1 INTRODUCTION

This paper describes a proposed procedure for comparing the quality of numerical simulation tools used for the prediction of effects of soil liquefaction. The development of this procedure is part of a collaboration of several universities and industry partners in the US, Japan, UK, Taiwan, and China. This project is called LEAP (Liquefaction Experiments and Analysis Projects). It is envisioned that there will be a series of asynchronous validation exercises (e.g., LEAP 1, LEAP 2, ..., LEAP n) hosted by different universities in the collaboratory. Each “LEAP” may focus on a certain issue related to liquefaction or a class of liquefaction problems (e.g., earth dams, lateral spreading, retaining structures, or waterfront structures). Ideally, the validation process will be similar from LEAP to LEAP, but evolution of this process is anticipated.

The LEAP workshop in Kyoto in September 2014 is one of the first LEAP exercises in this collaborative effort. This paper is presented at this workshop as a proposal for consideration for implementation in future LEAP exercises. An updated version of the proposed LEAP verification and validation process, along with a working draft of the evolving document is to be posted in a public folder on NEEShub, the data archiving website for the Network for Earthquake Engineering Simulation (NEES). This paper describes the current state of thinking about this verification and validation (V&V) procedure. Because the success of this project requires community consensus, it is important to present and discuss the V&V procedure prior to adoption in future LEAP projects.

2 CONCEPT OUTLINE FOR LEAP V&V PROCESS

The LEAP project documentation will include templates that the participants in numerical simulation exercise use to describe their constitutive and numerical models as well as laboratory and model test data used in the validation process. LEAP documentation will also describe the formats for submission of shared results that allow for meaningful and fair evaluation and side-by-side comparison of the capabilities and limitations of the models. The documentation will eventually attempt to establish metrics to score the quality of predictions on a variety of objective scales.

As described by Jeremic et al (2009) and others, the verification and validation process is differentiated as follows. Verification is the process of determining that a model implementation accurately represents the developer’s conceptual description and specification; verification provides evidence that the model is solved correctly. Validation is the process of determining the degree to which a model accurately models the intended real world behavior.

The evolving concept for LEAP verification and validation includes the above philosophy, but also requires the verification and validation results be clearly understood and fairly compared so that users
of these models can accurately and practically assess their capabilities and limitations. This is important for decision makers, such as leaders of regulatory agencies that determine the safety of infrastructure and allocate resources accordingly. It is envisioned that the verification and validation process not only verifies and validates – it also must illustrate the quality of the verification and validation to users and to decision makers that are not necessarily numerical modeling experts.

Thus, the suggested LEAP V&V process may be outlined as follows.

- **Constitutive model verification**: description of physics intended to be modeled with graphical presentations to illustrate predicted constitutive behavior for specific standard test paths.
- **Numerical model verification**: description of discretization and solution schemes with graphical comparisons to illustrate their stability and accuracy in selected numerical tests.
- **Element test validation**: predictions of constitutive model are compared to data from laboratory test data (e.g., triaxial, simple shear, and one-dimensional consolidation tests).
- **Class C validation** by comparing predictions of numerical models to data from:
  - Benchmark centrifuge model tests (relatively simple geometry such as 1-D saturated sloping ground)
  - System physical model tests (e.g., more complex problems such as dams, cases involving soil-structure interaction, etc.)
  - Benchmark simulations of liquefaction in the field (e.g., validation against observed liquefaction in the Christchurch earthquakes of 2010-2011)
- **Class A validation**:
  - New benchmark centrifuge model tests (relatively simple geometry such as 1-D saturated sloping ground)
  - New system centrifuge tests (e.g. dams, soil-structure interaction, etc.)

Similar to Lambe’s (1973) explanation, we differentiate the classes of validation as follows. **Class A** is a true prediction of an event made prior to the event. A **Class B** prediction is after the event, but with results unknown to the predictor. A **Class C** validation is after the event, with results known to the predictor. For class C validation, the individual(s) conducting the modeling (herein called “predictor(s)”) may or may not iteratively adjust the model parameters to improve the quality of the agreement between calculations and observations.

The verification steps (first two bullets in the above list) will identify which models have which capabilities, often with a yes/no or black/white binary classification. The verification steps do not require comparison to experimental data; but they do require demonstrating the model’s ability to provide consistent results. However the validation against experimental data (the last three bullets in the above list) requires metrics for assessing the quality of validation. Because validation involves comparison to experimental data, there will be inevitable questions about data consistency and accuracy. The LEAP process will endeavor to establish reliable calibration data and to develop reasonably robust validation metrics.

At the time of this paper submission, the verification steps (first two bullets in the above outline) have been developed to a greater extent than the validation steps (last three bullets). Hence this paper focuses more attention to verification than to validation.

## 3 CONSTITUTIVE MODEL VERIFICATION

Constitutive model verification is intended to develop a clear understanding of the range of behaviors that are intended to be simulated by the code and an illustration of the ability of the code to predict this behavior. Data will be collected in a standard format to allow systematic comparisons of the predictions from multiple constitutive models. Blank tables will be provided to encourage predictors to provide results in a format for easy cross-comparison of their constitutive model results with those of other researchers.

Verification of the constitutive models will involve the developers answering several key questions about their constitutive models (3.1) and conducting a series of specific verification tests (3.2) using a generic set of soil data like that found in Table 1.

### Table 1. Default hypothetical soil descriptions for verification of constitutive model

<table>
<thead>
<tr>
<th>Soil Type</th>
<th>Dense (D)</th>
<th>Medium Dense (M)</th>
<th>Loose (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Density</td>
<td>80</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>init. void ratio</td>
<td>0.55</td>
<td>0.7</td>
<td>0.85</td>
</tr>
<tr>
<td>Permeability</td>
<td>0.015 mm/s</td>
<td>0.03 mm/s</td>
<td>0.05 mm/s</td>
</tr>
<tr>
<td>Method of Placement</td>
<td>Dry pluviation</td>
<td>Dry pluviation</td>
<td>Dry pluviation</td>
</tr>
<tr>
<td>crit. State Friction Angle</td>
<td>32°</td>
<td>32°</td>
<td>32°</td>
</tr>
<tr>
<td>crit. Void Ratio at p’ = 100 kPa</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Gradation, D&lt;sub&gt;50&lt;/sub&gt;/D&lt;sub&gt;10&lt;/sub&gt;</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Gradation, D&lt;sub&gt;50&lt;/sub&gt;</td>
<td>0.2 mm</td>
<td>0.2 mm</td>
<td>0.2 mm</td>
</tr>
</tbody>
</table>
3.1 Constitutive Model Questionnaire

The numerical modelers will be asked to provide detailed papers and reports explaining their constitutive models. In addition, the following questions will be asked to enable tabulated comparisons of the capabilities of different models.

1. What are the mathematical equations for stress and strain parameters used in the constitutive model (e.g., p, q, Lode angle, b = (\(\sigma_2 - \sigma_3\))/(\(\sigma_1 - \sigma_3\)), volumetric strain, shear strain)?

2. Is the model formulated in the general stress-strain pace or is it designed for a special case such as plane strain condition?

3. Does the model account for rotation of the direction of the principle stresses? How?

4. Critical state:
   a. Does the model prescribe a critical state for which the void ratio or relative density at critical state is a function of the confining pressure?
   b. Does the model prescribe a critical state friction angle?

5. What are the dependent (embedded in code) and independent (user specified) parameters of the constitutive model and what behaviors they control?

6. What is the flow rule in the deviatoric plane and in various planes that contain the hydrostatic axis?

7. Does the model account for effects of grain crushing on the location of the critical state and normal compression lines in the e-p space?

8. Does the model include an ingredient to capture size effects and post-localization response?

9. Are strain-rate or aging effects included? If yes, describe the rate and aging effects.

    a. Does the constitutive model include damping or hysteretic energy dissipation at very small strains?
    b. Is it expected or typical to use numerical damping at the system equation level (e.g. Rayleigh damping) to account for energy dissipation at the constitutive level?

11. What are other important limitations and capabilities of the constitutive model?

3.2 Constitutive model verification test results

At least five classes of verification tests were identified and are described next with blank tables when appropriate to provide an example of how data may be input. As mentioned earlier, blank tables like the ones shown in this paper will be provided to predictors to encourage them to provide results in a format for easy cross-comparison of their constitutive model results with those of other researchers.

### Table 2. Example of results template for isotropic and 1-D consolidation verifications.

<table>
<thead>
<tr>
<th>Isotropic Consolidation</th>
<th>1-D Consolidation</th>
</tr>
</thead>
<tbody>
<tr>
<td>p' (kPa)</td>
<td>Void Ratio, e</td>
</tr>
<tr>
<td>10</td>
<td>0.7</td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

#### 3.2.1 Isotropic consolidation

Consider a medium density fine sand (see Table 1) that is pluviated in air and is normally consolidated to the initial state indicated in Table 2. The verification test is to subject the soil to isotropic consolidation for a sequence of mean effective stresses indicated in the Table 2. The blank columns in the table would be filled out by the predictors.

This sequence of stresses is designed to identify typical volumetric strain behavior; how the model behavior changes as it transitions from overconsolidated to normally consolidated states, to visually show how the model handles volumetric strains at very low effective stresses (some models introduce a minimum pressure parameter below which the constitutive behavior is simplified), and finally to show if the model is able to account for crushing.

#### 3.2.2 One-dimensional consolidation

Consider a medium density fine sand (see Table 1) that is pluviated in air and normally consolidated to the initial state indicated in Table 2, with the initial coefficient of lateral earth pressure, \(K_0 = 1\). The verification test is to show the evolution of \(K_0\) and void ratio under the sequence of vertical stresses indicated in Table 2.

#### 3.2.3 Monotonic drained test paths

This specific verification test is to calculate the monotonic drained test paths for conventional
drained triaxial conditions with constant horizontal stress, \( \sigma_r \), beginning with the initial void ratio described in Table 1, initially isotropically normally consolidated to \( p' = 100 \) kPa, will be calculated. Table 3 shows the format for submission of predicted test paths. For the strains specified, the void ratio and deviatoric stress \( q = (\sigma_a - \sigma_r) \) will be reported. Compression test paths will be applied to loose, medium and dense samples, and an extension path will be applied to medium dense sand. Note that ICDC stands for Isotropically Consolidated Drained Compression and ICDE stands for Isotropically Consolidated Drained Extension.

In addition to the cases listed in Table 3, the predicted paths for direct simple shear tests should also be predicted in a similar format. Instead of reporting values of \( q \) and axial strain, shear stress, shear strain and the evolution of the horizontal effective stress will be reported while the effective vertical stress is held constant. If a numerical procedure is not able to simulate one of these prescribed load paths, the table will be left blank.

### Table 3. Typical calculated test paths for default soils in triaxial tests.

<table>
<thead>
<tr>
<th>Soil and test path</th>
<th>Loose ICDC</th>
<th>Medium Dense ICDC</th>
<th>Dense ICDC</th>
<th>Medium Dense ICDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(</td>
<td>e_a</td>
<td>(%)</td>
<td>q</td>
<td>e</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.85</td>
<td>0</td>
<td>0.7</td>
</tr>
<tr>
<td>0.001</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
</tr>
</tbody>
</table>

The data collected in this way will allow assessment of the effective shear modulus as a function of shear strain as well as the ability of the constitutive model to predict dilatancy, peak and critical state friction angles and void ratios at critical state.

#### 3.2.4 Monotonic constant volume shearing

This specific verification test is to calculate the monotonic undrained test paths for conventional consolidated undrained triaxial conditions beginning with the initial void ratio described in Table 1, initially isotropically normally consolidated to \( p' = 100 \) kPa. A data template similar to that for drained tests will be provided, except that instead of reporting \( q \) and \( e \) as a function of strain, data regarding \( q \) and \( p' \) will be reported for triaxial compression (ICUC) tests and extension (ICUE) tests. Compression test paths will be applied to loose, medium and dense samples, and an extension path will be applied to a medium dense sand sample.

In addition to the cases listed in Table 3, the predicted path for simple shear tests should also be predicted in a similar format. Instead of reporting values of \( q \) and \( p' \), shear stress, shear strain and the evolution of the horizontal effective stress will be reported while the effective vertical stress is held constant.

The data collected in this way will allow assessment of undrained behavior for dilatant and contractive soil, and to demonstration of the influence of the test path (triaxial compression, triaxial extension, and simple shear) on the undrained test path.

#### 3.2.5 Cyclic undrained shearing

For cyclic undrained shear the verification test will be to determine the relationship between the CRR (cyclic resistance ratio) and number of cycles of loading for uniform loading. The CRR is defined as the cyclic stress ratio to cause development of a selected threshold strain. For simple shear conditions the cyclic stress ratio is defined as the ratio of cyclic shear stress to the initial consolidation stress prior to the undrained cyclic loading.

### Table 4. CRR curves to for verification

<table>
<thead>
<tr>
<th>CRR curve</th>
<th>Soil Density</th>
<th>Consol. stress ( kPa )</th>
<th>OCR</th>
<th>Test path</th>
<th>Threshold strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a,b,c D 0 100 1 DSS 1,3,10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 a,b,c M 0 100 1 DSS 1,3,10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 a,b,c L 0 100 1 DSS 1,3,10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 a,b,c D 1.5 100 1 DSS 1,3,10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 a,b,c M 1.5 100 1 DSS 1,3,10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 a,b,c L 1.5 100 1 DSS 1,3,10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 M 0 100 10 DSS 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 D 0 1000 1 DSS 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 M 0 1000 1 DSS 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 L 0 1000 1 DSS 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 M 0 100 1 Triax 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 M 0.5 100 1 Triax 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 M 1.5 100 1 Triax 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 shows a hypothetical effect of the soil density on the cyclic stress to cause 3% strain. In addition to the effects of density on the CRR curves, the effects of other factors listed in Table 4 on CRR curves will be demonstrated.
Figure 1. Hypothetical cyclic resistance ratio (CRR) curves predicted by numerical model.

The static shear stress is also known to have an important effect on cyclic loading. One of the factors affecting the importance of static shear stress is the reversal of shear stress direction after passing through zero shear stress. If the shear stress never actually reaches zero, then the soil particles must maintain effective stress contact. This factor is controlled by the ratio \( \frac{\tau_{\text{static}}}{\tau_{\text{cyclic}}} \); when this ratio is greater than 1, the shear stress does not pass through zero as illustrated by the dashed path in Figure 2. For the solid line, the cyclic shear stress is greater than the static shear stress.

As reconsolidation following cyclic loading causes volumetric strains and settlement, as a general practice, for every undrained cyclic loading test simulated, the cyclic loading should be terminated when the threshold strain is achieved, and then the soil should be reconsolidated back to the initial consolidation stress and static shear stress. The strains during reconsolidation will be reported as a function of the reconsolidation pressure.

3.2.6 Other features affecting liquefaction

Various constitutive models have unique features that will not be demonstrated in the above verifications. For example, some recent models (e.g., Manzari & Dafalias 1997, Dafalias & Manzari, 2004) account for the evolution of fabric that develops during cyclic loading.

Models that are true 3-D models will also have some capability to model the effects of multidirectional shaking. The predictors will be encouraged to provide specific simulation data to demonstrate the model’s other capabilities.

4 NUMERICAL MODEL VERIFICATION

Similar to the constitutive model verification, the numerical model verification will consist of first asking developers several key questions about their numerical model (4.1) and conducting a series of specific verification tests (4.2).

4.1 Numerical model questionnaire

The numerical modelers will be asked to provide detailed papers and reports explaining the detailed theory and implementation of numerical solutions. In addition, the following questions are asked to enable tabulated comparisons of capabilities of different models.

1. What is the discretization approach (Finite Element, Finite Difference, a meshless method, or otherwise)?
2. What types of elements (number of nodes and integration points are used in the simulation, u-p, u-p-U, etc)?
3. Is the approach capable of accounting for large deformations and large rotations? How?
4. What integration procedure is used for stress-strain relationship?
5. Is Rayleigh damping used? If so, is the damping matrix determined by the tangent stiffness, the initial stiffness, or both?

4.2 Numerical verification results

4.2.1 Time-rate of one-dimensional consolidation

Using a specified set of material properties (permeability, porosity, and elastic moduli the numerical so-
A solution to Terzaghi’s theory of consolidation should be presented.

4.2.2 Free vibration of an elastic column
A 1-dimensional column of soil 10 m thick will be subject to a shear stress step function applied at the ground surface. The shear wave will propagate to the base of the specimen and the time of arrival of the stress wave will be compared to that of the theoretical travel time \( t = H/V_s \), where \( H = 10 \text{ m} \) and \( V_s = (G/\rho)^{0.5} \). In addition the free vibration of the column will be displayed and the numerical damping during free vibration will be assessed.

4.2.3 Verification of Biot formulation and integration procedure
Following Jeremic et al. (2001), the implementation of the Biot formulation and accuracy of numerical integration will be illustrated by comparing the numerical model predictions to closed form solutions of shock wave propagation in poro-elastic material by Gajo (1995). Figure 3 illustrates the configuration of the validation scenario. A unit step function of displacement is applied to the surface, with amplitude 0.01 mm while the pressures and displacements of the fluid and solid phases are compared. Computed results by Jeremic et al. (2001) for this problem are compared to the closed form solution by Gajo (1995). Figure 4 shows the comparison between the numerical solution and the closed form solution.

4.2.4 Void migration beneath impermeable layer.
Malvick et al. (2006) presented centrifuge model tests of sloping liquefiable ground capped with an impermeable layer that traps the upward flowing water in the liquefying ground near the interface. They show that the soil loosens under the cap as illustrated in Figure 5. Researchers have attempted to capture this phenomenon numerically with some difficulty. The accumulation of water at this interface facilitates dilation (loosening of the soil near the interface); but the softening at the interface can lead to localization of strains and numerical instability. This appears to be a challenging issue that is not yet resolved; a generic submerged slope composed of medium dense sand, similar to that depicted in Figure 5 should be analyzed with the proposed numerical modeling frameworks.

4.2.5 Illustration of how shear strain localization is modeled
The sensitivity the relation between nominal stress-nominal strain to the mesh size shall be compared for a medium dense sand in plane strain biaxial compression loading. Results may be compared to
predictions from a single element, a 10 × 10 mesh of elements, 100 × 100 mesh of elements.

4.2.6 Illustration of large strains are modeling.
For a model of a DSS specimen consisting of a 10 × 10 mesh of elements, show the predicted path for

monotonic shearing to global apparent strain of 1000% strain. If a large strain formulation is used, show a comparison of a large strain simulation to a simulation with the large strain provision is turned off.

5 VALIDATION AGAINST ELEMENT TEST DATA

Up to this point all of the requested verification exercises may be completed without comparison to any experimental data. This section focuses on the validation of the constitutive models by comparing predictions of the constitutive model to the response measured in element tests such as triaxial, rotational shear, and DSS (direct simple shear) monotonic and cyclic loading tests. For each LEAP exercise, a different soil may be used. For the centrifuge or field experiments descried in the next section of this paper, element test data will be provided to help the predictors simulate the results of the centrifuge or field experiments.

This outline below describes the element test data that will ideally be made available to predictors and attempts to describe the metrics that will be used to evaluate the quality of the comparison between experiments and predictions:

1) Isotropic and 1-D consolidation comparison between experiment and calculation.
   a) Metrics for evaluation of comparison.

   i) Average difference between e measured and e predicted (over the important range of pressures)
   ii) Area of the hysteresis loop?

2) ICD monotonic
3) ICU monotonic
4) ICU cyclic
   i) p-q, q-ε_q, p-ε_p
   b) CSR vs Number of cycles to 3.5% strain
      i) Dr = 30, 50, 90?
      ii) Confining stress/P_0 = 0.2, 1, 5
   c) ACU cyclic
      i) Initial Static shear stress ratio = 0, 0.1
   d) Elastic behavior (V_s or G_max) as a function of confining pressure
   e) Reconsolidation volumetric strains following liquefaction

5) Experimentally developed shear modulus reduction curves.

6 CLASS C VALIDATION

Class C validation is considered to be comparison of numerically calculated response with known experimental data. Class C validation against one particular response to one particular loading case is not particularly illuminating because it is always possible for a predictor to adjust parameters to obtain a decent fit to a limited data set. On the other hand, validation against a series of cases can be very illuminating; for example, prediction of the response of a centrifuge model to a sequence of shaking events (small, intermediate and intense shaking, or shaking with different ground motions), while attempting to predict deformed shape, pore pressure contours, and spectral accelerations at multiple points. Even if the input properties and liquefying model responses are known, it may not be possible for any model to accurately predict all aspects of the response without changing the model parameters. If the results are known a priori, the ability of a model to duplicate a sequence of Class C validation events, without changing parameters from event to event will be much more meaningful than a simulation that predicts a single event.

It is envisioned that LEAP will archive (and make available to everyone) standard sequences of experiments archived especially for the purpose of validation of numerical models. Some of these experiments (Benchmark Model Tests) will be quite simple geometry, and others (System Model Tests) may attempt to model more realistic prototype situations with realistic geometries and heterogeneities.

6.1 Class C validation of benchmark tests

Benchmark tests will include, for example:
• uniform level ground sites with different densities modeled in laminar containers,
• uniform soil layers with sloping ground as in an infinite slope configuration,
• uniform submerged embankment on a rigid foundation.

Figure 6. Example of a LEAP benchmark tests performed for a workshop in Kyoto in 2013. (a) Top figure shows the model dimensions and instrument locations. (b) Bottom is a photograph of the actual container used for experiments at Kyoto University.

The details of the model and apparatus configuration along with all the input data will be archived and shared for free access for interested parties. This data includes a description of the experiment, model container, experimental apparatus, achieved input motions, soil properties measured in flight before shaking, dimensions, location and sensitivity of all sensors, and all other necessary information in more detail than can be explained in this paper.

Finally, each benchmark will have its own description and its own set of validation metrics.

6.1.1 Description of model container

Scrutiny of Figure 6 may beg the question: ‘What is the influence of the container on the results?’ Several researchers have considered model container effects on test results. Kutter (1995) discusses the importance of developing complementary shear stresses in laminar containers and on the damping that may be introduced by the shaker and container. Taboada and Dobry (1998) explain how the weight of the rings of a laminar container changes the effective ground slope -- the model slope represents a prototype slope perhaps two or three times steeper. Others have investigated other aspects of container effects. The properties of the model container (e.g., mass, stiffness, and friction) should be carefully documented for the finalized Class C validation benchmark tests. While there is some benefit to using complex containers designed to minimize boundary effects (such as that shown in Figure 6(b), complex containers have the disadvantage of being more difficult to model numerically. On the other simple containers (e.g., rigid box) are easy to model in an analysis, but they might be significantly constraining the model behavior. As each type of container has its advantages, validation benchmark tests will include tests in simple rigid containers as well as sophisticated, laminar containers. Detailed properties of each container along with their measured performance in satisfying their intended utilities will be documented and made available to predictors.

6.2 Example benchmark: Sloping ground in 1-D laminar container

6.2.1 Description of the benchmark

The benchmark will be described briefly with overview sketches similar to those in Figure 6, but the major source of information for predictors will be a complete data report archived in the project repository. This data report includes a complete description of the input data materials used, construction procedure, basic properties of materials and apparatus, sensor locations, calibrations, data processing techniques, and input ground motions. The report also summarizes some key results of the experiment and provides detailed instructions that allow predictors to find all of the output data, photographs, and video documentation of the experiments. Such documentation is already available for several experiments in NEEShub, the data archive for the Network for Earthquake Engineering Simulation. An example of such an archive is Allmond and Kutter (2014). The textual description of the experiment (Allmond and Kutter 2013), is also available on NEEShub or from the Center for Geotechnical Modeling at UC Davis.

6.2.2 Input motions

The benchmark will require predictions to be made on a sequence of ground motions. This will test the ability of the models (using one set of input parameters) to

• account for seismic history,
6.2.3 Validation metrics

The appropriate degree of complexity of validation metrics is difficult to describe. Instead of validating against the amplitude of response, the quality of capturing the important mechanisms of response is perhaps as important as matching the amplitude of response.

Figure 7 is an example from VELACS (Dobry & Taboada 1994) of an excellent way to demonstrate the quality of agreement between different simulations, in this case experimental simulations. As can be seen, for these particular experiments the deformed shape is reasonably similar, and this should receive a high validation grade.

To reduce the apparent agreement displayed in Figure 7 to a numerical score or grade may require subjective grading by a team of experts.

Possible comparisons to be made to perform the validation might include, for example:

1) Deformed shape contours at different stages of the test.
2) Surface settlement profile and lateral inclinometer profiles at the middle and at the end walls -- before shaking, after 1 pulse of shaking, at the end of shaking, after consolidation.
3) Time history of settlement, including magnitude and frequency content at centerline and near end walls of laminar box type container.
4) Time history of lateral displacement at centerline and end walls.
5) Time history of acceleration at base, mid-depth, and surface.
6) Time history of pore pressure at base, mid-depth, and surface.
7) Response spectrum of base motion, mid depth and surface.
8) Evaluation of how well the boundary conditions at the soil-container interface are modeled.

What has not yet been determined is to provide an objective score for each of the comparisons. While it may be difficult to provide an objective score, it might be necessary to resort to expert opinion and judgment. As class C simulations are done after the event, and after other simulations have been “scored”, the scoring criteria for class C simulations may be known prior to the simulation efforts. This knowledge will bias reported predictions.

6.3 List of potential benchmark experiments archived for class C simulations

- Sloping uniform ground with unidirectional shaking
  - D_r = 40, 60, 80%
  - Slope angle = 0, 0.03, 0.1 radians
- Level uniform ground, with bi-directional horizontal shaking
- Homogeneous trapezoidal embankment

6.5 List of potential system experiments for class C simulations

- Sloping ground with an impermeable layer, similar to Figure 5
- Embankment dam with zones of different soils
- Ground improvement options
- Waterfront earth retaining structures
- Soil-structure interaction in liquefiable soils

7 CLASS A AND B VALIDATIONS

We anticipate that most LEAP events will include at least one more formal class A or class B prediction contest. Class A predictions are based upon the planned experiment, not the actual experiment. Thus errors in comparisons between Class A predictions and the experiments may arise due to errors in the performance of the experiment in addition to errors in the simulation.

Class B predictions, conducted after the experiment is completed but without knowledge of the results, may be based upon as-built properties and measured input data. In cases where there are significant differences between the planned and performed experiments, much may be learned from Class B validations.

7.1 Complexity

Class A simulation exercises could attempt to match experiments of widely varying complexity. The complexity might vary from predicting the results of element tests to large scale centrifuge mod-
els of soil-structure interaction involving heterogeneous soil layers with multidirectional earthquake shaking.

### 7.2 Similarity to previous calibration exercises

Some Class A validation exercises might be quite similar to the previously reported Class C benchmarks. Class C benchmarks could include soils with relative density of 40% with a slope of 3% shaken by a sinusoidal motion while a later class A experiment involves layer of soil with relative density 60% and a slope of 3% shaken by a series of realistic earthquake motions.

### 7.3 Duplication of experiments

Although it may not be feasible for all prediction exercises, the experiments should be repeated at least once, and ideally the experiment should be repeated at a different facility or through modeling of models.

### 8 DISCUSSION AND CONCLUSIONS

Presently, there are many numerical techniques and constitutive models being used in research and practice, and validation standards are arbitrary and different for every numerical procedure. It is very difficult for owners and decision makers to have confidence in the results of numerical simulations if there is no standard for validation. The goal of LEAP is to attempt to establish a standard process for validation of numerical models associated with liquefaction. This will be hopefully be accomplished over a number of LEAP exercises occurring over the next several years.

This paper has outlined a concept for systematic comparison and validation of numerical simulation tools. This plan is in a state of evolution and we expect to modify the plan after each LEAP. The LEAP session at this GEDMAR 2014 conference is one of the first LEAPs.

One of the goals of this paper is to stimulate the discussion of how validation should be done. Following the GEDMAR 2014 conference, we expect to hold an additional workshop(s) to expand upon and add detail to the plans for systematic and accurate verification and validation.

### ACKNOWLEDGEMENT

The planning phase of the LEAP project has been funded by the US National Science Foundation NEES research program directed by Dr. Richard Fragaszy, through the grants CMMI-1344705, CMMI-1344630, and CMMI-1344619 to the George Washington University, the University of California Davis, and Rensselaer Polytechnic Institute, respectively. This support is gratefully acknowledged.

### REFERENCES


