Vassiliou, M. F., Broccardo, M., Cengiz, C., Dietz, M., Dihoru, L., Gunay, S., Mosalam, K., Mylonakis, G., Sextos, A., & Stojadinović, B. (2020). Shake table testing of a rocking podium: Results of a blind prediction contest. *Earthquake Engineering and Structural Dynamics*. <https://doi.org/10.1002/eqe.3386>

Peer reviewed version

Link to published version (if available):
[10.1002/eqe.3386](https://doi.org/10.1002/eqe.3386)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at <https://doi.org/10.1002/eqe.3386>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Shake Table Testing of a Rocking Podium: Results of a Blind Prediction Contest

M. F. Vassiliou¹, M. Broccardo², C. Cengiz³, M. Dietz³, L. Dihoru³, S. Gunay⁴, K. Mosalam⁴, G. Mylonakis^{3,5,6}, A. Sextos³, B. Stojadinovic¹

¹ Department of Civil, Environmental, and Geomatic Engineering, ETH Zurich, Switzerland

² Department of Civil, Environmental and Mechanical Engineering, University of Trento, Italy (formerly with ETH Zurich)

³ Department of Civil Engineering, University of Bristol, Bristol, United Kingdom

⁴ University of California at Berkeley (UCB), U.S.A.

⁵ Department of Civil Engineering, Khalifa University, U.A.E.

⁶ University of California at Los Angeles (UCLA), U.S.A.

Correspondence to: vassiliou@ibk.baug.ethz

Rocking motion is sensitive to the boundary and initial conditions of a rocking structure, making experiments non-repeatable. Thus, the claims that numerical rocking motion models are not only inaccurate, but that all rocking structures behave unpredictably. Hence, rocking is not used as a seismic design approach. This paper revisits the issue of rocking motion unpredictability. Seismic behavior of structures is inherently stochastic, because the loading is stochastic. Therefore, the question of interest is not whether models can predict the seismic response to a single ground motion, but if the statistical characteristics of the ensemble of responses to a set of ground motions that define the seismic hazard can be predicted. For this purpose, a rocking podium, which is a three-dimensional structure comprising an aluminum slab supported by 4 tubular steel columns, was tested on a shake table excited by two sets of 100 consistently generated ground motions. It was found that the Cumulative Distribution Function (CDF) of the experimentally obtained displacements is statistically stable. Next, a blind prediction contest was organized. The contestants were invited to predict the CDFs of the slab lateral displacement. They were able to predict the slab displacement CDF relatively well. Both finite element and discrete element modeling approaches were used, but no clear pattern emerged as it was found that the performance of either approach depends on the input parameters used and the assumptions made. It was also observed that the contestants who did not use Rayleigh damping in their models produced better predictions.

Keywords: rocking, blind prediction contest, shake table testing, FEM, DEM, model validation, Rayleigh Damping

INTRODUCTION

Rocking structures are the ones that can uplift from their base when subject to an earthquake. Such structures attracted the attention of the research community after the powerful 1960 Valdivia (Chile) earthquakes, when Housner presented his seminal work [1] on the dynamics of the archetype rocking block (Figure 1). The rocking block dynamic model is useful because it can also describe the seismic behavior of non-anchored equipment [2-12], masonry structures [13-21], and ancient Greco-Roman and Chinese temples [22-26].

Motivated by the increasing stability of larger rocking blocks with identical aspect ratios, researchers have suggested using rocking as a seismic design strategy because the uplift of the block acts as a mechanical fuse and limits the design forces of both the superstructure and the foundation. The idea is applicable to both buildings (Figure 2, left) [27-32] and bridges (Figure 2, right) [33-37].

However, seismic analysis of rocking structures is not straightforward because, during the uplift, such structures have negative stiffness (in terms of the relation between the lateral force F and displacement u , shown in Figure 1). This is why their seismic response cannot be described by any “equivalent linear” system such as those employed for ordinary yielding structures [38]. Even though attempts to create non-linear spectrum-based methods to present the maxima of the seismic displacement of rocking structures (not related to elastic response spectrum) have been made [39-40], these spectra were created by performing multiple time-history seismic response analyses using Housner’s model. Therefore, Housner’s model is as useful for rocking structures as the elastic SDOF oscillator is for the fixed-base ones.

THREE-DIMENSIONAL ROCKING – “WOBLING” UNDER 2D/3D GROUND EXCITATION

Housner’s model describes planar rocking of a rigid body excited by horizontal excitation. However, many real uplifting structures (e.g., bridges, statues, or ancient temples) would undergo a 3-dimensional motion characterized

Figure 1. Geometric characteristics of a rigid rocking body (a,b); Lateral force-deformation relation of an uplifted rigid rocking body (c).

Figure 2. Rocking isolation of a building [27-32] (Left); Rocking frame that can model a rocking bridge [33] (Right) (Both figures adapted from [31])

by simultaneous uplift from the ground (rocking) and change of the contact point with the ground (nutation) without twisting or sliding out of its original position, called “wobbling” [41] in this study. Such motion is similar to the motion of Euler’s disk, a coin spinning on a flat surface, above its initial position, without stepping out. Real uplifting structures may also twist and slide out of their positions, a phenomenon not investigated in this study.

There have been attempts to study wobbling using rigid body analytical dynamics. In [42-44], the motion of a rigid cylinder under seismic excitation was studied. Other researchers studied the 3D response of ancient conical or cylindrical columns numerically [45, 46], or experimentally [47-49]. Pappas et al. [49] numerically explored the behavior of an ancient cylindrical column with a height of 6m and a diameter of 0.66m with the intention of defining proper ground motion intensity measures to characterize the rocking response of such structures. Makris et al. [51] experimentally tested scaled models of uplifting bridges that exhibited wobbling. All the above studies conclude that wobbling motion is present, even under planar initial conditions and/or under planar ground excitation. Stefanou et al. [52] provided theoretical justification for the above observation.

The 3D behavior of non-cylindrical rocking bodies received attention recently. Konstantinidis and Makris [53] and Zulli et al. [54] studied the rocking motion of a 3D prism. Chatzis and Smyth [55] studied the motion of a 3D prism on a deformable base, taking sliding into account as well as the 3D dynamics of a rigid body with wheels on a moving base [56]. Mathey et al. [57] studied the influence of geometric defects on the 3D response of small-size blocks. They concluded that the blocks with imperfections are less stable than the theoretically perfect ones.

Figure 3. *Cylinder allowed to wobble without twist [58] (sketch adapted from [31] (Left); Wobbling frame [59] (Right).*

Vassiliou et al. [58] studied a 2DOF rigid cylinder constrained to wobble above its initial position without twisting or sliding, i.e. without stepping out of its original position (Figure 3, left). Subsequently, Vassiliou [59] extended this model to include a slab on top of a set of cylindrical columns (Figure 3, right). This is a model of a wobbling bridge, where the columns are physically constrained to wobble above their initial position without stepping out, under the assumption that wobbling is smooth and sliding on the ground does not occur. Response analyses conducted using these models showed that planar rocking model underpredicts the magnitude of 3D wobbling motion, even if two planar rocking response analyses are performed in the two orthogonal horizontal directions and the results are appropriately combined. The advantage of the models developed in [58] and [59] is that they are computationally inexpensive, thus suitable for parametric investigations and probabilistic estimates, at the cost of not considering the possibility of stepping out or the deformability of the system elements. These neglected phenomena can be modelled using finite element method (FEM) or discrete element method (DEM) models, which, however, are computationally expensive and often involve contact parameters that have no physical meaning or that cannot be experimentally quantified.

STATISTICAL SEISMIC RESPONSE MODEL VALIDATION

A major drawback for rocking seismic response models is that they are not validated - at least not by means of the conventional approach of deterministically comparing the displacement response of the model and the prototype to a given ground excitation. This is further hindered by the fact that the seismic response of a rocking block is particularly sensitive to all of the parameters that define it and as such it has been characterized as “chaotic” (i.e. non-reproducible and non-predictable). In fact, small model parameter perturbations lead to substantially different time-history responses. For this reason, experiments involving seismically or dynamically excited rocking specimens are seldom repeatable.

This lack of validation effectively means that Housner’s model should not be used as its ability to represent the physical reality has not been proven. Consequently, rocking cannot be used as a seismic design strategy, given that the seismic response of such structures would be effectively unpredictable, by analytical or numerical methods.

Bachmann et al. [60, 61] and Del Giudice et al. [62] suggested that the conventional seismic response model validation procedure focused on deterministic reproduction of the experimentally obtained response to a particular ground motion with acceptable accuracy is too strict of a test. Instead, they proposed a weaker, but sufficient, concept of statistical validation of structural seismic response models. Statistical seismic response model validation is a combination of general statistical model validation and simulation model validation and verification [63] applied to seismic response models of engineered structures. This novel model validation procedure comprises two steps. *First*, an experimental benchmark dataset, measuring the recorded dynamic response of the same specimen (or essentially identical specimens) subjected to ensembles of consistently generated and scaled ground motions that represent a given seismic hazard, is developed. *Second*, a model is used to produce a dataset comprising its dynamic responses to the same ground motion ensembles. Validation is conducted by comparing statistical distributions of the pertinent response quantities of the model and the benchmark using the two datasets. Statistical seismic response model validation is a weak method in that it involves the statistics of model predictions of the response to ensembles of ground motions, rather than predicting the responses to every single ground motion deterministically. Bachmann et al. [60], [61] applied the statistical seismic response model validation procedure to validate the Housner [1] rocking response model. They performed 600 shaking table tests using a well-defined and repeatable uplifting structure as well as 600 numerical

Figure 4. Comparison of the maximum rocking block (left) rotation obtained experimentally and numerically: Individual motion comparison (middle); CDF Comparison (right). Adapted from Bachmann et al. [60].

simulations for the same tests, and compared both the individual test responses and the statistical aggregates of these responses focused on predicting limit states such as overturning or maximum tilt angle (Figure 4). They showed that the 1963 Housner model passes the weak validation test (Figure 4, right), even though it fails the strong validation test (Figure 4, middle). In this case, the excitation uncertainty overshadowed the structural seismic response model errors, either systematic or random. Therefore, the Housner model was found to be good enough for the use in the scope of seismic design.

It is worth noting that the concept of comparing seismic analysis methods in terms of the statistics of the response to sets of ground motions rather than in terms of individual responses is not novel (e.g., Chopra used this process to evaluate different estimates of seismic structural response [64]). However, this involved comparisons between different analysis methods (usually a more approximate against a more refined), not comparisons between the dynamic responses of a model to an experimental tests benchmark dataset.

WOBBLING SEISMIC RESPONSE BENCHMARK DATASET

Wobbling (three-dimensional rocking without sliding) is even harder to predict than planar rocking. Thus, developing seismic wobbling models is challenging, but is necessary in order to further develop seismic design of rocking structures as well as of non-structural elements, such as equipment, as outlined in the introduction of this paper. Validation of wobbling models is a key step in this development process. To facilitate statistical seismic response model validation of wobbling response models, a series of shake table tests was performed at the University of Bristol to create the benchmark response dataset. The main objectives of the tests were three: (a) To observe whether wobbling tests are repeatable in the statistical sense; (b) To explore whether the CDF of the response is predictable by state-of-the-art methods; and (c) to recommend parameters and assumptions for the numerical modelling of wobbling structures.

Specimen Description

The wobbling specimen was designed at ETH Zurich and built in Bristol in the framework of the EU-funded research project SERA. It comprises an aluminum slab supported on four wobbling circular structural steel columns resting on the shake table platform (Figure 5). The four columns had a height, diameter and wall thickness of 1000 mm, 244.5 mm and 8 mm, respectively (Figure 6), with a corresponding slenderness (i.e. diameter to height) ratio of $\tan\alpha = 0.2445$. Note that the size of the wobbling specimen is not representative of the size of a prototype bridge or a prototype podium building. Given that planar and three-dimensional rocking response is size-dependent, the ground motion excitations used in the shake table experiments were also scaled to preserve acceleration scaling (as discussed later in this paper) so that the model represents a structure with a height similar to an actual bridge. The intent of the specimen detailing was to mimic the assumptions in Vassiliou's [59] "wobbling bridge" model, namely that the structure is rigid and that the columns wobble without sliding or twisting about their longitudinal axis. Namely, the ends of the columns were machined (lathed) to ensure they are right-angled to their longitudinal axis. Conical end restraints were placed at each end of each column to stop the columns from wobbling out of their original position (Figure 7). After the tests, scratches on the conical restrainers were observed, showing that the columns tried to "climb out" (Figure 7). Aluminum pegs were installed inside the columns (Figure 7) with enough clearance to enable

unrestrained wobbling up to overturning, but engage early enough to enable easy recovery of the specimen to its original position after a test that resulted in overturning. Moreover, padded frames were placed under and around the specimen to protect the shake table platform (Figure 5).

Figure 5. Wobbling specimen on the University of Bristol shake table.

Figure 6. Details of the wobbling specimen.

Figure 7. *Conical restraint and an aluminum peg at the bottom of the wobbling specimen column before installation (left). Metal shavings found after the shake table tests showing that the steel column scratched against the conical restrainer during the tests (right).*

Specimen Excitation Using A Shake Table

The wobbling response of the specimen was induced by a dynamic bi-directional (two orthogonal horizontal components) excitation of its support. This was achieved by placing the specimen on the top of the 6DOF 3 m×3 m shake table of the University of Bristol [65]. In both directions, the stroke of the table is ±150 mm and the maximum horizontal velocity is 1200 mm/s. The maximum acceleration of the table is 3.7 g and 1.6 g under zero and 100 kN of payload, respectively.

Ground Motions

To constrain the uncertainty in the ground motion excitation, ensembles of bi-directional ground motions were synthesized using a spectral version of the Rezaeian and Der Kiureghian stochastic ground motion model [66-68]. The generated synthetic motions are, thus, target-spectrum-compatible, in other words, consistent with a given elastic acceleration response spectrum. A stochastic method to generate ground motions with similar power spectral density was adopted as the best-available precondition for a statistical comparison of the benchmark and model response datasets, even though motions with similar elastic spectra could have very different overturning potential. Therefore, the objective is not to describe the full ground motion variability of a given site, but to limit the analysis to spectrum-compatible ground motions that have a degree of spectral consistency while retaining the necessary variability in terms of other pulse properties. In essence, the parameters of the stochastic ground motion model are fixed (and not random variables), and only the inherent (aleatory) uncertainty is considered. It is however acknowledged that, as ground motion selection for rocking is an open subject, different procedures for defining the excitation could have been chosen so that motion-to-motion variability is represented better.

Two recorded ground motions, the 1940 El Centro Array #9 record and the 1999 Chi-Chi CHY080 record, were used as the "seed" ground motions for the experimental campaign (Figure 8 a,b,g,h). Next, two bi-directional (two orthogonal horizontal components) ground motion ensembles, each comprising 100 synthetic ground motions, were generated from the two seed ground motions. These two ground motion ensembles were used to drive the shake table during the conducted tests. Vertical ground motion translational and the three rotational excitations were not considered.

Ground Motion Scaling

The mass of the rocking body originates both the stabilizing (self-weight) and the de-stabilizing (inertia) elements of Housner's rocking body model. Thus, the acceleration scale must be preserved to design the experiments, implying that the time scale should be equal to $S_T = \sqrt{S_L}$. Since the ground motion ensembles were generated as consistent with target elastic response spectra of two recorded ground motions, the size of the wobbling specimen was related to the size of columns in a bridge prototype structure (taken here as 9.67m) to determine a length scale S_L (equal to 1/9.67 in this case, as the height of the columns in the specimen is 1.0 m). This meant that the time component of the generated ground motions was scaled by a factor of $\sqrt{9.67} = 3.11$, i.e., the frequency of the ground motions was increased, without changing their acceleration amplitudes. The measured (as-applied) ground motions were used to compute the response of the rocking response model and conduct its statistical seismic response validation.

Figure 8. (a), (b), (g), (h): 1940 El Centro and 1999 Chi-Chi CHY080 originally recorded “seed” ground motions; (c), (d), (i) (j) Recorded ground motions scaled in time by 3.11; (e), (f), (k), (l): 5% damped elastic acceleration response spectra of the Measured, mean measured and originally recorded “seed” ground motions.

It should be noted that the wobbling specimen is only a distorted model of a prototype structure because the stress similitude is not preserved. Thus, the elastic modulus of the material and the natural frequencies of the prototype are not correctly scaled. In addition, the wobbling specimen columns are not perfectly rigid, as Vassiliou [59] assumes, nor are any realistic foundation condition modelled. However, it has been shown that the deformability of large structures does not qualitatively change their rocking behavior [69-73]. Furthermore, the ends of the columns of a real rocking podium structure would be protected from localized damage, and are likely to remain elastic during seismic response. Ultimately, the conducted experiments serve to generate a benchmark dataset for wobbling seismic response model validation and do not aim at representing the nuances of a prototype structure without distortion.

Motion Measurement

The response of the wobbling specimen was measured using displacement and acceleration sensors. The displacements were measured using an infrared tracking system, with six infrared passive markers installed on the top of the aluminum slab, as shown in Figure 9. These six targets can capture the six degrees of freedom of the slab (three translations and three rotations), assuming it remains rigid. The accelerations were measured using three-dimensional accelerometers. These sensors were placed on the shake table platform to record the actually applied excitation.

Relevant Wobbling Response Quantity

The horizontal displacement of the slab is the relevant response quantity in the created benchmark dataset, which was measured using the markers shown in Figure 9. The symmetry of the wobbling specimen would ensure that no rotation of the slab around the vertical axis should occur. However, imperfections did exist and such minor twisting rotations were observed. Based on this, the following maximum absolute average displacement (MAAD) of the wobbling specimen slab is defined as the relevant measure of the wobbling specimen response:

$$u = \max_t \left(\frac{u_1(t) + u_3(t) + u_4(t) + u_6(t)}{4} \right) \quad (1)$$

where u_i is the Euclidean norm of the horizontal displacement at points M_i of Figure 9, relative to the shake table. Note that using the displacement of the center of the slab would have obscured the effect of the imperfections and the resulting twisting of the wobbling specimen. MAAD is a convenient description of the motion of the wobbling specimen. In both extreme cases (i.e. pure rotation of the slab around the vertical axis, and pure horizontal translation along one axis), static overturning occurs when MAAD becomes equal to the diameter of the column. Figure 10 plots the Cumulative Distribution Functions (CDFs) of the specimen MAAD (i.e. of u) for the El Centro and the Chi-Chi ground motion ensembles using thick black lines. Even though the ground motions were generated so that they have consistent elastic acceleration response spectra, the dispersion in the response is large, confirming that the elastic spectrum is not a good measure of the overturning potential of a ground motion [38]. It is remarkable that the wobbling specimen did not collapse in any of the 200 conducted shake table tests, even though it experienced wobbling in all tests and developed large displacements in several tests (Figure 10). Noises heard during some of the large-

Figure 9. Marker positions on top of the aluminum slab used for wobbling specimen displacement measurements.

Figure 10. MAAD CDFs obtained from shake table tests (thick black lines) and CDFs of the same response quantity obtained using the 13 contesting models for the El Centro and Chi-Chi ground motion examples.

displacement tests may have been generated by contacts between the column and its restraining system (cone and peg, shown in Figure 7). However, displacement measurements do not support this observation. It is also noteworthy that the measured wobbling specimen response to repeated individual ground motions was not repeatable (same ground motion excitations were repeated 3 times in a row for this specific purpose). However, the important question is not whether a single test is repeatable, but whether the MAAD CDF from multiple tests is reproducible by a model that is statistically validated.

WOBBLING SEISMIC RESPONSE BLIND PREDICTION CONTEST

Statistical seismic response model validation was conducted for a number of wobbling response models in the scope of a blind prediction contest organized by the Pacific Earthquake Engineering Research center (PEER), the University of Bristol, and ETH Zurich, launched in October 2019 with a prediction submission deadline on November 20th, 2019. The purpose of the contest was to evaluate and assess the efficiency of computational models and to increase awareness in the scientific community about statistical validation in shake table testing. The contestants were asked to use their models to predict not the wobbling specimen seismic displacement (Equation 1) to each of the 200 ground motions in the two ground motion ensembles, but the experimental CDFs of the wobbling specimen MAAD u response, as shown in Figure 10 using thick black lines. For each set of ground motion ensembles, the competing models were ranked according to the maximum vertical distance, i.e. the Kolmogorov-Smirnov (K-S) distance [77-78], between their predicted and the experimentally obtained CDFs. Given that two ensembles of ground motions were used, the average Kolmogorov-Smirnov distance was taken as a best-fit performance indicator. The as-built geometry of the wobbling specimen and the recorded shake table accelerations were provided to the contestants, but no tests were performed to mechanically characterize the materials of the wobbling specimen (i.e., the steel of the columns and conical restrainers), as these were not essential in terms of affecting the response and use of common modulus of elasticity for steel and aluminum was sufficient.

It should be noted that the shake table accelerometer data acquisition system did not start for one ground motion in each of the two ground motion ensembles, leaving these two tests without the measured (as-applied) shake table excitation data. These two ground motions (El Centro #4 and Chi-Chi #58) were excluded from the evaluation of the submitted predictions, making each experimental benchmark dataset 99 entries long. The contestants were invited to submit results even if they did not compute the response of their model to all 198 ground motions. In such cases, CDFs of the model response were computed using only the submitted results.

Thirteen contestants participated in the blind prediction contest. The models they used can be grouped into three categories, as follows:

- a) 8 contestants developed models based on the Finite Element Method (FEM), out of which 4 used ABAQUS [79] (Teams #7,#9,#10, and #12), 2 used SAP2000 [80] (Teams #2 and #4), 1 used Midas GEN [81] (Team #1) and 1 used OpenSees [82] (Team #6);

- b) 4 contestants used the Discrete Element Method (DEM) to build their models, out of which 3 used 3DEC [83] (Teams #3, #5 and #8), and 1 used Code Aster [84] (Team #13);
- c) 1 contestant (#Team 11) used Vassiliou's [59] Rigid Body (RB) model (independently, as Vassiliou was one of the contest organizers and did not take part in the contest).

The contest rules required the submission of the wobbling specimen MAAD u for each simulation in millimeters. One team submitted extremely small wobbling specimen displacement predictions, seemingly in kilometers. The team subsequently clarified that the submitted model displacements were in meters. In the blind prediction contest, the data from this team was treated as if in meters, but in this paper, their data was processed using the units of kilometers.

The outcomes of the statistical seismic response validation of the contesting models are shown in Figure 10 and Table 1. Figure 10 graphically compares the predicted wobbling specimen MAAD response CDFs ($F(u)$) to the experimentally obtained wobbling specimen displacement response CDFs. Table 1 ranks the contestant predictions based on the K-S distance: Column 1 gives the ranking for the El Centro ensemble, column 2 for Chi-Chi ensemble, and column 3 gives the overall model ranking based on the average Kolmogorov-Smirnov distance for the two ensembles of ground motions.

Based on the K-S distance, the winning model is the one submitted by Team #9 (Myron Chiyun Zhong and Constantin Christopoulos from the University of Toronto). They used FE software ABAQUS 6.13 with C3D20R quadratic brick element, with reduced integration. They assumed a coefficient of friction of 0.3 and no Rayleigh damping [85]. The second best winning model was the one submitted by Team #3 (Daniele Malomo, Anjali Mehrotra, and Matthew DeJong from UC Berkeley). They used Distinct (or Discrete) element software 3DEC. The frictional mechanisms were modeled using a simplified Mohr-Coulomb criterion with tension cut-off. The blocks were modelled as rigid and the connection was provided by nonlinear springs with normal stiffness intensity $k_n = 10^6 \text{ kPa/m}$ and shear stiffness $k_s = 10^6 \text{ kPa/m}$ [86]. Zero Rayleigh damping was used.

Alternative Methods for Statistical Model Validation

The choice of the Kolmogorov-Smirnov distance to perform statistical model validation could be characterized as arbitrary and/or subjective. Therefore, two additional metrics to compare the competing models are explored in this paper (this was not done in the blind prediction contest). The metrics are based on comparing the horizontal distance between certain points on the experimental benchmark and model response CDFs. These points are the CDF median points and the CDF 90th percentile points, i.e., the 50% and the 90% quantile points of the MAAD CDFs.

The outcomes of the statistical seismic response validation of the contesting models using the median and the 90th percentile points of the MAAD experimental benchmark and model output CDFs are shown in Tables 2 and 3, respectively. For a more specific comparison, Table 4 reports the median and 90th percentile points of the experimental

Table 1. Contestant model ranking based on the Kolmogorov-Smirnov distance between the MAAD CDFs

	El Centro		Chi-Chi		Average	
	Team number (method used)	K-S Distance	Team number (method used)	K-S Distance	Team number (method used)	K-S Distance
1	12 (FEM)	0.25	3 (DEM)	0.06	9 (FEM)	0.19
2	9 (FEM)	0.27	9 (FEM)	0.11	3 (DEM)	0.21
3	6 (FEM)	0.32	13 (DEM)	0.11	5 (DEM)	0.24
4	5 (DEM)	0.34	5 (DEM)	0.13	4 (FEM)	0.30
5	3 (DEM)	0.36	4 (FEM)	0.16	13 (DEM)	0.30
6	4 (FEM)	0.43	8 (DEM)	0.26	6 (FEM)	0.39
7	13 (DEM)	0.48	10 (FEM)	0.27	12 (FEM)	0.48
8	2 (FEM)	0.65	6 (FEM)	0.46	10 (FEM)	0.53
9	1 (FEM)	0.69	1 (FEM)	0.48	8 (DEM)	0.58
10	11 (RB)	0.72	11 (RB)	0.62	1 (FEM)	0.59
11	10 (FEM)	0.79	12 (FEM)	0.71	11 (RB)	0.67
12	8 (DEM)	0.89	7 (FEM)	0.98	2 (FEM)	0.82
13	7 (FEM)	1.00	2 (DEM)	0.99	7 (FEM)	0.99

Table 2. Contestant model ranking based on the error of predicting the median MAAD CDFs

	El Centro		Chi-Chi		Average	
	Team number (method used)	$e_{50,EC} = \frac{ u_{50,m,EC} - u_{50,exp,EC} }{u_{50,exp,EC}}$	Team number (method used)	$e_{50,CC} = \frac{ u_{50,m,CC} - u_{50,exp,CC} }{u_{50,exp,CC}}$	Team number (method used)	$e_{50} = \frac{(e_{50,EC} + e_{50,CC})}{2}$
1	9 (FEM)	0.12	3 (DEM)	0.01	9 (FEM)	0.09
2	12 (FEM)	0.18	13 (DEM)	0.01	3 (DEM)	0.24
3	6 (FEM)	0.21	9 (FEM)	0.06	6 (FEM)	0.28
4	2 (FEM)	0.37	5 (DEM)	0.07	5 (DEM)	0.29
5	3 (DEM)	0.48	8 (DEM)	0.07	12 (FEM)	0.36
6	5 (DEM)	0.52	4 (FEM)	0.10	4 (FEM)	0.41
7	4 (FEM)	0.73	1 (FEM)	0.22	13 (DEM)	0.48
8	13 (DEM)	0.94	10 (FEM)	0.27	2 (FEM)	0.67
9	7 (FEM)	0.95	6 (FEM)	0.35	1 (FEM)	0.82
10	1 (FEM)	1.41	12 (FEM)	0.54	7 (FEM)	0.94
11	10 (FEM)	1.92	7 (FEM)	0.94	10 (FEM)	1.10
12	8 (DEM)	2.50	2 (FEM)	0.98	8 (DEM)	1.29
13	11 (RB)	2.87	11 (RB)	overturn	11 (RB)	n/a

Table 3. Contestant model ranking based on the error of predicting the 90th percentile of the MAAD CDFs

	El Centro		Chi-Chi		Average	
	Team number (method used)	$e_{90,EC} = \frac{ u_{90,m,EC} - u_{90,exp,EC} }{u_{90,exp,EC}}$	Team number (method used)	$e_{90,CC} = \frac{ u_{90,m,CC} - u_{90,exp,CC} }{u_{90,exp,CC}}$	Team number (method used)	$e_{90} = \frac{(e_{90,EC} + e_{90,CC})}{2}$
1	6 (FEM)	0.18	3 (DEM)	0.00	9 (FEM)	0.12
2	9 (FEM)	0.22	9 (FEM)	0.02	6 (FEM)	0.25
3	12 (FEM)	0.40	5 (DEM)	0.03	3 (DEM)	0.26
4	1 (FEM)	0.50	13 (DEM)	0.06	5 (DEM)	0.30
5	3 (DEM)	0.52	8 (DEM)	0.09	4 (FEM)	0.35
6	5 (DEM)	0.57	4 (FEM)	0.09	13 (DEM)	0.36
7	2 (FEM)	0.59	10 (FEM)	0.19	1 (FEM)	0.45
8	4 (FEM)	0.60	6 (FEM)	0.33	12 (FEM)	0.47
9	13 (DEM)	0.66	1 (FEM)	0.40	10 (FEM)	0.68
10	7 (FEM)	0.97	12 (FEM)	0.53	2 (FEM)	0.79
11	10 (FEM)	1.18	7 (FEM)	0.89	8 (DEM)	0.91
12	8 (DEM)	1.73	2 (FEM)	0.98	7 (FEM)	0.93
13	11 (RB)	2.25	11 (RB)	overturn	11 (RB)	overturn

Table 4. Median and 90th percentile points of the experimental benchmark dataset MAAD CDFs for the two ground motion ensembles, the corresponding 95% confidence intervals (CIs), and the Team #9 and #3 model-generated data

Dataset CDF points	El Centro	Chi-Chi
Median of the experimental benchmark (95% CI)	12mm (10-13mm)	98mm (88-108mm)
Median of the Team #9 model	10mm	104mm
Median of the Team #3 model	17mm	98mm
90 th percentile of the experimental benchmark (95% CI)	25mm (23-34mm)	159mm (151-181mm)
90 th percentile of the Team #9 model	31mm	162mm
90 th percentile of the Team #3 model	38mm	160mm

benchmark dataset CDFs, their 95% confidence intervals, and the same points computed using the model-generated datasets of first and second ranked teams #9 and #3. The confidence intervals are biased because they are in the right tail of the CDFs.

Discussion of the Statistical Seismic Response Model Validation Outcomes

The blind prediction contest provided a unique opportunity to conduct simultaneous statistical seismic response model validation for 13 different models. Ranking of the models depended on the employed MAAD CDF metric (Tables 1, 2 and 3). All metrics are mathematically well defined, but there is no apparent reason to pick one as the universally most important, as it often the case with statistical measures. However, it is worth mentioning that model of Team #9 scored the best in all metrics considered. Moreover, no model ranked in the top two positions in one metric simultaneously ranked below position 6 in any other metric.

The most intuitive structural engineering metrics are probably the median and the 90th percentile response, as they represent estimates of the maximum displacement of the structure. The relative error in predicting them is presented in Tables 2 and 3. It can be seen that four models (#3, #5, #6, and #9) manage to achieve an average (between the two ensembles) a less than 30% error in predicting both quantities. As the models of Teams #9 and #3 scored the best in the norm initially selected for the contest (the K-S distance), Table 4 compares these models in terms of their ability to predict the median and the 90th percentile of the response. Team #9 predicted the median response with an average accuracy (with respect to the experimental empirical estimate) of 9%, and the 90th percentile of the response with an average accuracy of 12%. Team #3 had 24% and 26% accuracies, respectively. Team #9 predictions of the median and the 90th percentile fell within the 95% Confidence Interval bounds of the measured corresponding quantities.

Because overturning is a clearly defined and objective limit state of a rocking structure, a meaningful metric would have been the prediction of overturning probability. However, no overturning events were observed in the experiments, because Vassiliou's Rigid Body dynamics model [59] was used for their design – a model that was proven to be very conservative. Thus, a direct evaluation of the ability of the contesting models to predict overturning cannot be performed using this experimental benchmark dataset. Nevertheless, it is reasonable to assume that the ability to predict overturning should be linked to the performance of the model in predicting large rotations. In that sense, out of all the measures discussed, the 90th percentile displacement error for the Chi-Chi ensemble should be the best index to evaluate a model. Based on this measure, the model of Team #3 performs the best, but, notably, six teams predict this response measure with less than 10% error.

Ranking of the models also depends on the ground motion ensemble characteristics. Generalized conclusions about which modelling methods and parameters are better for which ground motions cannot be drawn because only two ensembles were tested. However, the models of Teams #6 and #12 deserve special attention as they performed notably better in the El Centro ensemble than in the Chi-Chi one. The first comment that holds for both Team #6 and Team #12 models is related to the direction of the errors in the El Centro ensemble. The sign of the errors (obscured by the MAAD response quantity) indicates that both models approach the experimental benchmark MAAD CDF of the El Centro ensemble “from the left” (Figure 10), i.e. they underestimate the response. As the rocking podium has a pronounced softening behavior, it is expected that such underestimation of wobbling specimen displacements will only increase for ground motions in the more intense Chi-Chi ensemble. Therefore, models that underestimate the response for low-intensity ground motions are generally expected to underestimate the response even more in more intense ground motion response simulations, which is exactly what is observed from the predictions of Teams #6 and #12 for the Chi-Chi ensemble. Why do the models of Teams #6 and #12 underestimate the response in the first place? Team #6 used OpenSees and modelled the rocking surface with two independent rotational springs, meaning that the model responds independently in the two orthogonal directions. Neglecting such interaction has already been proven to underestimate the response ([59]). Not enough model details were provided to understand why model #12 underestimated the response.

Both FEM and DEM models used in this blind prediction contest performed well or poorly, depending on the modeling parameters. For example, of the two best contestants, Team #9 developed an FEM and Team #3 used a DEM model. Thus, there is no basis to recommend FEM or DEM to model wobbling structures seismic response. However, it should be mentioned that MAAD CDFs produced by DEM models were close to the experimental benchmark CDFs, except in the case of the model developed by Team #8 which clearly underpredicted the response for the El Centro ground motion ensemble. Similarly, the FEM models of Teams #2 and #7 grossly underestimated the response, as can be seen in Figure 10. This does not mean that FEM models cannot perform well, or that such models would underestimate the response; the FEM model of Team #9 performed the best in all metrics. This is merely an indication

that, when it comes to uplifting and rocking structures, modeling decisions should be made with care. This is more important for FEM models, as such models originate from continua, as opposed to DEM models that inherently assume the structure comprises of discrete parts.

The specific Rigid Block model proposed by Vassiliou [59] and used by Team #11 consistently overestimates the response. Hence it is conservative, but not accurate. It should be noted that the Rigid Block model formulation does not include any form of energy dissipation. Apparently, disregarding the energy dissipated in the tests through scratching of the conical restrainers leads to overestimation of the response. Therefore this model can only be used to obtain conservative estimates of the response.

Role of Energy Dissipation in Modeling of Wobbling Response

Modelling of energy dissipation in structures is an open problem, not only for rocking structures, but for fixed base structures as well. The widely used Rayleigh model, although numerically convenient (like viscous damping in SDOF oscillators), is not validated experimentally and can create implausible forces and moments, especially when the initial stiffness proportional component is included [87-89]. In the wobbling system discussed herein, energy is dissipated mainly through sliding (that created the metal shavings shown in Figure 7) and through radiation damping, through the interaction with the supporting wave transmitting structure [90]. Inherent material damping also exists, but since the stresses are relatively small, this form of energy dissipation is not expected to be significant when compared to the other mechanisms. Therefore, creating a mass and/or stiffness proportional damping matrix clearly has no physical meaning, especially if assigned to all degrees of freedom, including the rocking interface.

Based on the above, it is not a surprise that the best models by Teams #9 and #3 do not use Rayleigh damping and rely only on modelling friction to dissipate energy. On the contrary, Team #1 model that assumed a damping ratio of 1% at 0.05s and 1.5s clearly underestimated the response. Team #4 model used Rayleigh damping, but with damping ratio set to very low values (0.1%), which was low enough not to suppress the response.

Even though the Rayleigh damping model is not suitable for modeling the seismic response of rocking or wobbling structures, this does not imply that the model should not incorporate any energy dissipation. The tests showed that the input seismic energy is dissipated by friction and radiation, and that it should be modelled explicitly. Neglecting energy dissipation completely, like in the Team #11 Rigid Body dynamics model, grossly overestimates the response. The above observations on energy dissipation, corroborate that numerical models should respect the physics of the system, to the extent possible. This does not hold for Rayleigh damping applied in rocking.

COMPARISON OF STATISTICAL AND DETERMINISTIC MODEL VALIDATION APPROACHES

While statistical seismic response model validation admits to the inherent stochastic nature of seismic excitation and response, deterministic (motion-by-motion, point-by-point) comparisons of model and experimental benchmark data remain at the foundation of conventional validation procedures. The experimental benchmark dataset and the datasets generated by the winning models of Team #9 and Team #3 are used to compare the statistical and the deterministic model validation procedures.

The motion-by-motion comparison of the MAAD u of the wobbling specimens obtained experimentally and by modeling is shown with the scatter plots in Figure 11. The data generated by the models developed by Teams #9 and #3 are shown in the two rows in Figure 11, for the El Centro and the Chi-Chi ground motion ensemble datasets in the two columns. Had the model-generated displacement predictions been perfect, all of the plotted data points would have been on the diagonal. Clearly, the model-generated predictions are correlated to the experimental data, but there is a significant scatter, confirming that motion-by-motion model predictions are far from being perfect. The correlation coefficient ρ is used to quantify the correlation between motion-by-motion experimental and model-generated maximum average wobbling specimen displacement data (MAAD). The correlation coefficients $\rho=0.48$ and $\rho=0.59$ of Team #9 and $\rho=0.61$ and $\rho=0.56$ of Team #3 for the El Centro and the Chi-Chi ground motion ensemble datasets indicate that the correlation between the experimental and model-generated data is moderate.

In contrast, the CDFs of the maximum average wobbling specimen displacement data obtained from the experiments and from the models of teams #9 and #3, replotted in Figure 12 for clarity, shows that the models produce remarkably good statistical displacement estimate matches for the Chi-Chi ground motion ensemble and reasonably good estimates for the El Centro ground motion ensemble, where the discrepancy is large in the small-displacement range. This comparison corroborates the idea that a statistical seismic response model validation is appropriate, in particular

when the uncertainty of the excitation is large. In earthquake engineering, therefore, the rocking motion should be viewed as a stochastic simulator, where the seismic stochastic input is coupled with the physic-based rocking model. In this setting, the input are both the parameters of the ground motion model and of the rocking model, and the output the probability distribution of a quantity of interest.

Figure 11. Motion-by-motion comparison of the measured MAAD response to the prediction submitted by Teams #9 and #3: El Centro (left), Chi-Chi (right).

Figure 12. Statistical comparison of the measured MAAD response to the prediction submitted by Teams #9 and #3: El Centro (left); Chi-Chi (right).

STATISTICAL TESTING OF EXPERIMENT CONSISTENCY AND MODEL QUALITY

A comparison between the wobbling response model and planar rocking model validation is informative. Bachmann et al. [60] observed that even though the scatter of motion-by-motion comparison of planar rocking models and experimental data was larger than the one observed in this paper, the statistical comparison (using a CDF of the maximum rocking specimen displacements) was quite satisfactory. The wobbling specimen response comparisons shown in Figures 11 and 12 indicate a somewhat better (compared to [60]) motion-by-motion correlation, but a somewhat less satisfactory MAAD CDF match, especially for the El Centro ground motion dataset. The likely reason of these differences is the size of the planar rocking and wobbling specimens. The planar rocking specimen is 0.5m tall, while the wobbling specimen columns are 1 m tall, about twice as large and substantially heavier. The response of larger and heavier specimens is less sensitive to specimen imperfections and boundary condition imperfections.

The considerations above bring up two important issues relevant for statistical seismic model response validation. The first issue is related to the consistency of the experimental benchmark dataset. Namely, the experimental campaign must produce a dataset with consistent CDFs of the relevant response quantities. The second issue concerns the ability of a model to predict the CDFs of the relevant response quantities with some accuracy and precision, keeping in mind the modeling assumptions that reduce or increase model response randomness.

Both issues are investigated using the classical Kolmogorov-Smirnov (KS) [77, 78, 91] p -value hypothesis testing approach. In particular, two hypotheses, H_0 (which is tested) and its opposite H_1 , are considered. The tested null hypothesis H_0 is rejected when the p -value of the K-S hypothesis test is lower than a given statistical significance threshold α_s . A fairly large value of statistical significance of 0.1 is used to allow for a nuanced qualification of null hypothesis validity using an evidence classification scale shown in Table 5. Note that the computed p -values do not indicate the probability that H_0 is true, but rather the probability that, if hypothesis H_0 is true, the observed outcomes would conform with the opposite hypothesis H_1 . A detailed explanation of hypothesis testing procedures can be found in [90].

Test of Experimental Benchmark Dataset Consistency

The null hypothesis H_0 is that the CDF of the maximum average wobbling specimen displacement, obtained from shake table experiments converges to some, yet unknown, CDF. If this hypothesis is rejected, there is strong evidence that the conducted experiments do not produce a statistically consistent dataset of maximum average wobbling specimen displacements. The reasons for this behavior would need to be investigated (e.g., systematic or random experimental errors) and the experimental campaign would be redesigned accordingly.

To perform a K-S test of this null hypothesis, an experimental benchmark response dataset was split into two randomly selected data subsets D_1 and D_2 (this is valid, since all shake table tests were performed under the same conditions). The p -value is computed for the following hypotheses:

$$H_0: \bar{F}_{D_1}^{ex} = \bar{F}_{D_2}^{ex} \text{ vs } H_1: \bar{F}_{D_1}^{ex} \neq \bar{F}_{D_2}^{ex}$$

where $\bar{F}_{D_1}^{ex}$ and $\bar{F}_{D_2}^{ex}$ are the experimentally estimated conditional CDFs for the two data subsets. Since the p -value depends on the data subsets, the K-S test was repeated 10^5 times, each time with different and randomly selected data subsets D_1 and D_2 , each one including 49 elements. Table 6 reports the computed mean p -values for the El Centro and Chi-Chi ground motion ensemble experimental benchmark datasets. The obtained p -values are 0.53 for both ensembles. Therefore, referring to Table 5, there is small or no evidence against the consistency of the CDFs of the maximum average wobbling specimen displacement data in the experimental benchmark datasets. Furthermore, the fact that the same p -value is observed for both ground motion ensembles indicates that the shake table test data is consistent, regardless of the seed ground motion used to generate the excitation ensembles.

Test of Model Quality

The null hypothesis H_0 is that the CDF of the maximum average displacement of the wobbling specimen obtained by shake table testing is matched by the CDF of the maximum average wobbling specimen model prediction. If this hypothesis is rejected, there is strong evidence that the considered model cannot predict the CDF of the maximum average displacement of the wobbling specimen, i.e., that this model is a poor proxy of the wobbling motion. The following hypothesis is tested:

Table 5. Evidence classification p -value scale [92]

p-value	Evidence
<0.01	Very strong against H_0
0.01-0.05	Strong against H_0
0.05-0.10	Medium – weak against H_0
>0.10	Small or none against H_0

Table 6. Computed p -values for experimental benchmark data set consistency and the model quality K-S hypothesis tests.

Ground motion ensemble	El Centro	Chi-Chi
Benchmark dataset consistency test	0.53	0.53
Model #9 quality test	9×10^{-4}	0.43
Model #3 quality test	2×10^{-6}	0.95

$$H_0: \bar{F}_D^{ex} = \bar{F}_D^m \text{ vs } H_1: \bar{F}_D^{ex} \neq \bar{F}_D^m$$

where $\bar{F}_D^{ex}$ and $\bar{F}_D^m$ are the CDFs of the maximum average wobbling specimen displacement obtained from the experimental benchmark and the model response datasets, respectively.

The computed p -values are shown in Table 6. Referring to Table 5, the K-S hypothesis test reveals that there is small or no evidence against the hypothesis that Team #9 and Team #3 models produce good predictions of the CDF of the maximum average wobbling specimen displacement for the Chi-Chi ground motion excitation ensemble. However, there is very strong evidence that the models are not good predictors of the same CDF for the El Centro ground motion excitation ensemble. The reasons for this finding could be related to the different duration of the ground motions in these ensembles (the ground motions were consistently scaled to match the target elastic acceleration response spectra in terms of their amplitude, but their duration was not altered with respect to the seed ground motion), but are not investigated further in this study. Clearly, there is room for improvement of the seismic wobbling response models.

CONCLUSIONS

The seismic response of rocking structures is notoriously sensitive to all of the parameters that define it, to the point that shake table tests are commonly not repeatable. As expected, numerical models often fail to adequately predict the seismic response of a rocking structure to a single ground motion. Based on this, researchers have claimed that not only are current models inaccurate, but that a numerical model accurate enough to predict the seismic response of rocking structures is not feasible. Hence, rocking is not used as a seismic design approach, and virtually all engineered non-base isolated structures are designed as fixed to the ground.

Along these lines, this paper revisits the issue of seismic response predictability in seismic design. It claims that predicting the response to a single ground motion is too strong of an acceptance test for a model. Instead, the minimum precondition for a valid model to be used in seismic design is the ability to predict the statistical characteristics of the seismic responses to an ensemble of ground motions that are compatible to the seismic hazard of interest. Similarly, what needs to be repeatable is not the response to a single ground motion, but the statistical characteristics of the response to an ensemble of ground motions. This is a weaker, but sufficient model validation test.

The objective of this study is to answer two simple, yet important questions: 1) Are seismic tests of three-dimensional rocking structures repeatable? and 2) If yes, is the observed response predictable using numerical models? To answer the above questions in a statistical sense, an experimental campaign has been designed to obtain observations of the seismic response of a rocking podium structure that was able to sustain three-dimensional rocking motion without sliding. To constrain the uncertainty of the excitation, a stochastic model was used to generate two synthetic ground motion ensembles that match the physical characteristics of two recorded ground motions. These two ensembles were used to excite a rocking podium structure comprising 4 steel tubular columns supporting an aluminum slab and the corresponding response data was measured in the conducted 200 shake table tests.

It was found that the Cumulative Distribution Function (CDF) of the maxima of the recorded time history specimen responses was repeatable. This conclusion was formally tested by dividing each ensemble of results to two randomly selected subsets and performing a Kolmogorov – Smirnov test between them. Hence, the obtained datasets indicate that a response of the specimen is stable in the statistical sense.

A blind prediction contest was organized to evaluate the ability of several models submitted by the contestants to predict the response of the rocking podium structure statistically. To this end, the models were evaluated based on their ability to predict not the time history response to each single ground motion, but the CDF of the maxima of the specimen displacement response time histories. Contestants used FEM, DEM and Rigid Body Dynamics models.

The winning model, developed by means of the FEM, scored the best of all models under three different norms that measure the distance between the experimentally obtained and the computed specimen displacement CDFs. However, this is not sufficient to conclude that FEM models were uniformly better than the submitted DEM models, because some DEM submissions were also very accurate, plus some of the least accurate models were FEM. Therefore, both FEM and DEM can perform well or poorly in simulating the wobbling (three-dimensional rocking) seismic response, depending on the parameters used and the assumptions made. It was also found that the numerical models were able to predict better the specimen displacement CDF obtained using the Chi-Chi ground motion ensemble than the El Centro one. The Chi-Chi ensemble is more intense and induced larger specimen displacements, which is an indication

that the numerical models predict larger wobbling displacements better than small ones, hence they are expected to do well in predicting overturning. The ranking of the models depends also on the norm used to measure the distance between the experimental benchmark and the modeled specimen displacement CDFs. Nevertheless, models that excel according to one norm, behaved well according to the other norms, too. An objective decision on which norm is better depends on which structural performance limit state is sought in design. For example, if the design question concerns the probability of seismic overturning/collapse of the structure, then models should be compared using norms that emphasize large displacements. Notably 6 out of 13 contesting teams predicted the 90th percentile of the specimen maximum displacement CDF from the Chi-Chi ground motion ensemble with an accuracy of error less than 10%, indicating that the seismic overturning limit state using these models can be predicted with good confidence using these models. Therefore, it can be concluded that the response of a 3-dimensional rocking structure is predictable in terms of a CDF of the maxima of the time-history responses to a set of ground motions.

In terms of rocking and wobbling seismic response modelling guidelines, this blind prediction contest confirmed that energy dissipation is a key factor that needs to be modelled, though not by means of Rayleigh damping. The models that best predicted the response used zero Rayleigh damping, but modeled energy dissipation directly through friction elements. A model with 0.1% Rayleigh damping that explicitly modelled friction as well also performed relatively well, but a model with 1% Rayleigh damping clearly underestimated the response. The rigid body model which modelled no energy dissipation at all, overestimated the response and should only be used only to obtain conservative wobbling response estimates.

ACKNOWLEDGEMENT

This research was conducted under the auspices of the project “Statistical verification and validation of 3D seismic rocking motion models (3DROCK)”, funded by the EU H2020 programme under grant agreement number 730900 [SERA]. Additional funding from ETH Zurich, University of Bristol and the Pacific Earthquake Engineering Research Center, as well as the contribution of the personnel at EQUALS laboratory, University of Bristol, and Erika Donald and Amarnath Kasalanti of PEER Center and UC Berkeley are gratefully acknowledged.

We also acknowledge the participation of the following teams in the blind prediction contest, herein listed alphabetically (not following the team numbering in this paper for reasons of anonymity):

1) Turki Alalawi, Abduaziz S. Al-Hazmi, Majd R. Zain, Jahad I. Al-harbi, Maha M. Hassan, and Mohamed S. Laissy; 2) Yu Bao and Dimitrios Konstantinidis; 3) Shanshan Chen, JingCong Lin, Yanfei Cai, Hanbin Lin, Xiaomin Liang, Jiatun Deng, Haoyun Lin, Lei Zhang, Difang Huang, Zinan Wu, and Xiancai Chen; 4) Ioannis P. Christovasilis, Lorenzo Riparbelli, Dimitris Papaevangelou, Gabriele Pini; 5) Tamás Forgács, Daniel Joó, Sándor Ádány, Vasilis Sarhosis, and José V. Lemos; 6) Jincheng Liu and Xinxin Wei; 7) Royce Liu; 8) Daniele Malomo, Anjali Mehrotra, and Matthew J. DeJong; 9) Yi Peng, Jianian Wen, Zhenlei Jia, Yazhou Xie, Jian Zhang; 10) P. Ravi Prakash, Bora Pulatsu, Nuno Mendes and Paulo B. Lourenço; 11) Chiyun Zhong and Constantin Christopoulos, as well as the 2 participating teams that wished to remain anonymous.

REFERENCES

1. Housner, GW (1963): The behaviour of inverted pendulum structures during earthquakes, *Bull. Seismol. Soc. Am.* **53**, 404–417.
2. Konstantinidis, D., & Makris, N. (2010). Experimental and analytical studies on the response of 1/4-scale models of freestanding laboratory equipment subjected to strong earthquake shaking. *Bull. Earth. Eng.*, 8(6), 1457–1477.
3. Di Egidio, A., Alaggio, R., Contento, A., Tursini, M., & Della Loggia, E. (2015). Experimental characterization of the overturning of three-dimensional square based rigid block. *Int. J. Non Linear Mech.*, 69, 137–145
4. Wittich, C. E., & Hutchinson, T. C. (2015). Shake table tests of stiff, unattached, asymmetric structures. *Earthq Eng Struct Dynam*, 44(14), 2425–2443.
5. Dar, A., Konstantinidis, D., & El-Dakhakhni, W. W. (2016). Evaluation of ASCE 43-05 seismic design criteria for rocking objects in nuclear facilities. *J Struct Eng*, 142(11), 04016110.
6. Sextos, A. G., Manolis, G. D., Ioannidis, N., & Athanasiou, A. (2017). Seismically induced uplift effects on nuclear power plants. Part 2: Demand on internal equipment. *Nuclear Engineering and Design*, 318, 288–296.
7. Dar, A., Konstantinidis, D., & El-Dakhakhni, W. (2018). Seismic response of rocking frames with top support eccentricity. *Earthq Eng Struct Dynam*, 47(12), 2496–2518.
8. Voyagaki E, Psycharis IN, & Mylonakis G. (2013). Rocking Response and Overturning Criteria for Free Standing Rigid Blocks to Single – Lobe Pulses. *Soil Dynamics & Earthquake Engineering*, 46, 85–95.

9. Voyagaki E, Psycharis IN, & Mylonakis G. (2014). Complex Response of a Rocking Block to a Full-Cycle Pulse. *Journal of Engineering Mechanics ASCE*, 140(16)
10. Voyagaki, E., Kloukinas, P., Dietz, M., Dihoru, L., Horseman, T., Oddbjornsson, O., ... & Steer, A. (2018). Earthquake response of a multiblock nuclear reactor graphite core: Experimental model vs simulations. *Earthq Eng Struct Dynam*, 47(13), 2601-2626.
11. Di Sarno, L., Magliulo, G., D'Angela, D., & Cosenza, E. (2019). Experimental assessment of the seismic performance of hospital cabinets using shake table testing. *Earthq Eng Struct Dynam*, 48(1), 103-123.
12. Bao, Y., & Konstantinidis, D. (2020). Dynamics of a sliding-rocking block considering impact with an adjacent wall. *Earthquake Engineering & Structural Dynamics*, 49(5), 498-523.
13. Stefanou, I., Psycharis, I., & Georgopoulos, I. O. (2011). Dynamic response of reinforced masonry columns in classical monuments. *Constr. Build. Mater.* 25(12), 4325-4337.
14. DeJong, M. J. (2012). Seismic response of stone masonry spires: Analytical modeling. *Eng. Struct.*, 40, 556-565.
15. Tondelli, M., Beyer, K., & DeJong, M. (2016). Influence of boundary conditions on the out-of-plane response of brick masonry walls in buildings with RC slabs. *Earthq Eng Struct Dynam*, 45(8), 1337-1356.
16. Casapulla, C., Giresini, L., & Lourenço, P. B. (2017). Rocking and Kinematic Approaches for Rigid Block Analysis of Masonry Walls: State of the Art and Recent Developments. *Buildings*, 7(3), 69.
17. Kalliontzis, D., & Schultz, A. E. (2017). Characterizing the In-Plane Rocking Response of Masonry Walls with Unbonded Posttensioning. *J Struct Eng* 143(9), 04017110.
18. Mehrotra, A., & DeJong, M. J. (2018). The influence of interface geometry, stiffness, and crushing on the dynamic response of masonry collapse mechanisms. *Earthq Eng Struct Dynam*, 47(13), 2661-2681.
19. Shawa, O. A., de Felice, G., Mauro, A., & Sorrentino, L. (2012). Out-of-plane seismic behaviour of rocking masonry walls. *Earthquake Engineering & Structural Dynamics*, 41(5), 949-968.
20. Mendes, N., Costa, A. A., Lourenço, P. B., Bento, R., Beyer, K., de Felice, G., ... & Lemos, J. V. (2017). Methods and approaches for blind test predictions of out-of-plane behavior of masonry walls: A numerical comparative study. *International Journal of Architectural Heritage*, 11(1), 59-71.
21. Giresini, L., Sassu, M., & Sorrentino, L. (2018). In situ free-vibration tests on unrestrained and restrained rocking masonry walls. *Earthquake Engineering & Structural Dynamics*, 47(15), 3006-3025.
22. Mouzakis, H. P., Psycharis, I. N., Papastamatiou, D. Y., Carydis, P. G., Papantonopoulos, C., & Zambas, C. (2002). Experimental investigation of the earthquake response of a model of a marble classical column. *Earthq Eng Struct Dynam*, 31(9), 1681-1698.
23. Papantonopoulos, C., Psycharis, I. N., Papastamatiou, D. Y., Lemos, J. V., & Mouzakis, H. P. (2002). Numerical prediction of the earthquake response of classical columns using the distinct element method. *Earthq Eng Struct Dynam*, 31(9), 1699-1717.
24. Papaloizou, L., & Komodromos, P. (2009). Planar investigation of the seismic response of ancient columns and colonnades with epistyles using a custom-made software. *Soil Dyn. Earthquake Eng.*, 29(11), 1437-1454.
25. Vassiliou, M. F., & Makris, N. (2012). Analysis of the rocking response of rigid blocks standing free on a seismically isolated base. *Earthq Eng Struct Dynam*, 41(2), 177-196.
26. Nikolić, Ž., Krstevska, L., Marović, P., & Smoljanović, H. (2019). Experimental investigation of seismic behaviour of the ancient Protiron monument model. *Earthq Eng Struct Dynam*, 48(6), 573-593.
27. Cherepinskiy Y. Seismic isolation of buildings with application of the kinematics bases. *13th World Conference on Earthquake Engineering*, 2004.
28. Uzdin AM, Doronin FA, Davydova G V., Avidon GE, Karlina EA. Performance analysis of seismic-insulating elements with negative stiffness. *Soil Mechanics and Foundation Engineering* 2009; **46**(3): 15–21.
29. Bachmann, J. A., Vassiliou, M. F., & Stojadinović, B. (2018). Dynamics of rocking podium structures. *Earthq Eng Struct Dynam*, 46(14), 2499-2517.
30. Bachmann, J. A., Vassiliou, M. F., & Stojadinovic, B. (2019). Rolling and rocking of rigid uplifting structures. *Earthq Eng Struct Dynam*, 48(14), 1556-1574.
31. Bachmann, J. (2018). *Development of Self-Centering Systems with Geometry-Controlled Stiffness for Earthquake Hazard Mitigation* (Doctoral dissertation, ETH Zurich).
32. Ríos-García, G., & Benavent-Climent, A. (2020). New rocking column with control of negative stiffness displacement range and its application to RC frames. *Eng. Struct.*, 206, 110133.
33. Makris N. and Vassiliou M.F. (2013), Planar rocking response and stability analysis of an array of free standing columns capped with a freely supported rigid beam, *Earthquake Engng. Struct. Dyn.* 42(3), 431-449.
34. Makris, N., & Vassiliou, M. F. (2014). Are Some Top-Heavy Structures More Stable?. *J Struct Eng*, 140(5).
35. Dimitrakopoulos, E. G., & Giouvanidis, A. I. (2015). Seismic Response Analysis of the Planar Rocking Frame. *J. Eng. Mech.*, 141(7), 04015003.

36. Agalianos, A., Psychari, A., Vassiliou, M. F., Stojadinovic, B., & Anastasopoulos, I. (2017). Comparative Assessment of Two Rocking Isolation Techniques for a Motorway Overpass Bridge. *Front. Built Environ.* 3, 47.
37. Vassiliou, M. F., Mackie, K. R., and Stojadinović, B. (2017) A finite element model for seismic response analysis of deformable rocking frames. *Earthquake Engng Struct. Dyn.*, 46: 447–466.
38. Makris, N., & Konstantinidis, D. (2003). The rocking spectrum and the limitations of practical design methodologies. *Earthq Eng Struct Dynam*, 32(2), 265-289.
39. Reggiani Manzo, N., & Vassiliou, M. F. (2019). Displacement-based analysis and design of rocking structures. *Earthq Eng Struct Dynam*, 48(14), 1613-1629.
40. Reggiani Manzo, N., & Vassiliou, M. F. (2020). Dimensional analysis of negative stiffness bilinear systems under pulse like excitations. *Earthq Eng Struct Dynam*, DOI: 10.1002/eqe.3347
41. van den Engh, G., Nelson, P., & Roach, J. (2000). Numismatic gyrations. *Nature*, 408(6812), 540-540.
42. Koh AS and Mustafa G (1990): Free rocking of cylindrical structures. *J. Eng. Mech.*, 116(1), 35-54.
43. Koh AS and Hsiung CM (1991): Base isolation benefits of 3-D rocking and uplift. I: Theory. *J. Eng. Mech.*, 117(1), 1-18.
44. Koh AS and Hsiung CM (1991): Base isolation benefits of 3-D rocking and uplift. II: Numerical Example. *J. Eng. Mech.*, 117(1), 19-31.
45. Ambraseys, N., & Psycharis, I. N. (2011). Earthquake stability of columns and statues. *Journal of Earthquake Engineering*, 15(5), 685-710.
46. Psycharis, I. N., Fragiadakis, M., & Stefanou, I. (2013). Seismic reliability assessment of classical columns subjected to near-fault ground motions. *Earthquake engineering & structural dynamics*, 42(14), 2061-2079.
47. Krstevska, L., Mihailov, V., Boschi, E., & Rovelli, A. (1996). Experimental dynamic testing of prototype and model of the Antonina Column in Roma. *Proceedings of the 11th world conference on earthquake engineering*. Paper (No. 546).
48. Mouzakis, H. P., Psycharis, I. N., Papastamatiou, D. Y., Carydis, P. G., Papantonopoulos, C., & Zambas, C. (2002). Experimental investigation of the earthquake response of a model of a marble classical column. *Earthquake engineering & structural dynamics*, 31(9), 1681-1698.
49. Drosos, V., & Anastasopoulos, I. (2014). Shaking table testing of multidrum columns and portals. *Earthq Eng Struct Dynam*, 43(11), 1703-1723.
50. Pappas, A., Sextos, A., Da Porto, F., & Modena, C. (2017). Efficiency of alternative intensity measures for the seismic assessment of monolithic free-standing columns. *Bull. Earthquake Eng.*, 15(4), 1635-1659.
51. Makris N, Alexakis H, Kampas G, Strepelias I, Kolonas C and Bousias S (2015): *Seismic protection of bridges with rocking piers which recenter with gravity*, Report EEAM, 2015-01, University of Patras Dept of Civil Engng
52. Stefanou, I., Vardoulakis, I., & Mavraganis, A. (2011). Dynamic motion of a conical frustum over a rough horizontal plane. *Int. J. Nonlinear Mech.*, 46(1), 114-124.
53. Konstantinidis D and Makris N (2007): The dynamics of a rocking block in three dimensions, 8th Hell. Soc. Theor. Appl. Mech. Int. Congr. Mech. Patras, Greece 9/2015.
54. Zulli D, Contento A, and Di Egidio A (2102): 3D model of rigid block with a rectangular base subject to pulse-type excitation, *Int. J. Non. Linear. Mech.*, 47 (6), 679–687.
55. Chatzis, MN, and Smyth AW (2012). Modeling of the 3D rocking problem. *International Journal of Non-Linear Mechanics*, 47(4) 85-98.
56. Chatzis, M. N., & Smyth, A. W. (2012). Three-dimensional dynamics of a rigid body with wheels on a moving base. *Journal of Engineering Mechanics*, 139(4), 496-511.
57. Mathey, C., Feau, C., Politopoulos, I., Clair, D., Baillet, L., & Fogli, M. (2016). Behavior of rigid blocks with geometrical defects under seismic motion: an experimental and numerical study. *Earthquake Engineering & Structural Dynamics*, 45(15), 2455-2474.
58. Vassiliou, M. F., Burger, S., Egger, M., Bachmann, J. A., Broccardo, M., & Stojadinovic, B. (2017). The three-dimensional behavior of inverted pendulum cylindrical structures during earthquakes. *Earthq Eng Struct Dynam*, 46(14), 2261-2280.
59. Vassiliou, M. F. (2018). Seismic response of a wobbling 3D frame. *Earthquake Engineering & Structural Dynamics*, 47(5), 1212-1228.
60. Bachmann, J. A., Strand, M., Vassiliou, M. F., Broccardo, M., & Stojadinović, B. (2018). Is rocking motion predictable?. *Earthq Eng Struct Dynam*, 47(2), 535-552.
61. Bachmann, J., Strand, M., Vassiliou, M. F., Broccardo, M., & Stojadinovic, B. (2019) “Modelling of rocking structures: Are our models good enough?,” *Proceedings of 2nd International Conference on Natural Hazards & Infrastructure*, Crete, Greece.

62. Del Giudice, L., Wrobel, R., Leinenbach, C., & Vassiliou, M. F. (2020). Static Testing of Additively Manufactured Microreinforced Concrete Specimens for Statistical Structural Model Validation at a Small Scale. In *8th International Conference on Advances in Experimental Structural Engineering (8AESE)*.
63. Sargent, R. G. (1988, December). A tutorial on validation and verification of simulation models. In *Proceedings of the 20th conference on Winter simulation* (pp. 33-39).
64. Chopra A.K. (2012) *Dynamics of structures: theory and applications to earthquake engineering. 4th ed.*, Pearson, Prentice Hall
65. <http://www.bris.ac.uk/engineering/research/earthquakegeo/equip/tablespec/> (accessed on April 29th 2020)
66. Rezaeian S, Der Kiureghian A. (2008) A stochastic ground motion model with separable temporal and spectral nonstationarities. *Earthq Eng Struct Dyn* **37**(13):1565-1584.
67. Broccardo M, Dabaghi M. (2017) A spectral-based stochastic ground motion model with a non-parametric time-modulating function. In: 12th International Conference on Structural Safety and Reliability; Vienna; 2017:1-10.
68. Broccardo M, Der Kiureghian A. (2014) Simulation of near-fault ground motions using frequency-domain discretization. In: Proceedings of the 10th NCEE; Anchorage, Alaska; 2014:1-11.
69. Chopra AK, Yim SCS. Simplified earthquake analysis of structures with foundation uplift. *J Struct Eng.* 1985;111(4):906-930.
70. Psycharis IN. Effect of base uplift on dynamic response of SDOF structures. *J Struct Eng.* 1991;117(3):733-754.
71. Oliveto G, Calio I, Greco A. Large displacement behaviour of a structural model with foundation uplift under impulsive and earthquake excitations. *Earthq Eng Struct Dynam.* 2003;32(3):369-393.
72. Ma QTM. The mechanics of rocking structures subjected to ground motion (Doctoral dissertation, ResearchSpace@ Auckland) 2010.
73. Acikgoz S, DeJong MJ. The interaction of elasticity and rocking in flexible structures allowed to uplift. *Earthq Eng Struct Dynam.* 2012;41(15):2177-2194.
74. Vassiliou, M. F., Mackie, K. R., & Stojadinović, B. (2014). Dynamic response analysis of solitary flexible rocking bodies: modeling and behavior under pulse-like ground excitation. *Earthq Eng Struct Dyn.*, 43(10), 1463-1481.
75. Vassiliou, M. F., Truniger, R., & Stojadinović, B. (2015). An analytical model of a deformable cantilever structure rocking on a rigid surface: development and verification. *Earthq Eng Struct Dynam*, 44(15), 2775-2794.
76. Truniger, R., Vassiliou, M. F., & Stojadinović, B. (2015). An analytical model of a deformable cantilever structure rocking on a rigid surface: experimental validation. *Earthq Eng Struct Dynam*, 44(15), 2795-2815.
77. Kolmogorov A.N. (1933) Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari.*, 4, 83-91.
78. N Smirnov. (1939) Sur les écarts de la courbe de distribution empirique. *Matematicheskii Sbornik*, 48(1):3-26,
79. ABAQUS (2012). *Standard User's Manual*. Providence, RI: Dassault Systèmes Simulia Corp.,
80. CSI (2019), *SAP2000 Integrated Software for Structural Analysis and Design* Computers and Structures Inc., Berkeley, California.
81. MIDAS (2019), *Design of General Structures, Integrated Design System for Building and General Structures*, MIDASoft, Inc.
82. Mazzoni, S., McKenna, F., Scott, M. H., & Fenves, G. L. (2006). OpenSees command language manual. *Pacific Earthquake Engineering Research (PEER) Center*, 264.
83. Itasca, C. G. (2007). 3DEC Version 4.1 User's Guide. *Minneapolis, Minnesota USA*.
84. Aster, C. (2001). *General public licensed structural mechanics finite element software*.
85. Zhong M.C. and Christopoulos C. (2020), Finite element analysis of the seismic response of a rocking podium structure, 2020 PEER Annual Meeting, Berkeley, USA, 16-17.01.2020 (poster presentation)
86. Malomo D, Mehrotra A, DeJong M.J., Discrete Element Modeling of a rocking podium structure subjected to biaxial shake-table test, 2020 PEER Annual Meeting, Berkeley, USA, 16-17.01.2020 (poster presentation)
87. Chopra, A. K., & McKenna, F. (2016). Modeling viscous damping in nonlinear response history analysis of buildings for earthquake excitation. *Earthq Eng Struct Dynam*, 45(2), 193-211.
88. Cruz, C., & Miranda, E. (2017). Evaluation of the Rayleigh damping model for buildings. *Eng. Struct.*, 138, 324-336.
89. Hall, John F. (2018) *Performance of Viscous Damping in Inelastic Seismic Analysis of Moment-Frame Buildings*. EERL Report, 2018-01. California Institute of Technology, Pasadena, CA.
90. Stewart, J., Crouse, C. B., Hutchinson, T. C., Lizundia, B., Naeim, F., & Ostadan, F. (2012). *Soil-structure interaction for building structures* (No. Grant/Contract Reports (NISTGCR)-12-917-21).
91. Massey FJ. The Kolmogorov-Smirnov test for goodness of fit. *J Am Stat Assoc.* 1951;46(253):68-78.
92. Wasserman, L. (2013). *All of statistics: a concise course in statistical inference*. Springer Science & Business Media.