

An Existing Testing Center: CSEP

D. Schorlemmer (SCEC/USC)

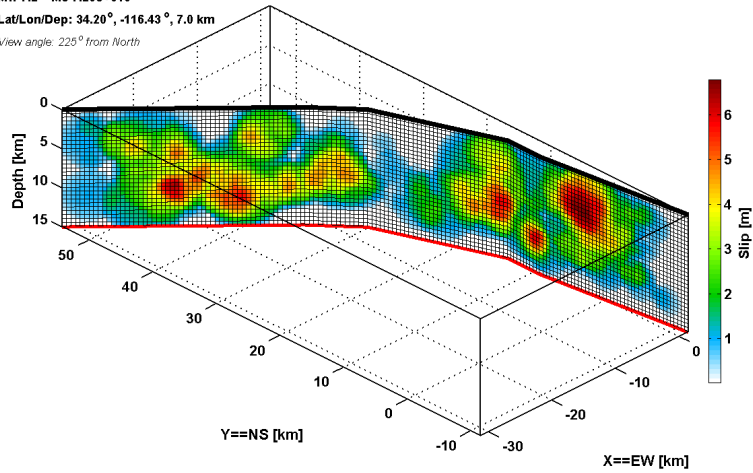
Problems in Assessing Predictions

- Scientific publications provide insufficient information for independent evaluation
- Active researchers are constantly tweaking their procedures, which become moving targets
- Standards are lacking for testing predictions against reference forecasts
- Data to evaluate prediction experiments are often improperly specified
- Infrastructure for conducting and evaluating long-term prediction experiments does not exist

Same Problem

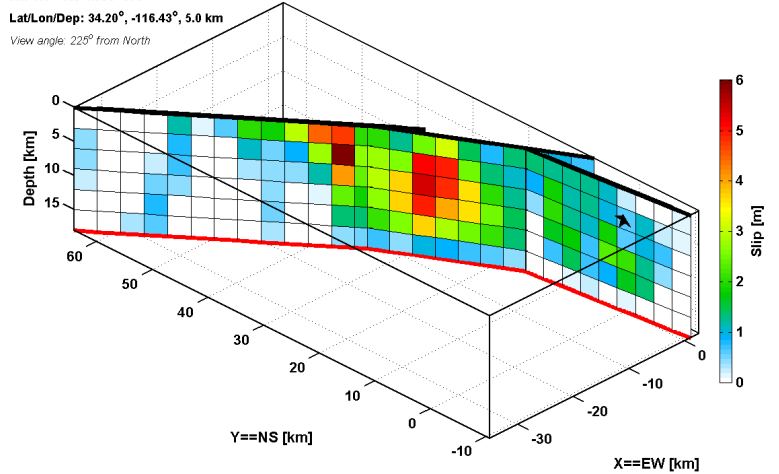
Landers (Calif.)

s1992LANDERzeng
Mw 7.2 Mo 7.20e+019
Lat/Lon/Dep: 34.20°, -116.43°, 7.0 km
View angle: 225° from North



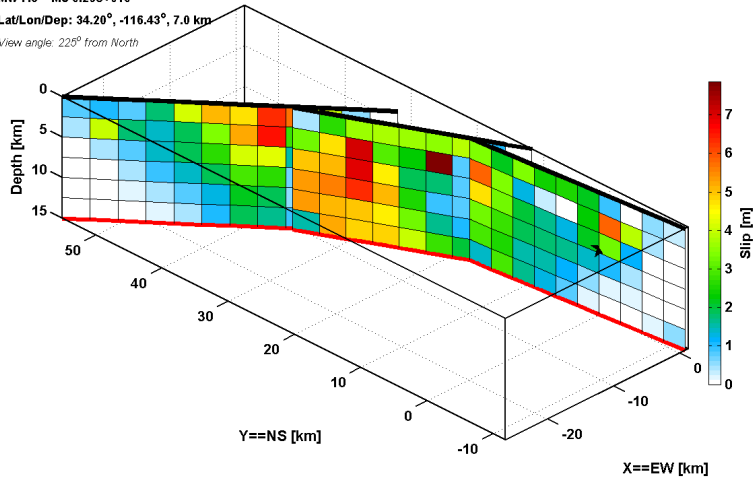
Landers (Calif.)

s1992LANDERcohe
Mw 7.1 Mo 4.66e+019
Lat/Lon/Dep: 34.20°, -116.43°, 5.0 km
View angle: 225° from North



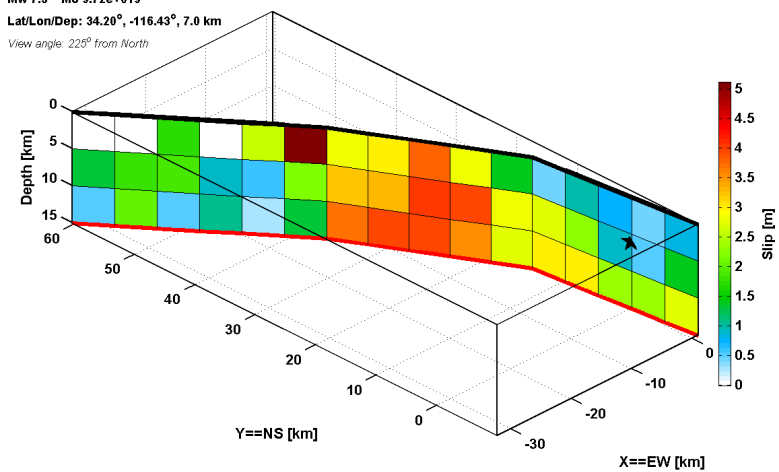
Landers (Calif.)

s1992LANDERwald
Mw 7.3 Mo 9.26e+019
Lat/Lon/Dep: 34.20°, -116.43°, 7.0 km
View angle: 225° from North



Landers (Calif.)

s1992LANDERcott
Mw 7.3 Mo 9.72e+019
Lat/Lon/Dep: 34.20°, -116.43°, 7.0 km
View angle: 225° from North



Strategy Change

Possible “problem solvers” (precursors):

- Seismic quiescence
- Foreshocks
- Radon emissions
- EM signals
- ...

None of the considered precursors can be used to predict earthquakes

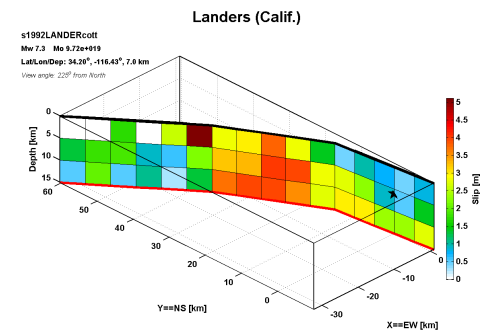
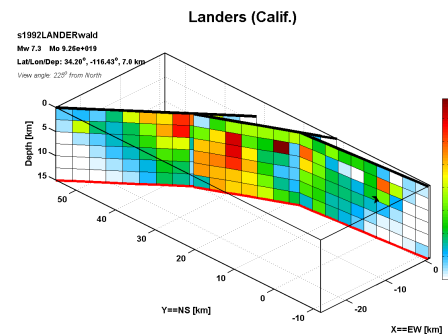
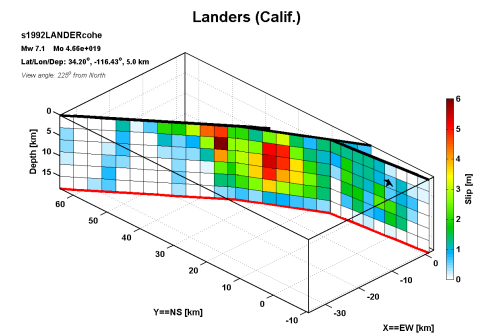
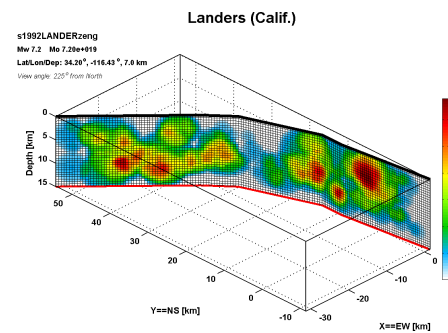
Let us first solve the 0-order problem!

Strategy Change

Develop generic forecast models, e.g., smoothed seismicity

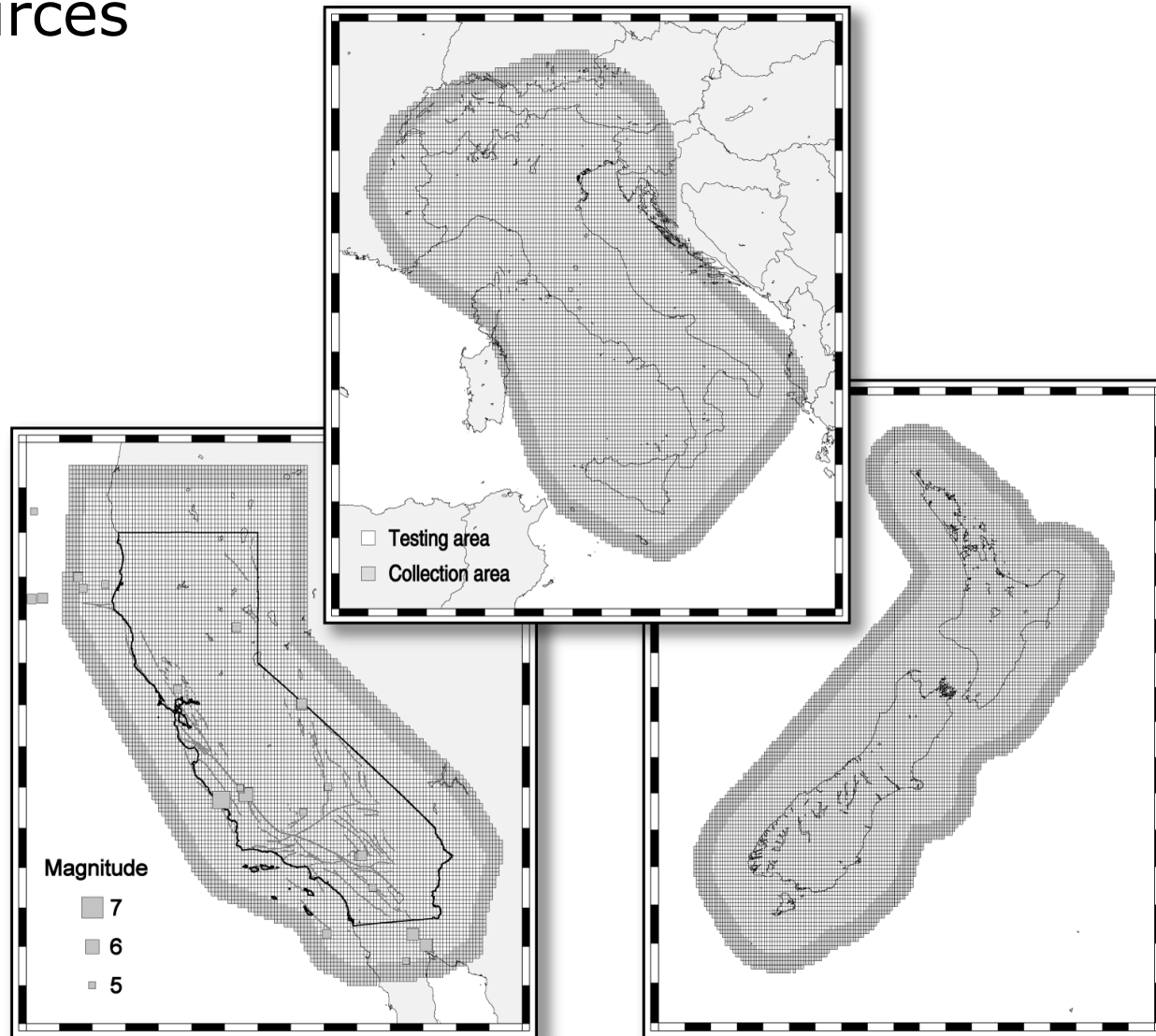
Test such models against future observations

Increase number of parameters and test for probability gain



Four Essential CSEP Components

- Regional natural laboratories with adequate, authorized data sources



Four Essential CSEP Components

- Regional natural laboratories with adequate, authorized data sources
- Community standards for registration and evaluation of scientific prediction experiments

Earthquake Likelihood Model Testing
 D. Schoffeleers^{1,2}, M. C. Gerstenberger^{1,3}, S. Wiemer⁴, D. D. Jackson⁵, and D. A. Rhoades⁶

INTRODUCTION

The Regional California Earthquake Model (RCM) project aims to produce and evaluate realistic models of earthquake hazard and seismic risk. The RCM project is a multi-disciplinary effort involving geologists, seismologists, and engineers. The RCM project is a multi-disciplinary effort involving geologists, seismologists, and engineers. The RCM project is a multi-disciplinary effort involving geologists, seismologists, and engineers.

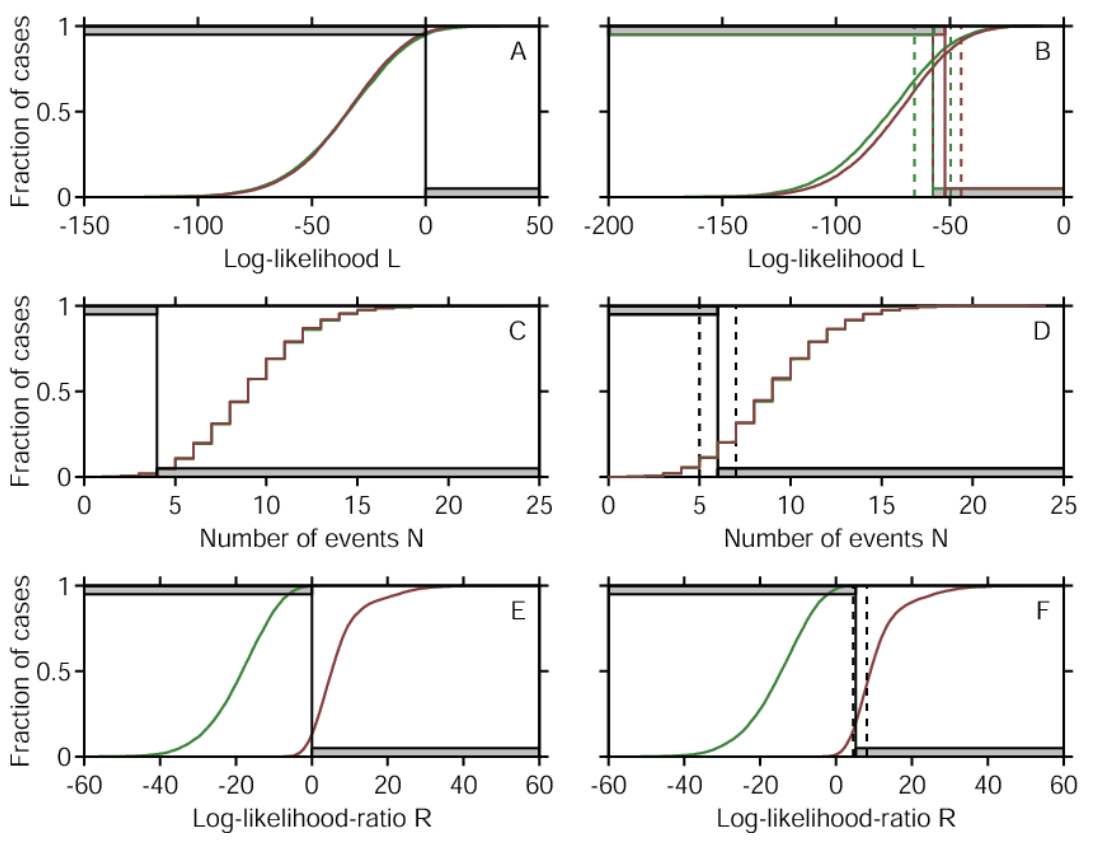
1. Southern California Earthquake Center, 2. Department of Earth and Planetary Science, 3. Department of Geological Engineering and Science, 4. Department of Earthquake Engineering and Structural Dynamics, 5. Department of Earthquake Engineering and Structural Dynamics, 6. Department of Earthquake Engineering and Structural Dynamics

RELM Testing Center
 D. Schoffeleers^{1,2}, M. C. Gerstenberger^{1,3}, S. Wiemer⁴, D. D. Jackson⁵, and D. A. Rhoades⁶

INTRODUCTION

The RELM Testing Center is a community-based effort to provide a standardized environment for the testing and evaluation of earthquake likelihood models. The RELM Testing Center is a community-based effort to provide a standardized environment for the testing and evaluation of earthquake likelihood models.

1. Southern California Earthquake Center, 2. Department of Earth and Planetary Science, 3. Department of Geological Engineering and Science, 4. Department of Earthquake Engineering and Structural Dynamics, 5. Department of Earthquake Engineering and Structural Dynamics, 6. Department of Earthquake Engineering and Structural Dynamics



Four Essential CSEP Components

- Regional natural laboratories with adequate, authorized data sources
- Community standards for registration and evaluation of scientific prediction experiments
- Testing facilities with validated procedures for conducting and evaluating prospective prediction experiments

Four Essential CSEP Components

- Regional natural laboratories with adequate, authorized data sources
- Community standards for registration and evaluation of scientific prediction experiments
- Testing facilities with validated procedures for conducting and evaluating prospective prediction experiments
- Communication protocols for conveying scientific results
 - the scientific community, including professional societies
 - government agencies responsible for risk management
 - the general public and other end-users



Four Goals of Testing Centers

1. Transparency
2. Controlled environment
3. Comparability
4. Reproducibility

RELm Testing Center
D. Scharlemmer
 ETH Zurich, Switzerland
M. C. Gerstenberger
 U.S. Geological Survey, Pasadena

INTRODUCTION

We describe the experimental testing and the use of data derived from test configurations developed at the Regional Earthquake Likelihood Models (RELm) centers. These procedures are designed to facilitate an efficient and reproducible way to compare test results for each model as well as to facilitate the interpretation of these results. We present the parameters, e.g., the testing area and grid layout, and the specific testing configurations provided to the models and used for testing and data collection. A comparison of these two parameters defines a class of models. Within RELm we distinguish between the test models, test area models, and the models, which use their own test area for the respective parameter. While the test models provide their location as coordinates, the test area models need to provide the testing area together with their own test area models, which are fully tested parameters. They will provide their own results for comparison or use.

The defined RELm goal of testing a multitude of test area models in a geographic (forward-looking) area is to use knowledge gained in earthquake risk to the first time that a group of modelers agree to release their models to a common, centrally agreed test area known as seismicity. We believe that this is an important milestone for forecasting and prediction-related research in seismicity because it offers an opportunity to compare test configurations and results. The primary goal of the current testing is to test the ability of models to release test results, which will be used to the regional development of the next earthquake event. However, this will lead to a small step in possibly gain input to better forecast models.

Procedural testing of models to determine if we have a test facility complete test (Singer and Jeffrey 1995, Brown and Bowers 1999, Hong et al. 2001). However, the testing of individual models for their parameters has three major advantages: (1) the availability of data is not guaranteed because models can be tested; (2) Testing procedures can be manually released or centrally accepted; (3) A comparison of results from different centers are not possible. A standard testing procedure can be fully tested against a model that has been tested in the past. The fully tested procedure has the advantage of having the same recent developments, but it has the disadvantage of being tested in the same region because they are different test results, regardless of testing choices. Regions are increasingly related research, but in our system have already happened by these factors, and it is for these reasons that we are testing the framework of the RELm software in a common, centrally agreed test area and the testing center.

Kochendorfer et al. (2005) also describe the ability of the RELm group (2004/2005) testing procedure for global prediction earthquake forecast. Although the definition of test of a fully implemented test is a very general description and its implementation depends on the specific test results to achieve with the test. A number of the parameters in the testing area to the model. They include the choice of models, testing area and grid, data collection, etc. The results "take of the game" test, the test that each model has to do the test in the test area model, is provided to allow for test procedures, reproducibility, and comparable testing. The test parameters need to be carefully chosen to compare or match with the development of the model, which is generally and to maintain the self-referential content in the results. Location-specific information, such as testing quality, testing area, etc., must be used.

In this paper we present the historical and the testing configurations of the RELm group to test California test results and use these data to test and compare the understanding of the underlying physics of earthquake generation and hazard calculation in general. The underlying principles of testing and the testing center facilities are general and could be applied to other regions.

The test parameters are following the test of multiple models in the definition of model class so that they can be tested in the same forward region. It would be able to test the test from the model, test or model, and then information only require different approaches in the testing center that allow of different values for regions and are likely to be influenced by different fundamental physical processes. Comparison of results from different centers are not possible. A standard testing procedure can be fully tested against a model that has been tested in the past. The fully tested procedure has the advantage of having the same recent developments, but it has the disadvantage of being tested in the same region because they are different test results, regardless of testing choices. Regions are increasingly related research, but in our system have already happened by these factors, and it is for these reasons that we are testing the framework of the RELm software in a common, centrally agreed test area and the testing center.

© Seismological Research Letters, Volume 76, Number 1, January/February 2007

Transparency

- All data cataloged
- All codes version-controlled
- Open-source codes (fully downloadable)
- Every change is documented

Controlled Environment

- Model codes cannot be changed (no moving targets)
- All data streams are authorized and fully specified
- All data streams come from independent providers and are not “interpreted” by the Testing Center (e.g., which fault ruptured during an earthquake)
- All models get identical input data
- All models are evaluated using the same testing codes

Comparability

- Models are tested against observation (consistency tests)
- Models are tested against each other (comparative testing)
- Avoids testing against overly simple null hypotheses

Reproducibility

- Any previous result can be recomputed:
 - All data used is cataloged
 - All codes are version-controlled
 - All system configurations are stored
 - All library packages (e.g., Python) are in a dedicated repository
- Tests can be rerun with alternative tests



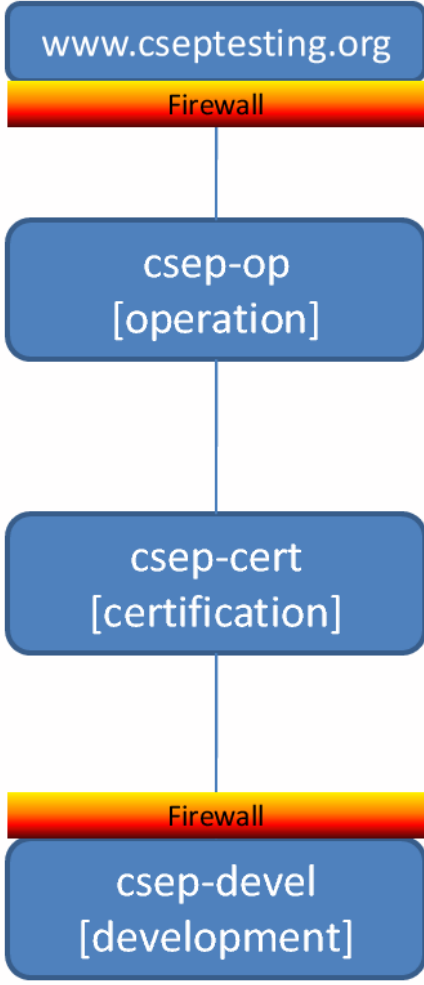
Testing Center

Hardware
 2 Dual-Core AMD Opteron 2220
 8 GB RAM
 3.4 TB RAID 5 disks

OS/Kernel
 Fedora Core 7 – 64 bit
 2.6.22.1-41.fc7
 gcc 4..1.2

Related Software
 apache ant-1.7 R-2.5.0
 elementtree-1.2.6-20050316
 matplotlib-0.90.1
 mpich2-1.0.3
 CruiseControl

Identical hardware & software



Webserver
 Restricted publication of results
 Restricted dispatcher download
 Test center setup information

Operational Machine
 Automated Nightly builds from dispatcher
 Publication of results to web server
 Back up results
 Quarterly Implementation of new models

Integration/Cert Machine
 Automated Nightly builds from dispatcher
 Emulates actual operation machine
 Fulfill software dependencies
 Implementation of new models

Development Machine
 Import new models
 Develop dispatcher
 Fulfill software dependencies

Summary

- Source inversion validation efforts may learn from CSEP
- CSEP partly has similar needs (e.g., input data)
- CSEP provides an excellent infrastructure for experiments
- Maybe this can be a starting point for an interdisciplinary effort?