

Applications Note

The Systems Biology Simulation Core Library

Hemil Panchiwala^{1†}, Shalin Shah^{2, 3†}, Hannes Planatscher⁴, Mykola Zakharchuk⁵, Matthias König⁶, and Andreas Dräger^{5, 7, 8, 9*}

¹Dept. of Computer Science and Engineering, Indian Institute of Technology, Roorkee, IN - 247667 ²Dept. of Electrical and Computer Engineering, Duke University, Durham, NC 27701, USA ³Bloomberg LP, New York, NY 10022, USA ⁴Signatope GmbH, 72770 Reutlingen, DE ⁵Dept. of Computer Science, University of Tübingen, 72076 Tübingen, DE ⁶Institute for Theoretical Biology, Humboldt University of Berlin, 10115 Berlin, DE ⁷Computational Systems Biology of Infections and Antimicrobial-Resistant Pathogens, Institute for Bioinformatics and Medical Informatics (IBMI), University of Tübingen, 72076 Tübingen, DE ⁸German Center for Infection Research (DZIF), partner site Tübingen, DE ⁹Cluster of Excellence 'Controlling Microbes to Fight Infections,' University of Tübingen, Tübingen, DE

† These authors contributed equally to the manuscript

*To whom correspondence should be addressed

Associate Editor: XXXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Summary: Studying biological systems generally relies on computational modelling and simulation, e.g., model-driven discovery and hypothesis testing. Progress in standardisation efforts led to the development of interrelated file formats to exchange and reuse models in systems biology, such as SBML, the Simulation Experiment Description Markup Language (SED-ML), or the Open Modeling EXchange format (OMEX). Conducting simulation experiments based on these formats requires efficient and reusable implementations to make them accessible to the broader scientific community and to ensure the reproducibility of the results. The Systems Biology Simulation Core Library (SBSCL) provides interpreters and solvers for these standards as a versatile open-source API in Java™. The library simulates even complex bio-models and supports deterministic Ordinary Differential Equations (ODEs); Stochastic Differential Equations (SDEs); constraint-based analyses; recent SBML and SED-ML versions; exchange of results, and visualisation of *in silico* experiments; open modelling exchange formats (COMBINE archives); hierarchically structured models; and compatibility with standard testing systems, including the Systems Biology Test Suite and published models from the BioModels and BiGG databases.

Availability: SBSCL is freely available at <https://draeger-lab.github.io/SBSCL/> and via Maven Central.

Contact: draeger@informatik.uni-tuebingen.de

Supplementary information: The material available at *Bioinformatics* online provides details on resources and availability, implementation, support of the SBML Test Suite, BioModels, and BiGG simulations with benchmark comparisons, and comparison to other simulators with SBML support.

1 Introduction

The Systems Biology Simulation Core Library (SBSCL) is an open-source, cross-platform pure Java™ programming library that numerically solves systems biology models in multiple mathematical frameworks. A popular file format for representing computational models in a standard way and facilitating the exchange of models between different tools is the Systems Biology Markup Language (SBML) [3]. SBML encodes biological models in a declarative form. The Simulation Experiment Description

Markup Language (SED-ML) format defines a workflow of simulation experiments. The combination of SED-ML and SBML facilitates reproducibility of typical model workflows in *in-silico* experiments, including the choice of interpretation framework and the postprocessing of the results [8]. SBSCL interprets the SBML models using the JSBML library [7] and simulates them according to dedicated API calls. Alternatively, it extracts an *in silico* experimental configuration from SED-ML to simulate the SBML models. To this end, SBSCL implements and ships several solvers for a wide range of mathematical frameworks, including ODEs [4], SDEs [2], and constraint-based analysis. SBSCL is designed as a

lightweight API and intended for use as a simulation backend within end-user software. This article introduces the SBSCL library, especially the new features introduced in version 2.1, along with a brief description of all other capabilities that figure 1 on page 2 pictorially summarises.

2 Description

Differential equation solver: The most fundamental feature of SBSCL is simulating ODEs. Version 2.1 adds interpreters and solvers for SDEs to support the latest SBML standards. SBSCL efficiently implements three deterministic numerical solvers [4], namely, Rosenbrock, Euler, and Runge-Kutta, as well as three stochastic solvers, namely, Gillespie, Gibson-Bruck, and Tau-Leaping [2].

Constrained optimisation solver: SBML Level 3 [3] combined with the `fbc` package added support for constrained-based models and their analysis. Typically, Flux Balance Analysis (FBA) is used for such time-invariant steady-state simulations. SBSCL performs FBA on SBML models using the SCPSolver (<http://www.scpsolver.org>), a linear programming API with support for various solver backends. This lightweight abstraction allows users to define model constraints and an objective function and solve the corresponding optimisation problem.

Result tables and plots: Since viewing is an essential aspect of understanding the results of a simulation experiment, SBSCL provides experiment output in graphical and tabular form, which it can export in conventional formats such as Comma-Separated Values (CSV).

Archival format support: Working towards exchangeability and reproducibility, SBSCL v2.1 supports the OMEX format as input. These archive files contain the information on running simulation experiments based on SBML and SED-ML [1]. SBSCL uses the COMBINE Archive Simulation Experiment Management for Systems Biology (SEMS) package to read and extract the required information from the OMEX files.

Hierarchical model simulations: The SBML extension package `comp` enables encoding complex and coupled biological systems that can be distributed or hierarchically structured. SBSCL v2.1 efficiently supports the simulation of this addition, including the automatic assembly of models from multiple and possibly remote input files.

Tests against benchmark suites: A crucial part of implementing new features is providing robust testing of the added functionality and use-cases. SBSCL tests all newly added features against the SBML Test Suite in a continuous integration approach. SBSCL provides full testing support against the genome-scale models from the BiGG Models database and kinetic models from the BioModels database.

3 Conclusion

The open-source library SBSCL simulates complex biological models in various frameworks specified in SBML format, optionally together with their *in silico* experiment definition SED-ML file or wrapped within OMEX archives. Benchmarks of SBSCL using the SBML Test Suite and a broad range of published models from relevant databases ensure its correctness and reliability (see supplement). With the support for exciting new features such as constraint-based model optimisation, hierarchical model decomposition, stochastic algorithms, archival input formats, this lightweight library is well suited as a simulation engine within any software with support for the Java Virtual Machine, e.g., Kotlin, Scala, or Groovy. The SBSCL project aims to provide a high-quality open-source simulation library to the scientific community to push frontiers and reproducibility in biology and related fields.

Acknowledgements

The authors acknowledge contributions by Nicolas Rodriguez, Alexander Dörr, Roland Keller, and the Harvey Mudd College for the ODEToolKit.

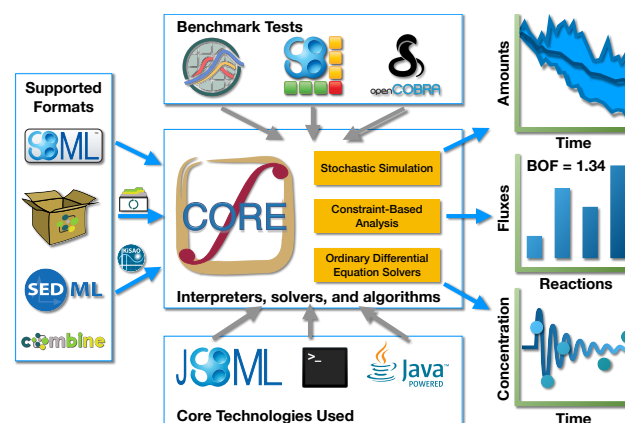


Fig. 1. The capabilities of the Systems Biology Simulation Core Library (SBSCL) as an overview. Supported input model definitions include SBML, possibly with an experiment configuration file (SED-ML) or bundled in a package file (OMEX). The model is parsed using the JSBML library, and solutions are numerically computed for the corresponding ODE or SDE system over time, following the specified constraints and algorithm (e.g., Rosenbrock, Euler, Gillespie) or via linear programming. Once the simulation completes, the model results are reported either graphically using a line plot or tabular form. The results can be exported to formats such as CSV for downstream use. For testing the library, its implementation, its robustness, reliability, and efficient reproducibility of the results, open model collections such as BiGG [6] and BioModels [5] are utilised, which comprise several hundred SBML models and their SED-ML configurations.

Funding: The National Resource for Network Biology (NRNB) and Google Inc. supported this work as part of their summer of code programs (GSoC). AD was funded by the German Center for Infection Research (DZIF), grant № 8020708703, and supported by infrastructural funding from the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation), Cluster of Excellence EXC 2124 Controlling Microbes to Fight Infections. MK is supported by the Federal Ministry of Education and Research (BMBF, Germany) within the research network Systems Medicine of the Liver (LiSyM, grant № 031L0054) and by the DFG within the Research Unit Programme FOR 5151 “QuaLiPerF (Quantifying Liver Perfusion–Function Relationship in Complex Resection—A Systems Medicine Approach)” by grant № 436883643.

Conflict of interest: None

References

- [1] Bergmann, F.T., Adams, R. et al. COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. *BMC Bioinformatics*, 15:369, 12 2014.
- [2] Erhard, F., Friedel, C.C. et al. FERN – a Java framework for stochastic simulation and evaluation of reaction networks. *BMC Bioinformatics*, 9:356, 8 2008.
- [3] Keating, S.M., Waltemath, D. et al. SBML Level 3: an extensible format for the exchange and reuse of biological models. *Mol. Syst. Biol.*, 16(8):e9110, 8 2020.
- [4] Keller, R., Dörr, A. et al. The systems biology simulation core algorithm. *BMC Systems Biology*, 7:55, 7 2013.
- [5] Malik-Sheriff, R.S., Glont, M. et al. BioModels—15 years of sharing computational models in life science. *Nucleic Acids Research*, 48:D407–D415, 1 2020.
- [6] Norsigian, C.J., Pesarla, N. et al. BiGG Models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree. *Nucleic Acids Res.*, 11 2019.
- [7] Rodriguez, N., Thomas, A. et al. JSBML 1.0: providing a smorgasbord of options to encode systems biology models. *Bioinformatics*, 6 2015.
- [8] Waltemath, D., Adams, R. et al. Reproducible computational biology experiments with SED-ML—the simulation experiment description markup language. *BMC Systems Biology*, 5(1):1–10, 11 2011.