differential element (the heat flux gage) viewing the sides of a tilted flame cylinder. One expression was for a differential element positioned downwind of the cylinder directly underneath the cylinder's axis (i.e., along $y = 0$ m), while the other was for a differential element positioned along a line intersecting the center of the base of the cylinder and oriented perpendicular to the wind direction (i.e., along $x = 0$ m).[6] Both were applied to the gages at $y = \pm 2$ m because the gages were 2.00 ± 0.12 m away from the fuel pan center in the y direction and also 2.64 m away from the fuel pan center in the x direction. Differences in view factor were greater for changes in position along the y direction, which is consistent with the notion that for small changes in the x direction, the differential element remained approximately the same distance from the cylinder, as long as the tilt angle of the cylinder was sufficiently large. Changes in view factor in the y direction were therefore used to estimate

[6]The latter expression contained an error that was identified by Howell [28] and corrected in the present analysis.

changes in incident radiative heat flux to the gages at $y = \pm 2$ m. These changes are reported in the first and last columns of Table 4.

Comparison of Predicted Differences to Measured Differences

To allow a direct comparison to the heat flux corrections in Table 4, Table 5 shows the percent differences between heat flux values measured by the DFT and those measured by the other two gages at all five locations in all four tests. Each difference is reported with a value of "uncertainty" estimated using the root-sum-square of the standard deviations listed with the measured heat flux levels in Table 2.

First considering the data for $y = \pm 2$ m only, Table 5 shows that, in general, the Gardon gage measured the highest heat flux level of the three gages in its group and the HFG measured the lowest level, despite the fact that the Gardon gage was closest to the fire at $y = 2$ m but furthest from it at $y = -2$ m, and vice versa for the HFG (Table 1). In the extreme, heat flux levels from the Gardon gage were higher than those from the DFT by 218 %, while heat flux levels from the HFG were lower by 85 %. Table 4 indicates that changes in view factor due to changes in gage location accounted for only approximately 10 % of the difference in heat flux. Therefore, the results suggest that in the thermal conditions present at $y = \pm 2$ m, there was an inherent bias in measured heat flux depending on the type of gage employed.

Part of this bias may be due to differences in temperature of the gage sensing elements. The Gardon gage was cooled by water at approximately 16°C, while the air above the gage ranged from 25 to 54°C (Table 3). Thus, a small amount of convective heating of the Gardon gage may have occurred in addition to radiative heating from the fire. (There may also have been conductive effects on the foil temperature due to the sensing foil being attached along its edges to a heat sink and at its centre to a metal wire [16,17].) In contrast, neither the

TABLE 5—*Percent differences in measured heat flux relative to DFT (based on values in Table 2).*

Gage	− 2 m, %	− 1 m, %	0 m, %	1 m, %	2 m, %
3 m/s Wind:					
Gardon	74 ± 43	−6 ± 19	−17 ± 7	32 ± 25	218 ± 71
HFG	−49 ± 8	−3 ± 9	−26 ± 5	…	−77 ± 4
5 m/s Wind:					
Gardon	71 ± 32	−7 ± 19	…	20 ± 22	180 ± 56
HFG	−50 ± 9	−4 ± 10	−20 ± 2	…	−78 ± 4
7 m/s Wind:					
Gardon	27 ± 25	0 ± 18	−17 ± 9	14 ± 17	137 ± 41
HFG	−61 ± 9	−8 ± 8	−16 ± 5	…	−85 ± 3
10 m/s Wind:					
Gardon	72 ± 42	−10 ± 13	−17 ± 6	2 ± 15	177 ± 58
HFG	−58 ± 13	1 ± 6	−18 ± 4	…	−85 ± 3

DFT nor the HFG was water-cooled and the temperatures of their sensor plates ranged from 41 to 225°C, resulting in convective cooling of the plates by the surrounding air (Table 3). In either case, such convective effects would be most pronounced at the low levels of heat flux measured at $y = \pm 2$ m and may explain why the Gardon gage consistently measured the highest values of heat flux at this position. This observation is consistent with Robertson and Ohlemiller [29], who found that the error caused by differences in convection due to a change in gage surface temperature could form a significant portion of the measured signal (~50 % for a surface temperature difference of 35°C) if the incident radiative heat flux was low (< 15 kW/m^2).

Similarly, differences between the HFG and DFT results at $y = \pm 2$ m may, in part, be due to the small differences in gage surface temperature (Table 3) that resulted from differences in thermal properties of the sensor plates and insulation backing in the gages. The values of heat flux from the HFG may be further influenced by significant conductive losses from the sensor plate to the cylindrical gage housing, which could cause differences of ~50 %, if not accounted for in the one-dimensional model used to reduce the data [12]. Since the 9 % decrease in heat flux, expected at $y = 2$ m because the HFG was further from the fire than the DFT (Table 4), would have been counteracted by lower convective cooling of the HFG (Table 3), only conductive losses could account for the large differences between the measured heat flux from the two gages (between -77 and -85 %). This reiterates the need to either improve the design of the HFG or to enhance the data reduction model to account for such losses.

At $y = 0$ m, where primarily radiative heat flux was expected, Table 5 indicates that the difference in measured heat flux between the Gardon gage and DFT was constant at -17 % in all tests. This value agreed reasonably well with the predicted radiation corrections in Table 4, which ranged from -10 to -17 %, and was consistent with the ~25 % uncertainty estimated by Nakos [23] for commercial heat flux gages and DFTs in high wind conditions. Most of the difference in measured heat flux at this location was therefore due to the difference in gage position.

For the HFG, the difference in measured heat flux compared to the DFT was between -16 and -18 % at the two highest wind speeds and increased in magnitude to -26 % as the wind speed decreased to 3 m/s (Table 5). This trend was consistent with the predicted corrections listed in Table 4, albeit of differing magnitudes (ranging from 10 to -14 %). The values listed in Table 4 for the highest two wind speeds (10 and -3 %) reflected a possible shift in the hottest portion of the plume in the $+y$ direction (thus closer to the HFG) during these tests. Figures 8(*d*) also suggest that the plume center shifted by approximately 0.2 m in the $+y$ direction along the $x = 3$ m measurement plane; however, better spatial resolution of the thermocouple data (less than 0.5 m) would be needed to confirm these results. The measured percent differences in Table 5 were consistent with such a shift, although the trend was not as pronounced as in Table 4, partly because the HFG was exposed to the flame over a field of view of 2π sr while the predicted corrections were determined using a single temperature to represent the flame immediately above the gage.

Differences in heat flux summarized in Table 5 for $y = 0$ m were reasonably consistent with the gage biases suggested by the data at $y = \pm 2$ m. The HFG

measured lower heat flux values than the DFT, even when the hottest portion of the plume shifted closer to the HFG at the higher wind speeds. At the same time, the lower levels of heat flux measured by the Gardon gage relative to the DFT suggested that the difference in position between the two gages counteracted any potential tendency for the Gardon gage to measure higher heat flux values than the DFT. Convective effects, particularly at the higher wind speeds, would have influenced the Gardon gage differently than the DFT, however, convection was expected to play a much smaller role than radiation at $y = 0$ m, so such differences would have much less impact on the data [29]. Changes in sensitivity of the Gardon gage may have occurred under the extremely high heat flux levels, higher than the manufacturer calibration range, experienced at this location (Table 2), contributing to additional difference between the Gardon and DFT measurements.

At $y = -1$ m, the Gardon gage and HFG measured heat flux levels similar to those from the DFT (within 10 % agreement), as seen in Table 5. Such agreement was better than that predicted in Table 4, which indicates an expected 12 to 29 % decrease in radiative heat flux and a 10 to 15 % decrease in convective heat flux for the Gardon gage. For the HFG, which was closest to the fire among the three gages (Table 1), the radiative heat flux was expected to be higher by 57 to 135 % and the convective heat flux lower by 30 to 36 %. Because the HFG sensing surface was hotter than the local surrounding gases (Table 3), the convective heat flux acted to cool the gage, so a lower convective heat flux would have contributed to a higher total heat flux. Although the relative ratio of radiative to convective heat flux at $y = -1$ m was not known, thus preventing the values for ΔQ_{rad} and ΔQ_{conv} from being combined into a single weighted value, comparison of the data in Tables 4 suggests that the expected difference in heat flux caused by the difference in gage location was fortuitously balanced by the biases involved in using the different types of gages. For instance, although Table 4 indicates that the Gardon gage was expected to measure lower values of heat flux than the DFT, the Gardon gage was biased to measure higher heat flux values than the DFT due to differences in temperature of the sensing element and thus convective response. When these two effects were of similar magnitude, they would cancel each other out, resulting in the Gardon gage measuring a heat flux level similar to that of the DFT. Similarly, for the HFG, Table 4 indicates that the measured heat flux levels should be higher than those from the DFT, however, this was balanced by a tendency for the HFG to measure lower heat flux values due to conduction losses from its sensor plate to its housing. The data in Table 5 therefore suggest that the conduction losses were, in this case, sufficiently large to counteract the difference in gage position and thus minimize the overall difference between heat flux measurements made by the HFG and DFT at $y = -1$ m.

The trends seen for the gages at $y = -1$ m were not observed for the data at $y = 1$ m. Here, the difference between the measured heat flux levels from the Gardon gage and DFT decreased from 32 to 2 % as the wind speed increased from 3 to 10 m/s (Table 5). The same trend was evident in the predicted corrections, although the values were much larger (Table 4). One possible cause was that at the highest two wind speeds, the fire plume tilted over and flattened against the ground so that

at $x = 2$ m, the region of the hot core (represented by the 900°C contour) remained below a height of $z = 0.5$ m and spread from approximately $y = -1$ to 1 m, as shown in Fig. 5(*c*). Under these high wind conditions, the gages were largely immersed in the hot plume and a change in the y position resulted in a smaller relative change in heat flux. In contrast, for lower wind speeds, the gages lay along the outer edge of the plume where they were influenced by steep temperature gradients. This was reflected in Eqs 2 and 4, in which T_f increased by 140 % from 235 to 565°C as the wind speed increased from 3 to 10 m/s (Table 3). With ΔT remaining between 100 and 150°C in all four tests, the increase in T_f with higher wind speed therefore corresponded to a decrease in ΔQ_{rad} and ΔQ_{conv}.

Although trends in the data for $y = 1$ m in Tables 4 were consistent with each other, the magnitudes at each wind speed greatly differed. Part of this difference was caused by the simplifications made in deriving the expressions for ΔQ_{rad} and ΔQ_{conv}. Additionally, the sensitivity of the Gardon gage has been shown to decrease when affected by convection [4,9,11]. Under normal conditions, the Gardon measurement is based on the temperature difference between the center and edge of the sensing foil [16,17]. With increased convection, the temperature distribution across the foil becomes asymmetric (rather than the temperature peaking at the center of the foil and decreasing towards the edges). When such a temperature gradient exists across the foil, convective heat transfer to the foil is non-uniform and the calibration response becomes nonlinear [9,11]. The data in Table 5, at $y = 1$ m (particularly at the highest wind speed), suggest that the above effects were significant and largely overcame the expected increase in total heat flux caused by both the change in gage position and the tendency of the Gardon gage to experience higher convective heat transfer to its water-cooled sensing element than the non-cooled DFT.

As a final note, it should be emphasized that, when convection is significant, water-cooled gages do not make representative measurements of the total heat flux to a surrounding surface unless the surface is also water-cooled. Having a water-cooled gage embedded in a surface that heats up during an experiment greatly affects both the surface temperature distribution and the thermal boundary layer, resulting in significant convective heat transfer effects. Conversely, a non-water-cooled gage will measure representative values of heat flux to a surrounding surface only if it has similar thermal properties and surface roughness such that perturbation of the surrounding thermal and flow fields is minimized [30]. Despite these practicalities and the various issues outlined in this paper, the measurement of heat flux remains of critical importance in assessing fire exposure and thermal hazard. Therefore, quantification of uncertainties in heat flux measurement remains a vital area of research.

Conclusions

When exposed to low levels of heat flux (≤ 15 kW/m^2), Gardon gage measurements at $y = \pm 2$ m were higher than those from the DFT by up to ~200 % due to differences in gage surface temperature and convective response. This variation was significantly larger than the 3 % calibration uncertainty specified by the manufacturer for radiation-dominated environments. On the contrary, when

exposed to high levels of radiation from optically thick flames at $y = 0$ m, differences between the Gardon and DFT measurement results could be mostly accounted for by the difference in gage position. Here, differences in convection due to changes in surface temperature did not have a significant impact on the measured total heat flux levels. In the mixed radiative-convective environments present at $y = \pm 1$ m, differences in heat flux caused by the difference in gage position appeared to be counteracted by differences in convective heat transfer due to different gage surface temperatures and by decreased sensitivity of the Gardon gage under convective flows.

For the HFG, measured values were up to ~80 % lower than those from the DFT for the low levels of heat flux experienced at $y = \pm 2$ m, due to differences in gage surface temperature and conduction losses from the HFG sensor plate to the gage housing. The effects of these losses need to be minimized by improving either the design of the gage or the one-dimensional model used to obtain heat flux results from the HFG temperature measurements. However, when exposed to optically thick flames at $y = 0$ m, the HFG results were much closer to those from the DFT after accounting for changes in gage position. At $y = -1$ m, the difference in heat flux levels due to the difference in gage position appeared to be of similar magnitude to conduction losses from the sensor plate of the HFG.

Based upon the aforementioned findings, heat flux gages should be selected based upon the thermal conditions expected to occur at the measurement locations of interest in the fire. All three types of gages can be used in optically thick, radiation-dominated fire environments. Non-water-cooled gages are preferred for use in environments where convection may be significant or where low levels of heat flux are expected, providing that all potential sources of heat loss from the sensor plate are either minimized or accounted for in the model of the gage. It should be noted that despite the difficulties involved in interpreting Gardon measurements, Gardon gages continue to be used in fires (e.g., see [4]) due to their fast time response, ease of use, and relatively low cost. Therefore, knowledge of when the use of Gardon gages is appropriate is important to improving the accuracy and interpretation of their heat flux measurements. Overall, in cases where a wide range of thermal conditions may be encountered or where thermal conditions are unknown, a non-cooled gage such as the DFT is recommended for more accurate results.

Acknowledgments

Contributions and funding were obtained from the Natural Sciences and Engineering Research Council of Canada, Canada Foundation for Innovation, and Ontario Innovation Trust. The experimental testing was conducted in consultation with researchers at Sandia National Laboratories of New Mexico, U.S.A.

References

[1] ASTM E1529-06, 2006, "Standard Test Methods for Determining Effects of Large Hydrocarbon Pool Fires on Structural Members and Assemblies," *Annual Book of ASTM Standards*, ASTM International, West Conshohocken, PA, pp. 1–15.

[2] Pitts, W. M., Murthy, A. V., de Ris, J. L., Filtz, J. R., Nygård, K., Smith, D., and Wetterlund, I., "Round Robin Study of Total Heat Flux Gauge Calibration at Fire Laboratories," *Fire Saf. J.*, Vol. 41, 2006, pp. 459–475.
[3] Bryant, R., Womeldorf, C., Johnsson, E., and Ohlemiller, T., "Radiative Heat Flux Measurement Uncertainty," *Fire Mater.*, Vol. 27, 2003, pp. 209–222.
[4] Silvani, X. and Morandini, F., "Fire Spread Experiments in the Field: Temperature and Heat Fluxes Measurements," *Fire Saf. J.*, Vol. 44, 2009, pp. 279–285.
[5] Lennon, P. F. and Silcock, G. W. H., "A Preliminary Investigation into the Partitioning of the Convective and Radiative Incident Heat Flux in Real Fires," *Fire Technol.*, Vol. 42, 2006, pp. 109–129.
[6] Jones, J. C., "Rates of Radiative and Convective Heat Transfer in a Cabin Fire," *J. Fire Sci.*, Vol. 17, 1999, pp. 103–110.
[7] Nakos, J. T. and Keltner, N. R., "The Radiative-Convective Partitioning of Heat Transfer to Structures in Large Pool Fires," *1989 National Heat Transfer Conference*, HTD-Vol. 106, Philadelphia, PA, August 6–9, 1989, American Society of Mechanical Engineers, NY, pp. 381–387.
[8] Holmberg, D. G., Womeldorf, C.A., and Grosshandler, W. L., "Design and Uncertainty Analysis of a Second-Generation Convective Heat Flux Calibration Facility," *Proceedings of the ASME Heat Transfer Division*, HTD-Vol. 364-4, Nashville, TN, November 14–19, 1999, American Society of Mechanical Engineers, NY, pp. 65–70.
[9] Borell, G. J. and Diller, T. E., "A Convection Calibration Method for Local Heat Flux Gages," *J. Heat Transfer*, Vol. 109, 1987, pp. 83–89.
[10] Gifford, A., Hoffie, A., Diller, T., and Huxtable, S., "Convection Calibration of Schmidt-Boelter Heat Flux Gauges in Stagnation and Shear Air Flow," *J. Heat Transfer*, Vol. 132, 2010, pp. 031601-1–031601-9.
[11] Kuo, C. H. and Kulkarni, A. K., "Analysis of Heat Flux Measurement by Circular Foil Gages in a Mixed Convection/Radiation Environment," *J. Heat Transfer*, Vol. 113, 1991, pp. 1037–1040.
[12] Lam, C. S. and Weckman, E. J., "Steady-State Heat Flux Measurements in Radiative and Mixed Radiative-Convective Environments," *Fire Mater.*, Vol. 33, 2009, pp. 303–321.
[13] Gritzo, L. A., Gill, W., Keltner, N., "Thermal Measurements to Characterize Large Fires," *Proceedings of the 41st International Instrumentation Symposium*, Aurora, CO, May 7–11, 1995, Instrument Society of America, Research Triangle Park, NC, pp. 337–346.
[14] Best, C., 2010, "Measurement of Fuel Regression Rate of a Pool Fire in Crosswind With and Without a Large Downwind Blocking Object," M.A.Sc. thesis, Univ. of Waterloo, Waterloo, ON, Canada.
[15] Gardon, R., "An Instrument for the Direct Measurement of Intense Thermal Radiation," *Rev. Sci. Instrum.*, Vol. 24, 1953, pp. 366–370.
[16] ASTM E511-07, 2007, "Standard Test Method for Measuring Heat Flux Using a Copper-Constantan Circular Foil, Heat-Flux Transducer," *Annual Book of ASTM Standards*, ASTM International, West Conshohocken, PA pp. 1–10.
[17] Keltner, N. "Thermal Measurements in Fire Safety Testing – Are We Playing With Fire?" *Fire Calorimetry, Report No.* DOT/FAA/CT-95/46, Gaithersburg, MD, July 27–28, 1995, Federal Aviation Administration Technical Center, Atlantic City, NJ.
[18] Keltner, N. R., Beck, J. V., and Nakos, J. T., "Using Directional Flame Thermometers for Measuring Thermal Exposure," *J. ASTM Int.*, Vol. 7, No. 2, 2010, pp. 1–12.
[19] Blanchat, T. K., Humphries, L. L., and Gill, W., "Sandia Heat Flux Gauge Thermal Response and Uncertainty Models," *Thermal Measurements: The Foundation of Fire Standards, ASTM STP1427*, L. A. Gritzo and N. J. Alvares, Eds., ASTM International, West Conshohocken, PA, 2002, pp. 81–110.

[20] Beck, J. V., 1999, *User's Manual for IHCP1D*, 7th ed, Beck Engineering Consultants Company, Okemos, MI.
[21] Lam, C. S., 2009, "Thermal Characterization of a Pool Fire in Crosswind With and Without a Large Downwind Blocking Object," Ph.D. thesis, Univ. of Waterloo, Waterloo, ON, Canada.
[22] Surfer 8, (2002). Golden Software, Inc., Golden, CO.
[23] Nakos, J. T., "Uncertainty Analysis of Steady State Incident Heat Flux Measurements in Hydrocarbon Fuel Fires," *Report No.* SAND2005-7144, Sandia National Laboratories, Albuquerque, NM, 2005.
[24] Blevins, L. G. and Pitts, W. M., "Modeling of Bare and Aspirated Thermocouples in Compartment Fires," *Fire Saf. J.*, Vol. 33, 1999, pp. 239–259.
[25] Nakos, J. T., Gill, W., and Keltner, N. R., "An Analysis of Flame Temperature Measurements Using Sheathed Thermocouples in JP-4 Pool Fires," *Proceedings of the ASME/JSME Thermal Engineering Joint Conference*, Reno, NV, March 17–22, 1991, American Society of Mechanical Engineers, NY, pp. 283–289.
[26] McCaffrey, B. J., "Purely Buoyant Diffusion Flames: Some Experimental Results," *Report No.* NBSIR 79-1910, National Bureau of Standards, Washington, D.C., 1979.
[27] Mudan, K. S., "Geometric View Factors for Thermal Radiation Hazard Assessment," *Fire Saf. J.*, Vol. 12, 1987, pp. 89–96.
[28] Howell, J. R., *A Catalog of Radiation Heat Transfer Configuration Factors*, 2nd ed., http://www.me.utexas.edu/~howell/index.html, accessed Nov 2010.
[29] Robertson, A. F. and Ohlemiller, T. J., "Low Heat-Flux Measurements: Some Precautions," *Fire Saf. J.*, Vol. 25, 1995, pp. 109–124.
[30] Hornbaker, D. R. and Rall, D. L., "Thermal Perturbations Caused by Heat-Flux Transducers and Their Effect on the Accuracy of Heating-Rate Measurements," *ISA Trans.*, Vol. 3, 1964, pp. 123–130.

Reprinted from JTE, Vol. 40, No. 1
doi:10.1520/JTE103913
Available online at www.astm.org/JTE

Timothy T. Earl[1] *and Marcelo M. Hirschler*[1]

Development of a Proposed ASTM Guide to Continued Applicability of Reports on Fire Test Standards

ABSTRACT: Fire test reports provide information relative to the fire-test-response characteristics of a material, product, or assembly at the time when it was tested and to the evaluative approach of the fire test used, including understanding of the fire test that was used. Such fire test reports remain valid for the particular point in time at which the fire test was conducted, as long as the test was conducted in full accordance with the fire test standard referenced in the test report. However, fire test reports may, at some time, cease being applicable to the material, product, or assembly currently being offered for use; for example, if there has been a technical change in the fire test protocol or if there has been a substantial change in the material, product, or assembly being offered for use. The ASTM Committee E05, and subcommittee E05.31 on Terminology and Services/Functions, is considering the development of an ASTM Guide to formalize the above ideas. The draft ASTM guide contains concepts which provide guidance for assessing the continued applicability of fire test reports. The concepts in the draft guide are intended for application by users of fire test reports to assess whether a particular fire test report continues to be an applicable representation of the fire-test-response characteristics of a material, product, or assembly which is required to be tested for a new assessment using the same fire test standard. The continued applicability of the fire test report will be a function primarily of two issues: (a) whether the material, product or assembly being offered for use is substantially the same as the one that was tested and (b) whether the test

Manuscript received April 15, 2011; accepted for publication August 23, 2011; published online October 2011.

[1] GBH International, 2 Friars Ln., Mill Valley, CA 94941.

Cite as: Earl, T. T. and Hirschler, M. M., "Development of a Proposed ASTM Guide to Continued Applicability of Reports on Fire Test Standards," *J. Test. Eval.*, Vol. 40, No. 1. doi:10.1520/JTE103913.

method being required is substantially the same as the test that was conducted for the report in question. Some concepts included in the draft guide may be a function of the type of fire test conducted and of the type of material, product, or assembly assessed. The draft guide considers in a separate category materials, products or assemblies listed by an accredited listing agency.

KEYWORDS: fire test, fire test response standard

Introduction

Fire tests are, sometimes, conducted for research and development, but they are very often conducted in order to obtain approval of a material or product by authorities having jurisdiction for use in a particular application. It is not unusual for a fire test to be conducted at a certain point in time and for the associated report to be continued to be used for long periods following its issuance. There was a time when that practice was logical and reasonable: materials and products remained virtually unchanged for decades and fire testing technology was stagnant. At that time, fire tests were conducted with wood or paper ignition sources (and perhaps a small gas burner) and the materials tested were all natural. However, that era has long since passed.

In the 21st century technology has a short life cycle and, therefore, materials and products are constantly being invented and updated.[2] The ever shortening lifespan of new materials and products is a recent development, although no date can be assigned to indicate a sudden transition.

Also, new fire tests are being developed every few years and existing fire tests are constantly being revised. Table 1 shows the fire test methods under the jurisdiction of the ASTM E05, Committee on Fire Tests, the dates when they were first developed, and the date of the latest edition. Figure 1 shows the date of first development of ASTM E05 fire tests. The figure and table clearly show the accelerating pace of standard fire test development, with more tests having been developed between 1990 and 2009 than in all the decades preceding that period. An analysis of newer fire test methods, and a comparison with traditional ones, illustrates the fact that newer tests tend to be technically more sophisticated than the older ones. For example, newer test methods often assess properties intended for use in fire safety engineering, such as heat release and/or critical fluxes for ignition or flame spread, while traditional tests generate data that can be used simply for material or product approval. It is also interesting to note that even the two oldest tests, ASTM E84 [1] and ASTM E119 [2], are constantly being renewed and updated, with 2011 editions of both standards.

[2]Note, for example, "Gordon Moore's law" (applicable to computer hardware and which has proven very accurate over a period of more than 30 years) stating that "The number of transistors that can be placed inexpensively on an integrated circuit doubles approximately every two years." Similar "laws" or analyses have shown; for example, such market facts as: (a) computer hard drive capacity has grown exponentially between 1980 and 2011 and (b) the number of "pixels" per dollar, in digital imaging has grown in a similar way between 1995 and 2005.

TABLE 1—*ASTM E05 fire test methods.*

Fire Test Method	Year Introduced	Year Latest Edition	E05 Subcommittee
D2859	1970	2006	E05.22
E84	1950	2011	E05.22
E108	1955	2010	E05.14
E119	1917	2011	E05.11
E136	1958	2011	E05.23
E162	1960	2011	E05.22
E648	1978	2010	E05.22
E662	1979	2009	E05.21
E814	1981	2011	E05.11
E906	1983	2010	E05.21
E970	1983	2010	E05.22
E1317	1990	2008	E05.22
E1321	1990	2009	E05.22
E1352	1990	2008	E05.15
E1353	1990	2008	E05.15
E1354	1990	2011	E05.21
E1474	1992	2010	E05.21
E1529	1993	2010	E05.11
E1537	1993	2007	E05.15
E1590	1994	2007	E05.15
E1623	1994	2011	E05.21
E1678	1995	2010	E05.21
E1725	1995	2008	E05.11
E1740	1995	2010	E05.21
E1822	1996	2009	E05.15
E1966	1998	2007	E05.11
E1995	1998	2008	E05.21
E2058	2000	2009	E05.22
E2102	2000	2011	E05.21
E2187	2002	2009	E05.15
E2257	2003	2008	E05.21
E2307	2004	2010	E05.11
E2336	2004	2009	E05.11
E2405	2005	2005	E05.21
E2652	2009	2009	E05.23
E2707	2009	2009	E05.11
E2816	2011	2011	E05.11

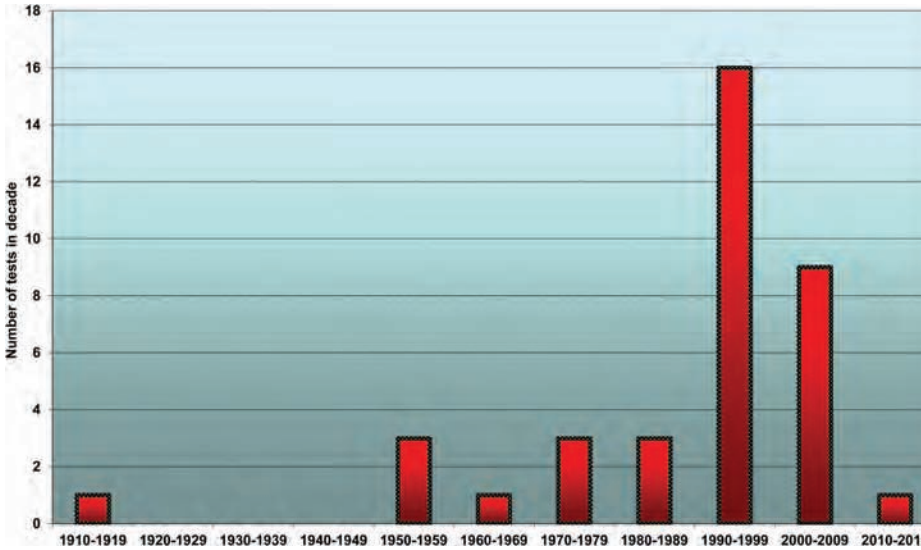


FIG. 1—*Development of ASTM E05 fire tests.*

The changes in both fire tests and materials and products being tested creates an interesting dilemma when considering fire test results: is it reasonable, given the changes in technology of material/product development and of fire test invention/revision, that there be no guidelines issued by ASTM E05 for the continued applicability of reports of fire test methods?

Test reports from fire test laboratories should be deemed to be a valid representation of the fire-test-response characteristics of the material, product, or assembly at the time that it was tested. The validity of such a test report is not within the scope of the guide being developed. In the remainder of this discussion it will be assumed that the fire test report in question was originally valid.

When considering the continued applicability of a fire test report, the following questions should be asked:

1. Is the time frame between the issuance of a fire test report and its new requested use so substantial that there is a need to question its continued applicability?
2. How does the material, product, or assembly that has been tested in the past compare with the material product or assembly being offered for use now?
3. How does the protocol of the fire test method that was used to perform the test described in the fire test report compare with the current test method protocol?
4. Are the inevitable differences in 2 and 3 small enough not to affect the performance and the safety of the material or product in its end-use application?

Those same questions can be addressed in a positive way to establish the continued applicability of an old fire test report. Summary of concept: a fire test

report will continue to be applicable to a material, product or assembly for long after the fire test was conducted unless: (a) either the test protocol has changed substantially or (b) the material, product or assembly offered for use has changed substantially, and a test conducted at the time the fire test report is needed again on the material, product or assembly offered for sale would result in a more unsafe fire test result.

Time Frame

When a U.S. code states that a material, product or assembly needs to meet a certain fire test requirement, it usually contains the date of the edition of the standard test method that must be complied with. The codes issued in the U.S. by the key code-making organizations (ICC, NFPA, and IAPMO) are renewed periodically, typically on a three year cycle. As a new edition of a code is issued, typically either the code development organization's staff or the organization's revision process will update all referenced standards to its most recent editions. This provides one time frame for the continued applicability of a fire test report: if the report is based on the edition of the fire test standard referenced in the code, the code deems it to continue to be applicable.

When a material, product, or assembly is subject to an evaluation report, and/or to a set of acceptance criteria which contains a fire test requirement, the acceptance criteria will also generally include the date of the edition of the standard test method that must be complied with. An evaluation report is a report issued by an organization such as the International Code Council (ICC) Evaluation Service. It is a public report that provides evidence, as certified by the issuing organization, that a material, product or system complies with certain requirements; for example, ones laid out in a code or regulation, based on its compliance with a variety of test methods. Evaluation report acceptance criteria are revised periodically, with a time frame typically based on the rules of the organization issuing the criteria or on requests by the holders of evaluation reports that are associated with those criteria. Once again, this provides one time frame for the continued applicability of a fire test report: if the report is based on the edition of the fire test standard referenced in the acceptance criteria, the acceptance criteria deem it to continue to be applicable.

When a certain regulatory body requires that a material, product, or assembly meet a certain fire test, the regulation will normally contain the date of the standard fire test method that must be complied with. Once again, a time frame has been provided.

When a fire test report is requested by an organization specifying such requirements without it being based on a code, a set of acceptance criteria or some regulation, the date of the edition of the standard test method to be used is likely to be contained in the specification, which may be an ASTM standard specification. Once more, a time frame has been provided.

In principle ASTM fire test standards must be revised, reapproved or withdrawn by a technical committee after no more than five years or they will be withdrawn automatically without committee action. In practice, ASTM standards are not automatically withdrawn until the ninth year after the date of their

last revision or reapproval/reconfirmation. Similarly, National Fire Protection Association (NFPA) fire test standards must also be revised, reapproved/reconfirmed or withdrawn by a technical committee after no more than five years from their last edition. This ensures that, for both organizations, a technical committee has considered whether revisions are needed and whether the test standard is technically suitable to continue to be used. Thus, if a fire test report based on an ASTM or NFPA test standard is being presented for use within the relevant time frame, as discussed above, the continued applicability of the fire test report will normally not be questionable.

The discussion above does not address the technical issues raised in questions 2 and 3 above but provides a background for when a fire test report may be substantially out of date.

Listing and Labeling

A particular issue that must be considered, and will be discussed later, is whether the material, product or assembly being considered is listed and labeled by an accredited certification agency.

The International Building Code (IBC) defines as follows:

Labeled: "Equipment, materials or products to which has been affixed a label, seal, symbol or other identifying mark of a nationally recognized testing laboratory, inspection agency or other organization concerned with product evaluation that maintains periodic inspection of the production of the above-labeled items and whose labeling indicates either that the equipment, material or product meets identified standards or has been tested and found suitable for a specified purpose."

Listed: "Equipment, materials, products or services included in a list published by an organization acceptable to the code official and concerned with evaluation of products or services that maintains periodic inspection of production of listed equipment or materials or periodic evaluation of services and whose listing states either that the equipment, material, product or service meets identified standards or has been tested and found suitable for a specified purpose."

Materials, Products, or Assemblies

Any person or organization who wishes to use a fire test report conducted substantially earlier than the time at which it is intended to be used should be able to provide assurances that the material, product or assembly that was tested does not differ substantially from the material, product, or assembly currently being offered for use. That concept is easy and technically valid. However, the problem is in the details, particularly in the interpretation of the term "substantially" and in the issue of how these assurances can be made within a modern industrial environment.

The rationale behind this basic concept is the assumption that any product being offered for sale in the 21st century is likely to have undergone significant changes in formulation in the period preceding its sale. In the case of plastic

materials; for example, it is not uncommon for formulation changes to occur several times per year. In the period during which the material/product is being developed, and the formulation is being changed, fire testing is likely to have occurred, either in-house or in an outside test laboratory. Such changes in formulation are likely to be justifiable trade secrets or proprietary information and may well have no effect on the fire-test-response characteristics of the material. However, it is essential for someone to bear the responsibility for assuring those who are asked to authorize a new use of a product based on an old fire test report that it is safe to do so. Thus, the burden of proof must start, at least, with the person or organization wishing to reuse an old fire test report.

If the information requested above is requested from the manufacturer, the information should be available. However, even though it is technically reasonable for a manufacturer to be asked to provide a detailed description of any changes in the ingredients or in the manufacturing process, it is not commercially reasonable for the requester of such a description to require the manufacturer to reveal proprietary information. A compromise must be reached between the parties that is technically valid, and ensures safety and commercial viability.

Fire tests are not only conducted on materials but also on composite products or assemblies. In such cases the person or organization wishing to use the old fire test report may be responsible for selling the final product but may not be the manufacturer of any of the components. Consequently, such a person or organization may not be in a position to provide detailed information on the individual components. However, it is probably still the responsibility of the person or organization wishing to use an old fire test report to provide enough information to assure the expected new user of the fire test report that the product being offered for sale is still adequately representative of the product that was tested.

If a material, product, or assembly is listed and labeled by an accredited certification agency, the test results associated with the listing should continue to be applicable as long as the listing remains in place, provided the listing organization has suitable procedures to validate the continuation of the listing. In the case of listed and labeled items no report from the manufacturer should be needed, since certification agencies, including listing organizations, are typically responsible for setting schedules, based on the type of fire test and on the type of material, product, or assembly, on which manufacturers will be asked to affirm that they believe there have been no changes of sufficient magnitude to warrant new fire test reports. Certification organizations need to have systems in place to address the continuation of listing based on the impact of standard changes on existing listed materials, products, assemblies. It should be the responsibility of the certification agency to provide suitable procedures to ensure that the listing is still applicable to the material, product, or assembly being offered for use.

The procedures used by certification organizations to ensure that the listing is still applicable could provide guidance for ways in which the fire test report submitter can provide assurances without either conducting a new fire test or divulging commercial secrets. These include chemical analyses or screening tests acceptable to both parties.

Test Method Protocol

In order to determine whether a standard fire test protocol has substantially changed since a report was issued, it is important for fire test reports to include detailed descriptions of the protocol used in the testing recorded in the report. It is common for "recent" fire test reports to be required to contain sufficient information for this purpose. However, that has not always been the case and it is not unusual for older fire test reports to contain relatively little descriptive information.

In particular, fire test reports should be reviewed for inclusion of the specific date of the edition of the fire test standard used for any assessment. The information on the date of the test standard used is necessary, but not sufficient, to help determine whether the fire test report being proposed for a new use is applicable to the intended use. A key reason that this information is not sufficient comes when fire test methods: (a) offer options, (b) mandate certain techniques (such as specimen preparation and/or mounting methods) that are changing or have changed over time or (c) develop newer, or varied, calculation methods.

In such cases, fire test reports should be reviewed for inclusion of sufficient information for identification of the exact test protocol used, particularly if the fire test standard includes options, whether mandatory or non-mandatory. An example of options contained in a test method would be the initial test heat flux used in heat release tests.

If a fire test standard includes specific specimen preparation or mounting methods for particular materials or products, the fire test report should be reviewed for inclusion of a detailed description of the specimen preparation and mounting methods. This is very important since specimen preparation and mounting methods can vary for many diverse test methods; details should be given in the report irrespective of whether these methods are mandatory or are provided as guidance. In the cone calorimeter; for example, mounting methods are a function of the expected or known behavior of the material tested. For example, mounting methods for the cone calorimeter take into account whether the test specimen material melts, curls, or intumesces. As another example, specific standard practices were developed for mandatory ways to test certain materials or products in the Steiner tunnel (ASTM E84 [1]). The following standard practices are available in 2011: ASTM E2231 [3], forpipe and duct insulation, ASTM E2404 [4], for wall and ceiling coverings, ASTM E2573 [5], for site-fabricated stretch systems, ASTM E2579 [6], for wood interior finish, ASTM E2599 [7], for reflective insulation materials, radiant barriers, and vinyl stretch ceiling materials, ASTM E2688 [8], for tapes, and ASTM E2690 [9], for caulks and sealants. In addition to the mandatory mounting methods, other guidelines for mounting methods also exist in the Steiner tunnel, primarily in a non mandatory appendix. In standard fire resistance tests, options include conducting a hose stream test, among others.

If a fire test contains varied, or optional, calculation methods or techniques, this information also needs to be included in the fire test report. This is usually a direct consequence of the use of the test protocol associated with a certain edition of the standard. However, it is important to ensure the information is clearly stated.

Fire test reports should be reviewed for inclusion of a description of any deviations between the test performed for the report and the published fire test standard. It is not unusual for test reports to be conducted with "slight" (or not so slight) variations from the standard test protocol. Such deviations may have been made to accommodate a certain material or a certain specifier or authority having jurisdiction. In that case it is essential that the information be known when the test report is intended to be reused for a different purpose.

Fire test reports should also be reviewed for inclusion of detailed descriptions of the test specimen, including such information as dimensions, density, thickness, color, and layers, as appropriate.

Fire test reports should also be reviewed for inclusion of particular observations of phenomena that have occurred during the test, including melting and dripping or burning away from the ignition source. This is an important consideration when the material, product or assembly is used under different conditions than it was intended to be used when the fire test report was originally produced. For example, if a material was tested with the original intention of using it behind a fire resistance-rated thermal barrier and the same test report is now being presented with the intention of using it exposed in a habitable environment, the information above could be an important decision tool.

As stated above, a code-writing body or a regulatory agency will almost invariably include the applicable edition date when it adopts a particular fire test standard. The same careful attention to fire test edition date is not necessarily used by individual specifiers.

If the fire test standard edition or test protocol referenced by the applicable regulatory document to which the material, product or assembly is to be evaluated is different from, and typically newer than, the edition or protocol used for the fire test report under consideration, the fire test report should be accompanied by a description of the differences and their implications to the applicability of the fire test report results. This may involve the provision of additional information to that present in the original report when it is prepared for reuse, if the referencing document has changed its references in the interim period.

It may be possible for the fire test laboratory which has issued the fire test report to be able to state, if requested, whether the fire test method used for generating the fire test report would have generated data that would still comply with the requirements, within the accuracy and precision of the fire test method, at the time of the enquiry. It may also be possible for the fire test laboratory which has issued the fire test report to provide guidelines or information that would indicate whether changes in test methods of the same designation have not resulted in substantially altering the anticipated fire test results should a similar test be conducted using the latest version of the fire test method.

A fire test laboratory should not necessarily be required to produce the above information and analysis for older fire test reports. An agreement must be reached between the parties that is technically valid and ensures safety and commercial viability.

As stated already, the adoption of a new edition of a specific code is usually accompanied by a change in the applicable edition of all fire test standards

included in that code. A fire test report based on the edition of the fire test standard referenced in the code should be considered to be suitable for reuse unless a particular issue arises that would indicate that further studies are needed. Examples of particular issues that might require reconsideration have been discussed above. A fire test report based on an edition of the fire test standard earlier than that cited in the regulatory document should be deemed to continue to be applicable only if it can be judged that a test to be conducted on the new edition is likely to result in fire-test-response characteristics that would still comply with the requirements.

In cases where a test standard has been slightly revised, one example of an agreement between the manufacturer and the fire test laboratory could include having the applicability of the fire test reports extended based on the indications of screening tests mutually agreed upon. This would then result in the issuance of a new or revised report that would indicate the laboratory's technical opinion on the extended applicability.

Individual fire test reports, or series of test reports, are often used to establish a product classification. Such fire tests may have been conducted to establish the performance of a product, or a family of products, that include a range of formulations. These reports, even if they become dated, should be considered a valid measure of the product, or family of products, as long as the product offered for use by the manufacturer continues to fall within the range that was established. Further testing, engineering analysis and/or chemical analysis may be used to extend or modify the ranges initially established, with the mutual agreement of the sponsor and the fire test laboratory.

The issue of test cost is often a key consideration. Even though this is not a technical issue it is a valid concern and a resolution must be reached as an agreement between the interested parties. Fire testing is an issue of public safety and it is not appropriate to determine that even if an old fire test report is no longer applicable; as a result of the analysis above, a new fire test will not be conducted because it is too costly. It is likely that negotiations between the person or organization wishing to use the old fire test report, the fire test laboratory that conducted the test and the requester of the fire test report will lead to an agreeable compromise.

It is essential to point out that any fire test report for which a concern has been raised about one or more of the issues discussed above would not necessarily have lost its applicability.

Additional Guidance

In the absence of either an analysis regarding the test protocol or regarding the composition of the material, product or assembly, it should not be automatically assumed that fire test reports for materials, products or assemblies continue to be applicable at the time they are being resubmitted.

Fire test reports referencing out-of-date editions of fire test standards should be considered by the user to continue to be applicable if there have been no significant changes either in the fire test standard or in the material, product or assembly tested since the test was conducted.

Test reports on fire tests conducted on materials, products or assemblies intended for use as tested should be presumed to remain applicable as long as the paragraph above continues to apply.

The continued applicability of fire test reports for reaction-to-fire tests conducted on component samples intended for use as part of a composite system should be considered to be a function of the continued presence and effect of the individual component on the fire-test-response characteristics of the system actually to be used.

Authorities having jurisdiction and code-writing bodies should be the appropriate groups to provide an interpretation of what predicted degree of change, or fractional change, in measured product performance would be considered significant for safety and should be used as a threshold for issuance of a new fire test report. That threshold can be defined as a one-sided criterion, whereby a change in the direction of better fire performance does not trigger the need for a new fire test report.

It is important to assume that manufacturers should somehow remain responsible for communicating to other parties when changes in the material, product, or assembly they offer for use are sufficiently large as to create a predicted change in product performance greater than a defined threshold, if such a threshold exists. It should probably also be the responsibility of the manufacturer to offer assurances to the user of fire test reports that the material, product, or assembly used for the original fire test remains substantially unchanged.

Proposed ASTM Guide

The ASTM committee on fire standards, Committee E05, has a subcommittee entitled Terminology and Services/Functions and designated E05.31. This subcommittee has been working for a few years to develop a guide to be entitled "Standard Guide for Assessment of Continued Applicability of Fire Test Reports." As of the date of this manuscript this draft standard has not been balloted beyond subcommittee level, where it received persuasive negative votes.

Summary

It is the responsibility of a new user of an existing fire test report to ensure its continued applicability. However, any fire test report will continue to be applicable to a material, product or assembly for long after the fire test was conducted unless: (a) either the test protocol has changed substantially or (b) the material, product or assembly offered for use has changed substantially, and a test conducted at the time the fire test report is needed again on the material, product or assembly offered for sale would result in a more unsafe fire test result.

References

[1] ASTM E84, 2011, "Standard Test Method for Surface Burning Characteristics of Building Materials," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[2] ASTM E119, 2011, "Standard Test Methods for Fire Tests of Building Construction and Materials," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[3] ASTM E2231, 2009, "Standard Practice for Specimen Preparation and Mounting of Pipe and Duct Insulation Materials to Assess Surface Burning Characteristics," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[4] ASTM E2404, 2010, "Standard Practice for Specimen Preparation and Mounting of Textile, Paper or Polymeric (Including Vinyl) Wall or Ceiling Coverings to Assess Surface Burning Characteristics," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[5] ASTM E2573, 2007, "Standard Practice for Specimen Preparation and Mounting of Site-Fabricated Stretch Systems to Assess Surface Burning Characteristics," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[6] ASTM E2579, 2007, "Standard Practice for Specimen Preparation and Mounting of Wood Products to Assess Surface Burning Characteristics," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[7] ASTM E2599, 2011, "Standard Practice for Specimen Preparation and Mounting of Reflective Insulation Materials and Radiant Barrier Materials for Building Applications to Assess Surface Burning Characteristics," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[8] ASTM E2688, 2010, "Standard Practice for Specimen Preparation and Mounting of Tapes to Assess Surface Burning Characteristics," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

[9] ASTM E2690, 2010, "Standard Practice for Specimen Preparation and Mounting of Caulks and Sealants to Assess Surface Burning Characteristics," *Annual Book of ASTM Standards*, Vol. 04.07, ASTM International, West Conshohocken, PA.

www.astm.org

ISBN: 978-0-8031-7527-3
Stock #: STP1541