Systematic validation of detailed models of hippocampal neurons based on electrophysiological data

Theses of the Ph.D. Dissertation

Sára Sárty

Supervisors: Dr. Szabolcs Káli, Dr. Tamás Freund

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics
Roska Tamás Doctoral School of Sciences and Technology

Budapest, 2021
1. Introduction and aims

The construction and simulation of anatomically and biophysically detailed models is becoming a standard tool in neuroscience [1]. Such models, which typically employ the compartmental modeling approach and a Hodgkin-Huxley-type description of voltage-gated ion channels, are capable of providing fairly accurate models of single neurons [2]–[4] and (when complemented by appropriate models of synaptic interactions) even large-scale circuits [5], [6]. However, building such detailed multi-compartmental models of neurons requires setting a large number of parameters (such as the densities of various ion channels in multiple neuronal compartments) that are often not directly constrained by the available experimental data. These parameters are typically tuned (either manually or using automated parameter-search methods [7], [8]) until the simulated physiological behavior of the model matches some pre-defined set of experimental observations.

For an increasing number of cell types, the available experimental data already provide diverse constraints on the expected physiological behavior of the neuron under a variety of conditions. Based on various (typically small) subsets of the available constraints, a large number of different models of several cell types have been developed to investigate diverse aspects of single-cell behavior, and for inclusion in realistic circuit models. As an example, there are currently 136 different models related to the hippocampal CA1 pyramidal cell (PC) in the ModelDB database [9]. However, even though these models are publicly available, it is still technically challenging to verify their behavior beyond the examples explicitly included with the model, and especially to test their behavior outside the context of the original study, or to compare it with the behavior of other models. This sparsity of information about the performance of detailed models may also be one reason why model re-use in the community is relatively limited, which decreases the chance of spotting errors in modeling studies, and may lead to an unnecessary replication of effort. In addition, even when models are re-used, they are often altered to fit a different subset of the available experimental data, and they may lose their ability to capture the behaviors that were used to constrain the original model. This phenomenon (whereby introducing new features breaks previously correct behavior) is known as a “regression” in software development, and is typically avoided by regularly applying a set of tests that comprehensively verify the correct behavior of the software under various circumstances. Such comprehensive checks are not routinely performed when neural models are developed – and this may be one of the reasons why the development of consensus (community) models, which would aim to capture a wide range of experimental observations by integrating diverse efforts, has rarely been attempted in neuroscience.

A collaborative approach to modeling, and even a systematic comparison of existing models built in different laboratories requires the development of a comprehensive validation
suite, a set of automated tests that quantitatively compare various aspects of model behavior with the corresponding experimental data. Such validation suites enable all modeling groups to evaluate their existing and newly developed models according to the same set of well-defined criteria, thus facilitating model comparison and providing an objective measure of progress in matching relevant experimental observations. Applying automated tests also allows researchers to learn more about models published by other groups (beyond the results included in the papers) with relatively little effort, thus facilitating optimal model re-use and co-operative model development. In addition, systematic, automated testing is expected to avoid regressions, aid the identification of problematic aspects of model behavior, and speed up model development in general by allowing researchers to easily evaluate models in relation to the relevant experimental data after every iteration of model adjustment. Finally, a comprehensive evaluation of model behavior appears to be critical for models that are then expected to provide useful predictions in a new context. A prime example of this is detailed single cell models included in network models, where diverse aspects of cellular function such as synaptic integration, intracellular signal propagation, spike generation and adaptation mechanisms all contribute to the input-output function of the neuron in the context of an active network. By comparing multiple different aspects of the behavior of the single cell model with experimental data, one can increase the chance of having a model that also behaves correctly within the network. The technical framework for developing automated test suites for models already exists [10], and is currently used by several groups to create a variety of tests for models of neural structure and function at different scales [11]–[13]. In the current study, our goal was to develop a validation suite for the physiological behavior of one of the most studied cell types of the mammalian brain, the pyramidal cell in area CA1 of the rat hippocampus.

CA1 pyramidal neurons display a large repertoire of nonlinear responses in all of their compartments (including the soma, axon, and various functionally distinct parts of the dendritic tree), which are experimentally well-characterized. In particular, there are detailed quantitative results available on the subthreshold and spiking voltage response to somatic current injections [14], [15]; on the properties of the action potentials back-propagating from the soma into the dendrites [16], which is a basic measure of dendritic excitability; and on the characteristics of the spread [17] and non-linear integration of synaptically evoked signals in the dendrites, including the conditions necessary for the generation of dendritic spikes [18]–[20].

The test suite that we have developed allows the systematic and quantitative comparison of the behavior of anatomically and biophysically detailed models of rat CA1 pyramidal neurons with experimental data in all of these domains. To demonstrate the utility of our approach, we applied our tests to compare the behavior of several different rat hippocampal CA1 pyramidal cell models from the ModelDB database against electrophysiological data available in the literature, and evaluated how well these models match experimental observations in different
We employed the test suite to aid the development of models within the European Human Brain Project (HBP), and integrated the tests into the Validation Framework [21] developed in the HBP, with the aim of facilitating more reproducible and transparent model building in the neuroscience community.

2. Methods

2.1. Implementation of HippoUnit

HippoUnit (https://github.com/KaliLab/hippounit) is a Python test suite based on the SciUnit [10] framework, which is a Python package for testing scientific models. During its implementation the NeuronUnit package [11] was taken into account as an example of how to use the SciUnit framework for testing neuronal models. In SciUnit tests usually four main classes are implemented: the test class, the model class, the capabilities class and the score class. HippoUnit is built in a way that keeps this structure. The key idea behind this structure is the decoupling of the model implementation from the test implementation by defining standardized interfaces (capabilities) between them, so that tests can easily be used with different models without being rewritten, and models can easily be adapted to fit the framework.

Each test of HippoUnit is a separate Python class that, similarly to other SciUnit packages, can run simulations on the models to generate model predictions, which can be compared with experimental observations to yield the final score, provided that the model has the required capabilities implemented to mimic the appropriate experimental protocol and produce the same type of measurable output.

Similarly to many of the existing SciUnit packages the implementations of specific models are not part of the HippoUnit package itself. Instead, HippoUnit contains a general ModelLoader class. This class is implemented in a way that it is able to load and deal with most types of models defined in the HOC language of the NEURON simulator (either as standalone HOC models or as HOC templates) [22]. It implements all model-related methods (capabilities) that are needed to simulate these kinds of neural models in order to generate the prediction without any further coding required from the user.

For the smooth validation of the models developed using parameter optimization within the Human Brain Project there is a child class of the ModelLoader available in HippoUnit that is called ModelLoader_BPO. This class inherits most of the functions (especially the capability functions) from the ModelLoader class, but it implements additional functions that are able to automatically deal with the specific way in which information is represented and stored in these optimized models. The role of these functions is to gather all the information from the metadata
and configuration files of the models that are needed to set the parameters required to load the models and run the simulations on them (such as path to the model files, name of the model template or the simulation temperature (the celsius variable of Neuron)). For neural models developed using other software and methods, the user needs to implement the capabilities through which the tests of HippoUnit perform the simulations and recordings on the model.

The capabilities are the interface between the tests and the models. The ModelLoader class inherits from the capabilities and must implement the methods of the capability. The test can only be run on a model if the necessary capability methods are implemented in the ModelLoader. All communication between the test and the model happens through the capabilities.

The methods of the score classes perform the quantitative comparison between the prediction and the observation, and return the score object containing the final score and some related data, such as the paths to the saved figure and data (JSON) files and the prediction and observation data. For simplicity, we refer to the discrepancy between the target experimental data (observation) and the models’ behavior (prediction) with respect to a studied feature using the term feature score. In most cases, when the basic statistics (mean and standard deviation) of the experimental features (typically measured in several different cells of the same cell type) are available, feature scores are computed as the absolute difference between the feature value of the model and the experimental mean feature value, divided by the experimental standard deviation (Z-score) [23]. The final score of a given test achieved by a given model is given by the average (or, in some cases, the sum) of the feature scores for all the features evaluated by the test.

## 2.2. Models from literature

In the dissertation I demonstrate the utility of the HippoUnit validation test suite by applying its tests to validate and compare the behavior of several different detailed rat hippocampal CA1 pyramidal cell models available on ModelDB [9]. For this initial comparison we chose models published by several modeling groups worldwide that were originally developed for various purposes. The models compared were the following: the Golding et al., 2001 model [16] (ModelDB accession number: 64167), the Katz et al., 2009 model [24] (ModelDB accession number: 127351), the Migliore et al., 2011 model [25] (ModelDB accession number: 138205), the Poirazi et al., 2003 model [26], [27] (ModelDB accession number: 20212), the Bianchi et al., 2012 model [15] (ModelDB accession number: 143719), and the Gómez González et al., 2011 [28] model (ModelDB accession number: 144450).

Models from literature that are published on ModelDB typically implement their own simulations and plots to make it easier for users and readers to reproduce and visualize the results
shown in the corresponding paper. Therefore, to be able to test the models described above using our test suite, we needed to create standalone versions of them. These standalone versions do not display any GUI, or contain any built-in simulations and run-time modifications, but otherwise their behavior should be identical to the published version of the models. We also added section lists of the radial oblique and the trunk dendritic sections to those models where this was not done yet, as some of the tests require these lists. To ensure that the standalone versions have the same properties as the original models, we checked their parameters after running their built-in simulations (in case they include any run-time modifications), and made sure they match the parameters of the standalone version. The modified models used for running validation tests are available in this GitHub repository: https://github.com/KaliLab/HippoUnit_demo.

2.3. Running the tests of HippoUnit

One convenient way of running a test on a model is to use an interactive computational notebook, such as the Jupyter Notebook [29], which enables the combination of program codes to be run (we used Python code to access the functionality of HippoUnit), the resulting outputs (e.g. figures, tables, text) and commentary or explanatory text in a single document. Therefore, we demonstrate the usage of HippoUnit through this method (See https://github.com/KaliLab/HippoUnit_demo).

3. Summary of novel scientific results

Thesis I: I proposed, elaborated and developed an open-source Python validation test suite (HippoUnit), which is the first test suite to make it possible to automatically and systematically test multiple properties of anatomically and biophysically detailed models of the hippocampal CA1 pyramidal cell by making quantitative comparisons between the models and electrophysiological data.

The tests of HippoUnit automatically perform simulations that mimic common electrophysiological protocols on single-cell models to compare their behavior with quantitative experimental data using various feature-based error functions. Current validation tests cover somatic (subthreshold and spiking) behavior as well as signal propagation and integration in the dendrites. These tests were chosen because they collectively cover diverse functional aspects of cellular behavior that have been thoroughly investigated in experimental and modeling studies.

Corresponding publications: [Th1], [Th3-Th7], [Th13-Th16]
Thesis II: I demonstrated the utility of my validation test suite by applying its tests to compare the behavior of several different hippocampal CA1 pyramidal cell models from the ModelDB database against electrophysiological data available in the literature. This way I also compared the models to each other and tested their generalization performance in paradigms that they were not originally designed to capture. I concluded and showed that each of these models provide a good match to experimental results in some domains but not in others. Thus automated, systematic testing is needed to reveal the weaknesses and strengths of neural models available in the literature and to evaluate their usefulness according to the needs of the user.

Corresponding publications: [Th1], [Th10-Th12]

Thesis III: I employed the HippoUnit test suite to validate the dendritic properties of models of hippocampal CA1 neurons that were developed within the Human Brain Project using parameter optimization methods, but considering only somatic features. This way I showed that these models are suitable for studying synaptic properties.

I validated the dendritic properties, especially the attenuation of synaptically induced EPSPs, of several versions of these models during their development process and showed that the v4 version of these models are suitable for being used in the in silico study of synaptic
physiology in the hippocampal CA1 region, as the attenuation of synaptic excitatory postsynaptic potentials (EPSPs) is consistent with experimental data [Th2].

Figure 2: Results from the PSP Attenuation Test of HippoUnit applied to the new (v4) version of the BluePyOpt [8] optimized models within the Human Brain Project. Soma/dendrite EPSP attenuation as a function of the synaptic input distance from the soma in the different models. The plot shows that the attenuation of synaptic EPSPs is consistent with experimental data [17].

Corresponding publication: [Th2], [Th7]

Thesis IV: I employed the test suite to aid the development of models of hippocampal neurons within the Human Brain Project by systematically validating and thus monitoring the performance of them at various stages of model development. Based on the validation results I proposed the direction in which they should be further developed and removed those models that did not meet our needs. These models were then used to build a network model of the hippocampal CA1 region.

I applied the tests of HippoUnit to the version of the models published in Migliore et al. (2018) [14], and to a later version (v4) described in Ecker et al. (2020) [Th2], which was intended to further improve the dendritic behavior of the models, as this is critical for their proper functioning in the network. The two sets of models were created using the same morphology files and similar optimization methods and protocols. These new optimizations differed mainly in the allowed range for the density of the sodium channels in the dendrites. For the pyramidal cell models a new feature was also introduced in the parameter optimization that constrains the amplitudes of back-propagating action potentials in the main apical dendrite, that were shown to be too high by the validation results. The new interneuron models also had an exponentially decreasing (rather than constant) density of Na channels, and A-type K channels with more hyperpolarized activation in their dendrites. With these modifications, that were
introduced based on previous validation results, the new (v4) version of the models showed dendritic behaviour matching experimental data significantly better.

Figure 4: Employing the tests of HippoUnit to monitor the behavior of a set of detailed data-driven models of hippocampal neurons at different stages of model development. Models of four different cell types (pyramidal cells and continuous accommodating (int cAC), bursting accommodating (int bAC) and continuous non-accommodating (int cNAC) interneurons) of the rat hippocampal CA1 region were developed within the Human Brain Project by automated optimization using BluePyOpt. The tests of HippoUnit were used to evaluate and compare the behavior of the older (Migliore et al 2018) version and the new (v4) version of these models. The median, the interquartile range and the full range of the final scores achieved by the models of each cell type were calculated and the results of the two versions of the model set are compared. Asterisks indicate significant differences (*: p<0.05; **: p<0.01). [Th1]

Corresponding publications: [Th1], [Th2], [Th7]

**Thesis V: HippoUnit was the first test suite to be integrated into the Validation Framework developed within the Human Brain Project, that makes it possible to permanently record, examine and reproduce validation results, and enables tracking the evolution of models over time, as well as comparison against other models in the domain. I also integrated the validation tests of HippoUnit into the Brain Simulation Platform of the Human Brain Project and developed online Use Cases, that allow to run the tests on different models in a browser without the need of locally installing the required packages. By making these tools widely available I facilitated more reproducible and transparent model building in the neuroscience community.**

Every test of HippoUnit has been individually registered in the Validation Framework [21]. The files containing the target experimental data for each test are stored (besides the HippoUnit_demo GitHub repository) in storage containers at the Swiss National Supercomputing Centre (CSCS), where they are publicly available. The location of the corresponding data file is associated with each registered test, so that the data are loaded
automatically when the test is run on a model via the Validation Framework. All the models that were tested and compared in this study (including the CA1 pyramidal cell models from the literature and the BluePyOpt optimized CA1 pyramidal cells and interneurons of the HBP) have been registered and are available in the Model Catalog of the Validation Framework with their locations in the CSCS storage linked to them. Moreover, the validation results discussed in the dissertation have also been registered in the Validation Framework, with all their related files (output figures and JSON files) linked to them. These can be accessed using the Model Validation app of the framework.

The Live Paper we created on the Brain Simulation Platform (https://humanbrainproject.github.io/hbp-bsp-live-papers/2021/saray_et_al_2021/saray_et_al_2021.html) shows the results of the study presented here in more detail. This interactive document provides links to all the output figures and data files resulting from the validation of the models. Moreover, as part of this Live Paper a HippoUnit Use Case is also available in the form of a Jupyter Notebook, which guides the user through running the validation tests of HippoUnit on the models, that are already registered in the Framework, and makes it possible to reproduce the results.

Corresponding publications: [Th1], [Th9], [Th15], [Th16]

4. Applications of the results

HippoUnit, together with its possible extensions and other similar tools, allows the rapid, systematic evaluation and comparison of neuronal models in multiple domains. By providing the software tools and examples for effective model validation, we hope to encourage the modeling community to use more systematic testing during model development, with the aim of making the process of model building more efficient, reproducible and transparent.

For anatomically and biophysically detailed data-driven neural models to be predictive, it is important that they are able to generalize beyond their original scope. However, most detailed biophysical models to date were built to capture only a few important or interesting properties of a given neuron type. Systematic testing and comparison of the behavior of these models is still rare, and thus it is often unknown how these models would behave when used under different circumstances, and to what extent they can be used to address different scientific questions. As a result, the modeling community still keeps building new models of the same cell type for various purposes, instead of reusing and further developing the already existing ones. On the other hand, in those cases when new models are based on previously published ones, model parameters are often adjusted to fit a new set of experimental data. These adjustments typically alter the ability of the model to capture the experimental data targeted by the original
model, but this remains unrecognized because of the lack of comprehensive testing. In addition, some publications on neuronal models simply state that the model has been validated against electrophysiological data, but the details of these validations (such as the methods used, the experimental data considered or even the results) are usually not shared.

Therefore, one important use case for the application of HippoUnit is the evaluation and comparison of existing models. We demonstrated this by using HippoUnit to test and compare the behavior of several models of rat CA1 pyramidal neurons available on ModelDB in several distinct domains against electrophysiological data available in the literature (or shared by collaborators). Besides providing independent and standardized verification of the behavior of the models, the results also allow researchers to judge which existing models show a good match to the experimental data in the domains that they care about, and thus to decide whether they could re-use one of the existing models in their own research.

Besides enabling the comparison of different models regarding how well they match a particular dataset, the tests of HippoUnit also allow one to determine the match between a particular model and several datasets of the same type. As experimental results can be heavily influenced by recording conditions and protocols, and also depend on factors such as the strain, age, and sex of the animal, it is important to find out whether the same model can simultaneously capture the outcome of different experiments, and if not, how closely it is able to match the different datasets.

HippoUnit is also a useful tool during model development. In a typical data-driven modeling scenario, researchers decide which aspects of model behavior are relevant for them, find experimental data that constrain these behaviors, then use some of these data to build the model, and use the rest of the data to validate the model. HippoUnit and similar test suites make it possible to define quantitative criteria for declaring a model valid (ideally before modeling starts), and to apply these criteria consistently throughout model development. We demonstrated this approach through the example of detailed single cell models of rat CA1 pyramidal cells and interneurons optimized within the HBP.

Furthermore, several authors have argued for the benefits of creating “community models” [30, 31] through the iterative refinement of models in an open collaboration of multiple research teams. Such consensus models would aim to capture a wide range of experimental observations, and may be expected to generalize (within limits) to novel modeling scenarios. A prerequisite for this type of collaborative model development is an agreement on which experimental results will be used to constrain and validate the models. Automated test suites provide the means to systematically check models with respect to all the relevant experimental data, with the aim of tracking progress and avoiding “regression,” whereby previously correct model behavior is corrupted by further tuning.
Finally, the tests of HippoUnit have been integrated into the recently developed Validation Framework of the HBP, which makes it possible to collect neural models and validation tests, and supports the application of the registered tests to the registered models. Most importantly, it makes it possible to save the validation results and link them to the models in the Model Catalog, making them publicly available and traceable for the modeling community.

5. The Author’s publications

5.1. Publications related to the theses:

5.1.1. Journal papers:

[Th1] Sára Sáray, Christian A. Rössert, Shailesh Appukuttan, Rosanna Migliore, Paola Vitale, Carmen A. Lupascu, Luca L. Bologna, Werner Van Geit, Armando Romani, Andrew P. Davison, Eilif Muller, Tamás F. Freund, Szabolcs Káli: HippoUnit: A software tool for the automated testing and systematic comparison of detailed models of hippocampal neurons based on electrophysiological data, PLOS Computational Biology; https://doi.org/10.1371/journal.pcbi.1008114

[Th2] András Ecker, Armando Romani, Sára Sáray, Szabolcs Káli, Michele Migliore, Joanne Falck Sigrun Lange, Audrey Mercer, Alex M. Thomson, Eilif Muller, Michael W. Reimann Srikanth Ramaswamy: Data-driven integration of hippocampal CA1 synaptic physiology in silico, Hippocampus https://doi.org/10.1002/hipo.23220

5.1.2. Posters at international conferences:


[Th6] Sára Sáray, Shailesh Appukuttan, Bence Bagi, Pedro E. Garcia-Rodriguez, Péter Kováč, Carmen A. Lupascu, Máté Mohácsi, Christian A. Rössert, Luca Tar, Márk P. Török, Andrew Davison, Michele Migliore, Eilif Muller, Tamás F. Freund, Szabolcs Káli:


[Th8] Luca Tar, Sára Sáray, Tamás Freund, Szabolcs Káli: Developing a detailed model of CA1 pyramidal neurons using automated optimization and validation tools, 11th FENS Forum of Neuroscience Conference, July 7-11, 2018, Berlin, Germany


5.1.3. Annual Proceedings of the PPCU Faculty of Information technology and Bionics Doctoral School:


5.1.4. Workshop presentation


5.2. Publications not related to the theses

5.2.1. Conference papers:


[Au5] Luca Tar, Zsuzsanna Bengery, Sára Sáray, Tamás Freund, Szabolcs Káli; The contribution of dendritic spines to synaptic integration and plasticity in hippocampal pyramidal neurons, 28th Annual Computational Neuroscience Meeting (CNS*2019), July 13-17, 2019, Barcelona, Spain https://doi.org/10.1186/s12868-019-0538-0

Bibliography


