BIOINFORMATICS APPLICATIONS NOTE Vol. 26 no. 13 2010, pages 1675–1676
doi:10.1093/bioinformatics/btq242

Systems biology

Advance Access publication May 5, 2010

SensSB: a software toolbox for the development and sensitivity analysis of systems biology models

Maria Rodriguez-Fernandez* and Julio R. Banga
(Bio)Process Engineering Group, IIM-CSIC C/Eduardo Cabello 6, 36208 Vigo, Spain

Associate Editor: Martin Bishop

ABSTRACT

Summary: SensSB (**Sens**itivity Analysis for **S**ystems **B**iology) is an easy to use, MATLAB-based software toolbox, which integrates several local and global sensitivity methods that can be applied to a wide variety of biological models. In addition to addressing the sensitivity analysis problem, SensSB aims to cover all the steps involved during the modeling process. The main features of SensSB are: (i) derivative and variance-based global sensitivity analysis, (ii) pseudo-global identifiability analysis, (iii) optimal experimental design (OED) based on global sensitivities, (iv) robust parameter estimation, (v) local sensitivity and identifiability analysis, (vi) confidence intervals of the estimated parameters and (vii) OED based on the Fisher Information Matrix (FIM). SensSB is also able to import models in the Systems Biology Mark-up Language (SBML) format. Several examples from simple analytical functions to more complex biological pathways have been implemented and can be downloaded together with the toolbox. The importance of using sensitivity analysis techniques for identifying unessential parameters and designing new experiments is quantified by increased identifiability metrics of the models and decreased confidence intervals of the estimated parameters.

Availability: SensSB is a software toolbox freely downloadable from http://www.iim.csic.es/~gingproc/SensSB.html. The web site also contains several examples and an extensive documentation.

Contact: mrodriguez@iim.csic.es

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on March 1, 2010; revised on April 28, 2010; accepted on April 29, 2010

1 INTRODUCTION

Sensitivity analysis (SA) is receiving an increasing attention in the field of Systems Biology (Sahle *et al.*, 2008) and several tools have been developed, namely SBML-SAT (Zi *et al.*, 2008), PottersWheel (Maiwald and Timmer, 2008), and Systems Biology Toolbox 2 (Schmidt and Jirstrand, 2006). However, neglecting any of the steps of the *model building cycle* as considered in the area of systems identification (Walter and Pronzato, 1997) can lead to wrong models or models with a low predictive capacity. That is why, in SensSB, the use of global optimization methods for parameter estimation (PE) and optimal experimental design (OED) has been coupled with other computational procedures based on SA for checking identifiability and other related measures. Another additional advantage is that the toolbox is able to handle deterministic non-linear dynamic models described by sets of ordinary differential equations (ODEs), differential algebraic equations (DAEs) or even black boxes.

Model building can be regarded as a cycle: starting from a goal definition (purpose of the model) and some a priori knowledge (e.g. preliminary data and initial hypothesis), a model framework is chosen and a model structure (system of equations) is proposed. The validity of this structure should be tested by means of an identifiability analysis (Srinath and Gunawan, 2010), e.g. making use of SA tools that can help to identify critical and negligible parameters and to establish a parameter ranking. If experimental data are available, PE is then carried out, leading to a first model. Otherwise a set of experiments must be designed and performed before the PE. The quality of these estimators should be assessed by checking the correlation between them and computing their confidence intervals. This initial model must be validated with new experiments, which in most cases will reveal a number of deficiencies. Thus, a new model structure and/or a new experimental design must be planned, and the process is repeated iteratively until the validation step is considered satisfactory.

All these tasks were implemented in MATLAB, generating gateways for calling external Fortran codes when needed. These Fortran codes were implemented as MATLAB MEX-files leading to an integrated environment able to perform robust PE, identifiability analysis and OED of dynamic experiments. Moreover, SensSB includes a novel methodology for performing OED based on global sensitivities with a great potential on the Systems Biology field where experiments can be very expensive and time consuming. In order to make the problem implementation easier, SensSB includes a user-friendly graphical interface where the model and other inputs (e.g. initial values, measurement times, control variables, etc.) can be specified. The MATLAB GUI (Graphical User Interface) generates an input file that SensSB uses for performing all the requested tasks as detailed below. Results are displayed on the screen and saved as an output file together with several graphic files.

2 SENSSB FEATURES

The main features of SensSB environment are depicted in Figure 1 and can be summarized as follows:

(i) Global sensitivity analysis (GSA): before any experimental data are obtained, a parameter ranking based on GSA provides insight about the model identifiability (Saltelli *et al.*, 2008). Negligible parameters (those with relatively low sensitivity measures) can be discarded from the model or fixed to literature values while special attention should be paid to the

*To whom correspondence should be addressed.

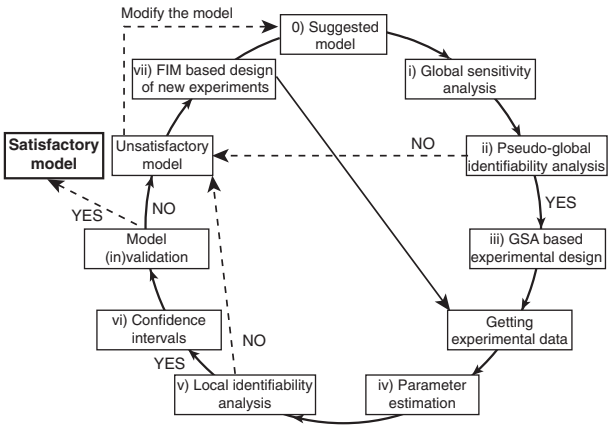

Fig. 1. SensSB covers all the steps involved on the modeling process.

highest ranked ones. It can also be used to identify the critical regions of the system and the most relevant state variables. By observation of the dynamics of the global sensitivities (also plotted by SensSB), the time span required for future experiments can be envisaged. Three GSA methods can be selected in SensSB: Sobol' indices, derivative-based global sensitivity measures (DGSM) and Morris indices.

(ii) Pseudo-Global a priori identifiability analysis: the computation of a pseudo-global correlation matrix based on simulated data is useful to identify highly correlated or redundant parameters that can be removed from the model with a minimum loss of information. Another matrix with the standard deviation of the correlations is also provided, giving information about the consistency of these correlations along the parameter space.

(iii) OED based on GSA: OED aims to identify the conditions and sampling schemes for a set of experiments that deliver data providing the maximum information with respect to the parameter identifiability and their estimation accuracy. No PE can be carried out before experimental data are collected, therefore, an OED based on local sensitivities has to be based on literature values for the parameters. Relying on these values can be misleading and so a GSA-based OED, consisting on averaging the selected scalar function of the Fisher Information Matrix (FIM) over the parameter space, is recommended. The features of the dynamic experiments that can be optimized through SensSB are: (i) initial concentration of the states, (ii) duration of the experiments, (iii) values for the control variables and their switching points and (vi) sampling points (i.e. when to measure the output variables).

(iv) Robust PE: once experiments are performed and data are collected, the aim of PE is to calibrate the model in order to fit the experimental data in the most accurate way. SensSB offers several global and local optimization methods that are able to handle noisy and partial measurements. Three global algorithms (`SSm`, `DE`, `SRES`) and two local ones (`fmincon`, `fminsearch`) can be selected. We strongly recommend the use of `SSm` as it has shown good efficiency and robustness in a wide number of benchmark problems (Rodriguez-Fernandez *et al.*, 2006). Depending on the information about the error of the experimental data, different objective functions can be selected, namely *least squares*, *weighted least squares* and *maximum likelihood*.

(v) Local sensitivity and identifiability analysis: after a first set of parameters has been identified, the computation of the parametric sensitivities, the *FIM*, its condition number and the correlation matrix is advisable. Particularly, for cases where the correlations are not very consistent along the parameter domain, this *a posteriori* local identifiability analysis will provide more precise information than the pseudo-global approach, always with the drawback of being only applicable to a specific set of parameters.

(vi) Confidence intervals: the computation of the 95% confidence intervals by means of the Cràmer-Râo method or the bootstrap approach offers an objective measure of the precision of the estimated parameters.

(vii) OED based on local SA: although a good option when the value of the parameters is unknown, OED based on GSA can be statistically pessimistic (considering sets of parameters that are unlikely to fit the data). Therefore, when the vicinity of the global optimum for the parameters has been found, local OED is preferred. Scalar functions of the *FIM* measuring the accuracy and/or decorrelation of the parameters are used as criteria.

An exhaustive description of the methodologies implemented in SensSB can be found in the software documentation.

3 CONCLUSIONS

The user-friendly interface together with the wide range of capabilities and the detailed documentation make SensSB ideal for helping the modeler along the integral modeling cycle. Several examples with all the necessary files to reproduce them are included in the toolbox and addressed in the documentation. These examples reveal GSA as a powerful tool for OED and for improving the accuracy of the estimated model parameters.

Funding: the EU ERASysBio and the the Spanish Ministry of Science and Innovation (SYSMO project KOSMOBAC, ref. MEC GEN2006-27747-E/SYS and MICINN project MultiSysBio ref. DPI2008-06880-C03-02).

Conflict of Interest: none declared.

REFERENCES

Maiwald,T. and Timmer,J. (2008) Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics*, **24**, 2037–2043.

Rodriguez-Fernandez,M. *et al.* (2006) Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. *BMC Bioinformatics*, **7**, 483.

Sahle,S. *et al.* (2008) A new strategy for assessing sensitivities in biochemical models. *Phil. Trans. R. Soc. A*, **366**, 3619–3631.

Saltelli,A. *et al.* (2008) *Global Sensitivity Analysis: The Primer*. John Wiley & Sons Ltd, England.

Schmidt,H. and Jirstrand,M. (2006) Systems biology toolbox for matlab: a computational platform for research in systems biology. *Bioinformatics*, **22**, 514–515.

Srinath,S. and Gunawan,R. (2010) Parameter identifiability of power-law biochemical system models. *J. Biotechnol.*, [Epub ahead of print, doi:10.1016/j.jbiotec.2010.02.019].

Walter,E. and Pronzato,L. (1997) *Identification of Parametric Models from Experimental Data*. Springer.

Zi,Z. *et al.* (2008) SBML-SAT: a systems biology markup language (SBML) based sensitivity analysis tool. *BMC Bioinformatics*, **9**, 342.