

Data-driven validation of neuroscience models

Richard (Rick) Gerkin, PhD
Arizona State University, USA

A2:

What I cannot create,
I do not understand.

I know how to solve every
problem that has been solved

Why const & sort etc

TO LEARN:

Bethe Ansatz Probs.

Kondo

2-D Hall

local Temp

Non linear classical Hydro

$$\textcircled{A} f = u(r, \alpha)$$

$$g = \psi(r, z) u(r, z)$$

$$\textcircled{B} f = 1/r \cdot \alpha |u| u$$

A2:

What I cannot create,
I do not understand

I know how to solve every
problem that has been solved

Why const & sort PC

TO LEARN:

Bethe Ansatz Probs.

Kondo

2-D Hall

local Temp

Non linear classical Hydro

$$\textcircled{A} f = u(r, \alpha)$$

$$g = \psi(r, z) u(r, z)$$

$$\textcircled{B} f = 1/r \cdot \alpha |u| u$$

MODEL VALIDATION

MODEL VALIDATION

- A scientific model has high *validity* if:

MODEL VALIDATION

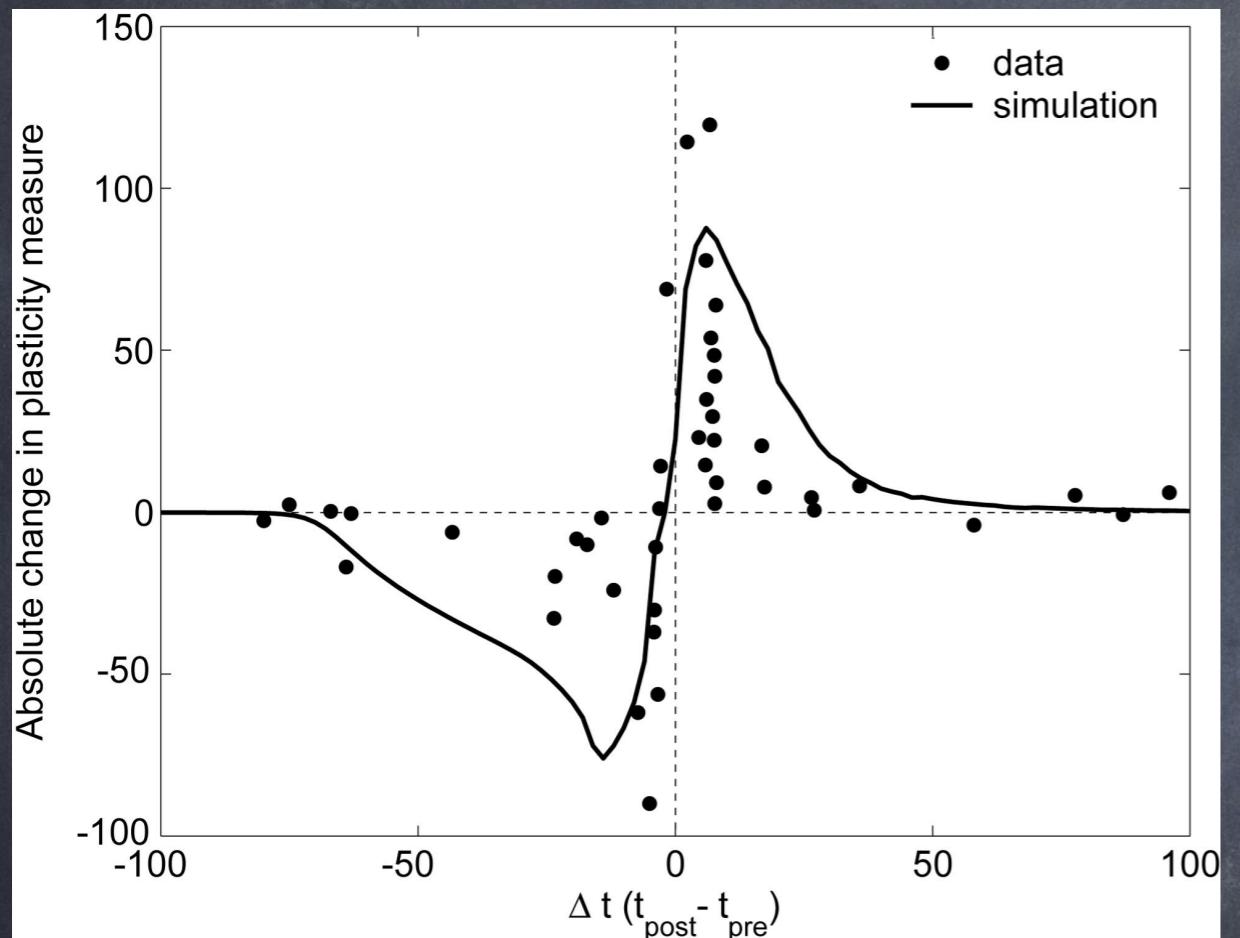
- A scientific model has high *validity* if:
 - It is *consistent with* a wide range of previously gathered data.

MODEL VALIDATION

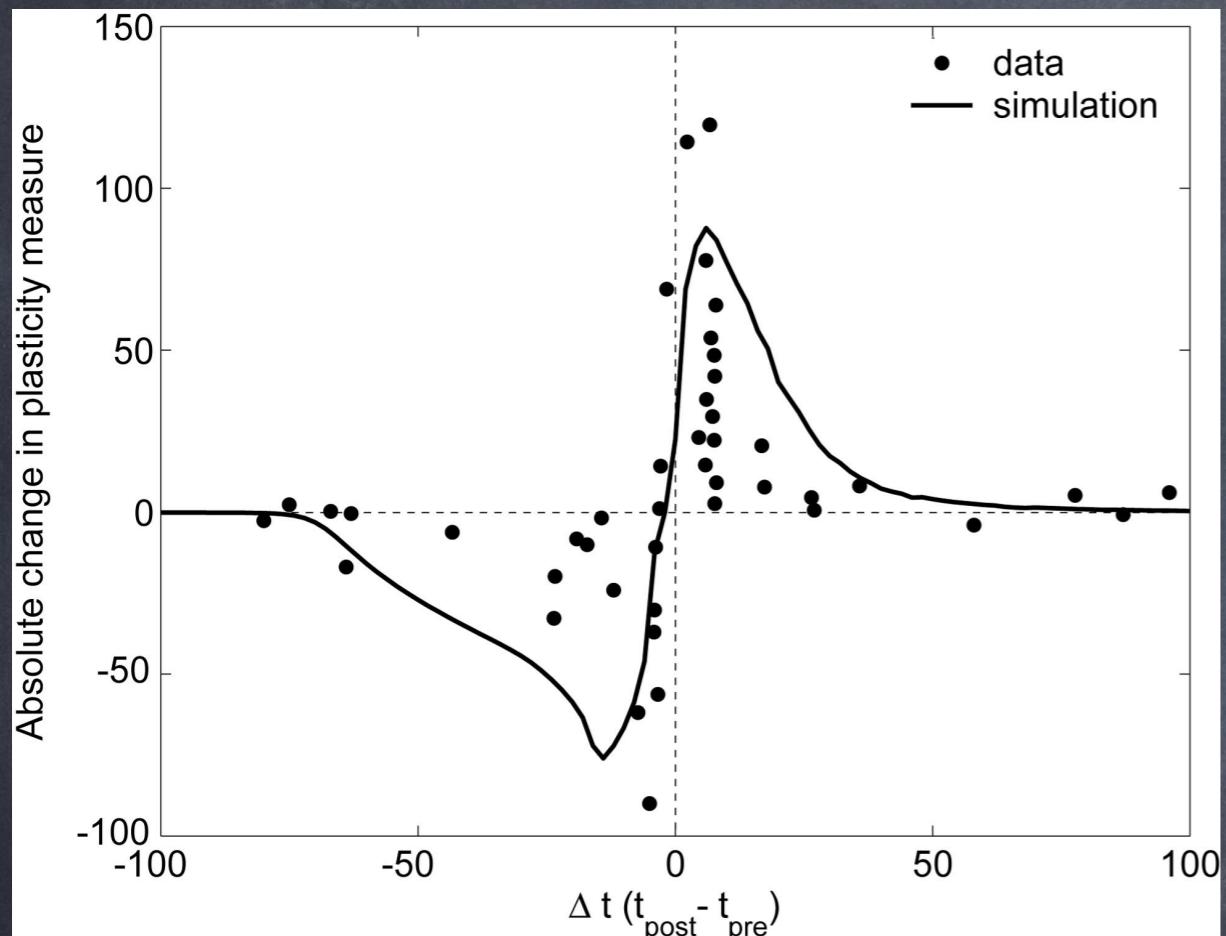
- A scientific model has high *validity* if:
 - It is *consistent with* a wide range of previously gathered data.
 - It can *predict* the results of many future experiments.

TESTING VALIDITY

TESTING VALIDITY

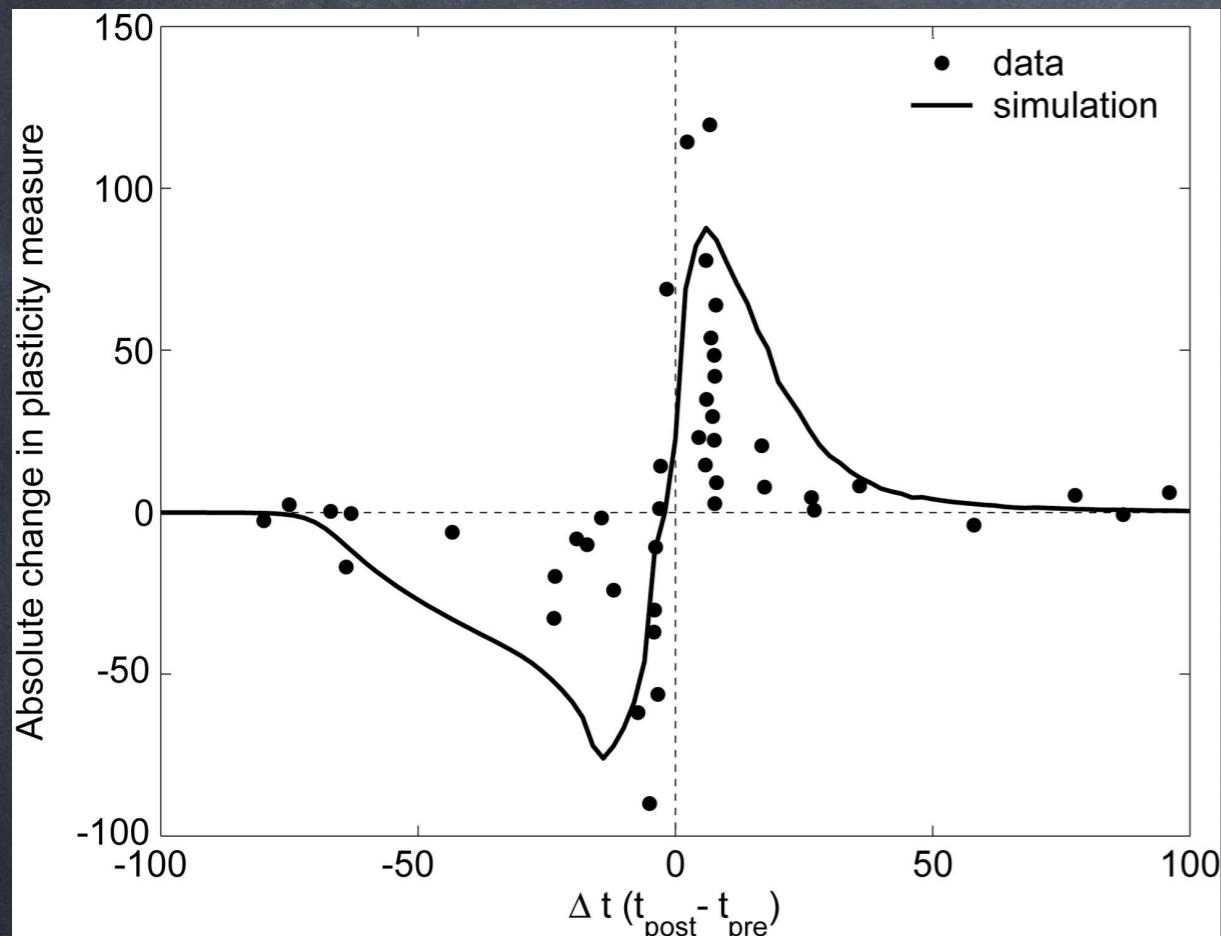


TESTING VALIDITY



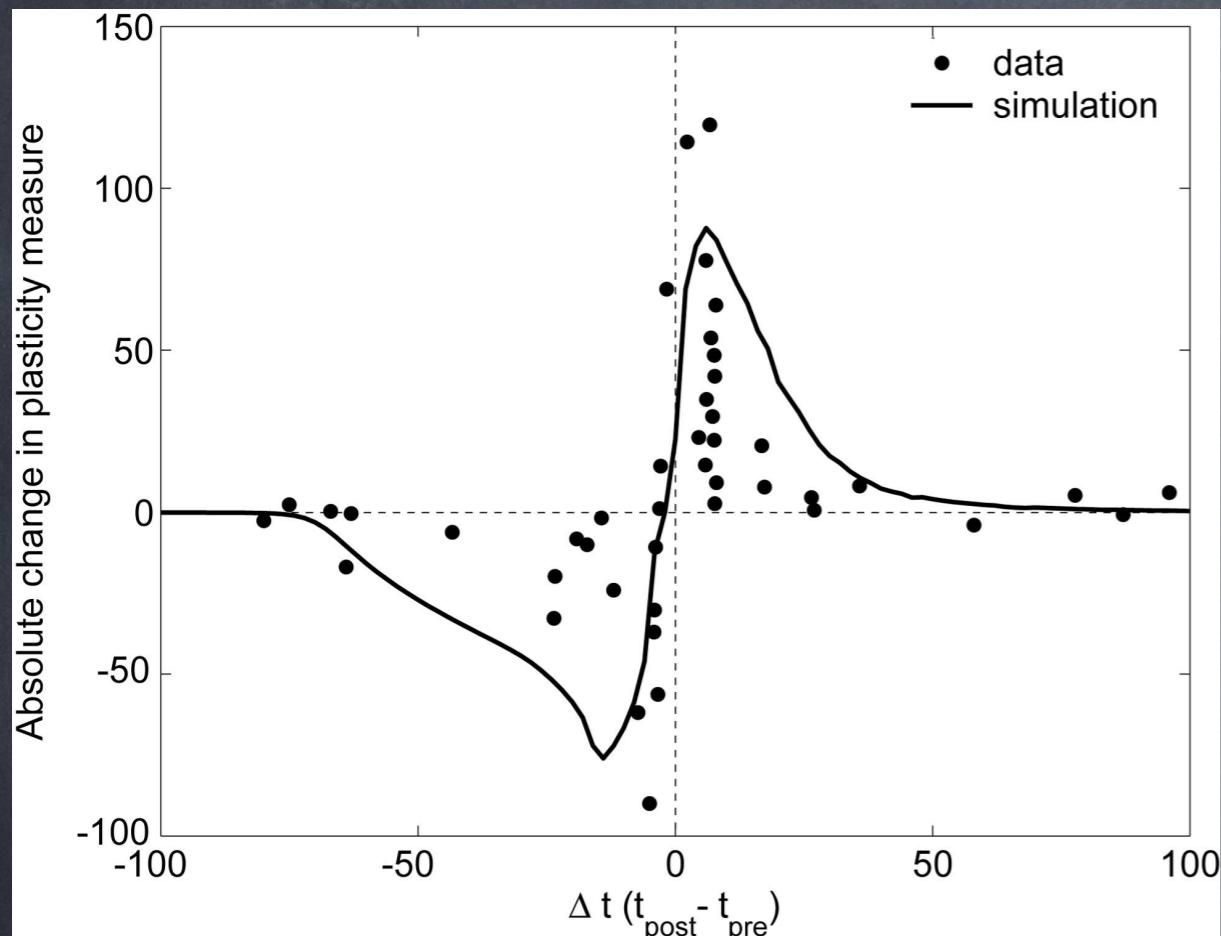
- Most claims about model validity are informal.

TESTING VALIDITY



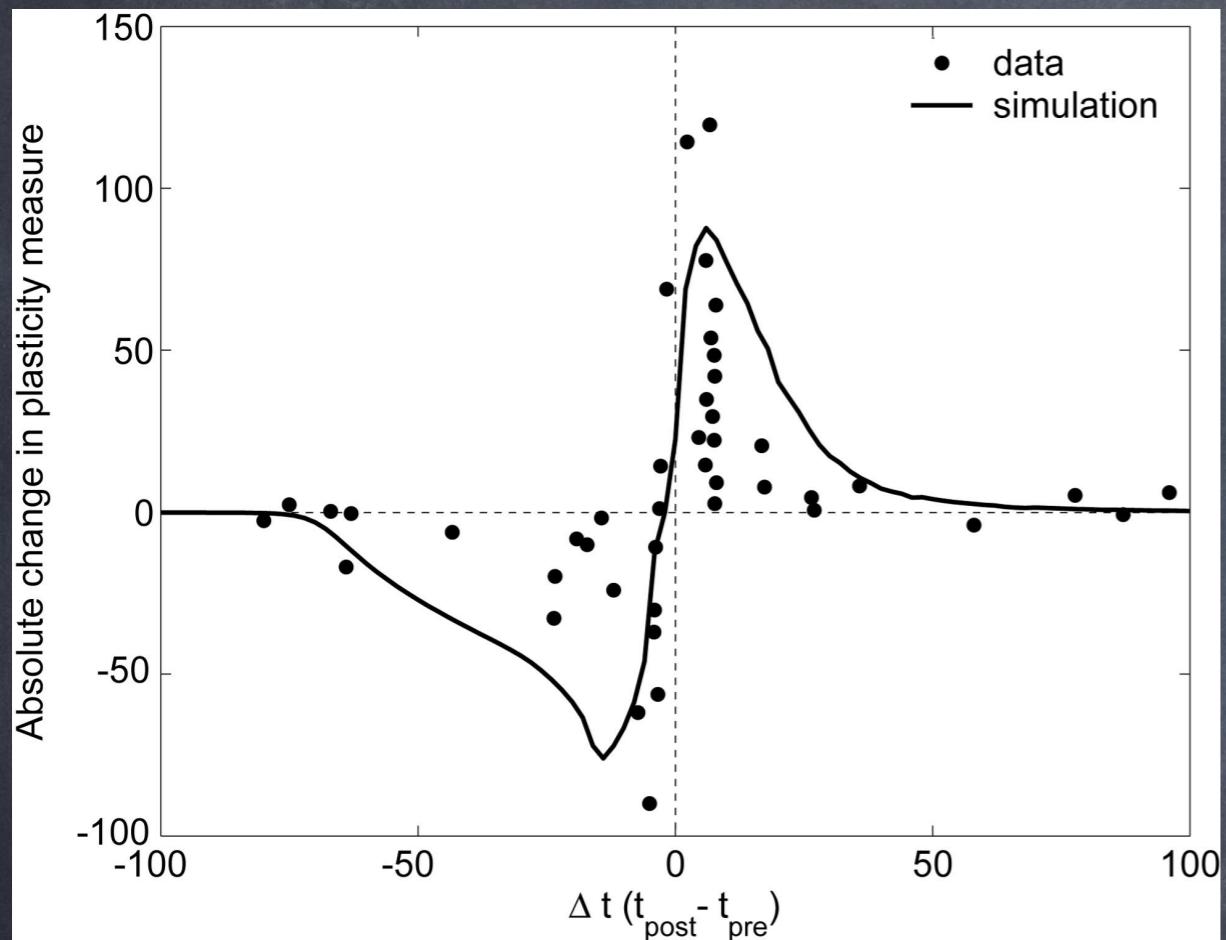
- Most claims about model validity are informal.
- These claims cannot be formally evaluated.

TESTING VALIDITY



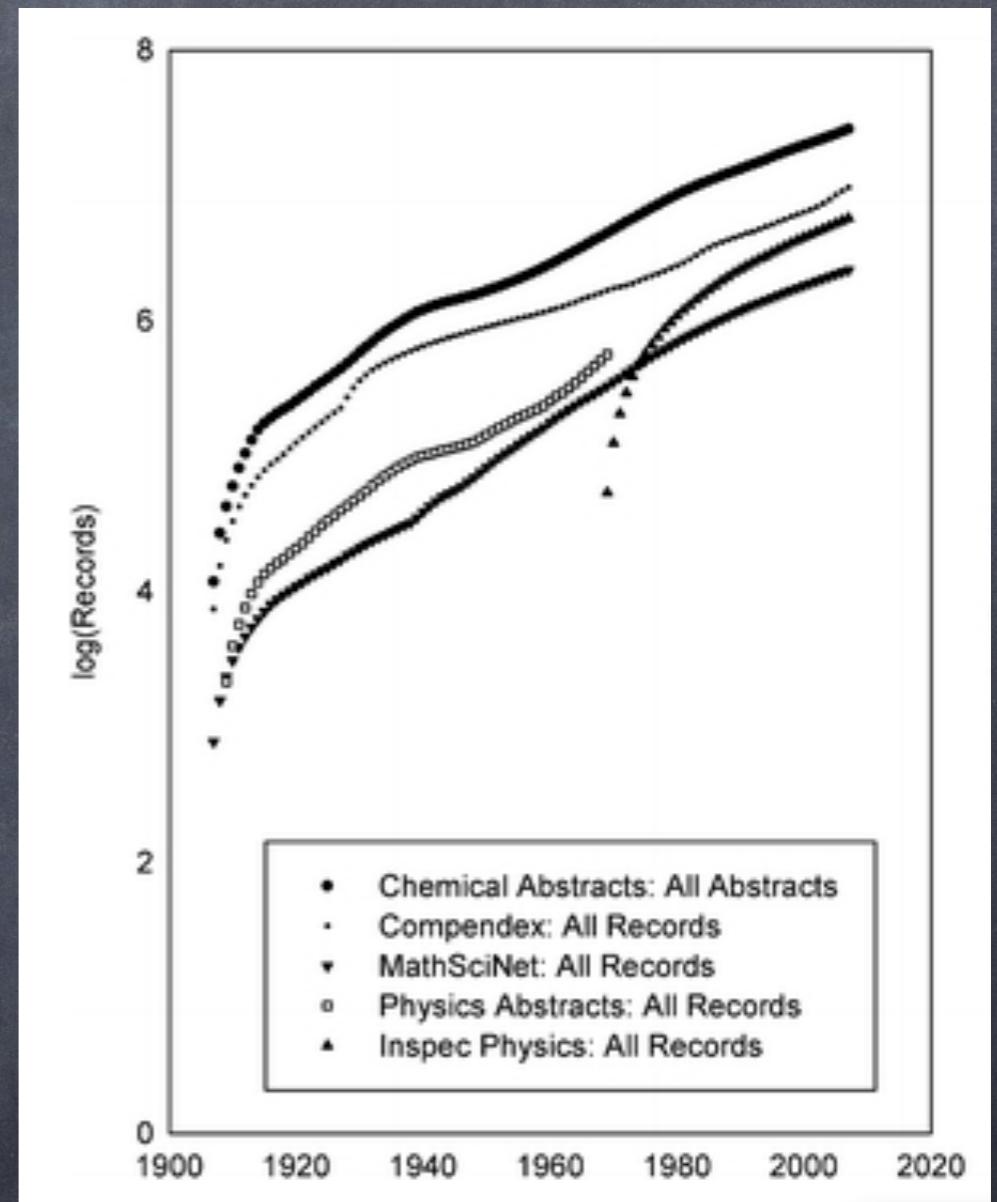
- Most claims about model validity are informal.
- These claims cannot be formally evaluated.
- Claims are difficult to locate.

TESTING VALIDITY

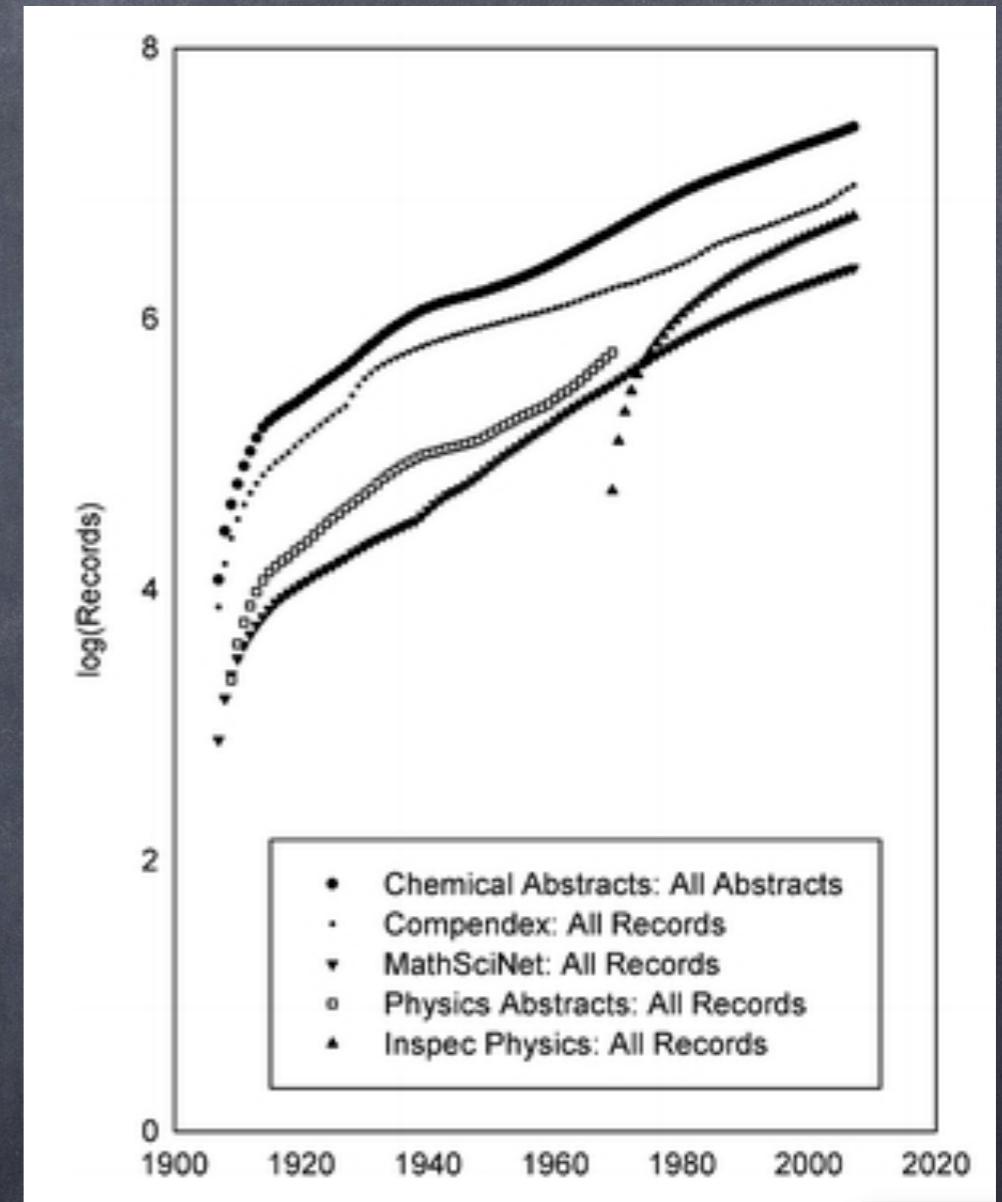


- Most claims about model validity are informal.
- These claims cannot be formally evaluated.
- Claims are difficult to locate.
- Falsifiable claims spanning many data sets are rare.

TESTING VALIDITY



TESTING VALIDITY



Huge amounts of data!
Even *informal validation* is becoming difficult.

TESTING VALIDITY

TESTING VALIDITY

- Not testing models rigorously leaves a field with an **unclear** view of:

TESTING VALIDITY

- Not testing models rigorously leaves a field with an **unclear** view of:
 - What has already been explained.

TESTING VALIDITY

- Not testing models rigorously leaves a field with an **unclear** view of:
 - What has already been explained.
 - What needs explaining.

TESTING VALIDITY

- Not testing models rigorously leaves a field with an **unclear** view of:
 - What has already been explained.
 - What needs explaining.
- Needed: a **framework** for validating scientific models, based on established techniques for **formally validating** software.

THE UNIT TEST

THE UNIT TEST

- “unit testing is a software testing method by which individual units of source code ... are tested to determine if they are **fit for use.**”

THE UNIT TEST

- “unit testing is a software testing method by which individual units of source code ... are tested to determine if they are **fit for use.**”
- “one can view a unit as the **smallest testable part** of an application.”

THE UNIT TEST

- “unit testing is a software testing method by which individual units of source code ... are tested to determine if they are **fit for use.**”
- “one can view a unit as the **smallest testable part** of an application.”
- “A unit test provides a **strict, written contract** that the piece of code must satisfy.”

SCIENTIFIC UNIT TESTS

SCIENTIFIC UNIT TESTS

- A **scientific model is just a function** that takes *observations* as inputs and produces *predictions* as output.

SCIENTIFIC UNIT TESTS

- A **scientific model is just a function** that takes *observations* as inputs and produces *predictions* as output.
- The observations include metadata about how they were obtained, to direct the model.

SCIENTIFIC UNIT TESTS

- A **scientific model is just a function** that takes *observations* as inputs and produces *predictions* as output.
- The observations include metadata about how they were obtained, to direct the model.
- Example: A *complete* model of a given neuron type should produce spikes that, given a specified description of the stimulus, **match physiologically observed features** such as spike shapes, rates, interval distributions, etc.

SCIENTIFIC UNIT TESTS

- A **scientific model is just a function** that takes *observations* as inputs and produces *predictions* as output.
 - The observations include metadata about how they were obtained, to direct the model.
- Example: A *complete* model of a given neuron type should produce spikes that, given a specified description of the stimulus, **match physiologically observed features** such as spike shapes, rates, interval distributions, etc.
 - Each such feature can be encoded in a unit test.

SCIUNIT

SCIUNIT

- Showing a match to empirical data in a journal article figure is nice, but not sufficient.

SCIUNIT

- Showing a match to empirical data in a journal article figure is nice, but not sufficient.
- What if we built a **collaborative collection** of *empirically-informed validation tests* and **characterized models** by the collection of **tests that they pass?**

SCIUNIT

- Showing a match to empirical data in a journal article figure is nice, but not sufficient.
- What if we built a **collaborative collection** of *empirically-informed validation tests* and **characterized models** by the collection of **tests that they pass?**
- <http://github.com/scidash/sciunit>



python™

VISUALIZING VALIDITY

VISUALIZING VALIDITY

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. A10, PAGES 22,375–22,388, OCTOBER 1, 1999

A synthesis of solar cycle prediction techniques

David H. Hathaway, Robert M. Wilson, and Edwin J. Reichmann

NASA Marshall Space Flight Center, Huntsville, Alabama

Table 3. Precursor Prediction Method Errors (Prediction - Observed) for Cycles 19-22

Prediction Method	19	20	21	22	RMS
Ohl's method	-55.4	19.1	21.8	4.4	31.3
Feynman's method	-42.8	9.6	26.9	3.6	25.8
Thompson's method	-17.8	8.7	-26.5	-13.6	17.9

VISUALIZING VALIDITY

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. A10, PAGES 22,375–22,388, OCTOBER 1, 1999

A synthesis of solar cycle prediction techniques

David H. Hathaway, Robert M. Wilson, and Edwin J. Reichmann

NASA Marshall Space Flight Center, Huntsville, Alabama

Table 3. Precursor Prediction Method Errors (Prediction - Observed) for Cycles 19-22

Prediction Method	19	20	21	22	RMS
Ohl's method	-55.4	19.1	21.8	4.4	31.3
Feynman's method	-42.8	9.6	26.9	3.6	25.8
Thompson's method	-17.8	8.7	-26.5	-13.6	17.9

Models

VISUALIZING VALIDITY

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. A10, PAGES 22,375–22,388, OCTOBER 1, 1999

A synthesis of solar cycle prediction techniques

David H. Hathaway, Robert M. Wilson, and Edwin J. Reichmann
NASA Marshall Space Flight Center, Huntsville, Alabama

Tests

Table 3. Precursor Prediction Method Errors (Prediction - Observed) for Cycles 19-22

Prediction Method	19	20	21	22	RMS
Ohl's method	-55.4	19.1	21.8	4.4	31.3
Feynman's method	-42.8	9.6	26.9	3.6	25.8
Thompson's method	-17.8	8.7	-26.5	-13.6	17.9

Models

VISUALIZING VALIDITY

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. A10, PAGES 22,375–22,388, OCTOBER 1, 1999

A synthesis of solar cycle prediction techniques

David H. Hathaway, Robert M. Wilson, and Edwin J. Reichmann
NASA Marshall Space Flight Center, Huntsville, Alabama

Tests

Table 3. Precursor Prediction Method Errors (Prediction - Observed) for Cycles 19-22

Prediction Method	19	20	21	22	RMS
Ohl's method	-55.4	19.1	21.8	4.4	31.3
Feynman's method	-42.8	9.6	26.9	3.6	25.8
Thompson's method	-17.8	8.7	-26.5	-13.6	17.9

Models

Goodness-of-Fit

VISUALIZING VALIDITY

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. A10, PAGES 22,375–22,388, OCTOBER 1, 1999

A synthesis of solar cycle prediction techniques

David H. Hathaway, Robert M. Wilson, and Edwin J. Reichmann
NASA Marshall Space Flight Center, Huntsville, Alabama

Tests

Table 3. Precursor Prediction Method Errors (Prediction - Observed) for Cycles 19-22

Prediction Method	19	20	21	22	RMS
Ohl's method	-55.4	19.1	21.8	4.4	31.3
Feynman's method	-42.8	9.6	26.9	3.6	25.8
Thompson's method	-17.8	8.7	-26.5	-13.6	17.9

Test Suite

Models

Goodness-of-Fit

SCIUNIT: CHALLENGES

SCIUNIT: CHALLENGES

- How to *interface* with a wide range of model scales, languages, and goals?

SCIUNIT: CHALLENGES

- How to *interface* with a wide range of model scales, languages, and goals?
- How to *minimize development time* for writing model validation tests?

SCIUNIT: CHALLENGES

- How to *interface* with a wide range of model scales, languages, and goals?
- How to *minimize development time* for writing model validation tests?
- How to adjudicate whether a test *score* really captures the functionality it claims to test?

SCIUNIT: CHALLENGES

- How to *interface* with a wide range of model scales, languages, and goals?
- How to *minimize development time* for writing model validation tests?
- How to adjudicate whether a test *score* really captures the functionality it claims to test?
- We aim to solve these with:

SCIUNIT: CHALLENGES

- How to ***interface*** with a wide range of model scales, languages, and goals?
- How to ***minimize development time*** for writing model validation tests?
- How to adjudicate whether a test ***score*** really captures the functionality it claims to test?
- We aim to solve these with:
 - ***unit testing philosophy***

SCIUNIT: CHALLENGES

- How to *interface* with a wide range of model scales, languages, and goals?
- How to *minimize development time* for writing model validation tests?
- How to adjudicate whether a test *score* really captures the functionality it claims to test?
- We aim to solve these with:
 - *unit testing philosophy*
 - *domain standards and tools*

SCIUNIT: CHALLENGES

- How to *interface* with a wide range of model scales, languages, and goals?
- How to *minimize development time* for writing model validation tests?
- How to adjudicate whether a test *score* really captures the functionality it claims to test?
- We aim to solve these with:
 - *unit testing philosophy*
 - *domain standards and tools*
 - *collaborative development*

SCIUNIT: CHALLENGES

- How to ***interface*** with a wide range of model types, languages, and goals?
- How to ***minimize development time*** for writing model validation tests?
- How to adjudicate whether a test ***score*** really captures the functionality it claims to test?
- We aim to solve these with:
 - ***unit testing philosophy***
 - ***domain standards and tools***
 - ***collaborative development***

SCIUNIT: CAPABILITIES

SCIUNIT: CAPABILITIES

Test

I need you to have a soma, to receive somatic current injection, and to produce action potentials... can you do that?

SCIUNIT: CAPABILITIES

Test

I need you to have a soma, to receive somatic current injection, and to produce action potentials... can you do that?

Model1

Sorry, I'm a non-spiking model.
I cannot produce action potentials.

SCIUNIT: CAPABILITIES

Test

I need you to have a soma, to receive somatic current injection, and to produce action potentials... can you do that?

Model1

Sorry, I'm a non-spiking model.
I cannot produce action potentials.

Model2

I can do all those things.

SCIUNIT: CAPABILITIES

Test

I need you to have a soma, to receive somatic current injection, and to produce action potentials... can you do that?

Model1

Sorry, I'm a non-spiking model.
I cannot produce action potentials.

Model2

I can do all those things.

Test

Model2, bring it on!

SCIUNIT: CHALLENGES

- How to *interface* with a wide range of model types, languages, and goals?
- How to ***minimize development time*** for writing model validation tests?
- How to adjudicate whether a test *score* really captures the functionality it claims to test?
- We aim to solve these with:
 - *unit testing philosophy*
 - ***domain standards and tools***
 - *collaborative development*

NEUROELECTRO

<http://www.neuroelectro.org>

Published literature

Novel subcellular distribution pattern of A-type K⁺ channels on neuronal surface.

Unique clustering of A-type potassium channels on different cell types of the main olfactory bulb.

Kollo M, Holderith N, Antal M, Nusser Z.
Theoretical and functional studies predicted a highly non-uniform distribution of voltage-gated ion channels on the neuronal surface. This was confirmed by recent immunolocalization experiments for Na⁺, Ca²⁺, hyperpolarization activated mixed cation and K⁺ channels. These experiments also indicated that some K⁺ channels were clustered in synaptic or non-synaptic membrane specializations. Here we analysed the subcellular distribution of Kv4.2 and Kv4.3 subunits in the rat main olfactory bulb at high resolution to address whether clustering characterizes their distribution, and whether they are concentrated in synaptic or non-synaptic junctions. The cell surface distribution of the Kv4.2 and Kv4.3 subunits is highly non-uniform. Strong Kv4.2 subunit-immunopositive clusters were detected in intercellular junctions made by mitral, external tufted and granule cells (GCs). We also found Kv4.3 subunit-immunopositive clusters in periglomerular (PGC), deep short-axon and GCs. In the juxtaglomerular region some calretinin-immunopositive glial cells enwrap neighboring PGC somata in a cap-like manner. Kv4.3 subunit clusters are present in the cap membrane that directly contacts the PGC, but not the one that faces the neuropil. In membrane specializations established by members of the same cell type, K⁺ channels are enriched in both membranes, whereas specializations between different cell types contain a high density of channels asymmetrically. None of the K⁺ channel-rich junctions showed any of the ultrastructural features of known chemical synapses. Our study provides evidence for highly non-uniform subcellular distributions of A-type K⁺ channels and predicts their involvements in novel

Physiology database

Olfactory Bulb Mitral Cell

Input resistance	200 MΩ
V _{rest}	-65 mV
Spike width	1 ms
...	

CA1 Pyramidal Cell

Input resistance	400 MΩ
V _{rest}	-70 mV
Spike width	.5 ms
...	

Extracted from Literature

NEUROELECTRO

```
In [1]: from neuronunit.neuroelectro import NeuroElectroSummary
```

```
In [2]: summary = NeuroElectroSummary(neuron={'name':'Hippocampus CA1 Pyramidal Cell'},  
...: ephysprop={'name':'spike width'})
```

```
In [3]: observation = summary.get_observation(show=True)
```

Getting data values from neuroelectro.org

http://www.neuroelectro.org/api/1/nes/?e__name=spike+width&n__name=Hippocampus+CA1+Pyramidal+Cell

```
{u'e': {u'definition': u'Duration of AP, not explicitly referred to as half-width',  
        u'id': 23,  
        u'name': u'spike width',  
        u'nlex_id': None,  
        u'norm_criteria': u'Values are unchanged from those reported. Values currently lump multiple measures of spike width which do not  
explicitly denote spike half-width. Refer to individual articles for definition and calculation methodology.'},  
u'n': {u'id': 85,  
       u'name': u'Hippocampus CA1 pyramidal cell',  
       u'neuron_db_id': 258,  
       u'nlex_id': u'sao830368389'},  
u'num_articles': 6,  
u'num_nedms': 10,  
u'value_mean': 2.602083333333333,  
u'value_sd': 0.766070080381394}
```

```
In [4]: from neuronunit.tests import SpikeWidthTest
```

```
In [5]: ca1_pyramidal_spike_width_test = SpikeWidthTest(observation = observation)
```

NEUROELECTRO

```
In [1]: from neuronunit.neuroelectro import NeuroElectroSummary
```

```
In [2]: summary = NeuroElectroSummary(neuron={'name':'Hippocampus CA1 Pyramidal Cell'},  
...: ephysprop={'name':'spike width'})
```

```
In [3]: observation = summary.get_observation(show=True)
```

Getting data values from neuroelectro.org

http://www.neuroelectro.org/api/1/nes/?e__name=spike+width&n__name=Hippocampus+CA1+Pyramidal+Cell

```
{u'e': {u'definition': u'Duration of AP, not explicitly referred to as half-width',  
u'id': 23,  
u'name': u'spike width',  
u'nlex_id': None,  
u'norm_criteria': u'Values are unchanged from those reported. Values currently lump multiple measures of spike width which do not  
explicitly denote spike half-width. Refer to individual articles for definition and calculation methodology.'},  
u'n': {u'id': 85,  
u'name': u'Hippocampus CA1 pyramidal cell',  
u'neuron_db_id': 258,  
u'nlex_id': u'sao830368389'},  
u'num_articles': 6,  
u'num_nedms': 10,  
u'value_mean': 2.602083333333333,  
u'value_sd': 0.766070080381394}
```

```
In [4]: from neuronunit.tests import SpikeWidthTest
```

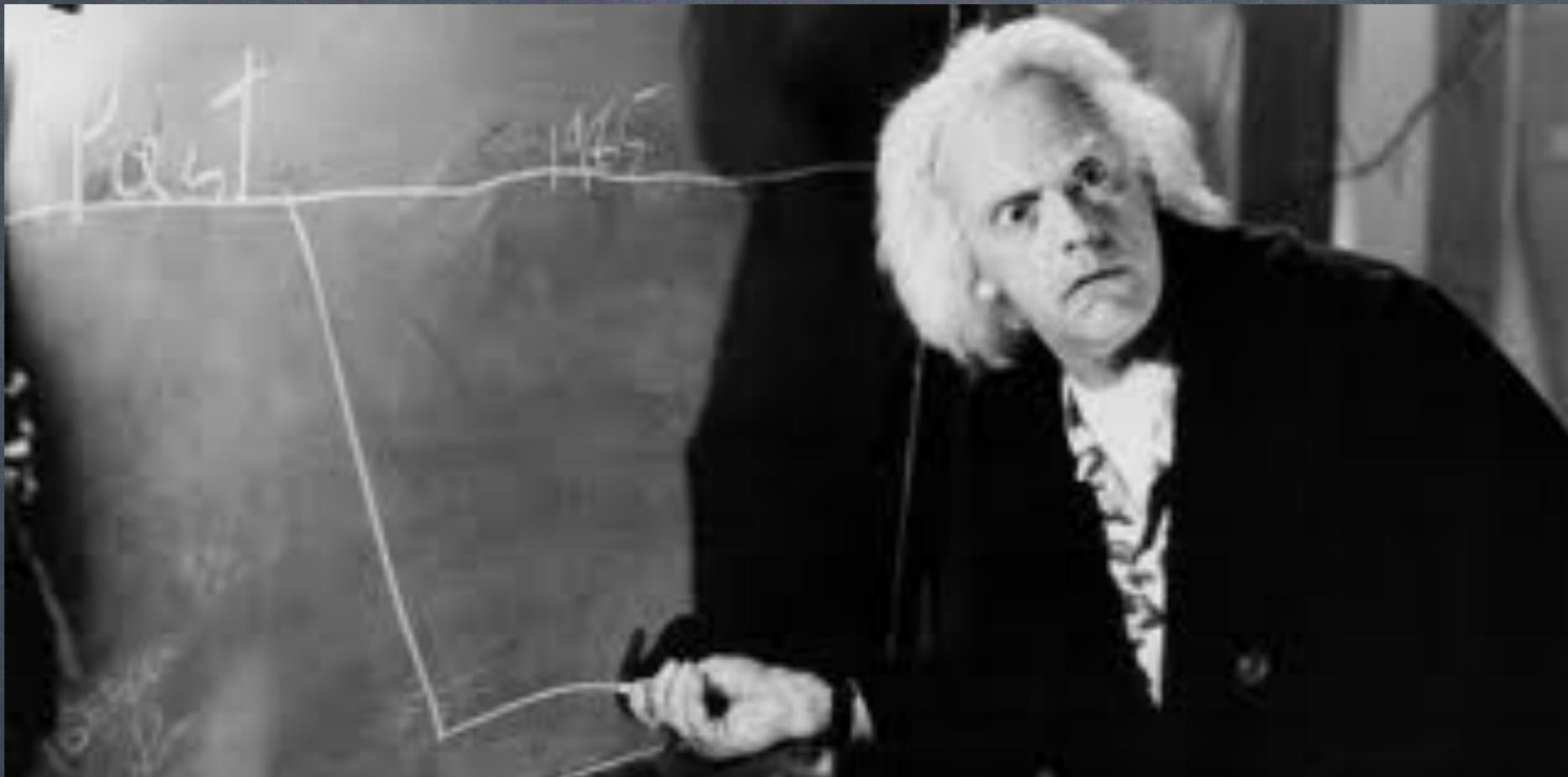
```
In [5]: ca1_pyramidal_spike_width_test = SpikeWidthTest(observation = observation)
```

- Uses the NeuroElectro API (and NeuroLex) to get collated data about a named *neuron type* for a specified *electrophysiological property* in order to *parameterize a test*.

SCIUNIT: CHALLENGES

- How to *interface* with a wide range of model types, languages, and goals?
- How to *minimize development time* for writing model validation tests?
- How to adjudicate whether a test **score** really captures the functionality it claims to test?
- We aim to solve these with:
 - *unit testing philosophy*
 - *domain standards and tools*
 - *collaborative development*

SCIUNIT: TESTS



SCIUNIT: TESTS



Fork it!



COMPETITION

```
from QSNMC.tests import tests
from QSNMC.models import models

for model in models:
    for test in tests:
        score = test.judge(model)
        score.summarize()
```

Quantitative Single-Neuron Modeling: Competition 2009

Richard Naud^{1*}, Thomas Berger¹, Brice Bathellier², Matteo Carandini³ and
Wulfram Gerstner¹

¹ Ecole Polytechnique Federal de Lausanne (EPFL), Switzerland

² University of Bern, Switzerland

³ University College London, United Kingdom

COMPETITION

```
from QSNMC.tests import tests
from QSNMC.models import models

for model in models:
    for test in tests:
        score = test.judge(model)
        score.summarize()
```

Quantitative Single-Neuron Modeling: Competition 2009

Richard Naud^{1*}, Thomas Berger¹, Brice Bathellier², Matteo Carandini³ and
Wulfram Gerstner¹

¹ Ecole Polytechnique Federal de Lausanne (EPFL), Switzerland

² University of Bern, Switzerland

³ University College London, United Kingdom

<https://github.com/incf/qsnmc>

COMPETITION

SciUnit can also provide a common evaluation and visualization framework for competitions between *algorithms or methods*:

- Spike-sorting algorithms
 - formalization and extension of Schmitzer-Torbert et al, 2005.
- Spike time extraction from fast calcium imaging.
 - Jason Kerr has a “ground truth” data set to enable this.
- Neural system prediction and identification challenge (*nuSPIC*)
 - Can we infer the function of neuronal networks given connectivity?
 - A latter day “Hopfield challenge”.
- Brain computer interface (BCI, PhyPa (<http://www.phypa.org>)
- Seizure prediction (Sejnowski challenge)

CHALLENGES

CHALLENGES

- A big challenge is handling **different implementation languages** with a common system.

CHALLENGES

- A big challenge is handling **different implementation languages** with a common system.
- We must be able to express **complex model capabilities and test requirements**.

CHALLENGES

- A big challenge is handling **different implementation languages** with a common system.
- We must be able to express **complex model capabilities and test requirements**.
- We must be able to handle **large, diverse data sets** and **complex models**.

CHALLENGES

- A big challenge is handling **different implementation languages** with a common system.
- We must be able to express **complex model capabilities and test requirements**.
- We must be able to handle **large, diverse data sets** and **complex models**.
- We aim to solve these with *unit testing philosophy + domain standards*

ACTIVE USE CASES

ACTIVE USE CASES

- Quantitative Single Neuron Modeling Competition
 - reduced, tractable neuron models
 - <http://github.com/incf/QSNMC>
 - tested using data collected by original authors

ACTIVE USE CASES

- Quantitative Single Neuron Modeling Competition
 - reduced, tractable neuron models
 - <http://github.com/incf/QSNMC>
 - tested using data collected by original authors
- Open Source Brain
 - biophysically detailed neuron/microcircuit models
 - <http://www.opensourcebrain.org>
 - tested using data from NeuroElectro

ACTIVE USE CASES

- Quantitative Single Neuron Modeling Competition
 - reduced, tractable neuron models
 - <http://github.com/incf/QSNMC>
 - tested using data collected by original authors
- Open Source Brain
 - biophysically detailed neuron/microcircuit models
 - <http://www.opensourcebrain.org>
 - tested using data from NeuroElectro
- OpenWorm
 - simulation of an entire organism
 - <http://www.openworm.org>
 - tested using data from neuron to behavior

Experiment Data Discovery



Experiment Data Format



Experiment Data Analysis



Data Summary



Test Instantiation



Test Score Visualization



Test Adjudication

Model Discovery



Model Description



Model Capabilities



Simulation Description



Model Interpreter

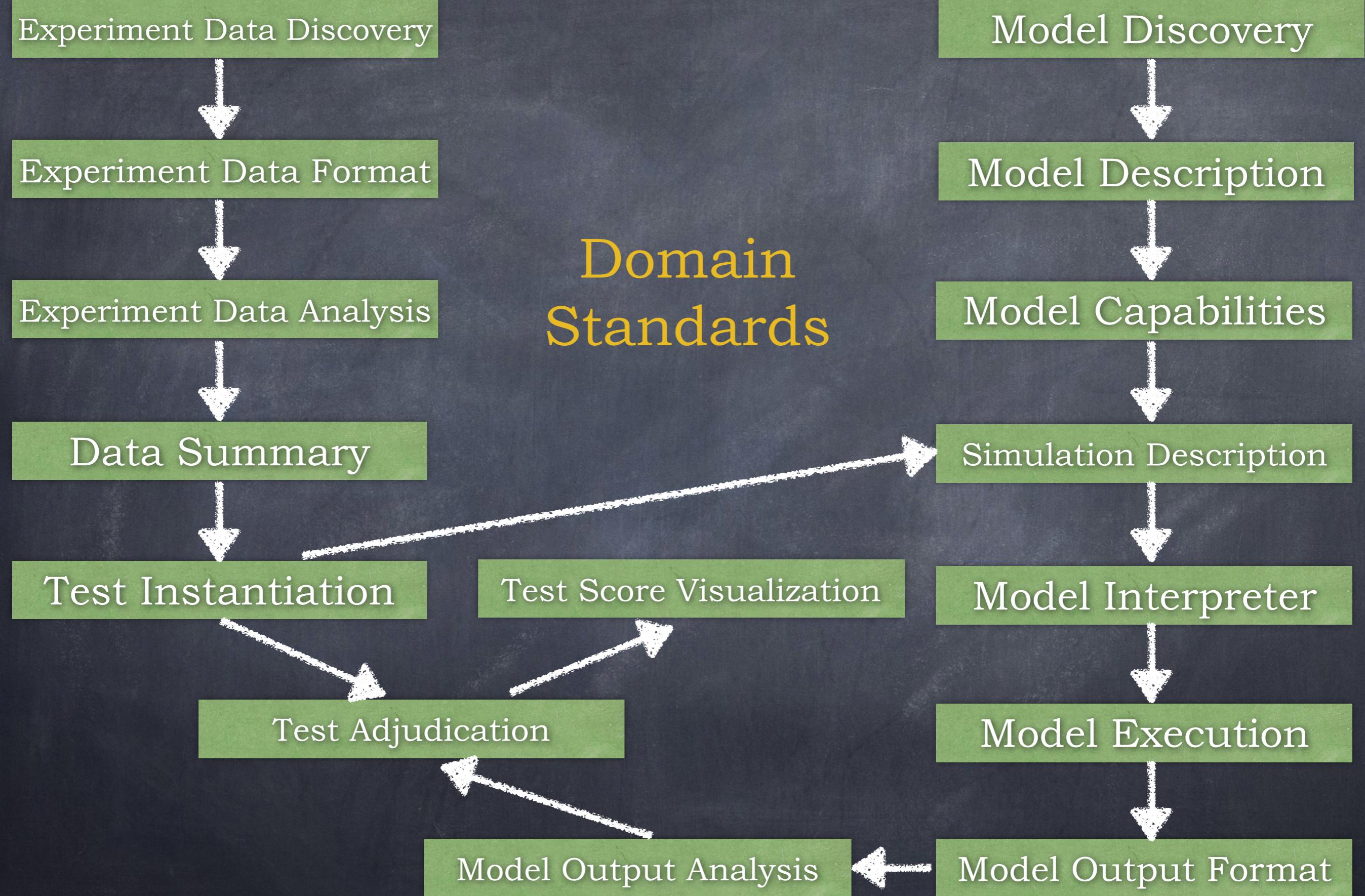


Model Execution



Model Output Format





Experiment Data Discovery

CRCNS.org

Experiment Data Format

NIX

Experiment Data Analysis

NEO/NeuroTools

Data Summary

NeuroElectro

Test Instantiation

NeuronUnit

Test Adjudication

NumPy/SciPy

Model Output Analysis

NEO/NeuroTools

Domain Standards

Model Discovery

Open Source Brain

Model Description

NeuroML

Model Capabilities

Comp. Neuroscience Ontology

Simulation Description

SED-ML

Model Interpreter

NeuroConstruct

Model Execution

NEURON

Model Output Format

NSDF

Experiment Data Discovery

CRCNS.org

Experiment Data Format

NIX

Experiment Data Analysis

NEO/NeuroTools

Data Summary

NeuroElectro

Test Instantiation

NeuronUnit

Test Adjudication

NumPy/SciPy

Model Output Analysis

NEO/NeuroTools

Linking Domain Standards

Model Discovery

Open Source Brain

Model Description

NeuroML

Model Capabilities

Comp. Neuroscience Ontology

Simulation Description

SED-ML

Model Interpreter

NeuroConstruct

Model Execution

NEURON

Model Output Format

NSDF

NEURONUNIT

NEURONUNIT

- *SciUnit* is practical with *domain-specific* libraries for test construction and model introspection / execution that utilize domain-specific *standards*.

NEURONUNIT

- *SciUnit* is practical with *domain-specific* libraries for test construction and model introspection / execution that utilize domain-specific *standards*.
- *NeuronUnit* is a such a library for single neuron electrophysiology.

NEURONUNIT

- *SciUnit* is practical with *domain-specific* libraries for test construction and model introspection / execution that utilize domain-specific *standards*.
- *NeuronUnit* is a such a library for single neuron electrophysiology.
 - <http://github.com/scidash/neuronunit>

NEURONUNIT

NEURONUNIT

- SciUnit *capability* classes crafted to span the things one would ask single cell neurophysiology models to do.

NEURONUNIT

- SciUnit *capability* classes crafted to span the things one would ask single cell neurophysiology models to do.
 - *Helper functions* for implementation of these capabilities to match common model/simulation platforms (*neuroConstruct*)

NEURONUNIT

- SciUnit *capability* classes crafted to span the things one would ask single cell neurophysiology models to do.
 - *Helper functions* for implementation of these capabilities to match common model/simulation platforms (*neuroConstruct*)
- SciUnit *model* classes that implement these capabilities as well as platform-specific model instantiation.

NEURONUNIT

- SciUnit *capability* classes crafted to span the things one would ask single cell neurophysiology models to do.
 - *Helper functions* for implementation of these capabilities to match common model/simulation platforms (*neuroConstruct*)
- SciUnit *model* classes that implement these capabilities as well as platform-specific model instantiation.
- SciUnit *test* classes that can be initialized with human-readable observation metadata.

NEURONUNIT

- SciUnit *capability* classes crafted to span the things one would ask single cell neurophysiology models to do.
 - *Helper functions* for implementation of these capabilities to match common model/simulation platforms (*neuroConstruct*)
- SciUnit *model* classes that implement these capabilities as well as platform-specific model instantiation.
- SciUnit *test* classes that can be initialized with human-readable observation metadata.
 - Observed data for test parameterization retrieved automatically from trusted data repositories.

BENEFITS (MODELER'S PERSPECTIVE)

BENEFITS (MODELER'S PERSPECTIVE)

- Know what other models can and can't do, and what a new model (if needed) should explain in order to be better.

BENEFITS (MODELER'S PERSPECTIVE)

- Know what other models can and can't do, and what a new model (if needed) should explain in order to be better.
- The ability to continuously test your model against the data it is supposed to explain/predict.

BENEFITS (MODELER'S PERSPECTIVE)

- Know what other models can and can't do, and what a new model (if needed) should explain in order to be better.
- The ability to continuously test your model against the data it is supposed to explain/predict.
 - Accelerates model development (towards some goal of realism).

BENEFITS (MODELER'S PERSPECTIVE)

- Know what other models can and can't do, and what a new model (if needed) should explain in order to be better.
- The ability to continuously test your model against the data it is supposed to explain/predict.
 - Accelerates model development (towards some goal of realism).
 - Gives you bragging rights in the arena of model competition.

BENEFITS (MODELER'S PERSPECTIVE)

- Know what other models can and can't do, and what a new model (if needed) should explain in order to be better.
- The ability to continuously test your model against the data it is supposed to explain/predict.
 - Accelerates model development (towards some goal of realism).
 - Gives you bragging rights in the arena of model competition.
 - Address reviewers (me) who demand that your model pass SciUnit tests.

BENEFITS (MODELER'S PERSPECTIVE)

- Know what other models can and can't do, and what a new model (if needed) should explain in order to be better.
- The ability to continuously test your model against the data it is supposed to explain/predict.
 - Accelerates model development (towards some goal of realism).
 - Gives you bragging rights in the arena of model competition.
 - Address reviewers (me) who demand that your model pass SciUnit tests.
 - Post-publication review of your model, *even as new data come to light.*

BENEFITS (MODELER'S PERSPECTIVE)

- Know what other models can and can't do, and what a new model (if needed) should explain in order to be better.
- The ability to continuously test your model against the data it is supposed to explain/predict.
 - Accelerates model development (towards some goal of realism).
 - Gives you bragging rights in the arena of model competition.
 - Address reviewers (me) who demand that your model pass SciUnit tests.
 - Post-publication review of your model, *even as new data come to light.*
 - Look your child in the face when they ask if you, an alleged scientist, used the scientific method in the development of your model.

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.
 - Is there a model that explains my data?

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.
 - Is there a model that explains my data?
 - Which model best explains my data?

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.
 - Is there a model that explains my data?
 - Which model best explains my data?
 - What other data does it explain?

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.
 - Is there a model that explains my data?
 - Which model best explains my data?
 - What other data does it explain?
- Pre-experiment, grant stage discovery of hypothesis implications.

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.
 - Is there a model that explains my data?
 - Which model best explains my data?
 - What other data does it explain?
- Pre-experiment, grant stage discovery of hypothesis implications.
 - If I do experiment E and get result Y, it will support model A.

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.
 - Is there a model that explains my data?
 - Which model best explains my data?
 - What other data does it explain?
- Pre-experiment, grant stage discovery of hypothesis implications.
 - If I do experiment E and get result Y, it will support model A.
 - If I do experiment E and get result Z, it will support model B.

BENEFITS

(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.
 - Is there a model that explains my data?
 - Which model best explains my data?
 - What other data does it explain?
- Pre-experiment, grant stage discovery of hypothesis implications.
 - If I do experiment E and get result Y, it will support model A.
 - If I do experiment E and get result Z, it will support model B.
- Increased community awareness of the data you collected.

BENEFITS

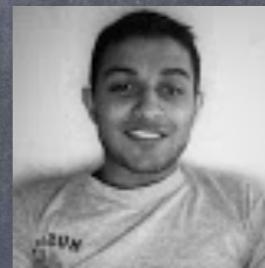
(EXPERIMENTALIST'S/TESTER'S PERSPECTIVE)

- The ability to put your observations in context.
 - Is there a model that explains my data?
 - Which model best explains my data?
 - What other data does it explain?
- Pre-experiment, grant stage discovery of hypothesis implications.
 - If I do experiment E and get result Y, it will support model A.
 - If I do experiment E and get result Z, it will support model B.
- Increased community awareness of the data you collected.
 - It could become the gold standard by which models are judged!

ACKNOWLEDGEMENTS

- **Cyrus Omar**

Carnegie Mellon Univ., USA



- **Shreejoy Tripathy**

Univ. of British Columbia, Canada



- **Sharon Crook**

Arizona State University, USA



- **Padraig Gleeson**

University College, London, UK

