Validity and Reliability of Forensic Engineering Methods and Processes

Joshua B. Kardon[1], Ph.D., S.E., Member, ASCE;
Robert G. Bea[2], Ph.D., Fellow, ASCE;
Robert Brady Williamson[3], Ph.D., P.E., Life Member, ASCE

[1]Principal Structural Engineer, Joshua B. Kardon + Company Structural Engineers, 1930 Shattuck Avenue, Suite C, Berkeley, California 94704; email: jbkse@jbkse.com
[2]Professor at Department of Civil and Environmental Engineering: Engineering & Project Management; University of California at Berkeley, Berkeley, California 94720
[3]Professor in the Graduate School, Department of Civil and Environmental Engineering: Structural Engineering, Mechanics and Materials; University of California at Berkeley, Berkeley, California 94720

Abstract

Under the Federal Rules of Evidence, and in accordance with case law, technical forensic evidence presented in court has to be valid and reliable—a judge may rule evidence inadmissable if it is shown to be invalid or unreliable. Engineering methods and processes operating in arenas other than litigation also must be valid and reliable in order that those methods and processes achieve their intended results. This paper discusses of types validity and reliability, and gives examples of expert witness evidence and of an engineering method that lacked validity and reliability.

Introduction

The Federal rule of evidence for expert testimony, Rule 702, is based on case law, such as *Daubert v. Merrill Dow Pharmaceuticals* (1993). It allows the judge to assess "whether the testimony's underlying reasoning or methodology is scientifically valid and properly can be applied to the facts at issue" (Federal, 2001). To determine that, the judge can consider whether the method "can be (and has been) tested, whether it has been subjected to peer review and publication, its known or potential error rate, the existence and maintenance of standards controlling its operation," as well as "whether it has attracted widespread acceptance within a relevant scientific community." The validity and reliability of a method generally have to do with the applicability of the method to the question asked, and the suitability of the method for the intended purpose.

Engineering methods and processes in areas other than forensics and litigation also must have validity and reliability. Engineering decisions that affect people's lives and livelihoods are made on the bases of methods and models that must be valid and reliable, or they will not result in acceptable levels of safety, durability, compatibility or serviceability.

Validity

There are two general types of validity: external and internal. External validity (Campbell & Stanley, 1963) is the extent to which the method is generalizable or transferable. A method's generalizability is the degree the results of its application to a sample population can be attributed to the larger population. A method's transferability is the degree the method's results in one arena can be applied in another similar arena.

In contrast to external validity, internal validity "is the basic minimum without which the method is uninterpretable" (Campbell & Stanley, 1963). Internal validity of a method addresses the rigor with which the method is conducted (e.g., the method's design, the care taken to conduct measurements, and decisions concerning what was and wasn't measured). There are different types of internal validity: face validity, content validity, criterion-related validity, and construct validity.

Face validity is the degree to which a method appears to be appropriate for measuring what it intends to measure (Fink, 1995). An example of face validity is the observation that a ruler appears to be an appropriate tool to measure length.

Content validity has to do with the degree to which the method measures the trait it is intended to measure. An example of a test which lacks adequate content validity is one which intends to measure a subject's mathematical ability by testing only addition (Carmines & Zeller, 1979).

Criterion-related validity has to do with the degree to which the method allows for assessment of a subject's performance in situations beyond the testing situation—in a different domain than the test. Criterion-related validity may be concurrent or predictive. That is, the test result may either be intended to assess a criterion independently measured at the same time (concurrent), or to predict achieving a criterion in the future (predictive). An example of predictive criterion-related validity is the extent to which a written driver's test accurately predicts how well the tested population will drive (Carmines & Zeller, 1979). The written driver's test does not involve physically driving a car on the road, but only involves answering several multiple choice questions. The extent to which good performance on that written driver's test correlates well with future good driving performance on the road is a measure of the test's criterion-related validity. The physical act of driving a car takes place in a different situation than the testing environment, and involves different skills and abilities. Threats to the predictive validity of the written driver's test include the possibility that a test subject can't read English but might be a good driver. That threat to criterion-related validity is addressed by having the test printed in several languages besides English. An inquiry of a test's predictive criterion-related validity asks the question, "How accurately does this test measure future performance in that setting?"

Construct validity has to do with the degree to which the results of the method can be accounted for by the explanatory constructs of a sound theory. A method's construct validity is understood by first specifying the theoretical relationships between the concepts, examining the empirical relationships between the measures of the concepts, and then interpreting how the observed evidence clarifies the concepts being measured (Carmines & Zeller, 1979). Construct validity is demonstrated when measures that are theoretically predicted to be highly interrelated are shown in practice to be highly interrelated. An inquiry of a test's construct validity focuses on the question of whether the results of the test are in fact a true measure of the construct, or theory, being tested, and not of some other phenomenon or process which might produce the same results. Such an inquiry asks, "Is this theory the best explanation for the results?"

Not all types of internal validity are applicable to any one method (Rossi & Freeman, 1979). For instance, a method may not be one which is intended to predict an outcome of a process; it may not seek to answer the question, "How accurately does the test measure future performance in a different setting?" so its criterion-related validity would not be an issue. The intended result of a method may not be something that is predicted by a theory; it may not ask, "Is the theory the best explanation for the results?" If the result of the method can't be measured and contrasted with a theoretically predicted result, its construct validity would not be an issue.

Reliability

A reliable method is one that yields consistent results upon repeated use; it is suitable for its intended purpose. However, when a reliable method is used in court by experts on two opposing sides in a dispute, the results will not necessarily be identical. After all, the reason disputes end up in court is just that there are good arguments for both sides. Expert witnesses are ethically obliged to help their attorney clients explain the case to the juries from the particular point of view of their client, within the bounds of truth (Kardon, Schroeder, & Ferrari, 2003). It is not unethical for an expert witness to explain technical aspects of the dispute from the particular point of view of their client. Issues end up in court because there are differences of opinions and shadows of doubt, and experts retained by attorneys representing both sides of a dispute often come to different and contrary opinions based on reasonable interpretations of the evidence they each review or develop. Each side's expert presents technical evidence for the purpose of aiding the trier of fact. The trier of fact is best served by the effective presentation of technical arguments from both sides of the dispute. The reliability of a method used as the basis of an opinion given as expert testimony, therefore, can not be evidenced by identical opinions being supported upon its repeated use in the dispute resolution process. Instead, the reliability of such a method will originate in the understanding that the method is suitable for supporting the opinion of the expert.

Example - Expert Testimony

An expert testified on behalf of an insurance company concerning the amount of structural movement that must have occurred in a house that was in the throes of a major remodel and seismic upgrade when it was allegedly damaged by the 1989 Loma Prieta earthquake. The expert did not observe the actual damage caused by the earthquake, but was asked by his client to determine whether the damage that was claimed by the homeowner could have been caused by the earthquake. The expert performed a computer-based analysis of the house, and relied on that analysis to come to his opinion regarding the amount of movement the house underwent during the earthquake, and therefore the amount of damage to the house that occurred as a result of the earthquake.

The structural engineer-of-record for the remodel and seismic upgrade, who performed structural observation during construction both before and after the earthquake, testified that most of the plywood on the exterior walls of the house at the time of the earthquake was attached with duplex nails (double-headed nails for easy removal) at a much wider spacing than the final shear wall nailing was to be, and the shear transfer clips and hold downs were not in place. The purpose of the plywood in place at the time of the earthquake was not to provide lateral load resistance, but to provide some jobsite security by preventing unauthorized access to the building after working hours.

The expert testified his computer model was based on the assumption that the plywood for all the shear walls was in place at the time of the earthquake, but was not nailed with all the nails specified in the design documents for the strengthening of the house. He testified he assumed the plywood was sufficiently nailed so that there was continuity between the plywood and the framing. He assumed the plywood was sufficiently attached to the framing at the time of the earthquake because that was the assumption of the computer model. He also testified his model assumed none of the hold down hardware was in place. He did not testify as to whether his model assumed the bottom plate nailing or the top plate shear transfer clips were in place.

There was no assurance the analysis by the insurance company's expert accurately replicated the actual behavior of the building in the earthquake. This was because in finite element analysis the real assembly of framing, plywood, nails, clips, hold downs, etc. is modeled using elements of assumed strength, stiffness and boundary conditions. These assumptions must be verified either by comparing the actual assembly to previously tested assemblies that have been shown to be accurately modeled, or by carrying out physical tests of the assemblies to compare their behavior with the model element's behavior.

Published models of wood-framed shear wall assemblies used in finite element analyses are based on assumptions of fully nailed walls, with competent, active shear and overturning transfer hardware in place. There was no testimony given by the insurance company's expert that he used a verified model of a wood-framed shear wall without hold downs and shear transfer hardware, and with plywood sheathing only lightly nailed at the panel edges.

In addition to the modeled wall assemblies, the interface of the house foundation and the ground, and the actual condition of the real foundation (cracks and all) must be accurately modeled in order for the analysis to be correct. There was no assurance the soil-structure interaction or the foundation were accurately modeled.

The same criticism can be brought against the modeled loads. The insurance company's expert testified that he estimated the earthquake ground motion at the subject house by examining published earthquake ground motion records from two nearby seismometer stations, and used that assumed ground motion as input for his computer analysis of his model of the house. The actual ground motions to which the building was subjected must be accurately modeled in order for the finite element analysis to produce results which reliably duplicate the actual behavior of the building. The insurance company expert relied on records of earthquake motion recorded at two stations which he stated were close to the building site, and which he stated were on similar soils. The insurance company's expert offered no testimony of any characterization of the soils at the subject site, or of the sites where the earthquake motions he used were recorded.

Interpolation of earthquake ground motion at a particular site from ground motions recorded at other sites is inaccurate. The United States Geological Survey, the source of one of the ground motion records used by the insurance company's expert, publishes maps of earthquake-induced ground shaking. They state, "ground motions and intensities typically can vary significantly over small distances, these maps are only approximate. At small scales, they should be considered unreliable." By the same token, any ground shaking at the subject house the insurance company's expert deduced from the two recording stations to which he referred must be viewed as unreliable. The only true measure of actual ground shaking at the site, absent a calibrated and functioning recording device, is the amount of damage which actually occurs at the site.

Because the model used by the insurance company's expert was not based on actual conditions at the building, and because no verification or justification of the model structure or loading was presented, the analysis was not valid or reliable. The computer model did not recreate an accurate depiction of the actual condition of the house or of the actual loads applied to the house by the Loma Prieta earthquake, it therefore lacked face validity. Because the model of the structure and of the loads was not representative of the real structure or loads, the method could not predict or describe the movement of the building, and therefore lacked criterion-related validity. Because of the absence of validity and reliability, it was argued that the expert evidence should be disallowed in the determination of damages.

Example - Engineering Method

A primary obligation of an engineer is to anticipate failure modes in the element, component, or system being engineered and then provide measures to prevent those failure modes from developing or from developing catastrophic results (Petroski 1985, 1994; Harr 1987; Wenk 1989). This obligation requires two primary elements: 1) anticipation of possible failure modes, and 2) provision of defenses in depth to prevent and/or mitigate those failure modes. The second element requires valid and reliable analytical models.

During Hurricane Katrina, a large segment of a drainage canal levee and floodwall lining the 17^{th} Street canal in New Orleans failed catastrophically before the design water elevations were realized. The Corps of Engineers Interagency Performance Evaluation Task Force analyses (Interagency Performance Evaluation Task Force 2006) of this failure concluded that a failure mode developed that was not recognized by the designers. This finding led to the official contention that this was a "design failure." Information developed by the Independent Levee Investigation Team (2006) clearly indicates that this failure was a result, not a cause.

The failure mode involved lateral deflection of the concrete floodwall and the sheet piles that supported that floodwall. This deflection resulted in separation between the stiff supporting sheet piling and the soft soil of the levee on the flood side of the wall. Water was then able to enter the gap and exert additional lateral forces on the remaining 'half' of the levee-floodwall. Now the levee only had about 'half' of its width able to transmit the lateral forces to the underlying soils. This combination resulted in lowering the lateral resistance with a commensurate lowering of the factor of safety.

This development was incorrectly reported as "unforeseen and unforeseeable" by the Interagency Performance Evaluation Task Force (Marshall 2006; Seed and Bea 2006). In 1985, the New Orleans district of the Corps of Engineers conducted a full scale instrumented lateral load test of a 200-foot long sheet pile flood wall in the Atachafalaya basin (U.S. Army Corps of Engineers 1988a). This particular location (south of Morgan City, Louisiana) was chosen because of the close correlation of the soil conditions in the New Orleans area with those at the test location. "The foundation soils are relatively poor, consisting of soft, highly plastic clays, and would be representative of near worst case conditions in the NOD (New Orleans District)." (U.S. Army Corps of Engineers 1988a).

Test data from the highly instrumented sheet pile wall and adjacent supporting soils indicated a gapping behavior (separation of the sheet piles from the soils). The test was designed to take an eight foot height of water (above the supporting ground level) with a factor of safety of 1.25. But the wall was already in a failure condition (increasing lateral displacements with no increase in loading) when the water level reached only 8 feet instead of the calculated 10 feet. Strain gage readings on the sheet piles indicated that they were well below the steel yield point, thus the yielding had to have been developing in the supporting soils. Two very important pieces of information developed by the E-99 sheet pile tests were that there was potential soil separation from the sheet piles (allowing water to penetrate below the ground surface between the piles and the soils) and that the calculated safety factor was not reached (it was over-estimated due to unanticipated deformations in the soils).

Additional reports and professional papers further developed the experimental information and advanced analytical models that could be used to help capture such behavior (U.S. Army Corps of Engineers Waterways Experiment Station 1989). Later developments in this work were reported by Oner, Dawkins and Mosher (1997):

> *As the water level rises, the increased loading may produce separation of the soil from the pile on the flooded side (i.e., a "tension crack" develops behind the wall). Intrusion of free water into the tension crack produces additional hydrostatic pressures on the wall side of the crack and equal and opposite pressures on the soil side of the crack. Thus part of the loading is a function of system deformations.*

These developments in technology were *not* reflected in the design guidelines used (U.S. Army Corps of Engineers 1988b, 1989, 1990). A traditional method of active and passive pressures acting along the length of the sheet piles embedded in the earth levee was used to determine stresses induced in the concrete wall - sheet pile joint and in the sheet piles. This traditional method did not incorporate the information developed from the E99 floodwall test. The traditional design guideline-based method used to design and engineer the floodwall system did not possess the required attributes of validity and reliability.

A second element in this development regarded characterizations of the soils that supported the earth levee and sheet piling in the vicinity of the 17th Street canal breach. The processes used at the time of design to analyze the soil types and engineering characteristics did not capture the unique characteristics of the soils. Higher soil strengths beneath the crest of the levee were used to characterize the strengths of the soils at and beyond the toes of the levees. In addition, the spatial averaging process (vertical and lateral) did not capture the unique soil characteristics in the vicinity. Soils in Southern Louisiana and other parts of the Gulf Coast have very complex histories due to past floods, hurricanes, the rise and fall of sea level, changes in vegetation, and other events. Far from being uniform in properties or geometry, they contained complicated and rapidly varying strata of different materials with very different characteristics.

A traditional design guideline-based method of planes with a prescribed geometry was used to model the failure surfaces (U.S. Army Corps of Engineers 1990, 2000, 2003). The shear resistance along these surfaces was based on averaged (laterally and vertically) soil shear strengths for soil units that did not represent the same depositional environments. The geometry of the soil units was assumed to be horizontal. The combination of these design guidelines and practices were used to evaluate the stability of the levee-floodwall system.

In 1964 - 1965 the Corps ran a full scale levee test in the Atachafalaya basin in which advanced studies were conducted regarding characterizations of the soil strengths and performance and stability characteristics of the levee (U.S. Army Corps of Engineers 1968; Kaufman and Weaver 1967). The levee test sections were thoroughly instrumented and their performance monitored during and after construction. Various analytical methods were used to evaluate the usefulness and reliability of the various methods. These developments clearly indicated the need to understand the geologic soil depositional processes and the associated variations in soil strengths (horizontal and vertical) in order to understand the performance and stability characteristics of levees. The importance of local soil conditions to the performance of the levee was clearly pointed out. Additional reports and professional papers were published that resulted in significant advances to the engineering knowledge (Duncan 1970, Ladd et al. 1972; Edgers et al. 1973; Foott and Ladd 1973, 1977).

In-depth background on the geologic and depositional environment of vital importance to understanding the characteristics of the Mississippi Basin soils were developed in the 1950s and 1960s (Kolb and Van Lopek 1958; Krinitzsky and Smith 1969) and the Corps of Engineers lead in development of this background. Of particular importance was recognition that the marsh and swamp deposits were "treacherous" and highly variable. It was repeatedly pointed out that "careful and detailed characterization of the soil properties was required." Further the studies that the method based on traditional Corps of Engineers soil characterization and stability analyses gave factors of safety that were too large (Foott and Ladd 1977). As in the first instance, these developments were not reflected in the design guidelines and practices that were used. Again, the traditional methods used in design and engineering the levee-floodwall system did not possess the required validity and reliability.

Important failure modes in the 17th Street canal levee-floodwall system components were not recognized. The combination of methods used to perform the design was neither valid nor reliable. When the system was tested, it failed.

Conclusion

Engineering must be based on valid and reliable methods and processes in order to yield its intended results. This is true for forensic engineering performed for the purpose of resolving disputes, and for engineering for the design and construction of the built environment. Lacking validity and reliability, engineering testimony may be inadmissible, and so would not serve to support the resolution of a dispute. Lacking validity and reliability, an engineering method or process intended to protect or enhance people's lives and livelihoods can fail to do so, possibly resulting in losses on a catastrophic scale.

References

Campbell, Donald T.; Stanley, Julian C., Experimental and Quasi-Experimental Design for Research, Houghton Mifflin Company, Boston, 1963.

Carmines, Edward G.; Zeller, Richard A., Reliability and Validity Assessment, Sage University Paper Series on Quantitative Applications in the Social Sciences, John L. Sullivan, Ed., Series No. 07-017, Newbury Park, California, 1979.

Daubert v. Merrill Dow Pharmaceuticals, Inc., 509 U. S. 579, 113 S. Ct. 2786, 125 L. Ed. 2d 469 1993.

Duncan, J.M. (1970). *Strength and stress-strain behavior of Atchafalaya foundation soils.* Research Report TE70-1, Department of Civil Engineering, University of California, Berkeley.

Edgers, L. et al (1973). *Undrained creep of Atchafalaya levee foundation clays.* Research Report R73-16, Soils Publication No. 319, Dept. of Civil engineering, Massachusetts Institute of Technology, Cambridge MA.

Federal Rules of Evidence, U. S. Government Printing Office, Washington, DC, 2001.

Fink, Arlene, Editor, The Complete Survey Research Kit, Volumes 1-9, Sage Publications, Inc., Thousand Oaks, California, 1995.

Foott, R., and Ladd, C. C. (1973). *The behavior of Atchafalaya test embankments during construction.* Research report R73-27, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge MA.

Foott, R., and Ladd, C. C. (1977). "Behaviour of Atchafalaya levees during construction." *Geotechnique*, 27(2), 137-160.

Harr, M. E. (1987). *Reliability-Based Design in Civil Engineering*, McGraw-Hill, New York.

Independent Levee Investigation Team (2006). *Investigation of the Performance of the New Orleans Flood Protection Systems in Hurricane Katrina on August 29, 2005*, University of California Berkeley, Report No. UCB/CCRM-06/01, May

Interagency Performance Evaluation Task Force (2006). *Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System,* Final Draft Report, U. S. Army Corps of Engineers, Washington, DC, May.

Kardon, Joshua B.; Schroeder, Robert A.; Ferrari, Albert J., "Ethical Dilemmas of Technical Forensic Practice," 3rd Forensic Congress, Technical Council on Forensic Engineering, American Society of Civil Engineers, San Diego, California, October, 2003.

Kaufman, R. I., and Weaver, F. J. (1967). "Stability of Atchafalaya levees." *J. Soil Mechanics and Foundations Division*, 93(4), 157-176.

Kolb, C.R., and Van Lopik, J.R. (1958). *Geology of the Mississippi River Deltaic Plain, Southern Louisiana.* Technical Report No. 3-483, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Krinitzsky, E. L., and Smith, F. L (1969). *Geology of backswamp deposits in the Atchafalaya basin, Louisiana.* Technical Report S-69-8, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Ladd, C. C. et al. (1972). *Engineering properties of soft foundation clays at two south Louisiana levee sites*. Research Report R72-26, Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge MA.

Marshall, R. (2006). "Floodwall failure was foreseen, team says." *Times Picayune*, New Orleans, LA, 03/14/2006, <http://www.nola.com> (May 1, 2006).

Oner, et al (1997). "Soil-Structure Interaction Effects in Floodwalls." *Electronic Journal of Geotechnical Engineering*, <http://www.ejge.com/1997> (Jan,1, 2006).

Petroski, H. (1985). *To Engineer is Human: The Role of Failure in Successful Design*, St. Martins Press, New York, NY.

Petroski, H. (1994). *Design Paradigms, Case Histories of Error and Judgment in Engineering*, Cambridge University Press, Cambridge, UK.

Rossi, Peter H.; Freeman. Howard E., Evaluation: A Systematic Approach, Sage Publications, Inc., Newbury Park, California, 1979.

Seed, R. B., and Bea, R. G. (2006). *Initial Comments on Interim (70%) IPET Study Report*, National Science Foundation-Sponsored Independent Levee Investigation Team (ILIT), University of California, Berkeley, Mar. 12, 2006.

U.S. Army Corps of Engineers (1988a). *E-99 Sheet Pile Wall Field Load Test Report.* Technical Report No. 1, U.S. Army Engineer Division, Lower Mississippi Valley, Vicksburg, MS.

U.S. Army Corps of Engineers, New Orleans District (1968). *Field tests of levee construction, test sections I, II, and III, EABPL, Atchafalaya Basin Floodway, Louisiana*. Interim Report, New Orleans, LA.

U.S. Army Corps of Engineers (1988b). *Lake Pontchartrain, LA., and Vicinity Lake Pontchartrain High Level Plan, Design Memorandum No. 19 Orleans Avenue Outfall Canal*. Three Volumes, New Orleans District, New Orleans, LA.

U.S. Army Corps of Engineers (1989). *Development of Finite-Element-Based Design Procedure for Sheet-Pile Wall.* U.S.Army Corps of Engineers Waterways Experiment Station, Tech. Report GL-89-14, Vicksburg, MS.

U.S. Army Corps of Engineers (1990). *Lake Pontchartrain, LA., and Vicinity Lake Pontchartrain High Level Plan, Design Memorandum No. 20, General Design, Orleans Parish, Jefferson Paris, 17th. Outfall Canal (Metairie Relief)*. Two Volumes, New Orleans District, New Orleans, LA.

U.S. Army Corps of Engineers (1994). *Design of Sheet Pile Walls*. Engineer Manual EM 1110-2-2504, <http://www.usace.army.mil/inet/usace-docs> (Apr. 1, 2006).

U.S. Army Corps of Engineers (2000). *Design and Construction of Levees*. Engineer Manual EM 1110-2-1913, <http://www.usace.army.mil/inet/usace-docs> (Apr. 1, 2006).

U.S. Army Corps of Engineers (2003). *Slope Stability*. Engineer Manual EM 1110-2-1902, <http://www.usace.army.mil/inet/usace-docs> (Apr. 1, 2006).

Wenk, E., Jr. (1989). *Tradeoffs: Imperatives of Choice in a High Tech World*, The Johns Hopkins University Press, Baltimore, MD

Learning From the Past Experiences of Practicing Engineers

F. N. Rad[1] and A. M. James[2]

[1]Department of Civil and Environmental Engineering, Portland State University, Portland, Oregon 97207-0751; PH (503) 725-4205; FAX (503) 725-5950; email: franz@cecs.pdx.edu
[2]425 SW Stark Street, Second Floor, Portland, OR 97204; PH (503) 445-8694; FAX (503) 273-5696; email: art.james@nishkiandean.com

Abstract

The authors initiated a course in Forensic Structural Engineering at Portland State University two years ago. The main goal of the course is to learn from the past experiences of practicing engineers, thus leading the students to a more critical, creative, and cautious thinking process. The course introduces the students to the basic principles and approaches of forensic engineering, along with studying several case histories. A portion of each case study is devoted to outlining the ways to help minimize potential failures of similar nature, as the students embark on their journey to professional practice. The authors have observed that teaching students by examining cases presented by the forensic engineers who actually investigated the cases, is an effective way to accomplish the goals established for this course. This paper describes the authors' experience conducting the course in the past two years.

Introduction

The course in Forensic Structural Engineering is designed to introduce the students to the basic principles of forensic engineering by utilizing actual case histories. The prerequisite for the course is senior or graduate standing in Civil Engineering and knowledge of concrete, steel, and timber.

The objectives of the course are to teach the students about forensic engineering by a methodical study of several case histories, and to help them better understand ways to prevent similar failures from recurring, thus leading to improved design and construction.

The format of the course is to first present what happened in a case that involved structural failure and the evidence of failure or non-performance. The students are then asked to consider issues that could have contributed to this failure. They are encouraged to think freely, and come up with a variety of possible causes of failure. The report from the structural engineer who investigated the failure is then reviewed.

Utilizing this format of teaching has helped the students learn to think more freely and creatively as the cases are discussed, analyzed, and the structural engineers' observations are presented.

Reference Book

The reference book used for the course is: "Locomotive in The River and Other Stories", by Art James. Mr. James is also the main speaker who presented the case studies to the class. The "book" includes 22 cases of forensic investigations conducted by the author over a span of fifty years. In order to give the reader of this paper an idea on the types of failure covered in the book and discussed in class, the cases have been cataloged according to the nature of the failure, summarized below, along with a brief description of each case.

1. Roof Failures

Link Beams in Las Vegas - A roof collapses with no known cause. No known live loads. What happened and why?

Walla Walla Cold Storage Roof Collapse - Nail laminated wartime trusses fail under ice loads on the lower chords.

Plugged Roof Drains - Several different failures that resulted from the water load caused by a plugged roof drain.

Warehouse Roof Collapses Shortly After Construction - Six inches of snow fell and five roof trusses failed at Warehouse; the effect of joint eccentricity in a Howestring truss. Why load tests need to be made safely.

Cracked Roof Slabs at the Airbase - The Colonel was insistent. "Get out here and look at your cracked roof slabs". It was a roofing failure and an overload from gravel.

Bowstring Roof Trusses Collapse Under Heavy Snow Load – The collapse destroyed a lot of pleasure boats; the insurance paid, then the law suits started.

2. Problems with Columns

Chiller Tank Columns - A Puzzler, the rebar were bulged out like they failed in compression, but that was deceptive.

What Is A Squaring Post and Why Did It Fail? - A major roof support is knocked off its pedestal. What caused this near collapse?

Removing a Main Column in a Portland Hotel – The engineers were very careful and did not expect what happened when their transit tilted.

3. Failures due to Storms, High Tides, Wind, and Strong Currents

The Old Salt, the Shark, and the Loose Barges – Case investigates major damage caused when a loaded rock barge and a fuel barge break their mooring lines and ride a big ebb tide into a dock in Astoria, Oregon.

What a Big Wind Did To Sally's School - Oregon's Famous October Storm did major damage to many structures, including the author's daughter's grade school. This leads to an effort to get the state legislature to pass a law requiring wind and seismic design for schools. Sounds easy, but it was not!

Battering Ram and the Freezer Dock Collapse - A heavy log driven by strong current knocks out a post and causes a domino effect failure.

Landslide Behind House on Montgomery Drive - Slides came up to the foundation, and the insurance company was prepared to call it a total loss!

Freeze Tunnel Collapses at Modesto, California - The tunnel imploded, the author was asked to investigate the cause, and ways to fix it.

Church Steeples- Be Careful What You Pull!- The author's curiosity causes a thrill ride during his inspection at the top of a church steeple.

4. Failures of Walls

The Big Jose' - The biggest retaining wall west of the Mississippi falls five days after backfilling.

My Retaining Wall Failed, Will Your Design Last? - The owner said "I'll sue you for everything you ever earn if your wall fails!"

5. Failures due to Construction-Related Causes

Down Pipe at the Teton Dam - Author investigates a fatal construction accident.

Why Did The Boom Collapse? - The operator was lifting a marine leg out of the hold of the barge when the boom collapsed. Why?

M V PacKing Deck Crane Collapse - A log loading deck crane toppled on a new vessel with no overload. Why?

The Wilsonville Bridge Cofferdam Failures - The cofferdams failed during high water and the State bridge engineer wanted extra seal concrete and larger cells in the rebuild. The contractor had a novel approach. Would it work?

Class Schedule

The course included ten sessions carried out in ten weeks, with each session running for two hours. Four sessions contained cases that dealt with similar topics and presented as a group by Art James, as described above. In another four class sessions, cases were presented by guest speakers. The class schedule thus took the following form.

Session No. 1	Introduction to Forensic Engineering, course syllabus and format; speaker: Franz Rad, PE, SE. An example of forensic investigation: Locomotive in the River - What happened to put the switch engine in 30 feet of water? Presented by Art James, PE, SE
Session No. 2	Roof Failures; speaker: Art James, PE, SE
Session No. 3	Failures of wood structures; speaker: Don Neal, PE, SE
Session No. 4	Problems with columns; speaker: Art James, PE, SE
Session No. 5	Four cases: Salem K-Mart roof failure, Portland International Airport parking structure collapse during construction, Clark County Square Dance Center, KGW-TV transmission tower collapse; speaker: Jack Talbott, PE, SE
Session No. 6	Failures due to storms, high tides, wind, and strong currents; speaker: Art James, PE, SE
Session No. 7	Two cases of failure due to foundation settlement; speaker: Gary Peterson, PE, SE
Session No. 8	Failures of Walls, and failures due to Construction-Related Causes; speaker: Art James, PE, SE
Session No. 9	Four cases: Church building roof failure, concrete slab excessive cracking, failure of a Bowstring truss, concrete tilt-up panel connection; speaker: Ray Miller, PE, SE
Session No. 10	Case studies prepared by students; speakers: student groups. The last class session is allocated to student presentations.

In session No. 1, part of the lecture relates to the principles of forensic engineering, course format, and a discussion of a series of books and reports brought to class as examples of references. The purpose of showing the reference materials to the students is to introduce them to the wealth of literature available in the library, and to allow students to "check out" one or more books from the instructor. Also, students are encouraged to consider selecting some of the cases described in the reference books as their cases to be fully described and presented at the last class session. A sampling of the reference books is shown below.

Forensic Structural Engineering Handbook, by Robert Ratay
Construction Failures, by Jacob Feld and Kenneth Carper
Building Failures, by Thomas McKaig
Why Buildings Stand up, by Mario Salvadori
Failure Mechanisms in Building Construction, Edited by David Nicastro
To Engineer is Human, by Henry Petroski
Dsign Paradigms, by Henry Petroski
Failures in Civil Engineering: Structural, Foundation and Geoenvironmental Case Studies, edited by Robin Sheperd and David Frost
Structural and Foundation Failures, by Barry LePatner and Sidney Johnson
Lessons Learned Over Time, Learning from Earthquake Series, Volumes I, II, III, and IV, published by EERI
Proceedings of ASCE 1st, 2nd, and 3rd Forensic Congress

General Guidelines for the Speakers

The guest speakers are provided with guidelines for their presentation. The text of the guidelines is shown below.

"Please consider presenting three or four cases in each session; 15 min for presentation, plus 15 min Q/A for each case. A one-or two-page summary of the facts about each case (pre-lecture material), including sketches or photos, should be given to students a week before your session. The 'Summary of Facts' will not contain your conclusions as to what caused the failure. The students will think about the items that may have contributed to the failure before hearing your lecture. When you present the seminar, you will describe the way you went about finding what caused the failure, including your detective work, calculations, references, etc. The students will write a report on each case and submit for grade."

The Class Room

The class room used for this course is a "distance learning lab" that contains several AV equipment. The lab contains equipment and the means to show on a large screen: 35-mm sides, transparencies, hard copy pages, photographs, digital images, videos, movies, and the image on the computer screen. The authors have found that these capabilities in the class room have improved instructional effectiveness.

The Teaching Process, Class Format, and Weekly Reports

As a part of the teaching methodology employed by the authors, for each case study the students are asked to first review the "pre-lecture" materials, consider what happened when the structure failed and how many potential weaknesses could have contributed to this failure; to study the structural engineer's verbal and written reports and to determine what could be learned in order to prevent similar failures in the future. The students are directed to follow the format below when reviewing the cases to be presented, and preparing for discussions. The text of the instructions given to the students follows.

"Before coming to class, study the assigned cases (pre-lecture materials) to learn what happened. In class, listen and take notes as the Forensic Engineer describes the case. Think freely as to the potential cause(s) for the failure. Think about how you may go about the process of finding the facts about this failure and potential causes. Are perceived and/or expressed facts really facts, or half-truths, or fiction? How would you find corroborating 'perceived facts' and/or evidence to point you to discovering the most probable cause?"

The course requirements include a report to be submitted by each student on each case study. One half of the final student's grade is based on his/her reports on the case studies. The student reports on the case studies follow a specific format. The text of the instructions to the students is shown below.

"For each case, describe what happened, the location, date, and other pertinent information. Describe what, in your judgment, may have either caused or contributed significantly to this failure. Briefly describe the reasons for the failure, as observed by the Forensic Engineer and reported to you verbally and in writing. Describe what you learned from this failure. Describe how you may use the lessons learned in your future design practice. Describe any other ideas and/or information, as related to other actual failures or potential failures that may be prevented by utilizing the lessons learned."

Examples of Case Studies

The reference book contained 22 case studies, investigated by Art James, and documented in the book. About 16 other cases are presented by four guest speakers. Three examples are presented below, to familiarize the reader with the types of cases presented in class.

The Old Salt, the Shark and the Loose Barges - The barge company had a minimal view of the under-dock damage. We felt it was more serious and the international consultants from Vancouver B.C. favored our view. The energy calculation and use of a table to "bracket" the exact solution provides a valuable tool. Comparing the impact energy to a major earthquake is useful.

Locomotive in the River - The longshoreman felt the trestle collapsed first. He said he had reversed and is heading inshore. We felt the engine hit the wheel stop first and the impact sheared the bolts in the track splices then the tracks pulled the trestle down. The underwater photo showed the throttle toward the river proving which way the train was headed.

The Battering Ram and the Cold Storage Dock Collapse - Two insurers contest the issue of what failed first, and why? The dispute goes to trial. Did the dock fail from an overload? Or did a heavy moving object dislodge a post and cause a "domino effect"? The case involved a court contest with "big bucks" at stake and a structural engineering solution.

Lessons Learned

At the end of each case study, a segment of the oral and written presentations are devoted to the lessons that can be learned. Lessons learned normally refers to a series of lessons and observations useful in the future practice of young engineers. For example, in the three cases summarized above, i.e., "The Old Salt...", "Locomotive...", and "The Battering Ram..." the students learn about the application of the energy method ($\frac{1}{2}mv^2 = Fd$, where m = mass, v = velocity, F = force, and d = distance) in finding the proximity of the solution for cases that involve impact of moving objects. By assuming a range of reasonable values for the parameters, it is possible to "bracket" the solution.

The other aspects of "lessons learned" refer to the processes used for:

- Gathering evidence
- Assessing or estimating loads
- Estimating the actual properties of materials, rather than those assumed in design
- Field testing
- Structural analysis, the assumptions, and the validity of assumptions
- Assessing different views of opposing sides on the probable cause of failure

Number of Cases Presented

As mentioned earlier, there are 22 cases in the reference book, plus usually four cases that are presented by each of the four guest speakers, for a total of about 38 cases covered in nine sessions. On an average, four cases are covered in each session, making the presentation time (including Q/A) for each case to be approximately 30 minutes. With the class format adopted, and the students having a chance to read about the cases in advance (pre-lecture materials), the presentation time allowed appeared sufficient. Of course, more complex cases take a bit more time.

Another instructional approach may be to cover a lower number of cases, but allow more time to present and discuss each case. For the available number of hours for this course, which is nine sessions of two-hour length, another possibility is to cover, say only 18 cases, with one hour presentation time per case. The authors have not experimented with this format, as of yet.

Speakers' Commentary to the Students

In addition to the specific cases presented, the salient points addressed and general advice given by the speakers to the students relate to the following items.

- Structural engineers learn from past failures, and the case studies covered in the course provide examples of valuable lessons from structural failures.

- Teaching by example exposes the students to the power of deductive reasoning.
- When you examine a failure, you must identify all possible causes. Then reason your way through them and discard the ones that defy logic.
- The possible causes must be considered and analyzed with approximate computations to gauge the probability of the causes. The ultimate answer sometimes is not the early favorite. Do not close your mind too soon!
- Make sketches of the critical parts, take photos and measurements and avoid early conclusions, before all the data are known.
- Avoid statements like "it might be this or that cause". Your duty is to eliminate the alternates and determine the "proximate cause".
- Sometimes when there are several equal possibilities it may be necessary to say: I can not say for certain, but here are the most likely causes of failure.
- Why does forensic engineering fascinate us? Because we are using a combination of our engineering knowledge and deductive reasoning, and knowledge is power!

Students' Final Projects

The last class session is allocated to student final projects and presentations. Four teams of students are formed early in the term. Each team selects one or more cases to investigate, write a report, and make a presentation to the class.

Below is a summary of the cases presented in the past two terms the course was offered.

Case 1, Building a cantilevered floor for a residence/winery. This is a local project, the student is the structural engineer on the project, and it involved a cantilevered beam with inadequate inboard length. The load on the cantilevered portion raised the inboard end excessively. The student described how he went about retrofitting the structure to minimize the uplift.

Case 2, The Britannia tubular bridge, a paradigm of tunnel vision in design. This is the story of the bridge designed by Robert Stephenson, who faced the challenge of building a bridge rigid and strong enough to carry a heavy train of many carriages. This is done by making the bridge out of two long iron tubes, rectangular in shape, through which the trains would travel. The bridge opened in 1850, with many problems that followed in the following 150 years.

Case 3, Roof collapse at the shoe store. This is a local case, the building constructed with timber roof joists and partially grouted CMU walls. The roof collapsed in winter of 2004 due to plugged roof drain, ice accumulation, and a heavy HVAC unit. The students described their detective work in finding the cause, and the structural repairs.

Case 4, Sagging roof at a commercial warehouse. This is a local building, built in the 1930s, made of timber roof trusses with bolted and nailed connections. The problem was a sagging roof, what caused the sag, and how it was repaired.

Case 5, The failing of Fallingwater House. This is one of Frank Lloyd Wright's most famous houses, located in southwestern Pennsylvania. The case describes the original design, and subsequent problems with cracking and leaking, and restoration of the building completed in 2002 at a cost of $11 million.

Case 6, The Good, the Bad, and the Galloping Gertie. The famous Tacoma-Narrows bridge failure is revisited, the history of design, details of design, and theories of failure.

Case 7, Collapse of parking structure, a case included in the Proceedings of the Second Forensic Congress. The students' report included a discussion of expansion/isolation joints, and the collapse of a parking structure. The structure included two sliding isolation joints which did not allow the required movement. The lessons learned that may be applied to other similar cases were presented.

Case 8, Kansas City Hyatt Regency Hotel walkways collapse of 1981. This notorious case is re-visited, describing what happened and why.

Case 9, Roof Investigations, January 2004 Portland Snow Storm. This project investigates the effects of heavy snow storm on the roofs of several buildings, with emphasis on snow drift. The students make observations on the need to include drift in snow load calculations, and the potential structural problems associated with snow drift.

Assessment and Future Plans

The written course assessment includes the following questions, assessing the level of achievement by students, as perceived by them.

1. Acquire knowledge about what forensic engineers do.
2. Be able to do (or help with) forensic investigative work.
3. Be able to identify potential "pitfalls" in design and construction.
4. Be more cautious about my own (and other people's) assumptions regarding analysis, properties of materials, construction quality, and inspection.
5. Be a better inspector.
6. Design ways to strengthen structures.
7. Become a better engineer and to minimize potential failures in structures that include my services.
8. Learn about topics that are not commonly addressed in other courses.

Most of the students have been graduate students, with some from the practicing sector. Course assessment indicates that the course format seems to be generally working well. As for the quantitative assessment of the course, with a rating of 1 for "poor" and 5 for "excellent", the course rating is about 4.5.

The speakers and the students seem to enjoy their interaction and sharing of knowledge. The students are especially fond of the course because it is "different" from the other courses; it provides a wider scope of information, and it teaches them about numerous situations where similar types of failure can be avoided in their practice.

The current course is a 2-credit course, and our plan is to expand it to 3 credits. Moreover, we have considered assembling a compendium of forensic cases, as the practicing engineers continue to contribute to the course. Most of the cases covered in the compendium will be local cases, involving local buildings, engineers and constructors.

Observations

The authors believe that the value of a course in Forensic Engineering as part of a students' lifetime learning of Structural Engineering is substantial. By showing the students real life examples of failures and the investigations that led to conclusions, the students are enabled in several ways, including an illustration of the tug of war between opposing points of view. The authors further believe that a significant part of educating engineering students in promoting design and construction integration is by understanding the causes of structural failures. As an added benefit, the course in Forensic Structural Engineering has brought about a closer working relationship between academia (students and faculty) and design professionals, the engineers who are willing to share their experiences in forensics with students. Course assessments by students as well as speakers indicate that this course is beneficial in rounding out the students' education in structural engineering.

Benchmarking Forensic Engineering Practice – A Philosophical Discussion

S.E. Chen[1], D. Young[2], D. Weggel[3], D. Boyajian[4], J. Gergely[5] and B. Anderson[6]

[1]Dept. of Civil Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-000; PH (704)687-6655; FAX (704)687-6953;email:schen12@uncc.edu
[2]Dept. of Civil Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-000; PH (704)687-4175; FAX (704)687-6953;email:dyoung@uncc.edu
[3]Dept. of Civil Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-000; PH (704)687-6189; FAX (704)687-6953;email:dweggel@uncc.edu
[4]Dept. of Civil Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-000; PH (704)687-3038; FAX (704)687-6953;email:dboyajia@uncc.edu
[5]Dept. of Civil Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-000; PH (704)687-4166; FAX (704)687-6953;email:ggergely@uncc.edu
[6]Dept. of Civil Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223-000; PH (704)687-6039; FAX (704)687-6953;email:jbanders@uncc.edu

Abstract

Recent series of unfortunate events resulted in thousands of damaged/deteriorated structures. To validate insurance claims and rehabilitation efforts and to assist disaster-worn citizens, a significant amount of forensic engineering work is currently on-going and will continue for a long time. In the midst of these activities, a significant amount of disputes/mis-judgements will occur and will cause further difficulties in settling claims and restoring normal operations. To ensure quality of forensic work, this paper attempts to address the more fundamental issue of current forensic science and engineering practices, which adopts an inverse engineering approach whereupon knowledge is accumulated from construction design and post-event observations. The reasoning process is based on pure deduction with very little in-between causality evidence. Current approach relies heavily on an engineer's interdisciplinary expertise, training, and reasoning ability, and lacks the fundamental scientific process of elimination of possibilities. It, therefore, often fails to produce complete multidisciplinary solutions to complex forensics problems. This paper attempts to establish quality quantification by suggesting forensic benchmarking such that the involved procedures can be standardized and eliminate the "guess work" still common in an otherwise rapidly developing and highly challenging field.

Introduction

Recent catastrophic events (2005 Hurricane Katrina, the 2005 Asian tsunami, 1990-2005 earthquakes and terrorist activities since 9/11) in the United States and around the world have resulted in millions of damaged or destroyed homes and structures. Hurricane Katrina alone destroyed more than 160,000 homes and generated 90 million tons of solid waste (Esworthy et al. 2005 and The White House, 2006). To assist disaster-worn citizens and other involved parties, forensic engineers are currently working to validate insurance claims and rehabilitation efforts, and this work may continue for a long time. Disputes result in difficulties in settling the claims, thereby delaying a return to everyday lives. To reduce lawsuits and pain and suffering for disaster victims, there is a need for more timely and reliable forensic practices. The ultimate goal for all forensic investigations should be to provide quality results that can substantiate a logical and valid conclusion to the case. Hence, the ability of forensic engineers to quantify forensic work qualities will be an important service to the clientele and society at large. This ability is also critical for future enhancement of engineering performance, including the increase in reliability and the reduction in forensic practice liability.

Forensic engineering is a highly versatile profession. It differs from other engineering and scientific practices in that there is very little obvious evidence available for establishing valid causality of failures. As a result, reliance on conclusions from past case studies becomes highly essential. A review of past publications in the ASCE *Journal of Performance of Constructed Facilities* shows majority papers are related to historical data or case studies. Literature published on forensic engineering pedagogy also tends to focus on the teaching of case studies (Bosela 1993, Rendon-Herrero 1993, Fowler et al. 1994, Pietroforte 1998, Rens et al. 2000 and Delatte et al. 2002). However, since failures in civil engineering studies are either rare or rarely reported, to establish a quantitative measure of failure probabilities is difficult.

To establish a reasonable measure of forensic investigation quality, a total shift of paradigm may be required, including the establishment of forensic benchmarking for generation of statistically sound causality samples. This approach would require significant experimentation efforts and instrumentation needs. With valid benchmarking, a probabilistic causality quantifier can then be established. This paper represents a first attempt in establishing forensic quality measures. It is the intent of the authors to open a dialog amongst forensic engineering professionals to determine ways to improve and moderate forensic investigations.

Rationale behind Current Forensic Practices

Forensic engineering is "the application of engineering in the jurisprudence system requiring services of legally qualified professional engineers. Forensic engineering include investigation of physical causes of accidents and other sources of claims and litigation, preparation of engineering reports, testimony at hearings in judicial proceedings, and rendition of advisory opinions to assist the resolution of disputes" (National Academy of Forensic Engineering,2005). Important elements for forensic work are professionalism, legal knowledge, and the capability to provide expert solutions to the judicial proceedings (Lewis 2003).

Forensic practice is a "fact finding mission" that provides the legal process with a doubt-free explanation of the causality for structural failures (Janney 1979). Current forensic engineering adopts an inverse approach, where knowledge is accumulated from construction design and post-event observations, instead of typical scientific approach of generating statistically reliable data to validate causal hypotheses. Current reasoning is based on pure deductive process with very little experimental support. Additionally, the field lacks "in-between data", thereby relying heavily on an engineer's interdisciplinary expertise, training, and reasoning ability. Therefore, current forensic practices often fail to produce complete multi-disciplinary solutions to complex forensics problems, resulting in unreliable forensic conclusions. The National Society of Professional Engineering has addressed this unreliability issue as a national concern (Lunch 1987). Cohen et al. (1992) indicated several areas of needs for forensic engineering that may able better practices including the collection of historical data, in-situ monitoring techniques, and experimental results.

Since forensic engineering does not rely on significant experimentation - this non-experimental approach (known as causal-comparative study, *ex post facto*) is characterized by reliance of collected structural failure observations and established causality based on past experiences (Patten 1997). Such an approach requires significant experience in deductive reasoning and complete compatibility of past and present data, including structural components, boundary conditions and failure observations. Since failures have already occurred, the causal relationship is typically established with a high level of speculation. Santamarina and Chameau (1989) showed that limitations in human cognitive capability with weak or limited evidences can induce bias.

Causality Quantification

Essential to establish reliable forensic studies is the promotion of deductive reasoning, hypotheses validation and false hypotheses elimination via reasonably-accurate experiments and modeling. *For forensic investigations, this approach means an accurate representation of the entire system including its components and boundaries.* Using a systems approach, Yao (1985) first suggested damage indices to quantify state conditions. Castaneda and Brown (1994) and Kaggwa (2005) suggested using fuzzy logic for causality investigations. However, both approaches required the causality relationships for a certain failure type are well defined from adequate failure cases. While this may not be true for structural components, failures of complex structures are definitely rare.

Assuming adequate data or case studies can be collected, test data from different measurements or observations can then be fused and integrated to explain causality. This process, though initially difficult and time-consuming, will gradually accelerate with accumulated knowledge. The results can be treated as random processes and probability quantifiers can be used to estimate causality likelihood. The probability of identifying causality can be defined as:

$$P(a_i) = P(A_{mi} > 0 \mid A = a_i) \tag{1}$$

where a_i is damage outcome due to specific causal relation i . A_{mi} is the measured outcome and A is the actual outcome. If a delta function is specified as

$$\delta(k_j) = \begin{cases} 1 \sim \text{when causality exists} \\ 0 \sim \text{otherwise} \end{cases} \tag{2}$$

Then probability of each causality, denoted using metric vectors a_i, where weighting coefficients b_i, are actual causalities and k_j defines unrelated causalities with weighing coefficients z_j, can be expressed as (with an error term, ε_m):

$$P(a_i)_m = \sum_{i=1}^{p} b_i a_{mi} + \sum_{j=1}^{q} Z_j \delta_m (k_j) + \varepsilon_m \quad (3)$$

To ensure positive causality, the square sum residuals, S are utilized

$$S = \sum_{m=1}^{n} \left[P(a_i)_m - \left\{ \sum_{i=1}^{p} b_i a_{mi} + \sum_{j=1}^{q} Z_j \delta_m k_j \right\} \right]^2 \quad (4)$$

The most likely causality can be established by optimization procedures:

$$\frac{\partial S}{\partial b_i} = 0 \quad (5)$$

$$\frac{\partial S}{\partial Z_j} = 0 \quad (6)$$

A set of simultaneous equations can then be assembled and decoupled into causal and non-causal variables. When adequate data have been collected, reliability of the causalities can be determined.

Adequate data sample is essential for all scientific research processes; this is because proper deductive reasoning and scientific research requires sufficient evidence to prove that a causal relationship is accurate. For most forensic engineering work, such as manufacturing and industrial engineering, a significant amount of data may be retrieved to establish the causes of failure. The investigation may start with problem definition, proceed through analysis and modeling, and conclude with testing and simulation of failure conditions. However, such databases usually do not exist for civil engineering forensic work. Databases such as the Forensic Anthropology Data Bank, where hundreds of human bodies were allowed to decompose for scientific study at the University of Tennessee at Knoxville, are impossible in civil engineering work if only because of the size of the evidence (The Forensic Anthropology Data Bank, 2005).

Forensic Benchmarking

In the cases, where limited data are available, additional data can be generated using extensive structural modeling and numerical simulation. If standard procedures can be established to ensure consistent data generation, then positive causality can be established for a wide variety of problems - **Forensic Benchmarking Experimentation**. Forensic benchmarking can be used as a forensic investigation procedure by generating extensive causal relationship data. The proper scientific procedures for a true experiment-based investigation typically involve the following steps (Blaxter et al. 2001):

- Observe some aspect of the structural failures;
- Establish a set of structural parameters for reproduction;
- Establish a set of hypotheses that is consistent with what has been observed about the causes of failure.
- Construct structural replicas;

- Use the hypothesis to make predictions.
- Test those predictions by experiments on the structural models ;
- Repeat the above steps until there are no discrepancies between theory and experiment and/or observation.

Such experimentation requires skilled structure model construction and extensive testing/instrumentation setups. Numerical simulation, especially for nonlinear phenomena, also requires extensive detailed modeling of realistic physical parameters. Experimentation-based forensic benchmarking differs from current forensic investigation methodologies in two notable ways:

1. **Unknown variables:** Typical engineering investigation focuses on only one or possibly a few specific variables, where other variables are eliminated or simplified. Forensic benchmarking allows consideration of a wide range of variables.

2. **Multiple processes**: Typical engineering investigations focus on specific damage process, whereas forensic benchmarking targets at capturing damage sequences in order to establish comprehensive explanations of failure mechanisms.

Figure 1 compares current forensic practices and a possible experiment-based forensic investigation. The objective of the experimentation is to generate databases that can be used in statistical inference to generate reliable causal relations. To accomplish such a premise using a structural model, certain design criteria must be established. It should be noted that depending on the application, a perfectly scaled-down model is not always essential; other small-scale relaxed models may be developed (Harris and Sabnis 1999).

Benchmarking Sensing and Nondestructive Testing

To ensure proper forensic benchmarking, it is essential that consistent measurements are made. Today, structural inspection is done mostly through visual assessment; however, the reliability of inspector skills remains an unsolvable issue (Phares et al. 2000). To address the causality issue of a failed structure, it is critical to ensure that collected data accurately capture the entire failure process. Such ambitious monitoring can be achieved only by using several sensors that measure different physical parameters, such as temperature, strain, deformation, ground motion, etc. Recent vigorous research and developments of Nondestructive Testing (NDT) techniques enabled the adoption of these techniques for forensic studies (Chase 1997, Rens et al. 1997, Davidson et al. 1998 and Washer 2000). In particular, nondestructive testing techniques such as X-Ray, Infrared thermography, ultrasound, ground penetration radars, impact-echo, fiber-optic strain sensing, ferromagnetic and geophysical testing methods can provide detailed spatial description of damage and are gaining increasing popularity as common testing methods (Azacedo et al. 1996, Cumming et al. 1997, Gucunski et al. 2000, Fu et al. 1996, Ballard, 1996, Chen et al. 1997, Kalinski 1997, Poston et al. 1997, Rens et al. 1998 and Sansalone et al. 1997).

It should be noted that NDT techniques can be differentiated into local and global methods. Methods such as acoustic emission, ultrasonics, impact echo, x-ray and radar are focused on damage studies on localized defects (Mannings 1985). Techniques, such as dynamic characterization of a structure, allow global damage quantification (Adams et al. 1975). By capturing the vibration signatures of a structure, determination of the deteriorating state of a structure can be assessed (Cawley et al. 1979, Mazurek et al. 1990, Hearn et al. 1991 and Aktan et al. 1997). Maser (1988) described the various physical parameters required during structural assessment to include: 1) inventory data, 2) condition data and 3) performance data. For material failure measurements such as corrosion, the range of interest could be only a fraction of a millimeter (Carper 1989, Day 1999 and Noon 2000).

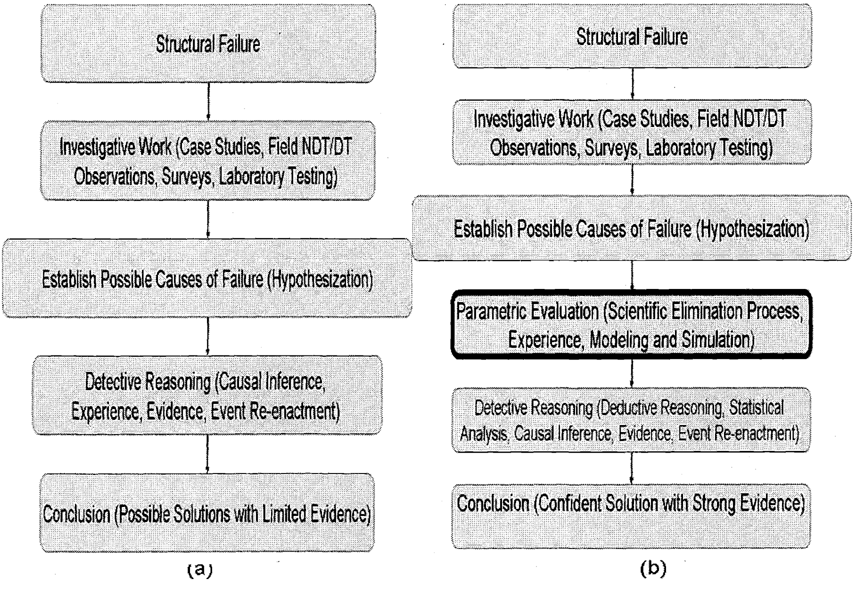

Figure 1 Forensic Investigation Processes: a) Current and b) Experiment-Based

Discussion

Good forensic engineering practices demand a valid measure of the relationship between failure and its causes. A true scientific-based investigation would require the generation of statistically sound sample population of failure cases. While real life data are hard to come by, using structural modeling, forensic benchmark studies can be conducted. To establish a valid structural model or numerical model, certain issues need to be addressed, including:

- Structural dimensions and in-place material verification - for many existing structures, complete as-built plans may not be available; even if plans are

available, they may not be correct or may not reflect errors or changes during construction;
- Damage mapping such as fatigue cracking; and
- Environmental effect determination and delineation from common load effects – some stresses are due to load and some are due to changes in moisture and temperature.

Since most of the above criteria are difficult to establish, forensic benchmarking is extremely difficult to establish for historical structures. Additional reasons why experiment-based studies have not been popular in forensic engineering include:
- Most structural systems are difficult to define (due to changes in surrounding conditions and challenges in scaling and reproducing structural aging and repetitive loading history);
- Few actual structures will allow extensive and repetitive experiments;
- There is currently no comprehensive instrumentation that allows synchronized multiple-level sensing and automated data fusion and self-calibration; and
- Instrumentation for such studies is too expensive for most educational institutions.

Another difficulty that faces a valid forensic benchmarking is that all measured observations should be synchronized: all collected data should be time/ location-stamped, such that a chronological sequence may be plotted out. Generally, output data from a sensor may have one of several possible formats:
- Continuous or sampled waveform complete with amplitude, frequency and phase information;
- Two-dimensional imagery with spatial amplitude or spectral data;
- Vector consisting of parametric positional data, state vector, target identity or characteristics; or
- Other useful outputs such as temperature, health index and background noise thresholds (Hall 1992).

Critical to forensic engineering is the quality of information that can be extracted from sensor measurements. Thus sensor reliability has a direct impact on the quality of the forensic work being performed (Joshi et al. 1992 and 1998). Faulty sensors need to be identified to ensure that bad data do not continue to be collected. Currently sensor reliability issues are not commonly considered when data acquisition systems are designed and developed, although the reliability of collected information is highly dependent on the sensors used. In fact, most of the existing schemes used in industry for fault detection are confined to simple strategies for pre-determined small fault-sets.

Example

The causality calculation is demonstrated using an example where a truss bridge (the Haupt bridge) with 21 members is used (Chen et al. 2002). The West Point Bridge Designer (Ressler 2000) is used to simulate the truss bridge behavior under a standard AASHTO H20-44 truckload. To illustrate the effect of individual member stiffness as a random variable for a structure, static analysis was conducted to determine if the bridge will fail in the load test using the program. The span of the bridge is 24 m long. The original design consists of members of 120 mm ×120 mm cross-section, except for members No. 11 and 21, which are 140 mm×140 mm cross-sections. A Monte Carlo simulation is then conducted using randomly generated members and sections. The cross-sections are limited to 10 mm × 10 mm to 140 mm × 140 mm range. The computation of causality quantification is then conducted. Figure 2 shows the WPBD program with the standard truck crossing over the bridge and caused bridge failure. The members under high stresses are shown in red color. A total of 207 data is generated for the example problem. Figure 3 shows the bridge member numbers. Table 1 shows the outcome probabilities for each of the members.

In this example, causality is defined as a function of the bridge members. Without considering the measurement techniques, the weighting coefficient b_i is considered 1 for all damage cases. For the non-causal terms, coefficients Z_j and errors ε_m are considered zero. The maximum S for the problem is at members 6 and 14, indicating the most probable failure cases. However, the closeness of the percentage S values for all members indicating that they have equal likelihood of failure, which is expected for the simple problem. The accuracy of the causality study will improve with more data sets.

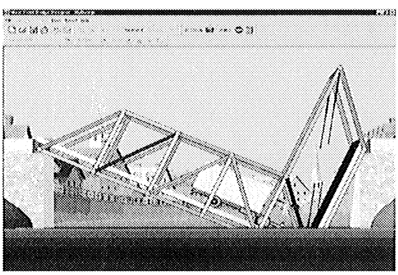

Figure 2 WPBD showing failed bridge during truck crossing

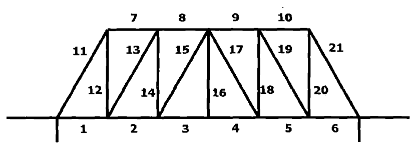

Figure 3 Member numbering system

Conclusion

This paper suggests the possibility of improving forensic causalities by massive generation of case data using structural modeling and forensic benchmark studies. This approach provides the in-between data that can help create direct conclusions of causal relationships between failures and their causes. However, such approach requires significant capabilities in structural modeling as well as extensive instrumentation and nondestructive sensing. To establish such approach may require an industry-wide joint effort and sharing of information.

A simple probabilistic model is suggested in order to identify the most likely causality to a failure, which is illustrated by a simple truss problem. A total of 207 failure cases have been generated to demonstrate that under the same test scenarios, all members should have the same likelihood of failures.

References

Adams, R.D, Walton, D., Flitcroft, J.E.and Short, D. (1975). "Vibration Testing as a Non-Destructive Test Tool for Composite Materials," *Composite Reliability*, ASTM STP 580, 159-175.

Aktan, A.E., Farhey, D.N., Helmicki, A.J., Brown, D.L., Hunt, V.J., Lee, K.L. and Levi, A. (1997) "Structural Identification for Condition Assessment: Experimental Arts." *J. Struct. Engrg.*, 123(12), 1674-1685.

Azacedo, S.G., Mast, J.E., Nelson, S.D., Rosenbury, E.T., Jones, H.E., McEwin, T.E., Mullenhoff, D.J., Hugenburger, R.E., Stever, R.D., Warhus, J.P. and Wieting, M.G. (1996). "HERMES: A high-speed radar imaging system for inspection of bridge decks," *Nondestruc. Eval. of Bridges and Highways*, SPIE # 2946, 195-205.

Ballard, C.M. and Chen, S.S. (1996). "Automated Remote Monitoring of Structural Behavior via the Internet," *Smart Sys. for Bridges, Struct. and Highways*, SPIE # 2719, 90-101.

Blaxter, L., Hughes, C. and Tight, M. (2001). *How to Research*, 2^{nd} Ed., Open University Press, Buckingham, UK.

Bosela, P. (1993). "Failure of Engineered Facilities: Academia Responds to the Challenge," *J. Perform. Constr. Facil.*, 7(2), 140-14.

Carper, K.L. (1989). *Forensic Engineering*, Elsevier Science, New York, N.Y.

Castaneda, D. and Brown, C. (1994). "Methodology for Forensic Investigations of Seismic Damage," *J. Perform. Constr. Facil.*, 120(12), 3506-3524.

Cawley, P. and R.D. Adams (1979). "The location of defects in structures from measurements of natural frequencies," *J. Strain Anal.*, 14(2), 49-57.

Chase, S. and Washer, G. (1997). "Nondestructive Evaluation for Bridge Management in The Next Century," *Public Roads*, 61(1), 16-25.

Chen, S., Pong, W., Chen, P. and Nishihama, Y. (2002) "Teaching Reliability to Civil Engineers," *Proc. Amer. Soc. Engrg. Educ. Southeast Section Annual Conference*, Gainesville, Fl., 2002.

Chen, Z., Cudney, H.H., Giurgiutiu, V., Rogers, C.A., Quattrone, R. and Berman, J. (1997). "Full-scale ferromagnetic active tagging trdting of C-Channel Composite Elements," *Smart Sys. for Bridges, Structures, and Highways,* SPIE # 3043, 169-180.

Cohen, J.M., Corley, W.G., Wong, P.K. and Hanson, J.M. (1992). "Research Needs Related to Forensic Engineering of Constructed Facilities," *J. Perform. Constr. Facil.*, 6(1), 3-9.

Cumming, N.A. and Ooi, O.S. (1997) "Locating Delaminations and Other Defects in Concrete Silo Walls Using the Impact-Echo Procedure," in Pessiki, S., and Olson, L., Ed., *Innovations in Nondestructive Testing of Concrete,* SP-168, ACI, Chicago, IL.

Day, R.W. (1999). *Forensic Geotechnical and Foundation Engineering*, McGraw-Hill, New York.

Davidson, N.C. and Chase, S.B. (1998). "Radar Tomography of Bridge Decks," *Proc. Structural Materials Technology III: An NDT Conference*, SPIE v.3400, San Antonio, TX, NJ, 250-256.

Delatte, N. and Rens, K.L. (2002). "Forensics and Case Studies in Civil Engineering Education: State of the Art," *J. Perform. Constr. Facil.*, 16(3), 98-109.

Esworthy, R., Schierow, L.J., Copeland, C. and Luther, L. (2005). *Cleanup after Hurricane Katrina: Environmental Considerations,* Congressional Research Service Report for Congress, RL33115, Washington DC.

Fowler, D. W., and Delatte, N. J. (1994). "Graduate Course in Forensic Engineering," *Proc. ASCE Texas Section Spring Meeting*, Corpus Christi, Texas.

Fu, X., and Chung, D.D.L. (1996). "Self-Monitoring Concrete," *Smart Sys. for Bridges, Structures and Highways*, SPIE # 2719, 62-68.

Gucunski, N., Vitillo, N. and Maher, A. (2000). "Pavement Condition Monitoring by Seismic Pavement Analyzer (SPA)", *Proc. Structural Materials Technology IV – An NDT Conference*, Feb. 28-Mar. 3, Atlantic City, NJ, 337-342.

Hall, D.L. (1992). *Mathematical techniques in Multisensor Data Fusion*, Artech House, London, UK.

Harris, H.G. and Sabnis, G.M. (1999). *Structural Modeling and Experimental Techniques*, 2nd edition, CRC Press, Boca Raton, FL.

Hearn, G., and Testa, R. B. (1991). "Modal Analysis for Damage Detection in Structures." *J. Struct. Engrg.*, 117 (10), 3042-3063.

Janney, J.R. (1979). *Guide to Investigation of Structural Failures*, ASCE, New York, NY.

Joshi, B. and Seyed H. (1992). "Reliability Analysis of Self-Diagnosable Multiple Processor Systems," *Proc. 23rd Pittsburgh Conf. Model. & Simul.*, Pittsburgh, PA, 1993-1999.

Joshi, B. and Hosseini, S. (1998). "Diagnosis Algorithms for Multiprocessor Systems," *Proc., IEEE Workshop on Embedded Fault-Tolerant Systems,* Boston, MA, 112-116.

Kaggwa, W.S. (2005). "Probability-Based Diagnosis of Defective Geotechnical Engineering Structures," *J. Perform. Constr. Facil.*, 19(4), 308-315.

Kalinski, M.J., (1997) "Nondestructive Characterization of Damaged and Repaired Areas of a Concrete Beam Using the SASW Method," Pessiki, S. and Olson, L., Ed., *Innovations in Nondestructive Testing of Concrete,* SP-168, ACI, Chicago, IL.

Lewis, G.L. (2003). *Guidelines for Forensic Engineering Practices*, ASCE Pub., Reston, VA.

Lunch, M.F. (1987). "Liability Crisis – Where do We Go from Here?" *J. Perform. Constr. Facil.,* 1(1), 30-33.

Manning, D.G. (1985). *Detection of Defects and Deteriorations in Highway Structures*, NCHRP No.118, TRB.

Maser, K.R. (1987). "Sensors for Infrastructure Assessment," *J. Perform. Constr. Facil.*, 2(4), 226-241.

Mazurek, D. F., and DeWolf, J. T. (1990). "Experimental study of bridge monitoring technique." *J. Struct. Engrg.*, 116 (9), 2532-2549.

National Academy of Forensic Engineering (2005). http://www.nafe.org/NafeMainDef.htm>12/25/05.

Noon, R. (2000). *Forensic Engineering Investigation,* CRC Press, Boca Raton, FL.

Patten, M.L. (1997). *Understanding Research Methods – An Overview of the Essentials,* Pyrczak Pub., Los Angeles, CA.

Phares, B.M., Rolander, D.D., Graybeal, B.A., Washer, G.A. and Moore, M. (2000). "Visual Inspection Reliability Study," *Proc. Structural Materials Technology IV – An NDT Conference*, Feb. 28-Mar. 3 Atlantic City, NJ, 14-22.

Pietroforte, R. (1998). "Civil Engineering Education Through Case Studies of Failures," *J. Perform. Constr. Facil.*, 12(2), 51-55.

Poston, R., and Sansalone, M. (1997). "Detecting Cracks in the Beams and Columns of a Post-Tensioned Parking Garage Using the Impact-Echo Method," in Pessiki, S., and Olson, L., Ed. *Innovations in Nondestructive Testing of Concrete,* SP-168, American Concrete Institute, Chicago, IL.

Rendon-Herrero, O. (1993). "Too Many Failures: What can Educators do?" *J. Perform. Constr. Facil.*, 7(2), 133-139.

Rens, K.L., Rendon-Herrero, O. and Clark, M.J. (2000). "Failure of Constructed Facilities in Civil Engineering Curricula," *J. Perform. Constr. Facil.*, 14(1), 27-37.

Rens, K.L. and Transue, D.J. (1998). "Recent Trends in Nondestructive Inspections in State Highway Agencies," *J. Perform. Constr. Facil.,* 12(2), 94-96.

Rens, K.L., Wipf, T.J. and Klaiber, F.W. (1997). "Review of Nondestructive Evaluation Techniques of Civil Infrastructure," *J. Perform. Constr. Facil.*, 11(4), 157-160.
Ressler, S.J. (2000). *West Point Bridge Designer, Dept. Civ. and Mech. Engrg.*, US Military Academy, West Point, NY, v.4.06.
Sansalone, M., and Streett, W. (1997). *Impact-Echo: Nondestructive Evaluation of Concrete and Masonry*, Bullbrier Press, Ithaca, NY.
Santamarina, J.C. and Chameau, J.L. (1989). "Limitations in Decision Making and System Performance," *J. Perform. Constr. Facil.*, 3(2), 78-86.
The White House (2006). *The Federal Response to Hurricane Katrina: Lessons Learned,* The White House, Washington DC.
The Forensic Anthropology Data Bank (2005). http://web.utk.edu/~anthrop/FACdatabank.html, last accessed July 2005.
Washer, G. (2000). "Developing NDE Technologies for Infrastructure Assessment," *Public Roads*, 63(4), 44 – 50.
Yao, J.T.P. (1985). *Safety and Reliability of Existing Structures*, Pitman Advanced Pub., Marshfield, MA.

Table 1 The failure probability for each bridge member for 207 cases

Member number	Occurrences	Failed cases	Causality probability	Square sum residuals S	% ratio to total S
1	11	4	0.364	105.5	4.89%
2	13	7	0.538	102.0	4.73%
3	14	4	0.286	107.1	4.97%
4	10	5	0.5	102.8	4.76%
5	5	1	0.2	108.9	5.05%
6	8	1	0.125	110.5	5.12%
7	5	4	0.8	96.8	4.49%
8	11	8	0.727	98.2	4.55%
9	10	5	0.5	102.8	4.77%
10	7	2	0.286	107.1	4.97%
11	11	10	0.909	94.6	4.39%
12	18	9	0.5	102.8	4.77%
13	10	8	0.8	96.8	4.49%
14	8	1	0.125	110.5	5.12%
15	8	7	0.875	95.3	4.42%
16	10	4	0.4	104.8	4.86%
17	12	5	0.417	104.4	4.84%
18	21	5	0.238	108.1	5.01%
19	7	5	0.714	98.5	4.57%
20	3	1	0.333	106.2	4.92%
21	5	5	1	92.9	4.31%