Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online
CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-26 (2016) 063–070

New Concepts for Evaluating the Performance of Computational Methods ⋆

Clemens Kreutz

Center for Biosystems Analysis (ZBSA),
Freiburg Center for Data Analysis and Modelling (FDM),
Albert-Ludwigs-University Freiburg
79104 Freiburg, Germany (e-mail: ckreutz@fdm.uni-freiburg.de).

Abstract: Research in Systems Biology is currently entering a new era. After a decade characterized by adopting existing experimental protocols and theoretical approaches to the requirements of Systems Biology, there is now a variety of tools and approaches available. However, many statistical and modeling concepts are not well-tested in application settings and their applicability is often seriously delimited. Therefore, a major challenge for the transfer of theoretical approaches to applications is the assessment and optimization of the methods' performance for supporting experimental research.
In this paper, new concepts for assessing methods which were developed for analyzing experimental data in the context of systems biology will be introduced. Some ideas are illustrated by evaluating the impact of the logarithmic transformation for parameter estimation. A strong benefit of the log-transformation was observed for five different ODE models. The suggested framework enables less biased and more reliable and valid assessment and comparison of competing approaches than currently performed in the literature. The presented concepts could serve as basis for developing decision guidelines for optimal selection of analysis methods and thereby enhancing the transfer of systems biological procedures and reverse engineering methods to industrial applications like drug development.

Keywords: Assessment, mathematical modelling, method comparisons, ordinary differential equations (ODEs), parameter estimation, performance, statistics, Systems Biology.

1. INTRODUCTION

Systems Biology is a strongly interdisciplinary field. Besides biological and experimental expertise, it requires expert knowledge about statistical and computational methods. Application of an inappropriate or suboptimal approach for analyzing data can lead to erroneous and misleading results and conclusions, thereby hampering scientific advance. Therefore, standards for the selection, assessment, and optimization of systems biological analyses would accelerate the progress in many applications. Such decision guidelines are an indispensable requirement for applying Systems Biology in clinical and industrial applications.

For annotation, storage, and documentation of data and models, standards are already available (Bruhn and Burton, 2003). As an example, for a broad range of experimental techniques, the *Bioscience reporting guidelines and tools* (MIBBI Project Consortium, 2008) and the so-called *Minimal Information* standards (Brazma et al., 2001), have been developed. As a further example, the *Systems Biology Markup Language (SBML)* has been established (SBML Forum, 2003) as a standard for definition and exchange of computational models of biological processes. However, standards and tests for the design of experiments, for the selection of appropriate techniques for statistical analysis and mathematical modeling as well as for assessing the quality and performance of measurements and their subsequent analyses are still missing. This is a major reason why it is currently not feasible to formulate precise statements about selection of appropriate methodology and about the expected performance. The absence of such reliable guidelines is a serious obstacle for the transfer of Systems Biology to pharmaceutical industry.

To fill this gap, some concepts and ideas can be adapted from related fields like bioinformatics or statistics. In clinical studies, as an example, a typical question is the comparison and assessment of two drugs or treatment strategies. The difficulty is that many other so-called nuisance or confounding effects like the patient's gender, age, genetic disposition, or habits like smoking, etc. are not of primary interest, but have a remarkable impact on the efficiency of treatment strategies. The statistical methodology to design and analyze studies in such settings can be adopted to a considerable extent for the purpose of method comparison studies in Systems Biology. Here, beside the effect of interest, namely the impact of the decision between two alternative analysis methods, also a large number of nuisance or confounding effects exist

⋆ The author acknowledge financial support by the by the German Ministry of Education and Research (BMBF) via e:Bio Grant No. 031L0080.

10.1016/j.ifacol.2016.12.104

because the performance of a method also depends on a variety of other factors and decisions made for the analysis, e.g. selection of an optimization method, stopping rule or tolerances controlling the precision of ODE integration.

The scientific progress in the Systems Biology exhibits many parallels to the development of methodology and standards in bioinformatics, e.g. for the analysis of omics data. As an example, the development of gene expression microarrays as one of the first modern omics technique initially attracted a large fraction of researchers in statistics and computer science. Similar challenges and perspectives as they occur in Systems Biology also led to a huge variety of tools and approaches, each of them published with certain justification for its necessity or superiority.

Because of missing guidelines and gold standards for data analysis, it had been realized that for microarrays comprehensive method evaluation studies were required to be able to reliably assess the quality of the data and to establish mature standards for data processing and analysis. Since such standards and quality assessments are indispensable for applications in clinical and pharmacological practice, the FDA, as an example, launched the three huge *Microarray Quality Control (MAQC)* funding initiatives (MAQC Consortium, 2006; MAQC-II Consortium, 2010; SeQC/MAQC-III Consortium, 2014) for assessment of methodology and creation of analysis standards.

Two major results of these studies which also hold for Systems Biology are that (1) some approaches exhibit superiority over other methods in very general settings and are therefore suitable as gold standard for analysis. However, (2) there is no unique optimal analysis strategy for all circumstances since the performance depends on characteristics of the data, e.g. on effect size, biological variability, size of the models and data sets. Statements of both kinds as well as the identification of broadly applicable analysis gold standards and decisive properties of the data for the choice between analysis methods are major requirements in Systems Biology and are provided by the presented methodology.

2. GENERAL METHODOLOGY

The setting considered in this paper is comparing the performance of computational approaches. Although the concept is not restricted to a specific field of application, the discussions in the following focus on a typical setting in Systems Biology, namely establishing an ODE-based model of a biological process based on experimental data as it is frequently done for cellular signalling pathways, metabolic-, or gene reguatory networks. In this section, general aspects for reliable method assessment studies are introduced and discussed.

2.1 Experimental design

In applications, a first step in planning experiments and subsequent analysis is selecting informative experimental conditions for addressing the open questions. This comprise selection of measurement techniques, observed compounds, treatments, measurement times, the whole cell preparation procedure, as well as parallel quantification of standards, normalizers, positive or negative controls

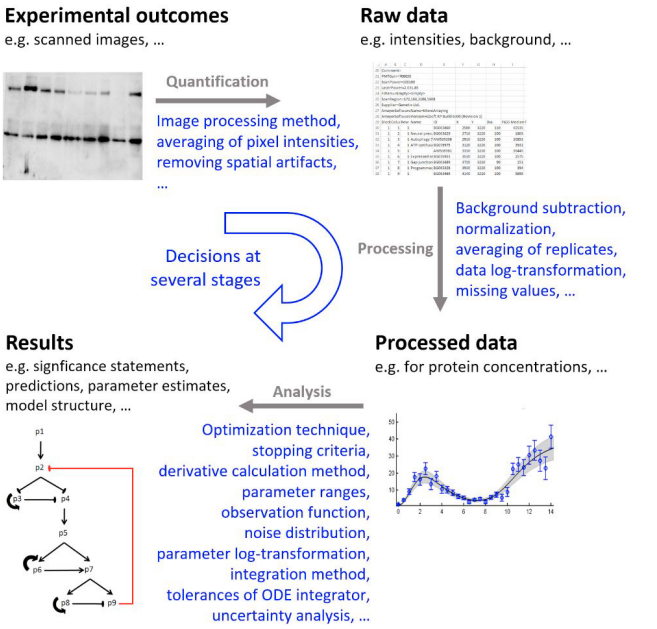

Fig. 1. The setup of the data analysis workflow from experimental data generation to fitting and interpreting a model involves many decisions. Each alternative data processing step could possibly have an impact on the outcome of the whole analysis pipeline.

(Kreutz and Timmer, 2009). Although in practice, experimental planning is often tailored to technical feasibility, remaining experimental design decisions are highly important in biomedical research because experiments with living cells have large variability and large risk for systematic errors.

The experimental design is also of importance for data analysis because the performance of the analysis depends on the information provided by the experimental design, and, vice versa, the demand on a design depends on the requirements of subsequent analysis. Moreover, assessing the performance of algorithms provides only valuable results, if the analyzed data provides meaningful information. Here, it is assumed that the analysis strategy has already been chosen to such a level of detail that also the experimental design could be selected reasonably and independently from the exact choice between competing algorithms whose performance is demanded. This means that performance assessment involving both, selection of the analysis method and combination with a specifically tailored experimental design is beyond the scope of this article.

2.2 Decisions for data analysis

For improvement of an analysis workflow, it has to be evaluated which decisions, or, which alternative strategies of analyzing the data have an impact on the quality of the results. Fig. 1 shows, as an example, several decisions which have to be made if western blot data is processed and used for fitting a dynamical model. Each decision could possibly affect the performance of the comprehensive analysis and has potentially to be considered for optimization of the entire analysis. Then, a comprehensive evaluation and optimization of such an analysis pipeline