16TH EUROPEAN CONFERENCE ON EARTHQUAKE ENGINEERING
THESSALONIKI 18 - 21 JUNE 2018

PERFORMANCE-BASED DESIGN PROCEDURES: BEWARE UNCHARTED WATERS

Terrence PARET[1], Andrew SHUCK[2]

ABSTRACT

Performance-based design (PBD) is a term used to describe techniques that are commonly promoted as rational improvements --- relative to prescriptive codified procedures --- for evaluating and upgrading existing buildings, and for designing new buildings, to resist earthquakes. Recent examples of such procedures are included in Eurocode 8 - Part 3 (Assessment and Retrofitting of Buildings), ASCE/SEI 41-13 (Seismic Evaluation and Retrofit of Existing Buildings) and PEER Guidelines for Performance Based Seismic Design of Tall Buildings. Without any PBD tall buildings having been field-verified by design level earthquake shaking, confirmation that PBD actually delivers reliable performance for tall buildings is limited to analysis simulations and laboratory testing. PBD remains, therefore, a hypothesis founded on engineering theory but not empirics. Development of an empirical basis for PBD, even if indirect, is therefore essential before major cities become saturated with PBD structures, especially those with critical function and high-occupancies. Accurate retroactive prediction of damage to non-PBD structures during prior earthquakes using PBD methods is an example of indirect empirics.

Structures that have been damaged by earthquake ground motion have been studied by many engineers in recent decades, often using methods whose specifics pre-date the publication of the most recent PBD procedures, but whose underlying concepts and analysis techniques have been carried forward into current PBD methodologies. These earlier studies demonstrate that post-earthquake "predictive" studies regularly failed to match the damage observed in the field, often by wide margins, casting doubt on the use of these methods if reliability is the *raison d'être* for employing PBD methods. More recent attempts, including well-subscribed blind prediction studies demonstrate similar chasms between predicted and actual performance, highlighted by extreme variance.

In light of this, logical questions that the engineering community should be asking are:

- "What justifications exist for using PBD to design and assess tall buildings other than PBD's theoretical bases, and how sound are they?"
- "Given the apparent limitations of the available PBD methodologies to accurately predict performance, what are the risks of using PBD as the primary methodology for designing tall buildings?"
- "Are engineering professionals adequately communicating the benefits and shortcomings of PBD to others (e.g., engineering colleagues or students, jurisdictional authorities, and laypersons)?"

The scope of this paper is to explore these subjects and begin to address these questions from various perspectives with the goal of encouraging engineers to revisit commonly held assumptions regarding PBD. Considered issues include whether PBD analysis procedures, as currently articulated in the most recent published standards and guidelines, are likely to meet with better success in preventing life-threatening damage from earthquakes than was possible earlier. In addition to the stated objectives, particular attention will be given to the statistical significance of the latest PBD procedures, which is widely-misunderstood. The paper will also touch on the significance of various differences between widely used analysis/design procedures and tools; and on the influence of variations in ground motion, from recorded to design to actual ground motion at the building site.

Keywords: Performance-based design, response prediction, design uncertainty, ground motion dispersion

[1]Senior Principal, Wiss, Janney, Elstner Associates, Inc., Emeryville, U.S.A., TParet@wje.com
[2]Senior Associate, Wiss, Janney, Elstner Associates, Inc., Emeryville, U.S.A., AShuck@wje.com

1. INTRODUCTION

Over the years, we have consistently urged caution with regard to the growing application of formalized performance-based design (PBD) methods for design and assessment of important or significant structures (Searer et al. 2008; Freeman et al. 2008; Paret et al. 2011). With the skylines of major cities in regions of high seismicity quickly becoming dominated by PBD structures, many of which have, given the absence of field verification during extreme events, “experimental” lateral force resisting systems, the need for caution is quite urgent. Yet as the construction of PBD structures is proliferating, the engineering profession appears to now be so enamored with the PBD paradigm that it can neither remember its past nor see beyond the glare of PBD.

From its humble beginnings and through very recent times, structural engineering has always been rooted in empirics. At its heart, structural engineering is applied empirics even more than it is applied science. Structural advances throughout the last millennium were largely founded on incrementally extending long-established precedent, essentially relying on past performance of what worked and what did not, i.e. empirically-derived proof-of-concept. This is especially true for structures whose primary role was to support gravity loads. Laboratory testing, another avenue for empirically deriving proofs-of-concept wherein designed structural components or assemblages are subjected to controlled loading regimens and their behavior is observed and measured, became central to the advancement of engineering for gravity, wind and earthquake loading in the last century. Such laboratory experimentation was the backbone that enabled, for example, the development of the diversity of structures and structural types that fostered the growth of twentieth century cities. Also in the last century, empirical study of structural performance in the aftermath of earthquakes, wherein the primary characteristics of the structures under scrutiny are more or less known but the loading to which they were subjected was uncontrolled, came to underlie the late twentieth century development of improvements in engineering and construction technology for earthquake resistance. Without question, empiricism is not just at the root of engineering, it is the thread from which the cloth of engineering is woven. Yet today, the earthquake engineering profession is rushing headlong away from empirics, the only proven means for establishing engineering truth to a reasonable degree of certainty, toward PBD. We continue to advise caution: when we stray from empirics, we stray into uncharted waters.

The “experimental” earthquake-resisting systems on which so many of today’s performance-based structures rely often have little or no empirical basis; they are theoretical constructs made real by investors, developers, engineers and contractors who, other than the engineers, are all largely unaware of the lack of empirical underpinning of those constructs and their explicit reliance on highly idealized simulation to show they work. Outside of this inner circle, few know that the “experiment” that makes these structures truly “experimental” is expected to come long after they are constructed --- when they are someday subjected to a design-level earthquake shaking, long after the engineers responsible have left the scene. In this way, we are unfairly experimenting with the lives of the individuals who will occupy these structures as well as with the communities that rely on them --- communities whose planning and permitting-agencies are as seemingly enamored with derring-do and growth as they are ignorant of these structures’ lack of empirical underpinning. To make matters even worse, a disproportionate share of these skyline-dominating performance-based structures have high occupant loads and/or critical lifeline functions, like stadiums or bridges. While laboratory studies of the nonlinear behavior of some of the components of these systems under idealized conditions have been conducted in many cases, validation of system-wide performance of the majority of these has never been attempted in any laboratory, neither private nor public nor mother nature’s, whose earth-shaking potential we estimate with only wispy basis. It would be a stretch to characterize many of these systems as mere incremental steps built upon prior experience with similar proven strategies for earthquake resistance: many of these systems, since they do not resemble anything that has ever been relied on before to resist earthquakes, would have been beyond imaginable to nearly all of us 20 years ago. Under these circumstances, it is hard to argue that extra precaution is unmerited, yet from the authors’ perspective, extra precaution is being meted out in miniscule doses. What will it take to bring to the forefront of the PBD discussion that there is a wide chasm between what we are actually doing as a profession and what we say to society we are doing?

2. IT IS IMPRUDENT TO INSINUATE THAT WE CAN FORECAST HOW OUR BUILDINGS WILL RESPOND TO FUTURE LARGE EARTHQUAKES --- FIVE REASONS WHY

2.1 Reason One: *Published studies using common industry analysis methods routinely fail to generate reliable and accurate predictions of earthquake damage to real structures.*

Over the decades, leading practitioners and researchers have conducted studies, in hindsight, to try to analytically post-predict damage that had actually occurred to real buildings during real earthquakes (Alali et al. 1995; ATC 1996; Bayhan, Gülkan 2011). The goal of these studies was to see if damage that was observed to have occurred during an earthquake could have been predicted analytically with sufficient advance knowledge of the structure and with recordings of actual earthquake ground shaking and structural response. Because structural engineers are not prescient, we will never have such knowledge in advance of the design of any actual PBD structure, so these studies represent a better-than-best-case scenario with respect to the information available at the time of design. Design of a PBD structure requires engineers to be able to generate accurate blind predictions, not merely predictions in hindsight using data that does not exist during the design process.

One of the earliest recordings, the El Centro earthquake, was used as a basis for seismic "predictions" for many years until the San Fernando, Loma Prieta and Northridge earthquakes of 1971, 1989 and 1994 provided a wealth of new recordings. Notable buildings that have been evaluated and published include the following: Imperial Service Building (Rojahn, Mork 1981), Olive View Hospital (Murphy et al. 1973), Holiday Inn (PEER 2005), Oakland building (Freeman, Paret 2000), and similar buildings in three cities in Turkey (Bayhan, Gülkan 2011), but numerous other published studies exist documenting attempts by academics and practitioners to predict real earthquake damage after-the-fact (ATC 1996; Alali et al. 1995).

In essentially every one of these studies the computer analysis was not able to accurately corroborate the observed damage, although the primary damage mechanisms were, in many cases, identifiable. Some promising results were achieved, but the promising aspects resulted only from iteratively generating and refining the analyses to achieve predictions that matched observations made after the earthquake. The iterative process was assisted by close examination of the recorded building response which revealed certain features of the building response characteristics that were not obvious a priori. These characteristics included an initial fundamental period that was shorter than initial modelling indicated, and which resulted from the structural model having neglected to account for certain nonstructural elements, closing of gaps between masonry infill and columns, as well as portions of the structure that contribute significantly to lateral stiffness (e.g. the contribution of concrete slabs to the stiffness of the horizontal framing elements). Examination of the recorded response motion, which would of course not be available during design of a PBD structure, was therefore necessary to identify the lapses in the model. When these and other observations were accounted for, the analysis could be interpreted as generating predictions that were reasonably close to the observations, but this highlights the difficulty of predicting the actual performance of a building without having the actual input and building response time histories.

The authors know of no published instance in which an a priori analysis of a building was able to generate results that closely matched damage generated by a prior earthquake, despite these published studies being generated by leaders of the profession. Moreover, the studies most often employed structural analysis models that were assembled with reasonable care in a commercial engineering office that from time to time might be engaged in performance-based engineering design; thus, they are representative of what must be presumed to be a standard of care equal or exceeding that which would be exercised during the design of an important PBD structure. Each participant in these studies most typically had access to reasonable representations of the ground motions to which the actual structures had been subjected, with the ground motion input taken directly from nearby recorded motion being or developed by a leading seismologist specifically to translate available recorded motions to the site of the buildings being studied, and the analytical methods employed were either those selected by the leading engineering professionals who were conducting the studies, or specific methods assigned by the studies' expert panels for the purpose of vetting state-of-the-art procedures. The studies employed the construction drawings for the building, often supplemented by field measurements and observations of the actual construction in the development of structural models. Most importantly, these studies had advance detailed knowledge of the actual behavior and damage to these structures as documented after the earthquakes. Gülkan's study outshines the others in its elegance in that its subjects were three nearly identical buildings in different parts of Turkey that had each been subjected to relatively intense ground shaking from three different earthquakes, and nearby recordings of ground motions from those earthquakes were available and used in the study, but still the accuracy of the predictions of even global performance left much to be desired (Bayhan, Gülkan 2011).

In short, none of these analytical studies accurately predicted the documented structural behavior of their subject buildings, thus failing the dispositive test of whether the profession is up to the task of predicting earthquake damage in real structures, either blindly or with full advance knowledge of the ground shaking input, the structural configuration and details, and the structure's response. To be clear, and fair, all of the studies referenced displayed some degree of predictive ability, but those abilities were always wanting. For example, in some cases the critical story levels were identified, but often not consistently, and the specific elements damaged most critically, or the correct failure modes, were rarely identified. However, as the engineering profession is prone to do, the outcomes were typically characterized in the studies themselves as "glass half full" success stories rather than "glass half empty" failures, though the predictive outcomes were often dismal.

Characterizations of this type suggest that as a profession we have lost sight of the big picture and strayed from our mission, which is to provide the public with earthquake-safe structures. While it is true that the poor predictive ability demonstrated by the cited studies sets before us an imperative to direct our intellectual might to improve the accuracy of our damage predictions, this situation also presents us with a moral dilemma. The question before us is not really whether the sanctioned PBD methods (as employed in a practitioner's environment) are sufficiently robust to be able to proclaim that the methods have some modest predictive ability, but rather, whether we can morally justify building cities of gloriously architected and performance-based engineered structures --- all the while proclaiming that the structures and their occupants are earthquake-safe. If our designs are based on the same analysis methods as the studies cited, but are without benefit of the full advance knowledge that benefited the aforementioned studies (which themselves proved our <u>inability</u> to predict what had already happened), then how do we justify selling the false promise of earthquake safety that is embodied by the PBD paradigm?

2.2 **<u>Reason Two:</u>** ***Blind prediction contests are objective evidence of the failings of our predictive abilities***

Over the past decade, blind prediction contests of a variety of structural types have been sponsored by a number of entities. In principle, blind prediction contests involve laboratory-built structures that are subjected to known excitation on a shake table. Prior to subjecting the structure to the excitation, the contest sponsors release all known relevant design parameters to the subscribed expert prediction teams, including the time histories of the shake table motions that the structures will be subjected to. Although there have only been relatively few such contests whose results have been well-disseminated, these have been well-advertised and well-subscribed with a significant number of expert practitioners and academic teams participating. The expert teams have typically employed some of the most sophisticated seismic response prediction software available and predictive methods that are far more detailed than those typically employed in common design practice. It would therefore not be an overstatement to characterize these contests as representing the very best attempts at response prediction that the structural engineering profession can offer. It would also not be an overstatement to characterize these contests as the most objective measure available of our collective ability to predict seismic response since the contest is "blind", i.e. the participants have no opportunity to adjust their model to more closely match the performance of the laboratory specimen prior to submitting their predictions. In the sense that the contests require prediction of performance before the fact, they simulate the design profession's real assignment. They are, of course, also far simpler than what the design profession is charged with because the specific input motion is known in advance, the construction quality of the test specimen is more tightly controlled than for structures built for commercial purposes, the material properties of the test specimen are known in detail, and in some contests, there are no nonstructural interactions that need to be considered because the structures are bare. Moreover, the structures are far simpler and far more regular than is the vast majority of commercial-purpose structures.

Two of the better publicized blind prediction studies were organized by the Pacific Earthquake Engineering Research Center (PEER) in 2010 and by the E-Defense Steel Building Project in 2007 (Terzic et al. 2015; Pavan 2008; Ohsaki et al. 2008). In the PEER contest, a single full-scale reinforced concrete bridge column was subjected to six consecutive unidirectional ground motions and forty-one teams from fourteen different countries using 12 different analysis engines participated, with 17 of the teams registered as professional engineers and the balance registered as researchers. In the E-Defense contest, a full-scale regular four-story building with composite floors and a reinforced concrete roof deck was subjected to scaled time histories of the Takatori record from the 1995 Kobe Earthquake and many professional and research teams also attempted to predict the structure's behavior.

In the PEER contest, there was wide dispersion in the results that varied according to the parameter being tracked. Predictions of displacement were significantly biased on the low side relative to the experimental results while predictions of response acceleration were significantly biased on the high side. (Fig. 1) These prediction errors suggest that nearly all the participants significantly underestimated damage and nonlinear response, which is particularly strong evidence of why performance-based engineering is more myth than reality.

Dismayingly with respect to PBD, residual displacement at the end of each ground motion was found to not be predicted with an acceptable degree of accuracy and had the greatest dispersion compared to all other response quantities, with the mean bias for the majority of the earthquakes exceeding 40 percent (Fig. 2). Moreover, the mean predictions for residual displacement were consistently lower than the experimentally measured residual displacement. Residual displacement being one of the best quantities for characterizing performance, this indicates a poor ability of the best in our field using the best available tools and methods to predict performance even of a single-degree of freedom system.

In the E-defense test, submissions were accepted from 47 analysis teams, consisting of 30 researchers and 17 practicing engineers. When the model was analyzed and the results compared to the shake table test result, the analysis results were also found to poorly represent the testing of the physical structure, with errors more than 100 percent. (Table 1)

In producing their results summary, the organizers of the contest removed "exceptional" data that was "obviously wrong" in order to produce useful statistics; this resulted in almost 13 percent of the results for certain parameters (e.g., overturning moment) being discarded offhand, an inauspicious start. Despite this culling of real participants' responses, the results are still appallingly dispersed. For example, the interstory drift angle has a mean, maximum, and minimum RMS of 1.36, 4.39, and 0.192, respectively, which suggests that each analysis is nothing more than a mathematics-based guess. This blind prediction study and others like it provide incredibly compelling evidence that we should not be pretending that the analysis results we generate via PBD methods correctly simulate the interaction between engineered structures and earthquakes. The necessity for physical validation of a hypothesis prior to it being declared to be factual is one of the first lessons taught in any fundamentals course in science, but contrary to all objective evidence, today's engineers have prematurely proclaimed PBD analysis tools to be sufficiently predictive for design, selling a "dream" to the general public and profiting handsomely while not letting on that as a profession we are not capable of accurately or reliably predicting performance of even simple structures with known ground motions *a priori*.

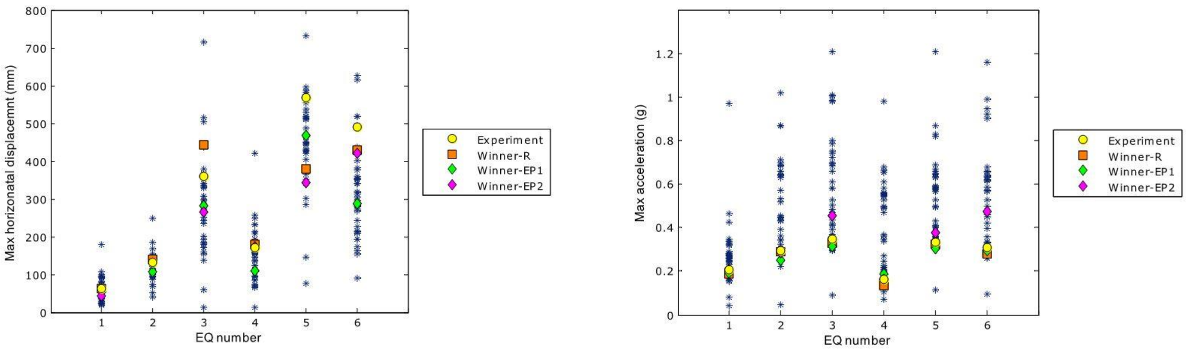

Figure 1 - Plots of predicted maximum displacement (left) and maximum acceleration (right) at the top of the column for each earthquake in the PEER contest (Terzic et al. 2015).

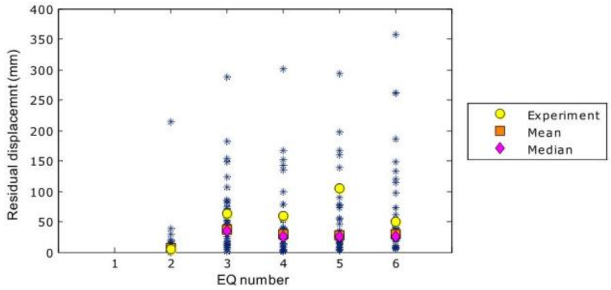

Figure 2 - Plot of predicted residual displacement at the top of the column after each earthquake in the PEER contest (Terzic et al. 2015).

Table 1 - Error evaluation in predicting relative displacements in different analyses (Pavan 2008).

TEST	LEVEL	Displ. NS (mm)	error %	Displ. EW (mm)	error %
MEASURED	1 (GL)	0		0	
	2	75		47	
	3	135		95	
	4	177		134	
	5 (Roof)	200		157	
PRE TEST	1	0		0	
	2	128	71%	105	123%
	3	218	61%	185	95%
	4	282	59%	247	84%
	5	321	61%	292	86%
POST TEST	1	0		0	
	2	171	128%	114	143%
	3	259	92%	175	84%
	4	329	86%	232	73%
	5	373	87%	272	73%

2.3 **Reason Three:** ***Demand side uncertainty makes for bad science***

Like most other natural phenomena, no two earthquakes, past or present, are exactly the same. Each earthquake involves a portion of a fault interface violently releasing energy and shaking the ground in a distinct way. Nonlinear time history analysis conducted under the PBD umbrella has, until recently, typically involved subjecting a building to no fewer than seven different ground motions that are loosely derived from some previously recorded event on a faraway fault. Thus, in the exercise of PBD, structures are designed for ground motions that they are guaranteed to never experience. We do not suggest that this nod to experiential engineering is entirely misplaced; it is possible to learn a great deal about a building by analyzing its response to ground shaking that corresponds in some probabilistic sense to site seismicity. The flaw is that while this process may well reveal much about the structure that is relevant, it also may well leave the structure with a critical inestimable vulnerability to the actual ground shaking that will eventually occur -- which is not how PBD is being sold to society. The proposition that we can design structures to specific performance targets with reasonably high confidence when, for example, we are relying on analytical models that are subjected to a handful of ground motions having no known relationship to the design event that may actually occur, and when that handful of motions may simply be relatively small ground motion recordings that have been scaled up and adjusted to the analytically-derived building period, would likely not pass as “good science” in any other field (ASCE 2017). However, the industry has plowed ahead with implementing a methodology that does not appear to us to be ready for “prime time”, assuring those outside the industry that designs created using ground motions that have been scaled, matched, and synthesized will be safe in large earthquakes when no one really knows that safety is assured.

Examples of the hubris that underlies ground motion prediction and probabilistic seismic hazard analysis (PSHA) as a whole abound in recent history. The 2010 Darfield Earthquake, with a moment magnitude of 7.1, roughly realized the New Zealand Standard (NZS) design spectral acceleration demands for downtown Christchurch and the 2011 Lyttleton Earthquake, with a moment magnitude of 6.2 significantly exceeded the NZS spectral acceleration for the “maximum considered earthquake,” or 2,475-year event, as seen in Fig. 3 (i.e., the design earthquake and more than 1.5 times the design earthquake, respectively). That these two allegedly rare events occurred only five months apart on previously-unknown fault structures should give us all pause (Bradley et al. 2014). In addition, many other modern events, such as the 1994 Northridge earthquake also occurred on previously “blind”, or unknown fault segments, which is problematic, due to the necessity of identifying all contributing seismicity sources if the PSHA methodology is to have any validity. It is not as though earth scientists are simply putting the finishing touches on a database of the characteristics of all seismic structures in existence - they are discovering new sources of earthquakes all the time, as well as earthquakes with characteristics never before observed. The uncertainty associated with this evolving understanding of seismicity does not go away simply because the structural community chooses to assume that the inputs to their design process are known, and chooses to characterize their performance-based designs as earthquake-safe even knowing that the demand side of the equation is at best incomplete. If we are going to represent seismicity and structural performance probabilistically, and tell the world that our methods are sufficiently robust that we have reduced the uncertainty with respect to our ability to design earthquake-safe structures to an acceptable level, we would do well recognize and to publicly acknowledge both the underlying fallacy of the approach and the plain meaning of the word uncertainty. We cannot fairly represent to the general public that we can quantify uncertainty for data sets of unknown size and characteristics, regardless of what the statisticians tell us.

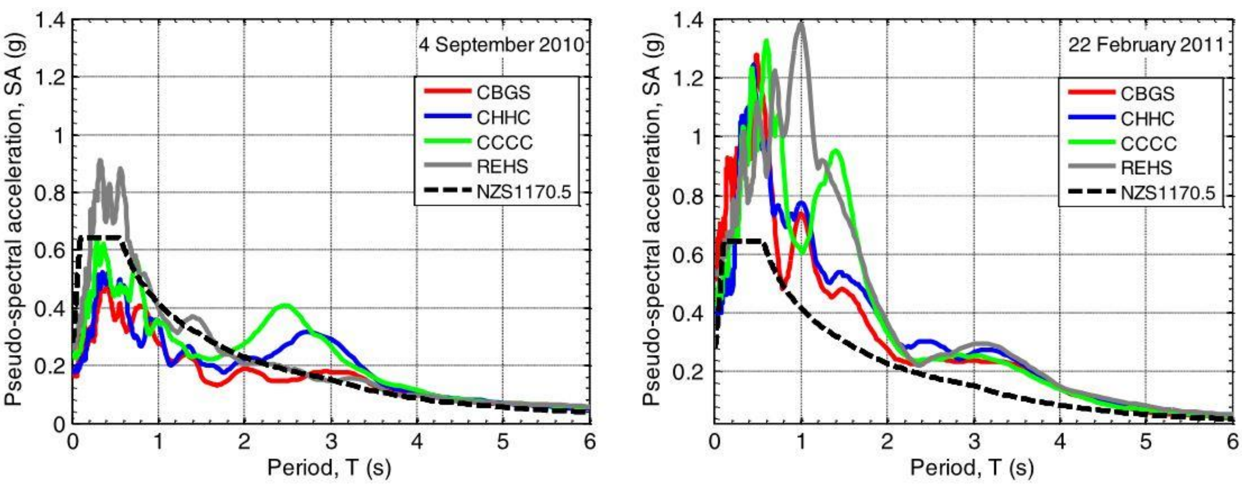

Figure 3 - Geometric mean recorded ground motion spectra for the September 2010 event (left) and February 2011 event (right) relative to the New Zealand Standard 1170.5 500-year return period design spectrum (the black dashed lines) (Bradley et al. 2014).

The cornerstone of the state-of-the-art of PSHA, the NGA West 2 Ground Motion Prediction Equations (GMPE), were developed as regressions of more than 21,000 acceleration response spectra of recorded ground motions. However, nearly 40 percent of these spectra have been found to not reflect rational ground shaking properties: namely, the spectra do not asymptotically approach the peak ground displacement at very long periods as they should --- if we agree that basic structural mechanics must be satisfied (Malhotra 2015). As a result, the industry's latest-and-greatest means of selecting ground motions is based on a flawed population of ground motions, which relegates all nonlinear time-history analyses conducted under the umbrella of PBD to a "garbage in-garbage out" reality. As a result, these widely-used GMPEs generate spectral velocities and displacements that fail to match the basic physics of the problem, so how can we expect them to generate realistic predictions of earthquake response?

For the vast majority of design engineers, ground motion acceleration time histories are an abstraction, something handed to them by the geotechnical engineer on their project. This is, hopefully, less true for PBD designers, but they too exhibit a lack of critical introspection regarding ground motions themselves (and seismic hazard issues in general) that would certainly give an outside-the-industry stakeholder pause --- if only they knew and understood what is common in our industry. For instance, peak ground displacements (PGD) in near-fault zones are significantly biased by common signal processing techniques: this bias can create a PGD ratio of two-to-one between unprocessed and processed records, respectively (Kamai, Abrahamson 2015). Even more alarming to the authors is the abject neglect of peak ground velocity (PGV) and PGD within the PSHA methodology which is beholden mostly to PGA. Even where PGV and PGD demands are the ground motion parameters of greatest structural concern, these are generally byproducts of a PGA-based methodology, single and double integrals of ground acceleration, rather than drivers of the ground motion selection process. The fundamental flaw is that PSHA-based ground motion selection is rooted in the assumption that PGA, PGV and PGD have similar relationships to earthquake magnitude, but this assumption is demonstrably false. There is little correlation between PGA and PGD for large magnitude earthquakes, especially at near-source locations; while PGA saturates with magnitude, PGD does not. In essence, there is a great deal more uncertainty associated with predicting PGD for near-fault sites than PGA (Yamada et al. 2009). We are greatly concerned that the buildings most affected by large ground displacements are long-period, near-source structures, like high-rises and seismically-isolated structures in cities with nearby faults, which is the same inventory of structures most likely to be designed under the PBD paradigm. Currently, we see that ground motion selection for some of these structures is targeting PGD, but the above-noted poor correlation across magnitudes is still not being accounted for when records are scaled.

In sum, we believe that the myriad sources of seismic demand uncertainty, which are only broached in passing above, seriously undermine the representation that engineers are today able to develop accurate quantifications of seismic risk and structural performance, which are bedrock PBD concepts.

2.4 Reason Four: *Embracing the absurdity of it all - different answers mean no one really knows*

The engineering profession is obsessed with analysis-based numerical simulations, employing them as surrogates for empirical evidence at every turn. The reason is obvious: there is a dearth of empirical data on the actual performance of structures designed using performance-based methods during strong earthquake shaking, so no one really knows how any performance-based building will perform. Simulations are our only recourse, the only feasible, cost-effective alternative available. But the absence of empirical data and the feasibility of conducting analyses does not justify the construction and occupancy of cities whose premise of earthquake safety is founded almost entirely on simulations. We now develop the designs for the grandest most daring structures ever built by relying entirely on simulations, then we validate that those designs will work using more simulations. Then we peer review them, sometimes using more simulations. Yet the one proven element in all of this is that simulations are neither accurate nor good performance predictors, even when we know the outcome. While it seems that the old twentieth century saying, "Better living through chemistry" has been transmuted by engineers today to "Better living through simulations," it was a sales pitch then and it remains a sales pitch now and, unfortunately, the engineering profession has swallowed too much of its own kool-aid.

The nature of the problem should be readily apparent to anyone who is looking, anyone who cares to look. In the US, ASCE 41 has become the de facto standard for performance-based assessment and design and the PEER Tall Buildings guidelines are now referenced regularly for the design of tall buildings in regions of high-seismicity on the west coast (ASCE 2014; PEER 2010). Other design guidelines like the Los Angeles Tall Buildings procedure are accepted in Los Angeles (LATBSDC 2015). Each of these documents sets forth different "rules" for simulation performance and thus for validation that a design is appropriately earthquake-safe. ASCE-41 itself sets forth four different analytical methods. Eurocode 8 as well allows for a suite of analysis methods to be used (BSI 1996). Obviously, none of these different methods could be expected to generate precisely the same predictions of behavior for any specific building, and practitioners well-versed in the use of these documents know that they do not. In the discussion below, we focus on this issue at several levels: document-to-document methodological disparities; assessment technique-to-assessment technique disparities within each document; modeling technique-to-modeling technique disparities between the users of the methodologies and assessment techniques; and analysis software-to-analysis software disparities between different analysis engines in common use across the globe.

At its simplest, it should be clear that each of these documents sets forth similar methodologies, though there are clear differences between them. Indeed, there would be no need for these different documents if the methodologies were precisely the same. The differences in the methodologies are significant enough, or thought to be by the various groups that authored these various documents, that there is regular professional tussling over the various provisions, and it might be said that a friendly competition exists between some of them. Moreover, to some degree, the documents fill different niches and are not interchangeable. While ASCE-41 is taken to have general purpose applicability, PEER and LATBSDC are defined for tall buildings only. Yet the latter are not simply excerpts from ASCE-41; they have provisions that the authors believe are especially needed to accurately assess and design tall buildings for earthquake resistance. How is the profession to reconcile this? Is it appropriate to use ASCE-41 for assessment and design of very tall buildings, or is it not? What is the added risk to society of using one of these documents over another? Does one of the documents really produce "better" assessments and designs? Returning to our theme, it is clear that these questions are all rhetorical because we will never be able to answer them until the big experiment takes place when a city full of function-critical structures designed using these various documents is subjected to strong earthquake shaking.

As one strong indication that our profession is blindly plowing ahead with design of function-critical structures without paying heed to the most obvious analytical limitations and their lack of readiness for prime-time, one only needs to examine the findings of researchers and professionals who have studied and compared damage predictions generated by various assessment techniques (ATC 2005). We view some of their findings to be nothing short of terrifying. As only one example and as summarized in FEMA 440, various researchers have reported that one commonly used nonlinear analysis technique set forth in ASCE-41, the Capacity Spectrum Method (CSM) overestimates displacement response --- while other researchers have found that it underestimates response. How can it be that researchers cannot even agree as to whether an assessment method employed worldwide overestimates or underestimates response? This inconsistency is not an indictment of CSM but rather is an indictment of the ability of engineering research community even to competently evaluate the efficacy of one of its sanctioned performance-based analysis methods. *If in assessing CSM we cannot even achieve a consensus as to whether it generates overly conservative or unconservative estimates of building displacement response, what business have we designing structures using any performance-based method*?

Assessment technique-to-assessment technique disparities that exist within each of the cited published documents are also an issue. The techniques within each document typically extend across a wide range of complexities from linear static to nonlinear dynamic. One need not ask if these different techniques would produce different designs or different results when a given structure is being assessed for adequacy; we all know they produce inconsistent predictions of performance. Indeed, how could a linear dynamic method or a nonlinear static method generate the same results as a nonlinear dynamic analysis method? If either did, no one would ever opt to undertake the more expensive, more time-consuming analysis types. To explain the fact that each method yields different answers, most professionals will argue that the linear methods are merely "more conservative", but this is a cheap out that projects a mis-begotten air of confidence when none of us really knows if either method will generate reliable predictions of performance when applied to a structural system that has never been field-tested.

Modeling technique-to-modeling technique disparities between the users of the various methodologies and assessment techniques, as well as analysis software-to-analysis software disparities between different analysis engines in common use across the globe, also create a quandary for anyone examining this issue. Simply put, practitioners executing the most state-of-the-art methods at their disposal, in circumstances where cost is virtually no object, generate radically dissimilar predictions of performance. The dissimilarities are often not trivial; they rise to the level of life-and-death level significance, for hundreds if not thousands of future occupants. For example, construction of a 48-story, mixed use, reinforced concrete shear wall building in Las Vegas was abruptly terminated at the 26th level when the global financial crisis occurred and defects in the as-built reinforcement were identified. After the work stoppage, numerous expert teams -- many of whom are luminaries in the development of codes and standards pertaining to earthquake engineering and who played prominent roles in the development of the performance-based methodologies mentioned above, were retained to study the building and its seismic adequacy. Their simulations all relied on nonlinear dynamic simulation methods and all used software, such as PERFORM® 3-D by Computers and Structures, Inc., which is widely recognized as an industry-leading performance-based analysis tool. Some used ASCE-41, some used the PEER Tall Buildings Guidelines and some used the Los Angeles Tall Buildings document. In support of the theme of this paper, the breadth of the predictions generated from these simulations was staggering and ranged from total collapse, including collateral damage to adjacent structures from the collapse, to modest damage that would normally fall well-within expectations for engineered structures subjected to a design event. Some experts' simulations led them to recommend complete and immediate demolition of the building while other argued that virtually nothing needed to be done except to correct some localized weaknesses. Even understanding that these experts were representing different parties, their simulations all relied on the same basic model geometries and material strengths, and all could be tied to modelling methods supported by the documents and guidelines on which each simulation was based. At a minimum, the scatter in these results should give pause to all who promote performance-based engineering methods as a methodology that reduces the uncertainty of the outcome to tolerable levels.

2.5 **Reason Five:** ***PBD yields results with poorer than a 70% confidence level***

Members of the National Earthquake Hazards Reduction Program (NEHRP) Provisions Update Committee (PUC) issue team that developed the change proposals for Chapter 16 of ASCE/SEI 7-2016 published a series of articles intended to educate engineers about the reasons for the changes and their practical effects. One of the changes involved increasing the minimum number of required simulations, from seven to eleven. The first article in the series, which covers the methods for ground motion selection and scaling, states that, "results for fewer [than eleven] ground motions demonstrated significantly more variability. We recommend the use of a minimum of eleven motions based on the FEMA (2012) findings and the judgment of the team." As a point of reference, the paper cites a FEMA study (FEMA P-58): "when eleven motions are used, mean response parameters (primarily story drift) are predicted within 30% at a 70% confidence level" (Haselton et al. 2017). While not identical, FEMA P-58 states, "when ground motions are selected without consideration of spectral shape, analyses using eleven pairs of properly scaled motions can provide a reasonable estimate of median response, defined as 75% confidence of being within plus-or-minus 20% of the median (Huang et al., 2011)" (FEMA 2012). Who among you are today using PBD to design a high-rise are aware of and comfortable with this statistical confidence? Who among you has any information that the confidence is any better?

If the above is not enough to concern PBD practitioners, we call attention to certain caveats and assumptions behind the above-noted statistics, in the words of the researchers citing FEMA P-58:

> Simple bilinear models were used [to generate the statistical conclusions cited] although the writers recognize that deterioration of strength and stiffness will affect distributions of displacement and acceleration response. Yield strengths were set at infinity, 0.40W, 0.20W, 0.10W, and 0.06W to represent, albeit simplistically, conventional and isolated (0.06W)

> construction, where W is the reactive weight of the structure. For the oscillators with yield strengths of 0.10W and larger, the elastic period ranged between 0.05 s and 2 s; for the oscillator with a yield strength of 0.06W, the post-yield (isolated) period ranged between 2 and 4 s… Using bilinear SDOF oscillators will limit the utility of the results and observations presented subsequently to low-rise, code-conforming, regularly configured buildings whose displacement response (and thus damage) is dominated by one mode of response. (Huang et al. 2011)

It ought to be troubling to all of us that the statistical basis for the PBD methodologies set forth in FEMA P-58 and ASCE 7 is the cited research, to wit:

- <u>The achieved statistical confidence achieved when 11 suites of ground motions are used, (e.g., results "predicted within 30% at a 70% confidence level") is absurdly low relative to common engineering expectation that 90 or 95 confidence is achieved for most codified design</u>. Using the equations and uncertainty in the Huang et al. research, an analysis would require 81 ground motions to achieve 90 percent confidence of being within plus-or-minus 10 percent of the "true" median response --- and even then would be achieved only within the limits of the study, "*low-rise, code-conforming, regularly configured buildings whose displacement response (and thus damage) is dominated by one mode of response*." Given this, how can anyone justify using 11 suites of ground motions as the basis for PBD for a high-rise?
- <u>These statistics and the stated limits of their applicability are not common knowledge among engineering practitioners who are using or considering using PBD.</u> We admit to being unaware of the statistical confidence levels stated in the reviewed research despite experience with multiple PBD projects and involvement in design provision and building code development. Information like this has not been but must be widely disseminated to practicing structural engineers and to stakeholders so that everyone can make informed decisions about using PBD.
- The research by Huang, et al. employed a ground motion selection and scaling technique called distribution-scaling. To the authors' knowledge the distribution-scaling ground motion technique is not commonly used in practice. Huang, et al. make it clear, however, that other ground motion selection and scaling techniques cannot provide a reliable estimate of "true" response because they do not explicitly consider the distribution of spectral acceleration present within the hazard. The ground motion selection methods of Chapter 16 of ASCE/SEI 7-16 do not even include distribution-scaling. It is therefore entirely unknown if the statistical confidence associated with ground motion selection methods that are in common use even reach the very low confidence threshold noted by the Huang et al. research.

In sum, we are alarmed and mystified at the use of PBD to justify the design of high-rise buildings, as well as the continued push to broaden the use of PBD, given that there is neither empirical nor acceptable statistical basis for it. Beware uncharted waters!

3. REFERENCES

Alali A, Anderson J, Beck J, et al (1995): Report No. SAC 95-04, Part 2, Technical Report: Analytical and Field Investigations of Buildings Affected by the Northridge Earthquake of January 17, 1994. *SAC Joint Venture,* Sacramento, California.

American Society of Civil Engineers (2014): Seismic Evaluation and Rehabilitation of Existing Buildings. *ASCE Standard ASCE/SEI 41-13,* Reston, Virginia, USA.

American Society of Civil Engineers (ASCE) (2017). "Minimum design loads and associated criteria for buildings and other structures." *ASCE/SEI 7-16*, ASCE, Reston, VA.

Applied Technology Council (1996): Seismic Evaluation and Retrofit of Concrete Buildings, *California Seismic Safety Commission, Report No. SSC 96-01, ATC-40, Applied Technology Council,* Redwood City, California.

Applied Technology Council (2005): FEMA 440 - Improvement of Nonlinear Static Seismic Analysis Procedures, *Applied Technology Council,* Redwood City, California.

Bayhan B, Gülkan P (2011): Buildings Subjected to Recurring Earthquakes: A Tale of Three Cities. *Earthquake Spectra, 27 (3), 635-659*, *Earthquake Engineering Research Institute,* Oakland, California.

Bradley, BA, Quigley, MC, Van Dissen, RJ, Litchfield, NJ (2014): Ground Motion and Seismic Source Aspects of the Caterbury Earthquake Sequence. *Earthquake Spectra, 30 (1), 1-15, EERI,* Oakland, California.

British Standards Institution (1996). Eurocode 8: Design of Structures for Earthquake Resistance. *British Standards Institution,* London, England.

Federal Emergency Management Agency (FEMA). (2012). "Seismic performance assessment of buildings." *Rep. No. P-58-1*, FEMA, Washington, D.C.

Freeman SA (1998): Development and Use of Capacity Spectrum Method, Paper *No. 269, Proceedings of 6th US Conference on Earthquake Engineering, EERI,* Seattle, Washington.

Freeman SA (2004): Review of the Development of the Capacity Spectrum Method. *Performance-Based Seismic Design Special Issue March 2004, Paper No. 438, ISET Journal of Earthquake Technology, India,* 41(1), 1-13.

Freeman SA, Paret TF (2000): Several Case Studies of Performance-Based Seismic Engineering for Evaluation and Upgrading. *Proceedings of Ninth US-Japan Workshop on the Improvement of Structural Design and Construction Practices.* Report ATC 15-8, Applied Technology Council, Victoria, British Columbia, Canada.

Freeman SA, Paret TF, Searer GR (2008): ASCE-31 and ASCE-41: What Good Are They? *Proceedings of Structures Congress 2008, SEI/ASCE,* Vancouver, British Columbia, Canada.

Haselton, C.B., Baker, J.W., et al (2017). "Response history analysis for the design of new buildings in the NEHRP provisions and ASCE/SEI 7 standard: part I - overview and specification of ground motions." *Earthquake Spectra, Vol. 33 (2), 373-395, EERI,* Oakland, California.

Huang, Y.-N., Whittaker, A., Luco N., and Hamburger R. (2011). "Scaling earthquake ground motions for performance-based assessment of buildings." *Journal of Structural Engineering*, Vol. 137, No. 3, 311-321.

Kamai, R, Abrahamson, N (2015): Are Near-Fault Fling Effects Captured in the New NGA West2 Ground Motion Models? *Earthquake Spectra, 31 (3), 1629-1645, Earthquake Engineering Research Institute,* Oakland, California.

Los Angeles Tall Buildings Structural Design Council (2015): An Alternative Procedure for Seismic Analysis and Design of the Tall Buildings Located in the Los Angeles Region. *2014 Edition with 2015 Supplements, LATBSDC,* Los Angeles, California.

Malhotra, PK (2015): Normalized Response Spectrum of Ground Motion*, The Bridge and Structural Engineer, Volume 45, Number 1: March 2015, ING-IABSE,* New Delhi, India.

Murphy, LM, Coffman, JL, Benfer, NA (1973): San Fernando, California, Earthquake of February 9, 1971. *U.S. Department of Commerce, National Oceanic and Atmospheric Administration,* Washington, D.C.

Ohsaki, M, Kasai, K, Hikino, T, Matsuoka, Y (2008): Overview of 2007 E-Defense Blind Analysis Contest Results. *Proceedings of The 14th World Conference on Earthquake Engineering.* Beijing, China.

Pacific Earthquake Engineering Research Center (October 2005): Van Nuys Hotel Building Testbed Report: Exercising Seismic Performance Assessment. *Report No. 2005/11,* Berkeley, California.

Pacific Earthquake Engineering Research Center (November 2010): Guidelines for Performance-Based Seismic Design of Tall Buildings. *Tall Buildings Initiative, Report No. 2010/05,* Berkeley, California.

Paret TF, Searer GR, Freeman SA (2011): ASCE 31 and 41: Apocalypse Now. *Proceedings of Structures Congress 2011, ASCE,* Las Vegas, Nevada, USA.

Pavan, A (2008): Blind Prediction of a Full-Scale 3D Steel Frame Tested Under Dynamic Conditions. *A Dissertation Submitted in Partial Fulfilment of the Requirements for the Master Degree in Earthquake Engineering,* Instituto Universitario di Studi Superiori di Pavia, Pavia, Italy.

Rojahn C, Mork PN (1981): An analysis of strong-motion data from a severely damaged structure, the Imperial County Services Building, El Centro, California. *Open-File Report 81-194. U.S. Geological Survey.*

Searer GR, Paret TP, Freeman SA (2008): In Name Only: Peer Review of Performance-Based Engineering Projects. *Proceedings of 14th World Conference on Earthquake Engineering,* Beijing, China.

Terzic, V, Schoettler, MJ, Restrepo, JI, Mahin, SA (2015): Concrete Column Blind Prediction Contest 2010: Outcomes and Observations. *PEER, PEER Report No. 2015/01,* Berkeley, California.

Yamada, M, Olsen, AH, Heaton, TH (2009): Statistical Features of Short-Period and Long-Period Near-Source Ground Motions. *Bulletin of the SSA, Vol. 99, No. 6, pp. 3264-3274, SSA,* Albany, California.