THE APT MODEL FOR ASSESSMENT OF PROSPECTIVE SHINGLE PERFORMANCE TEST METHODS

AARON R. PHILLIPS, RELLA JEFFRIES, W. KENT BLANCHARD, STANLEY P. FRANKOSKI, and HELENE HARDY PIERCE
TAMKO Roofing Products, Inc.
Joplin, Missouri

Heightened interest in performance of asphalt shingles in place on residential and commercial structures has led to numerous proposals of test methods intended to predict the performance of asphalt shingles. In the absence of a common definition of performance, many difficulties have been experienced in establishing and agreeing on the viability of individual methods as predictors of shingle performance.

This paper incorporates the concepts of logic, statistics, science, and engineering to provide a fresh perspective on the subject of asphalt shingle performance. Building on the foundation of a clear definition of performance, the Assessment of Performance Tests (APT) model is presented comprising two portions: 1) a framework that identifies the stages through which a shingle passes from manufacturing to in-service use and 2) five criteria for screening potential performance tests. Finally, three test methods are assessed using the APT model.

KEYWORDS

APT model, asphalt shingles, criteria, customer, model, performance tests, shingle performance.

INTRODUCTION

Beginning in 1990,[1] questions about field performance of asphalt shingles arose from problems experienced mainly in the states of Arizona and California,[2] a market area served by a few manufacturers. The reported problems, associated primarily with commodity-grade three-tab fiberglass-reinforced shingles, were described as cracking and splitting, wherein shingles developed horizontal, vertical, or diagonal separations during in-service life on the roof.

These initial splitting reports focused attention on the performance of asphalt shingles. When roofing contractors faced an increasing number of splitting and cracking complaints from homeowners, shingle performance became an issue. This economic stimulus[3] precipitated a search for test methods that would predict asphalt shingle performance. Because the initial problem reports in Arizona and California generally involved very few manufacturers, no immediate response materialized from the shingle manufacturing industry. This exacerbated the situation as contractors and consultants began a hasty search for tests believed to be predictive of shingle performance.

In 1992, the Asphalt Roofing Manufacturers Association (ARMA) formed the shingle performance task force.[4] The task force immediately initiated a survey of ARMA members concerning complaint rates of fiberglass-reinforced shingles. Results from the survey provided data for estimating a complaint rate for the industry during the ten-year period from 1981 to 1990. This estimated complaint rate is perhaps the most relevant measurement of shingle performance presently available. Survey results indicated less than 0.1 percent of fiberglass-reinforced shingles shipped during the ten-year period from 1981 to 1990 generated any type of complaint to the manufacturer. Cracking and splitting complaints were estimated to have occurred on 0.03 percent of the fiberglass-reinforced shingles shipped during the same ten-year period.[5]

In spite of the real-world performance record indicated by the survey, the outcry for "performance" test methods continued. Ultimately, the ARMA shingle performance task force recommended several tests thought to be meaningfully related to shingles and their performance.[6] Incorporation of these tests into a nationally recognized fiberglass-reinforced shingle specification, American Society for Testing and Materials (ASTM) D 3462, was initiated and continues to date. Discussions at ASTM D 3462 task group meetings, however, are hampered by lack of a clear definition of "performance."

This paper proposes a concise definition of performance, from which a model for assessing prospective shingle performance tests is derived. The APT model includes criteria for assessing prospective performance tests and a framework within which those methods can be properly discussed. The paper concludes by evaluating presently accepted and proposed test methods, using the model to determine whether the methods can be properly identified as performance tests.

SHINGLE PERFORMANCE DEFINED

Clear identification of the customer is essential when defining performance. The wide variety of shingle customers, including distributors, designers, roofing contractors, builders, do-it-yourself applicators, and individual building owners, complicates the development of a shingle performance definition. Each customer group has its own idea about shingle performance.

Roofing distributors are concerned with neat and effective packaging, competitive pricing, reasonable transportation costs, and minimal manufacturing defects that can generate complaints from their customers. The roofing contractor shares some of these interests but is also concerned with handling, application characteristics, and short- and long-term durability. The issue of durability is crucial for protecting the contractor's professional reputation. The building owner is interested in appearance, color, overall cost of the installation, prevention of water penetration into the building, and durability. The customer's frame of reference is extremely important when discussing shingle performance and must be clearly expressed.

The Merriam-Webster dictionary[7] includes six definitions of the word performance, with the most pertinent for this discussion being "fulfillment of a claim, promise, or request." O'Halloran echoing this definition, states that the product must "satisfy the requirements of the application for which it is intended."[8] A common definition of performance must be agreed upon and understood before performance test methods can be established.

Extending the concepts discussed previously, the authors define performance as **fulfillment of the expectations of the customer**. It is the customer who establishes the intended end-use. Therefore, the specific customer to whom a test method is relevant must be identified to facilitate clear communication about the method's relationship to performance.

THE NEED FOR A MODEL TO ASSESS PROSPECTIVE PERFORMANCE TESTS

It is critical to understand the distinction between tests that are used for classifying or for assessing product attributes and those that are predictive of performance. When a test is inappropriately considered a "performance" test, customers can be misled in product selection. Suppose, for example, that a customer desires shingles that keep water out of the building. Further assume that the industry commonly reports weight and length as the relevant product information. What effect do weight and length have on water shedding capability? Could shingles of more or less weight or different lengths prevent water infiltration? No one questions that measurements of weight and length are needed for maintaining product uniformity and will continue to be used for that purpose. However, weight and length probably do not predict performance (fulfillment of the customer's expectation to keep water out of the building) and, therefore, should not be inappropriately or unknowingly categorized as performance tests.

Performance tests constitute a special class of test methods that must satisfy some criteria beyond opinion. Tests not classified as performance tests should not be referred to as such, with the connotation attached to that word.

THE APT MODEL

The APT model for evaluating shingle performance tests, derived from the customer-oriented definition of performance, is illustrated in Figure 1. It is composed of two major elements: 1) the supply chain and 2) five criteria for evaluating potential performance tests.

The Supply Chain

Clearing up the confusion surrounding performance testing begins with establishment of a framework that identifies several stages through which a shingle passes from manufacturing through application and in-service use. Figure 2 is a graphical representation of the supply chain describing the four major stages through which shingles pass. Examples of customer expectations at each stage are included.

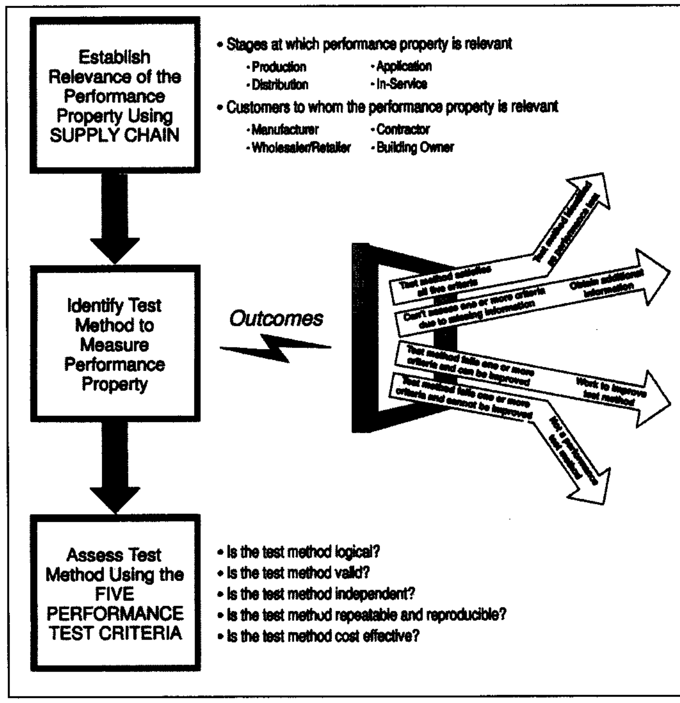

Figure 1. The APT model for assessment of prospective shingle performance test methods.

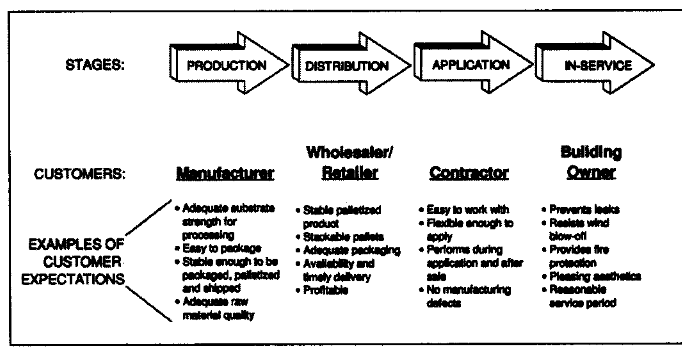

Figure 2. The supply chain.

Stage 1, production, encompasses conversion of raw materials into an asphalt shingle. The primary customer during this stage is the roofing manufacturer. Stage 2 is referred to as distribution. During this stage, shingles are delivered from the manufacturing facility to the distributor and retail sales outlets in the geographical area where the product will be sold. Included in this stage is the period of time during which the product resides in the distributor's warehouse. The primary customer is the roofing wholesaler/retailer. Stage 3 encompasses application of the product in a roof system, with the roofing contractor as the primary customer. The in-service life of the shingle, from completion of the installation until replacement of the roof, comprises the fourth stage. The primary customer is the building owner.

The supply chain illustrated in Figure 2 is a crucial element of the APT model. Test methods being evaluated as performance tests must be associated with one or more stages in the supply chain to provide an accurate frame of reference. Discussion of methods without this frame of reference leads to confusing and conflicting ideas. Association with a specific stage in the supply chain implicitly defines the customer group whose expectations must be satisfied for the shingle to perform.

Two examples are provided to illustrate the utility of the supply chain concept. First, consider the distribution stage, during which the wholesaler/retailer is the primary customer. Performance of individual shingles has limited meaning during this stage because the shingles are rarely removed from the package. They are usually stored in bundles and are subjected to transportation mainly via fork lift. Performance of the palletized load of shingles is of more concern than performance of the individual shingles.

As a second example, consider the final two stages. During the application stage, individual shingles are removed from the package and applied to the building. They must resist the rigors of handling. After application, they must endure a wide variety of weather conditions to prevent water infiltration into the building. During these two stages, performance of individual shingles becomes most critical.

Criteria for Evaluating Prospective Performance Tests

Five interrelated criteria for assessing prospective shingle performance tests comprise the second element of the APT model. Criteria originally developed by Christensen[9] have been adopted and extended by the authors. Tests that satisfy all five criteria can be properly identified as shingle performance tests. Failure to satisfy one or more criteria does not mean the method under consideration is useless; rather, it simply does not qualify as a performance test. The five criteria for assessing prospective performance tests are: logic, validity, independence, repeatability and reproducibility (R&R), and cost-effectiveness.

Logic—The first criterion is logic. The test method must measure the property under consideration, and that property must clearly relate to an end use defined by the customer. Christensen[9] takes the position that a performance test "must preferably be of a nature which makes it immediately evident that there is a reasonable link to the end use conditions." The test must be obviously relevant even to the casual observer. If the method does not display a clear link to an end-use requirement of the customer and cannot be easily explained to the layman the test does not satisfy the criterion of logic.

Validity—Validity is closely related to the criterion of logic. Logic addresses the issue of relationship between results of a test and a specific end-use requirement of the product. Validity involves verification of this perceived relationship. A test method fully satisfies the criterion of validity when a statistical correlation is demonstrated between results from the test and the end-use performance characteristic of interest.

Unfortunately, it is not always possible to demonstrate strict statistical correlation. In the absence of statistical correlation, one is forced to rely on the inferred relationship described in the section on logic. In instances where correlation has not been established, validity of the test method is questionable.

Independence—The third criterion is independence. The method should be free from the influence of the material or construction being tested. If an asphalt shingle test is related to performance, it should be relevant, as well, for other sloped roofing materials that perform the same function. Tests that are applicable regardless of product construction and raw material components satisfy the criterion of independence.

As in the case of the criteria of logic and validity, independence focuses on the end-use defined by the customer. Performance tests should apply to all materials that fulfill the same function. In the case of roofing, asphalt shingles, tile, wood shakes, and roll roofing are all utilized on sloped roofs. Strict application of the criterion of independence can be difficult, however, because of the different nature of the product classes used to fulfill a given end-use.[9]

Repeatability and Reproducibility—Wheeler states that a "proper understanding of how much uncertainty is attached to any single measurement is crucial to a proper use of measurements. Without such an understanding, one will be misled by the apparent objectivity of the measurements and, in consequence, may make incorrect decisions."[10] This "proper understanding" is developed from consideration of two statistical issues.

First, the test method must be repeatable. Repeatability means that test results obtained with the same test method in the same laboratory by the same operator using the same equipment are statistically similar, within an acceptable limit of variation.[11] This is a key measure of the stability of the test method itself. If the method produces statistically different results under these conditions, it fails the test of repeatability.

Second, the method must demonstrate reproducibility. Reproducibility means the test produces similar data when conducted on the same material at different laboratories by different operators using different equipment.[11] If a test method produces statistically different results on the same material tested in different laboratories by different operators using different equipment, it fails the test of reproducibility.

Discrimination ratio is one method available for determining acceptable repeatability and reproducibility (see Appendix 1). It compares total variation with R&R variation and quantifies the number of product categories among which a given measurement system can differentiate. For example, a discrimination ratio of 2.0 indicates the measurement system is able to distinguish two statistically different product levels. In general, the more levels that can be distinguished, the better the measurement system. "When the Discrimination Ratio shows that a particular measurement cannot detect product variation it will be ... best to work on improving the measurement process. This latter course of action is certainly recommended when the Discrimination Ratio gets down in the region of 1.0 to 2.0."[10]

Cost-effectiveness—Assessment of economic benefit to users of a test method is crucial. If the value of the test results to the customer exceeds the cost of performing the testing, the method is considered to be cost-effective. The importance of cost-effectiveness is emphasized by ASTM in their guidelines for standards development:[12]

> "When a standard is being developed, the costs associated with its development and subsequent use generally should be considered. The prime objective should be the optimum use of resources to achieve satisfactory definition of the product or service. Some standards ... that include numerous and extensive requirements, can entail significant expense to users of the standard. The requirements to be included should, therefore, be those that are technically relevant and yield benefits commensurate with the cost of their determination."

In summary, the APT model incorporates two elements:

1) the supply chain and 2) five criteria for assessing prospective test methods intended to measure a particular performance property. The criteria should be evaluated within the context of the supply chain, which provides a clear picture of the stage of relevancy and defines the customer to whom the test method is relevant. Proposed tests that satisfy the five criteria can be properly identified as shingle performance tests. In the sections that follow, the utility of the APT model is demonstrated by applying it to three prospective shingle performance test methods.

APPLYING THE APT MODEL: TEAR RESISTANCE

Most users of shingles agree that tear resistance is an important performance property, and the authors concur with this position. In the sections that follow, the APT model is utilized to assess tear resistance, and specifically the Elmendorf tear test as a prospective shingle performance test.

Where is Tear Resistance Relevant?

Using the supply chain illustrated in Figure 2, application of the model begins by identifying those stages during which the property of tear resistance is relevant. During the production stage, tear resistance of the substrate is more important than tear resistance of the finished shingle. Tear resistance also has little meaning during the distribution stage, because the shingle rarely leaves the package.

During application, however, shingles are subjected to forces that may generate tears in the material. Tear resistance has obvious importance at this stage. It also has applicability during the in-service stage for the period of time prior to shingle sealing. Prior to sealing, shingle tabs can be lifted and torn by wind forces. After the shingles seal, tear resistance is of limited importance.

Evaluating The Elmendorf Tear Strength Test Using the Performance Test Criteria

D 3462, one of two ASTM fiberglass-reinforced shingle specifications, includes a tear resistance requirement as determined by the Elmendorf tear tester. In 1992, this same tear test was proposed for inclusion in D 225, the ASTM specification for organic-reinforced shingles. In both contexts, the Elmendorf tear test is frequently described as a "shingle performance test."

Logic—Conformance of the Elmendorf tear test to the criterion of logic is questioned on the basis of two arguments. First, the Elmendorf tear test subjects shingle samples to nonplanar forces exclusively,[13] with no capability to assess planar tearing forces. During application, shingles are removed from the bundle by grasping one end and pulling, an action that concentrates the majority of the forces in the plane of the shingle rather than perpendicular to the plane of the shingle. After removal from the bundle, shingles must resist both planar and nonplanar forces prior to being fastened to the deck. The Elmendorf tear test addresses only one component of these forces.

Second, the Elmendorf tear test measures the force required to continue tearing rather than the force required to initiate tearing.[14] Most shingles do not have a preexisting tear from which tearing can continue during handling or high winds. It may be argued that three-tab shingle cutouts provide preexisting tears. Careful comparison of the shape of cutouts with the slit created during the Elmendorf tear test reveals these to be dramatically different. The Elmendorf slit concentrates forces at a point from which tearing continues. The continuous arc of most cutouts does not provide this point for concentration of tearing forces.

The authors conclude that the Elmendorf tear test departs from most end-use conditions present during application and service and, thus, only partially satisfies the criterion of logic.

Validity—It is natural to expect shingles with higher Elmendorf tear strength values to "perform" better during application and during high winds prior to sealing, the time periods during which tear resistance is relevant. No work has been completed to correlate Elmendorf tear test results to actual experience with shingles during the application and in-service stages. Some questions that need to be answered to establish validity of the Elmendorf tear test include the following: Do shingles with higher Elmendorf tear strengths have less tendency to tear during handling and application and less tendency to tear during high winds after application? Conversely, do shingles with lower Elmendorf tear results tend to tear more frequently when subjected to these conditions? Until this work is completed, validity of the Elmendorf tear test is questionable. The relationship between the test and end use conditions described in the section on logic provides the only basis for validity. Because logic of the Elmendorf test is not satisfied in all cases, validity is also suspect.

Independence—Ideally, performance tests should be applicable to all materials that serve similar functions, such as wood shakes, clay and concrete tiles, and slate roof coverings. The different nature of these materials makes application of a single tear resistance method to all materials difficult, if not impossible. Application of the Elmendorf tear test to materials such as shakes and tiles is clearly not plausible, suggesting that the Elmendorf tear test does not satisfy a strict interpretation of the criterion of independence. It does, however, apply to all asphaltic shingles presently available in the marketplace. The Elmendorf tear test satisfies independence within the context of nonrigid roofing materials but not within the broader context of all sloped roof coverings.

Repeatability and Reproducibility—Repeatability and reproducibility of the Elmendorf tear strength method has been debated vigorously in recent years. Using data from an ARMA Round Robin study (see Table 1), a discrimination ratio of 1.4 was calculated (Appendix II), indicating less than two product categories can be discerned.[15] Shingles included in the ARMA study were selected to represent **three** product categories—high (Material H), medium (Material M), and low (Material L) tear strength. Assignment of the three products to these categories was based on results of prior Elmendorf tear testing conducted for the Western States Roofing Contractors Association (WSRCA).[2] The large variability in the repeatability and reproducibility of the measurement system prevented product distinction.

A soon-to-be-published study identifies within-shingle variability, rather than equipment and/or operator variability, as the major source of variability in repeatability of Elmendorf tear test results of fiberglass-reinforced shingles.[16] The study refutes the opinion that the Elmendorf tester is not repeatable. The instrument is adequate when applied to relatively homogeneous materials, such as paper, for which it was designed.[17] Application of the instrument to complex composites, such as shingles, produces R&R numbers that are so large, meaningful decisions cannot be made from the results. The high variation of Elmendorf test results on fiberglass-reinforced shingles makes it extremely difficult to distinguish among inferior, acceptable, and superior products. Users of the method risk identifying "superior" shingles as "inferior," or vice versa.

Cost-effectiveness—Cost of applying a method must be carefully weighed against the benefit received by the user. This is an area in which the Elmendorf tear test has an advantage when conducted with sets of ten specimens, as prescribed by existing standards. Unfortunately, the large number of specimens required to produce repeatable results increases the cost of the test greatly. Even with fifty specimens, however, the cost of the test remains reasonable.

Summary of Assessment

Tear resistance is clearly an important functional requirement for asphalt shingles, being especially relevant during the application stage and the in-service stage prior to sealing. The Elmendorf tear test, however, is a poor choice to evaluate this property. Although it satisfies the criterion of cost-effectiveness, serious limitations in the areas of logic, validity, and repeatability and reproducibility exist (Figure 3). Because of these limitations, the Elmendorf tear test should not be classified as a shingle performance test. Alternative means of assessing the property of shingle tear resistance should be investigated.

Table 1. Elmendorf Round Robin Data (grams)

LAB	MATERIAL H	MATERIAL M	MATERIAL L
1	1522 1562	1794 1992	1666 1510
2	1312 1360	1811 1790	1798 1516
3	1739 1739	1664 1843	1791 1739
4	1560 1517	1645 1728	1578 1477
5	1501 1456	1846 1472	1501 1344
6	1555 1577	1885 1785	1550 1495
7	1306 1315	1811 1683	1392 1622
8	1677 1542	1862 1642	1651 1578
9	1549 1530	1978 2125	1619 1645
10	1558 1373	1571 1747	1853 1539
11	1504 1683	1293 1402	1302 1395
12	1472 1475	1882 1798	1990 1664
13	1644 1456	1748 1713	1734 1392

APPLYING THE APT MODEL: ABILITY TO RESPOND TO IMPOSED LOADS

Shingles must have the ability to respond to loads imposed in the plane of the shingles. This property is widely identified as relevant to shingle performance.

Where is Ability to Respond to Imposed Loads Relevant?

Referring once again to the supply chain illustrated in Figure 2, evaluation of the ability to respond to imposed loads begins by identifying the stages during which, and the customers to whom, this property is relevant. During manufacture, sufficient strength must be present in the substrate (fiberglass mat, organic felt) to permit processing without breaking. This requirement is related to performance of the substrate, rather than performance of the finished shingles. Ability to respond to imposed loads is of minor importance during the distribution stage, because shingles are rarely removed from the package.

Ability to respond to imposed loads achieves greater importance as a shingle performance property during the application stage. As discussed previously, shingles are subjected to in-plane and out-of-plane forces during removal from the package and application on the roof. If shingles do not have the ability to withstand these forces, the roofing contractor's expectations may not be fulfilled.

Ability to respond to imposed loads is also important during the service life of the shingle. After application, thermally induced expansion and contraction of the deck and shingles can occur. The shingles must be able to accommodate these imposed loads to perform during service on the roof.

Evaluating The Proposed ASTM D 3462 Tensile Strength Test Using the Performance Test Criteria

Recent ASTM D 8.02 subcommittee ballots have included a proposed tensile strength test for fiberglass-reinforced shingles.[18] The method being considered for inclusion in D 3462 is conducted at a rate of extension of 51 mm/minute (2 inches/minute) with a gauge length of 76 mm (3 inches) and a 25- by 127-mm (1- by 5-inch) shingle specimen. The express purpose for incorporating this method within D 3462 is to include an additional "performance" test in the standard. Shingles respond to tensile forces by resisting them (tensile strength) and accommodating the deformations associated with them (elongation). This test method addresses one aspect of a shingle's ability to respond to imposed loads. At present, it does not incorporate determination of the elongation of the material.

Logic—Logical arguments supporting the proposed tensile strength test are strong. A difficulty with logic of the test method, however, is the fact that it includes no measurement of elongation. Failure of the method to include both aspects of a shingle's response to imposed loads is a limitation in logic. The proposed tensile test partially satisfies the criterion of logic.

Validity—Experience suggests that shingles with higher tensile strength should resist loads imposed in the plane of the shingle better than those with lower tensile strength. In order to validate the proposed tensile strength test, work to establish a correlation between shingle tensile strengths measured using the proposed method and rooftop performance must be completed.

One area that should be carefully considered in any effort to validate the method is matching of the method's rate of extension (crosshead speed) with real-world strain rates. The authors have been unable to locate published research that quantifies real-world strain rates experienced by shingles during the application and in-service stages. Further study in this area is necessary.

Because statistical correlation has not been established between measurements using the proposed method and real-world experience, validity of the tensile test has not been demonstrated. Once again, the inferred relationship discussed in the section on logic provides the only basis for validity.

Independence—All roof coverings must be able to respond to imposed loads during the application and in-service stages. The proposed tensile strength test can be easily applied to a wide variety of roofing materials and, therefore, satisfies the criterion of independence.

Repeatability and Reproducibility—A repeatability and reproducibility study of the proposed ASTM tensile strength test has been completed by ARMA for the ASTM task group addressing this method.[19] Machine-direction (MD) and cross-direction (CD) data from this work are summarized in Tables 2 and 3. Calculation of the discrimination ratios (Appendices III and IV) reveals that the tensile measurement system is capable of discriminating up to four product levels successfully for samples oriented in both machine-direction and cross-direction. In the study, three levels of product were anticipated. The tensile measurement system is adequate in repeatability and reproducibility.

Cost-effectiveness—Tensile strength measurements are relatively inexpensive to conduct, making the proposed method very cost-effective.

Summary of Assessment

Ability to respond to imposed loads is an important functional requirement for shingles. Resistance or accommodation of forces imposed during the application and in-service stages is imperative for the product to perform satisfactorily. If the shingle falls apart when it is being handled, the contractor's expectations are not fulfilled.

The proposed ASTM tensile strength test satisfies the criteria of independence, repeatability and reproducibility, and cost-effectiveness (Figure 3). Additional work is necessary to establish validity of the method. Its failure to consider the other means by which shingles can accommodate imposed loadings and associated deformations through elongation is a logic limitation. Because of limitations in logic and validity, the proposed ASTM tensile test does not satisfy all criteria for identification as a shingle performance test.

APPLYING THE APT MODEL: FIRE RESISTANCE

At times, shingles are required to resist the effects of fire. Although not a daily performance requirement, ability to resist fire is a critical characteristic when fire occurs. Of primary interest is the ability of the roofing material to prevent the structure from catching fire from external sources of ignition.

Where is Fire Resistance Relevant?

Fire resistance is a relevant shingle performance requirement during the in-service stage. The customer for this functional characteristic is the building owner.

Evaluating UL 790 Tests for Fire Resistance of Roof Covering Materials Using the Performance Test Criteria

The Underwriters Laboratories (UL) fire test[20] has been used to evaluate roofing materials since 1903.[21] It was developed by Underwriters Laboratories to provide "standardized evaluations of roof coverings in terms of their ability to withstand ignition, fire spread, and fire penetration from exterior fires."[22] The method incorporates three fire resistance tests regularly applied to asphalt shingles. These tests, designed to simulate various external fire hazards, include the intermittent-flame, spread-of-flame, and burning-brand tests. Essentially identical versions of these procedures are incorporated in ASTM E 108.[22]

Logic—Each segment of the fire test is designed to evaluate performance of the roof covering when subjected to a specific type of fire condition originating outside the building.[22] The intermittent-flame and spread-of-flame tests evaluate resistance to ignition and surface spread-of-flame. The burning-brand test determines resistance to fire penetration from the exterior through the roof deck. The relationship of these tests to fire ignition and propagation, the performance requirements of interest, are easily comprehended and explained. The fire test satisfies the criterion of logic.

Validity—Although no statistical correlation has been established between each element of the fire test and rooftop performance, it is known that organic substrates are not as fire-resistant as fiberglass substrates.[23, 24] UL 790 fire testing of shingles manufactured with these substrates consistently reveals better fire resistance for fiberglass-reinforced shingles. News reports of large multiple building fires have documented superior fire resistance of roof coverings generally classified by UL 790 as Class A.[25, 26, 27, 28] This evidence, combined with the logical basis for the method, makes a strong argument for validation of the fire test.

Independence—The UL fire test can be used to evaluate relative fire resistance of any roofing material. The criterion of independence is satisfied.

Repeatability and Reproducibility—Although the fire test has been used at UL for years, data on repeatability and reproducibility has not been developed. In the absence of data, no statement can be made about its conformance to the repeatability and reproducibility criterion.

Cost-effectiveness—Although conducting the fire test is relatively expensive, the importance of the fire resistance property to the end user provides sufficient justification.

Summary of Assessment

Importance of fire-resistant roof coverings during the in-service stage is emphasized when widespread fires occur. The UL 790 fire tests satisfy the criteria of logic, validity, independence, and cost-effectiveness (Figure 3). Work to establish the repeatability and reproducibility of this important shingle test method is strongly encouraged. Demonstration of acceptable repeatability and reproducibility would classify the UL 790 fire test as a shingle performance test method.

CONCLUSION

Performance is properly defined from the perspective of the customer. Identification of specific customers is essential for

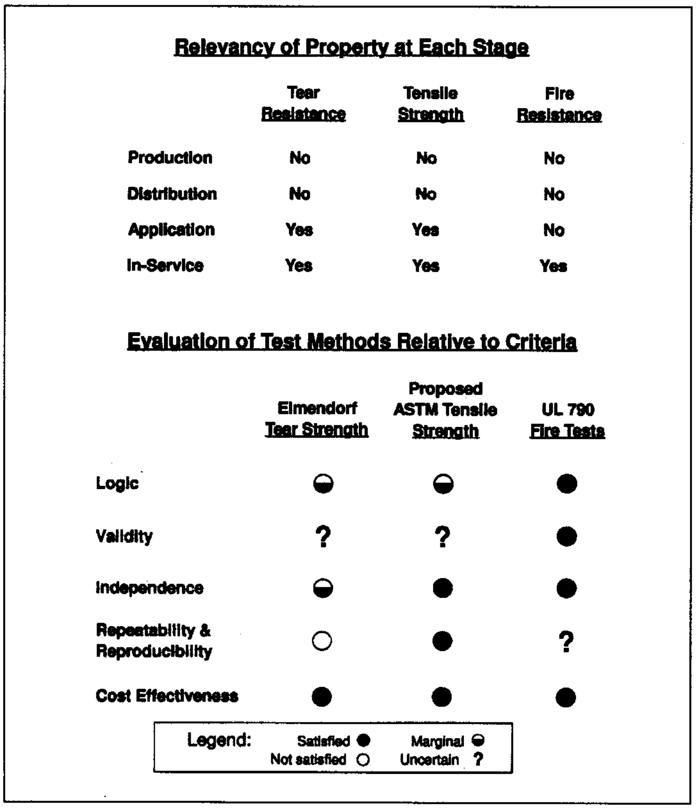

Figure 3. Application of the APT model.

clear communication of performance expectations. The APT model provides a useful approach for assessing prospective "performance" tests, built upon this customer-oriented definition of performance.

None of the shingle tests assessed using the APT model satisfy all the criteria necessary for classification as a performance test. The UL fire test fulfills most of the requirements. Demonstration that the fire test is repeatable and reproducible would allow it to satisfy all conditions of the APT model. The proposed ASTM tensile strength test also satisfies many of the requirements but needs additional work in the areas of logic and validity. The Elmendorf tear test is revealed by the model to be an inappropriate method for evaluating the important performance characteristic of tear resistance.

Those accountable for development and use of shingle tests must remember that the "expert in the subject matter holds the responsibility for the use of the data from a test."[29] Consequently, a responsible approach is necessary to ensure performance measurements are logical, valid, independent, repeatable and reproducible, and cost-effective. The APT model provides such an approach and is recommended as the basis for future discussions about shingle performance tests.

Table 2. MD Tensile Strength Round Robin Data lbf/in (kN/m)

LAB	MATERIAL A	MATERIAL B	MATERIAL C
1	86.17 (15.08) 82.48 (14.43) 85.03 (14.88)	47.30 (8.28) 51.50 (9.01) 49.58 (8.68)	90.10 (15.77) 81.56 (14.27) 87.39 (15.29)
2	112.90 (19.76) 111.90 (19.58) 103.10 (18.04)	60.40 (10.57) 62.70 (10.97) 53.50 (9.36)	100.00 (17.50) 114.40 (20.02) 108.00 (18.90)
3	93.40 (16.35) 91.70 (16.05) 91.30 (15.98)	48.50 (8.49) 44.70 (7.82) 50.20 (8.79)	96.90 (16.96) 92.80 (16.24) 91.60 (16.03)
4	99.60 (17.43) 94.47 (16.53) 100.35 (17.56)	54.61 (9.56) 53.33 (9.33) 54.19 (9.48)	97.91 (17.13) 104.71 (18.32) 109.64 (19.19)
5	105.80 (18.52) 102.40 (17.92) 100.00 (17.50)	54.00 (9.45) 58.50 (10.24) 55.10 (9.64)	106.00 (18.55) 102.60 (17.95) 105.80 (18.52)
6	116.40 (20.37) 118.80 (20.79) 106.10 (18.57)	63.01 (11.03) 62.29 (10.90) 59.49 (10.41)	107.00 (18.73) 106.90 (18.71) 108.60 (19.01)
7	98.44 (17.23) 99.47 (17.41) 92.18 (16.13)	53.91 (9.43) 57.80 (10.12) 55.88 (9.78)	105.09 (18.39) 97.57 (17.07) 106.51 (18.64)
8	106.91 (18.71) 114.88 (20.10) 105.04 (18.38)	56.97 (9.97) 60.63 (10.61) 62.64 (10.96)	107.84 (18.87) 108.55 (19.00) 110.69 (19.37)

U.S. Customary Units are shown first with SI Units in parentheses because the data was originally generated in U.S. Customary Units.

REFERENCES

1. "WSRCA to Reveal Results of Asphalt Shingle Study," *RSI*, May 1993, 38.
2. "The Asphalt Fiberglass Shingle Study," *Proceedings of Western States RCA 18th Annual Convention*, June 1992.
3. Ribble, R., R. Olson, D. Summers, and J. Goodman. "From Generation to Generation: Issues and Problems Facing the Steep-Slope Roofing Industry," *Proceedings of the 10th Conference on Roofing Technology*, 1993, 1-5.
4. Snyder, R.D. "ARMA Committed to Asphalt Shingle Performance Criteria," *RSI*, June 1992, 26-28.
5. Asphalt Roofing Manufacturers Association (ARMA), personal communication of information collected by the Asphalt Shingle Performance Task Force, April 1992.
6. "ARMA Recommends Upgraded Standard," *Professional Roofing*, July 1994, 36.
7. *Merriam-Webster's Tenth Collegiate Dictionary*, Merriam-Webster, Springfield, Massachusetts, 1996.
8. O'Halloran, M. R. "Experience with Performance Standards and Criteria for Wood-Based Structural Panels in Building Construction," *Proceedings of the 9th CIB Congress*, Vol. 4, 1983, 447-456.
9. Christensen, G. "The Performance Approach—Some Criteria for Working with Performance Testing Methods," *Proceedings of the 9th CIB Congress*, Vol. 4, 1983, 291-298.
10. Wheeler, D. J. and R. W. Lyday. *Evaluating the Measurement Process*, 2nd Edition, SPC Press, Knoxville, Tennessee, 1989.
11. "Standard Terminology Relating to Quality and Statistics," ASTM E 456-92, *Annual Book of ASTM Standards*, Vol. 14.02, 1995, 256-263.
12. *Form and Style for ASTM Standards*, 10th Edition, ASTM, West Conshohocken, Pennsylvania, 1996, 11.
13. Van Den Akker, J. A., W. A. Wink, and R. H. Van Eperen. "Instrumentation Studies LXXXIX. Tearing Strength of Paper. III. Tearing Resistance by the In-Plane Mode of Tear," *Tappi*, Vol. 50, No. 9, 1967, 466-470.

Table 3. CD Tensile Strength Round Robin Data lbf/in (kN/m)

LAB	MATERIAL A	MATERIAL B	MATERIAL C
1	65.22 (11.41) 66.84 (11.70) 60.00 (10.50)	38.35 (6.71) 41.01 (7.18) 39.50 (6.91)	51.63 (9.04) 55.76 (9.76) 55.78 (9.76)
2	72.40 (12.67) 73.00 (12.78) 74.90 (13.11)	44.30 (7.75) 45.40 (7.95) 49.10 (8.59)	58.70 (10.27) 58.00 (10.15) 67.20 (11.76)
3	68.80 (12.04) 61.40 (10.75) 67.10 (11.74)	37.40 (6.55) 39.00 (6.83) 38.00 (6.65)	55.40 (9.70) 53.80 (9.42) 59.50 (10.41)
4	68.93 (12.06) 72.12 (12.62) 68.42 (11.97)	42.34 (7.41) 43.84 (7.67) 44.84 (7.85)	57.12 (10.00) 58.83 (10.30) 58.44 (10.23)
5	65.30 (11.43) 76.30 (13.35) 77.10 (13.49)	44.00 (7.70) 41.40 (7.25) 40.10 (7.02)	56.90 (9.96) 61.90 (10.83) 57.50 (10.06)
6	73.01 (12.78) 80.09 (14.02) 79.32 (13.88)	45.73 (8.00) 43.72 (7.65) 48.39 (8.47)	67.54 (11.82) 67.05 (11.73) 69.15 (12.10)
7	73.40 (12.85) 68.23 (11.94) 63.90 (11.18)	41.33 (7.23) 39.98 (7.00) 40.56 (7.10)	53.93 (9.44) 62.55 (10.95) 62.47 (10.93)
8	78.91 (13.81) 72.19 (12.63) 79.09 (13.84)	48.02 (8.40) 42.80 (7.49) 46.64 (8.16)	65.97 (11.54) 59.13 (10.35) 65.69 (11.50)

U.S. Customary Units are shown first with SI Units in parentheses because the data was originally generated in U.S. Customary Units.

14. Cohen, W. E. and A. J. Watson. "The Measurement of Internal Tearing Resistance," *Proc. Australian Pulp Paper Ind. Tech. Assn.*, Vol. 3, 1949, 212-244.
15. Asphalt Roofing Manufacturers Association (ARMA), personal communication of data from round robin study of Elmendorf tear strength test, May 1993.
16. Phillips, A. R., R. Jeffries, J. Schneider, and S. P. Frankoski. *Quality Engineering*, in press.
17. Elmendorf, A. "The Principle of the Elmendorf Paper Tester," *Paper*, Vol. 28, No. 23, 1921, 15-17.
18. ASTM, personal communication, Subcommittee Ballot D 802000396, 1996.
19. Asphalt Roofing Manufacturers Association (ARMA), personal communication of data from round robin study of proposed ASTM tensile strength test, July 1995.
20. *UL 790: Standard for Tests for Fire Resistance of Roof Covering Materials*, Fifth Edition, Underwriters Laboratories, Chicago, Illinois, 1983.
21. *History of Underwriters Laboratories, Inc.*, Underwriters Laboratories, Chicago, Illinois.
22. "Standard Test Methods for Fire Tests of Roof Coverings," ASTM E 108-93, *Annual Book of ASTM Standards*, Vol. 04.07, 1995, 424-435.
23. *Encyclopedia of Polymer Science and Engineering*, Vol. 7, John Wiley & Sons, New York, 1987, 165.
24. Loewenstein, K. L. *The Manufacturing Technology of Continuous Glass Fibres*, Elsevier Scientific Publishing Company, Amsterdam, 1973.
25. Claiborne, W. "Southern California Blazes Send Thousands From Homes." *The Washington Post*, October 23, 1996.
26. Ingrassia, R. "Fire Destroys 4 Houses in North Dallas." *Dallas Morning News*, August 11, 1995.
27. "Are Wood Roofs Fire Resistant?" *RSI*, February 1991, 60-64.
28. Best, R. "Houston Conflagration Spreads Across Rooftops," *Fire Journal*, January 1980.
29. Deming, W. E. In forward to Shewhart, W. A., *Statistical Method from the Viewpoint of Quality Control*, Dover Publications, Mineola, New York, 1986.

APPENDIX I.

Discrimination Ratio Formula

$$\text{Discrimination Ratio} = \sqrt{\frac{2s^2_{TM}}{s^2_{R\&R}} - 1}$$

where s^2_{TM} = Variance of total measurements

and $s^2_{TM} = s^2_{Material} + s^2_{Lab} + s^2_{Error}$

and $s^2_{R\&R}$ = Variance of repeatability and reproducibility

APPENDIX II.

Calculation of Discrimination Ratio for Elmendorf Tear Strength

Nested Random Effects Analysis of Variance for Variable ELMENDORF TEAR STRENGTH

Variance Source	Degrees of Freedom	Sum of Squares	F Value	Pr > F	Error Term
TOTAL	77	2586263			
MATERIAL	2	732586	9.844	0.0004	LAB
LAB	36	1339499	2.822	0.0009	ERROR
ERROR	39	514180			

	Variance Source	Mean Square	Variance Component	Percent of Total
	TOTAL	33588	37853	100.0000
	MATERIAL	366293	12657	33.4373
R&R	LAB	37208	12012	31.7333
R&R	ERROR	13184	13184	34.8295

Mean	1619.68
Standard error of mean	68.53

Variance estimates used in calculating the Elmendorf tear strength discrimination ratio were determined by analysis of variance (ANOVA) of the ARMA round robin data shown in Table 1. Results of the ANOVA are shown above. The discrimination ratio is calculated using the formula in Appendix I, as follows:

s^2_{TM} = Variance of total measurements = 37853

$s^2_{R\&R}$ = Variance of lab + Variance of error = 12012 + 13184

$s^2_{R\&R}$ = Variance of R&R = 25196

$$\text{Discrimination Ratio} = \sqrt{\frac{2s^2_{TM}}{s^2_{R\&R}} - 1}$$

$$= \sqrt{\frac{2(37853)}{25196} - 1}$$

$$= 1.4$$

APPENDIX III.

Calculation of Discrimination Ratio for MD Tensile Strength

Nested Random Effects Analysis of Variance for Variable MD TENSILE STRENGTH

Variance Source	Degrees of Freedom	Sum of Squares	F Value	Pr > F	Error Term
TOTAL	71	38309			
MATERIAL	2	33798	92.681	0.0000	LAB
LAB	21	3829.027813	12.830	0.0000	ERROR
ERROR	48	682.153000			

	Variance Source	Mean Square	Variance Component	Percent of Total
	TOTAL	539.564849	766.778694	100.0000
	MATERIAL	16899	696.526128	90.8380
R&R	LAB	182.336658	56.041046	7.3085
R&R	ERROR	14.211521	14.211521	1.8534

Mean	86.08
Standard error of mean	15.32

Variance estimates used in calculating the MD tensile strength discrimination ratio were determined by analysis of variance (ANOVA) of the ARMA round robin data shown in Table 2. Results of the ANOVA are shown above. The discrimination ratio is calculated using the formula in Appendix I, as follows:

s^2_{TM} = Variance of total measurements = 766.778694

$s^2_{R\&R}$ = Variance of lab + Variance of error = 56.041046 + 14.211521

$s^2_{R\&R}$ = Variance of R&R = 70.252567

$$\text{Discrimination Ratio} = \sqrt{\frac{2s^2_{TM}}{s^2_{R\&R}} - 1}$$

$$= \sqrt{\frac{2(766.778694)}{70.252567} - 1}$$

$$= 4.6$$

APPENDIX IV.

Calculation of Discrimination Ratio for CD Tensile Strength

Nested Random Effects Analysis of Variance for Variable CD TENSILE STRENGTH

Variance Source	Degrees of Freedom	Sum of Squares	F Value	Pr > F	Error Term
TOTAL	71	11382			
MATERIAL	2	9792.007853	93.821	0.0000	LAB
LAB	21	1095.869608	5.064	0.0000	ERROR
ERROR	48	494.600733			

	Variance Source	Mean Square	Variance Component	Percent of Total
	TOTAL	160.316594	226.090029	100.0000
	MATERIAL	4896.003926	201.825819	89.2679
R&R	LAB	52.184267	13.960028	6.1745
R&R	ERROR	10.304182	10.304182	4.5576

Mean	57.94
Standard error of mean	6.25

Variance estimates used in calculating the CD tensile strength discrimination ratio were determined by analysis of variance (ANOVA) of the ARMA round robin data shown in Table 3. Results of the ANOVA are shown above. The discrimination ratio is calculated using the formula in Appendix I, as follows:

s^2_{TM} = Variance of total measurements = 226.090029

$s^2_{R\&R}$ = Variance of lab + Variance of error = 13.960028 + 10.304182

$s^2_{R\&R}$ = Variance of R&R = 24.26421

$$\text{Discrimination Ratio} = \sqrt{\frac{2s^2_{TM}}{s^2_{R\&R}} - 1}$$

$$= \sqrt{\frac{2(226.090029)}{24.26421} - 1}$$

$$= 4.2$$