Model calibration and uncertainty analysis in signaling networks

Tim Heinemann and Andreas Raue

For a long time the biggest challenges in modeling cellular signal transduction networks has been the inference of crucial pathway components and the qualitative description of their interactions. As a result of the emergence of powerful high-throughput experiments, it is now possible to measure data of high temporal and spatial resolution and to analyze signaling dynamics quantitatively. In addition, this increase of high-quality data is the basis for a better understanding of model limitations and their influence on the predictive power of models. We review established approaches in signal transduction network modeling with a focus on ordinary differential equation models as well as related developments in model calibration. As central aspects of the calibration process we discuss possibilities of model adaptation based on data-driven parameter optimization and the concomitant objective of reducing model uncertainties.

Address
Merrimack, One Kendall Sq., Suite B7201, Cambridge, MA 02139, USA

Corresponding author: Raue, Andreas (araue@merrimack.com)

Introduction

Cellular decision making is governed by a complex interplay of signal transduction and gene regulation networks. The genome is well-researched and provides a comprehensive list of network components that contribute to this interplay. In addition large scale transcriptome and proteome data characterizing the abundance of each of these components is available. However due to the complexity of the possible interactions and crosstalk between different signaling pathways, purely experimental approaches are limited in depicting the functional behavior emerging from those signaling networks. In recent years computational models have been used to help describe signaling networks and their behavior more quantitatively [1–3].

Building and calibrating computational models is a challenging task in itself. Many aspects of signaling networks, such as the possible interactions between the parts/proteins/nodes in the network are not well described or discussed controversially in literature. Depending on the biological context, signaling networks can rewire and change their function. Therefore a computational scientist has to rely on well curated literature as well as experimental data acquired in the particular biological context of the study to complete two tasks: firstly, obtaining the appropriate network structure, and secondly, calibrating model parameters such as the strength of protein–protein interactions or protein abundances to the biological context of interest. Depending on the biological question and on the available data, different mathematical models can be employed, each with particular strengths and weaknesses [4]. Given the chosen model structure and calibration, the next step is to validate the model with further experimental data and/or literature. If the model does not perform well, assumptions, model structure and parameterization have to be revised. Otherwise, it is important to analyze the degree to which the model is constrained by the available data/literature and the extent to which model predictions can be trusted. If the model is insufficiently constrained or predications are uncertain, experimental design techniques can be used to plan additional experiments that will provide new data with optimal information content. Figure 1 illustrates how these fundamental tasks are repeated until the biological questions can be answered with sufficient statistical support.

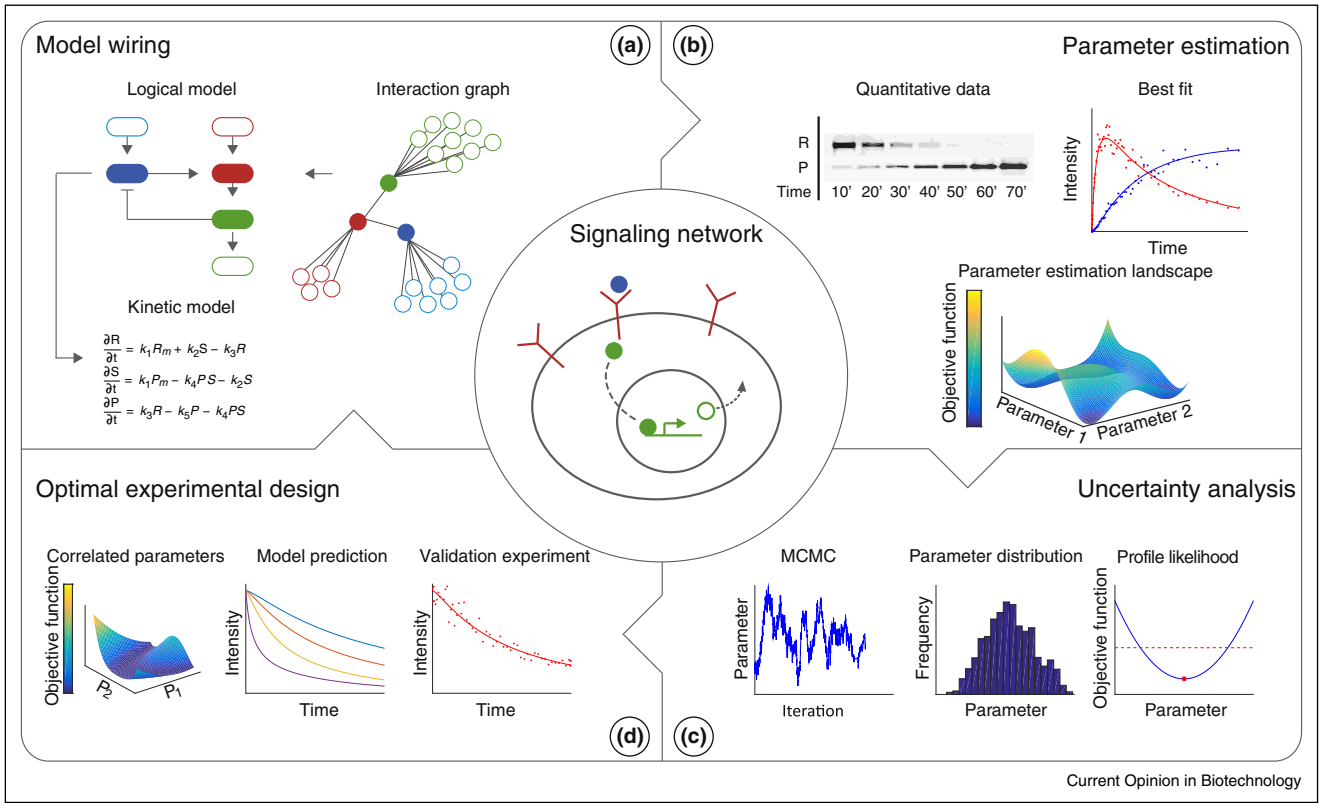
We will review model calibration and uncertainty analysis approaches used for the modeling of signaling networks.

Modeling strategies for signaling networks

The spectrum of modeling approaches for signaling networks covers a broad range of abstraction levels. The choice of a particular model-class is often predetermined by the system's complexity, by the type and quality of available biological knowledge and by the particular biological question that should be addressed.

For integrating information of larger signal transduction networks, in particular in cases where detailed mechanistic data on the pathway components is missing, logic-based models are powerful approaches [5,6]. In various recent attempts, fragmentary qualitative knowledge of regulatory interactions was successfully integrated using Boolean modeling, leading to predictive mathematical models suitable for network analysis and even to simulation of the network dynamics [7–9]. Despite these advances in Boolean network modeling, the predictive power of those

Figure 1



Model calibration for cell signaling networks. **(a)** Depending on the available data and prior knowledge, the signaling pathway can be translated into mathematical models with different abstraction levels. The corresponding model classes range from undirected interaction graphs over direction-specific logical models to quantitative ODE models. **(b)** Time resolved quantitative data allows for a strong interlinking of the signaling dynamics and their mathematical representation in the model. A carful choice of the objective function is critical for a successful fitting result. **(c)** The quality of the model fit and hence model predictions can be affected by structural and practical non-identifiabilities due to limitations in the model wiring or the data. A variety of numerical approaches for the analysis of model uncertainties are available, that is, Markov-chain Monte Carlo (MCMC) sampling of the posterior distribution or likelihood based methods. **(d)** Uncertainties in model parameters lead to ill-defined model predictions. These, however, may give rise to dynamic ranges where additional experimental validation would significantly improve model calibration.

models is intrinsically limited because it casts biology into a binary format. For a more adequate rendering of biological behavior, signaling networks and their underlying reactions can be translated into algebraic equations describing protein concentration over time and space. In this setting, mechanistic models consisting of ordinary differential equations (ODE) are often used. They rely on quantitative data but, if applied successfully, produce realistic predictions of the system's dynamics [10]. In the context of signal-transduction networks, mechanistic modeling provides powerful tools for the analysis of regulatory motifs governing the relationship between stimuli and responses such as feedback mechanisms, cycle and cascade motifs or spatial gradients and their effects on location-specific signaling [3]. If signaling is directly influenced by random fluctuations in molecular numbers or due to the underlying discreteness of some of the subsystems, stochastic effects need to be considered explicitly. Over the last decades addressing this aspect contributed to the development of multiple exact, approximate and hybrid stochastic simulation methods [11], which were successfully used to model experimental data [12].

We will focus on model calibration techniques that are used in conjunction with mechanistic models. Box 1 gives a short primer on the mathematics behind this type of model.

Model calibration by parameter estimation

Many models, in particular mechanistic models, will contain parameters which values are unknown. These parameters will introduce uncertainty in the model behavior and thus have to be calibrated. Estimating unknown parameters is essential for determining if a model appropriately fits the available experimental data and is in line with prior knowledge. In some cases the model

Box 1 Introduction to mechanistic modeling

Network dynamics in space and time can conveniently be modeled by ordinary or partial differential equations [58]. We will briefly introduce the former approach here using a simple binding reaction: $A + B \leftrightarrow C$. The reaction is governed by five parameters, the association rate k_{on}, the dissociation rate k_{off} and the initial concentrations A_0, B_0 and C_0. Assuming a well-mixed environment, spatial distribution can be neglected and mass action kinetics define the reaction flux of the forward $\boldsymbol{v}_1 = \boldsymbol{k_{on}} \cdot \boldsymbol{A} \cdot \boldsymbol{B}$ and backward $\boldsymbol{v}_2 = \boldsymbol{k_{off}} \cdot \boldsymbol{C}$ reactions. The ODE for the concentration dynamics of this system are:

$$\frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{A} = -\boldsymbol{k_{on}} \cdot \boldsymbol{A} \cdot \boldsymbol{B} + \boldsymbol{k_{off}} \cdot \boldsymbol{C} \quad \text{with} \quad \boldsymbol{A}(\boldsymbol{t} = 0) = \boldsymbol{A}_0$$

$$\frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{B} = -\boldsymbol{k_{on}} \cdot \boldsymbol{A} \cdot \boldsymbol{B} + \boldsymbol{k_{off}} \cdot \boldsymbol{C} \quad \text{with} \quad \boldsymbol{B}(\boldsymbol{t} = 0) = \boldsymbol{B}_0$$

$$\frac{\boldsymbol{d}}{\boldsymbol{dt}}\boldsymbol{C} = +\boldsymbol{k_{on}} \cdot \boldsymbol{A} \cdot \boldsymbol{B} - \boldsymbol{k_{off}} \cdot \boldsymbol{C} \quad \text{with} \quad \boldsymbol{C}(\boldsymbol{t} = 0) = \boldsymbol{C}_0$$

The general structure of this type of model is $(\boldsymbol{d}/\boldsymbol{dt})\boldsymbol{x} = \boldsymbol{N} \cdot \boldsymbol{v}(\boldsymbol{x}, \boldsymbol{\theta})$ with initial conditions $x(t = 0) = f_0(\boldsymbol{\theta})$. Here, x are modeled species such as proteins, $\boldsymbol{\theta}$ model parameters such as reaction rates or initial concentrations, $\boldsymbol{N}$ is the stoichiometric Matrix and $\boldsymbol{v}$ are the reaction rate equations such as mass action kinetics, Hill or Michaelis–Menten kinetics. The above equation system can be solved using standard numerical solver packages, a range of specialized software for signaling networks [59–61] and commercial tools such as the SimBiology toolbox for MATLAB (The Mathworks Inc., Natick MA, USA).

wiring can be encoded in terms of the model parameters. Therefore the optimal network wiring can be inferred as well. For instance, in an ODE model binding affinities encode reactions of proteins binding to each other. Only if the association rate between two proteins is larger than zero is there an effective link in the network between the two. Here we will deal with the generic problem of finding a set of parameters that lead to the best agreement of model and data. Box 2 highlights some of the related mathematical concepts.

For estimating unknown parameters, an objective function that is dependent on the unknown parameters can be formulated. It ensures that the model is calibrated to all data simultaneously. The set of parameters that minimizes the objective function provides the best possible fit of model and data and can be found using numerical procedures. A simple and commonly used objective function is the sum-of-squared-residuals. A residual is the difference between a model output and its corresponding experimentally observed value. Model outputs can for instance be concentrations of proteins at a particular point in time, or sums or ratios of those concentrations. In general, every observable quantity described by the model can be an output, however for signaling models the experimentally accessible outputs are often limited.

Box 2 Mathematical concepts in parameter estimation

Model output: $y_i(\boldsymbol{\theta})$

Corresponding experimental observation: $\boldsymbol{y_i^d}$

Residual: $\boldsymbol{r_i}(\boldsymbol{\theta}) = \boldsymbol{y_i^d} - \boldsymbol{y_i}(\boldsymbol{\theta})$

Sum-of-squared-residuals: $\boldsymbol{SSR}(\boldsymbol{\theta}) = \sum_{\boldsymbol{i}} \boldsymbol{r_i}(\boldsymbol{\theta})^2$

Standard deviation of experimental observation: $\boldsymbol{\sigma_i}$

Weighted sum-of-squared-residuals: $\boldsymbol{WSSR}(\boldsymbol{\theta}) = \sum_{\boldsymbol{i}} \frac{\boldsymbol{r_i}(\boldsymbol{\theta})^2}{\boldsymbol{\sigma_i}^2}$

Measurement noise model: $\boldsymbol{\sigma_i}(\boldsymbol{\theta})$

Likelihood for normally distributed measurements: $\boldsymbol{L}(\boldsymbol{\theta}) = \prod_{\boldsymbol{i}} \frac{1}{\boldsymbol{\sigma_i}(\boldsymbol{\theta})\sqrt{2\pi}} \boldsymbol{e}^{-(\boldsymbol{r_i}(\boldsymbol{\theta})^2 / 2\boldsymbol{\sigma_i}(\boldsymbol{\theta})^2)}$

Prior probability distribution for jth parameter: $P(\boldsymbol{\theta_j})$

Posterior probability distribution: $\boldsymbol{P}(\boldsymbol{\theta}|\boldsymbol{y^d}) \propto \boldsymbol{L}(\boldsymbol{\theta}) \cdot \prod_{\boldsymbol{j}} \boldsymbol{P}(\boldsymbol{\theta_j})$

By weighting the residuals by the standard deviation of a measurement, a slightly more advanced objective function also takes the quality of experimental data into account. This approach facilitates the use of more general measurement noise models. The noise models can be determined by independent methods [13] or by a joint estimation together with the signaling model [14••]. Assuming normally distributed measurements sum-of-squared-residuals objective functions are special cases of a much more general type of objective function called likelihood. For cases where the appropriate probability distribution of the measurements is unknown or multiple objectives have to be optimized simultaneously the method of moments and set-based methods have their respective advantages [15–17]. Here, we concentrate on the likelihood, which is the conditional probability of the experimental data given the model parameters [18]. The optimal parameter values can be ascertained by maximization of the likelihood, which is referred to as maximum-likelihood estimation. For practical reasons the objective function is chosen to be minus the logarithm of the likelihood so that the resulting objective function can again be minimized. As the amount of appropriately selected data increases, the maximum-likelihood estimator of the parameters converges to the true parameter values [19]. In addition, no other asymptotically unbiased estimator has lower asymptotic mean squared error. The likelihood also provides the most general way of dealing with measurement noise models. When the distribution of the measurement noise is different from normal (e.g. log-normal), distribution parameters (e.g. the variance) can be estimated together with the remaining model parameters [14••].

Prior knowledge from literature can provide important additional information to calibrate a model properly. It can come in many different forms, such as previously reported parameter values and their estimation uncertainties, known facts about protein concentrations at a certain point in time or general plausibility assumptions on the model dynamics. Following Bayes' theorem, combining

the likelihood with the prior probability distribution yields the posterior probability distribution. Like in maximum-likelihood estimation, the maximum a posteriori probability estimate of the parameters can be determined.

As opposed to studying the mean behavior of the signaling network, one might be interested in creating and calibrating signaling networks across heterogeneous populations. Modeling signaling networks of heterogeneous populations such as single cells or different cancer cell lines is an especially challenging task. One mathematical concept to deal with this is known as mixed-effect modeling [20]. It can potentially be used with different types of signaling models but is especially useful for mechanistic models. Mixed-effect modeling distinguishes between fixed effects and random effects. The former affects all instances of a signaling model in the same way. The latter introduces a certain amount of randomness across different instances of the signaling model that follows a particular distribution assumption. Technically, independent instances of the original model are initialized, with individual parameters for the random effects. The random effect parameters are then constrained by assuming a certain distribution of their values. The distribution constraint is added to the objective function similar to prior knowledge. As an example, assuming that the biochemical properties are conserved across the population, the reaction rate constants might be modeled as fixed effects. Total protein concentrations on the other hand could fluctuate across the population. They could be modeled as random effects following a log-normal distribution. Mixed-effect modeling is not often applied in modeling of signaling networks but has a very strong foundation in pharmacokinetic modeling [21].

Parameter estimation by numerical optimization

After the objective function has been specified, numerical optimization algorithms are applied to minimize the objective function. For signaling networks this can however be a difficult task. All of the discussed models are non-linear and may contain a large number of unknown parameters. At the same time, the available prior knowledge and experimental data is often limited. As a result, the objective function can have multiple (local) minima, for illustration see the parameter estimation landscape in Figure 1b. In the remainder of this section we will focus on the task of finding a set of parameters that minimizes the objective function. The implications of multiple minima on model uncertainty will be discussed in the next section.

Analytical solutions of this optimization problem are not feasible and numerical solvers that iteratively refine solutions are applied. The corresponding numerical methods can be subdivided into deterministic approaches, stochastic approaches or hybrid approaches. Benchmark properties of all of these methods are how quickly they converge to a solution and how reliably the best solutions can be obtained if the method is repeated multiple times.

Starting with an initial guess of the parameters, deterministic approaches make use of derivative information to iteratively reduce the value of the objective function. The simplest approach is known as gradient descent method. It uses only first order derivatives but has poor convergence. The well-known Newton's method makes use of second order derivatives as well, thus improving convergence [22]. In many scenarios it is advantageous to be able to switch between gradient decent and Newton's method. Examples are the Levenberg–Marquardt algorithm or more recently published trust-region approaches [23]. For non-linear models, one fundamental limitation of deterministic approaches is that they can get stuck in local minima. Therefore, these approaches have to be combined with stochastic elements, such as multiple starts from randomized initial parameters. For the calculation of the derivatives of the objective function with respect to the parameters finite difference approximation should not be used, because it leads to inaccurate derivatives and poor estimation performance. The methods of choice for derivative calculation are the sensitivity equations [14••] or complex step methods [24].

Stochastic approaches apply heuristic search strategies to find solutions. They are designed to circumvent convergence to local minima. Popular tools in this category include genetic algorithm, differential evolution or particle swarm optimization. Methods in this class of algorithms typically do not make use of derivative information. This simplifies their implementation but comes at the cost of considerable convergence disadvantages [14••]. Finally, hybrid algorithms have been proposed. They initially make use of stochastic algorithms before switching over to deterministic strategies; as, for instance, in the case of the scatter search algorithm [25]. All of the optimization methods mentioned above are limited to continuous optimization variables. More general approaches such as the software suite MEIGO [26•] can also account for discrete variables.

A quality control for model fitting has been devised that is based on the analysis of the recurrence of multiple independent runs of parameter estimation [14••]. In this quality control, using two benchmark signaling networks of different complexity, a deterministic multi-start trust region method outperformed all other tested optimization methods by orders of magnitude.

Uncertainty in model calibration and prediction

The previous two sections discussed strategies and challenges for finding the so-called best model fit and the corresponding set of optimal parameter values. If model behavior will be simulated, it is important to determine

how much uncertainty is contained in the model prediction.

Experimental uncertainties are propagated to uncertainties in the estimated parameters, which in turn are propagated to the model predictions. Depending on the amount and quality of the experimental data as well as of the prior assumptions, the set of optimal parameters might not be uniquely defined [27••]. Besides the possible existence of local minima, structural non-identifiability can arise. It indicates that due to the structure of the model, there may be infinitely many sets of parameters that fit the data equally well. Various different methods are available to detect structural non-identifiability for signaling networks [28]. Some model parameters, although structurally identifiable, might still be practically non-identifiable. This means that there is a unique set of best fit parameters but the parameters can still not be estimated with finite uncertainty and are hence still undetermined [29].

Uncertainty in parameter estimates is usually quantified in terms of confidence intervals. For signaling models and non-linear models in general, asymptotic confidence intervals based on the Fischer information matrix are not reliable [29]. A method that is based on likelihood profiles has been proposed. By using this method, both structural and practical non-identifiability may be determined and confidence intervals can be calculated [29]. Alternatively, Bootstrap or Monte Carlo methods can be used to sample from the likelihood or the posterior probability distribution [30,31,32••]. The reliability of bootstrapping-based confidence intervals in the presence of non-identifiabilities was recently studied [33]. Uncertainty in model predictions can also be quantified directly, without the need to quantify uncertainty in model parameters first [34,35]. The results of uncertainty analyses can be used to design novel experiments tailored to reduce specific uncertainties in the model calibration [36,37].

A complementary approach for the further reduction of uncertainties in model calibration is termed optimal experimental design. It involves the systematic adaptation of experimental protocols in order to restrict the range of possible parameter values [38]. This concept extends several qualitative aspects of thorough experimental planning, such as the minimization of sources of technical variability due to inhomogeneity of the experimental units, or the determination of the number of reasonable replications that reveal the system's intrinsic variability [39]. Optimal experimental design comprises a number of numerical optimization methodologies designed to minimize properties of the parameter covariance matrix, which describes the uncertainty region in the parameter space close to a particular solution of the dynamical system. A recent example of how this theory can be successfully applied in a cell signaling context can be found elsewhere [40•].

Conclusion

Building useful models of signaling networks on the cellular level is a challenging task considering the high degree of complexity inherent to these pathways and the limitations of experimental techniques. The key steps consist of the inference of the core signal transmitters and modulators; a careful choice of the level of detail that will be considered in the model; the translation of presumed physical interactions into a mathematical model; the acquisition of highly resolved quantitative data as readouts for the signaling dynamics; and, as we discussed in this review, the calibration of the model to maximize its predictive power. A proper balance between those steps is essential, otherwise a poor fitting result could disqualify a particular model unnecessarily [41]. At the same time, a well-balanced setup should restrict regions in the parameter space where the model behavior is insensitive to changes of the parameter values [42].

As opposed to the deterministic ODE models discussed here, the methodology for other models such as partial differential equation, stochastic [43] or agent-based models [44] is less well established. Stochasticity of the model output is a considerable problem for many of the commonly used methods that are based on objective functions. A trivial but computationally expensive solution is to average over many independent stochastic model simulations. There is a growing amount of literature on the use of Bayesian approaches for the inference of stochastic models [45•]. To circumvent the aforementioned issues related to computational cost, several recent attempts try to approximate stochastic models [46,45•,46–49]. Instead of taking the full distribution of the measured population into account, these methods use only lower-order moments and thereby sacrifice some of the underlying information. However, it was shown that in some cases already mean and variance are sufficient for parameter inference even if the measured distributions are not well determined by the lower-order moments [46,50]. In addition to the simultaneous inference of the model parameters and their confidence bounds, Bayesian methods allow for the dissection of intrinsic, extrinsic and technical noise [51,52].

When hybrid models or large multi-scale models [53–56] are considered, reliable model calibration gets even more complicated [4]. A recent study showed that currently available methods do not perform better than random in these settings [57]. We conclude that there is a large body of well-established methods for medium-scale mechanistic modeling of signaling networks.

Acknowledgments

We thank Carla Roth and Dr. Wendy Qiao for reviewing the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Cho K-H, Wolkenhauer O: **Analysis and modelling of signal transduction pathways in systems biology**. *Biochem Soc Trans* 2003, **31**:1503-1509.

2. Gunawardena J: **Signals and systems: towards a systems biology of signal transduction**. *Proc IEEE* 2008:96.

3. Kholodenko BN: **Cell-signalling dynamics in time and space**. *Nat Rev Mol Cell Biol* 2006, **7**:165-176.

4. Hasenauer J, Jagiella N, Hross S, Theis FJ: **Data-driven modelling of biological multi-scale processes**. *J Coupled Syst Multiscale Dyn* 2015 http://dx.doi.org/10.1166/jcsmd.2015.1069.

5. Albert R, Wang R-S: **Discrete dynamic modeling of cellular signaling networks**. *Methods Enzymol* 2009, **467**:281-306.

6. Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, Klamt S, Sorger PK: **Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction**. *Mol Syst Biol* 2009, **5**:331.

7. Samaga R, Klamt S: **Modeling approaches for qualitative and semi-quantitative analysis of cellular signaling networks**. *Cell Commun Signal* 2013, **11**:43.

8. Saadatpour A, Wang R-S, Liao A, Liu X, Loughran TP, Albert I, Albert R: **Dynamical and structural analysis of a T cell survival network identifies novel candidate therapeutic targets for large granular lymphocyte leukemia**. *PLoS Comput Biol* 2011, **7**:e1002267.

9. Saez-Rodriguez J, Alexopoulos LG, Zhang M, Morris MK, Lauffenburger DA, Sorger PK: **Comparing signaling networks between normal and transformed hepatocytes using discrete logical models**. *Cancer Res* 2011, **71**:5400-5411.

10. Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK: **Physicochemical modelling of cell signalling pathways**. *Nat Cell Biol* 2006, **8**:1195-1203.

11. Pahle J: **Biochemical simulations: stochastic, approximate stochastic and hybrid approaches**. *Brief Bioinform* 2009, **10**:53-64.

12. Wilkinson DJ: **Stochastic modelling for quantitative description of heterogeneous biological systems**. *Nat Rev Genet* 2009, **10**:122-133.

13. Kreutz C, Bartolome Rodriguez MM, Maiwald T, Seidl M, Blum HE, Mohr L, Timmer J: **An error model for protein quantification**. *Bioinformatics* 2007, **23**:2747-2753.

14. •• Raue A, Schilling M, Bachmann J, Matteson A, Schelke M, Kaschek D, Hug S, Kreutz C, Harms BD, Theis FJ *et al.*: **Lessons learned from quantitative dynamical modeling in systems biology**. *PLoS ONE* 2013, **8**:e74335.

The paper introduces novel conceptual approaches to deal with model calibration challenges and presents benchmarks for parameter estimation algorithms.

15. Anandkumar A, Hsu D, Kakade SM: **A method of moments for mixture models and Hidden Markov Models**.**Twenty-Fifth Annual Conference on Learning Theory**. 2012:1-31: arXiv:1203.0683.

16. Lillacci G, Khammash M: **A distribution-matching method for parameter estimation and model selection in computational biology**. *Int J Robust Nonlinear Control* 2012, **22**:1065-1081.

17. Zitzler E, Thiele L, Bader J: **On set-based multiobjective optimization**. *IEEE Trans Evol Comput* 2010, **14**:58-79.

18. Fisher RA: **On the mathematical foundations of theoretical statistics**. *Philos Trans R Soc A Math Phys Eng Sci* 1922, **222**: 309-368.

19. Radhakrishna Rao C: **Information and accuracy attainable in the estimation of statistical parameters**. *Bull Calcutta Math Soc* 1945, **37**:81-91.

20. Baayen RH, Davidson DJ, Bates DM: **Mixed-effects modeling with crossed random effects for subjects and items**. *J Mem Lang* 2008, **59**:390-412.

21. Huang X, Li J: **Pharmacokinetic–pharmacodynamic modeling and simulation**. *Am J Pharm Educ* 2007, **71**:129-145.

22. Meza JC: **Newton's method**. *Wiley Interdiscip Rev Comput Stat* 2011, **3**:75-78.

23. Yuan Y: **A review of trust region algorithms for optimization**. *ICIAM* 2000, **99**:271-282.

24. Martins JRRA, Sturdza P, Alonso JJ: **The complex-step derivative approximation**. *ACM Trans Math Softw* 2003, **29**:245-262.

25. Egea Ja, Rodríguez-Fernández M, Banga JR, Martí R: **Scatter search for chemical and bio-process optimization**. *J Glob Optim* 2007, **37**:481-503.

26. • Egea JA, Henriques D, Cokelaer T, Villaverde AF, MacNamara A, Danciu D-P, Banga JR, Saez-Rodriguez J: **MEIGO: an open-source software suite based on metaheuristics for global optimization in systems biology and bioinformatics**. *BMC Bioinformatics* 2014, **15**:136.

27. •• Vanlier J, Tiemann CA, Hilbers PAJ, van Riel NAW: **Parameter uncertainty in biochemical models described by ordinary differential equations**. *Math Biosci* 2013, **246**:305-314.

This paper presents a detailed comparison between different approaches to uncertainty analysis, in particular between frequentist and Bayesian approaches.

28. Raue A, Karlsson J, Saccomani MP, Jirstrand M, Timmer J: **Comparison of approaches for parameter identifiability analysis of biological systems**. *Bioinformatics* 2014, **30**:1440-1448.

29. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J: **Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood**. *Bioinformatics* 2009, **25**:1923-1929.

30. Hug S, Raue A, Hasenauer J, Bachmann J, Klingmüller U, Timmer J, Theis FJ: **High-dimensional Bayesian parameter estimation: case study for a model of JAK2/STAT5 signaling**. *Math Biosci* 2013, **246**:293-304.

31. Eydgahi H, Chen WW, Muhlich JL, Vitkup D, Tsitsiklis JN, Sorger PK: **Properties of cell death models calibrated and compared using Bayesian approaches**. *Mol Syst Biol* 2013, **9**:644.

32. •• Liepe J, Kirk P, Filippi S, Toni T, Barnes CP, Stumpf MPH: **A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation**. *Nat Protoc* 2014, **9**:439-456.

This study presents the approximate Bayesian computation framework ABC-SysBio. It is particularly suitable for the estimation and discrimination of stochastic networks.

33. Fröhlich F, Theis FJ, Hasenauer J: **Uncertainty analysis for non-identifiable dynamical systems: profile likelihoods, bootstrapping and more**. *Comput Methods Syst Biol* 2014 http://dx.doi.org/10.1007/978-3-319-12982-2_5.

34. Kreutz C, Raue A, Timmer J: **Likelihood based observability analysis and confidence intervals for predictions of dynamic models**. *BMC Syst Biol* 2012, **6**:120.

35. Vanlier J, Tiemann CA, Hilbers PAJ, van Riel NAW: **An integrated strategy for prediction uncertainty analysis**. *Bioinformatics* 2012, **28**:1130-1135.

36. Raue A, Becker V, Klingmüller U, Timmer J: **Identifiability and observability analysis for experimental design in nonlinear dynamical models**. *Chaos* 2010, **20**:045105.

37. Steiert B, Raue A, Timmer J, Kreutz C: **Experimental design for parameter estimation of gene regulatory networks**. *PLoS ONE* 2012, **7**:1-11.

38. Hagen DR, White JK, Tidor B: **Convergence in parameters and predictions using computational experimental design**. *Interface Focus* 2013, **3**:20130008.

39. Kreutz C, Timmer J: **Systems biology: experimental design**. *FEBS J* 2009, **276**:923-942.

40. • Bandara S, Schlöder JP, Eils R, Bock HG, Meyer T: **Optimal experimental design for parameter estimation of a cell signaling model**. *PLoS Comput Biol* 2009, **5**:e1000558.
This work presents a OED strategy exemplary on a PIP3 signaling process. It suggests to iteratively minimize the trace of the parameter covariance matrix by optimized experiments in order to confine parameter confidence regions.

41. Banga JR: **Optimization in computational systems biology**. *BMC Syst Biol* 2008, **2**:47.

42. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP: **Universally sloppy parameter sensitivities in systems biology models**. *PLoS Comput Biol* 2007, **3**:1871-1878.

43. Gillespie DT, Gillesple DT: **Exact Stochastic Simulation of couple chemical reactions**. *J Phys Chem* 1977, **81**:2340-2361.

44. Hoehme S, Drasdo D: **A cell-based simulation software for multi-cellular systems**. *Bioinformatics* 2010, **26**:2641-2642.

45. • Villaverde AF, Banga JR: **Reverse engineering and identification in systems biology: strategies, perspectives and challenges**. *J R Soc Interface* 2014, **11**:20130505.
This work reviews convergent ideas concerned with the identification of interaction structures in biological systems and translating these into dynamical models.

46. Zechner C, Ruess J, Krenn P, Pelet S, Peter M, Lygeros J, Koeppl H: **Moment-based inference predicts bimodality in transient gene expression**. *Proc Natl Acad Sci U S A* 2012, **109**:8340-8345.

47. Golightly A, Wilkinson DJ: **Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo**. *Interface Focus* 2011, **1**:807-820.

48. Toni T, Ozaki Y, Kirk P, Kuroda S, Stumpf MPH: **Elucidating the in vivo phosphorylation dynamics of the ERK MAP kinase using quantitative proteomics data and Bayesian model selection**. *Mol Biosyst* 2012, **8**:1921-1929.

49. Liepe J, Taylor H, Barnes CP, Huvet M, Bugeon L, Thorne T, Lamb JR, Dallman MJ, Stumpf MPH: **Calibrating spatio-temporal models of leukocyte dynamics against in vivo live-imaging data using approximate Bayesian computation**. *Integr Biol* 2012, **4**:335-345.

50. Ruess J, Lygeros J: **Identifying stochastic biochemical networks from single-cell population experiments: a comparison of approaches based on the Fisher information**. *Proc IEEE Conf Decis Control* 2013 http://dx.doi.org/10.1109/CDC.2013.6760291.

51. Lillacci G, Khammash M: **The signal within the noise: efficient inference of stochastic gene regulation models using fluorescence histograms and stochastic simulations**. *Bioinformatics* 2013, **29**:2311-2319.

52. Zechner C, Unger M, Pelet S, Peter M, Koeppl H: **Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings**. *Nat Methods* 2014, **11**:197-202.

53. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW: **A whole-cell computational model predicts phenotype from genotype**. *Cell* 2012, **150**:389-401.

54. Martins ML, Ferreira SC, Vilela MJ: **Multiscale models for biological systems**. *Curr Opin Colloid Interface Sci* 2010, **15**:18-23.

55. Dada JO, Mendes P: **Multi-scale modelling and simulation in systems biology**. *Integr Biol (Camb)* 2011, **3**:86-96.

56. Walpole J, Papin JA, Peirce SM: **Multiscale computational models of complex biological systems**. *Annu Rev Biomed Eng* 2013, **15**:137-154.

57. Karr JR, Williams AH, Zucker JD, Raue A, Steiert B, Timmer J, Kreutz C, Wilkinson S, Allgood BA, Bot BM *et al.*: **Summary of the DREAM8 parameter estimation challenge: toward parameter identification for whole-cell models**. *PLoS Comput Biol* 2015, **11**:e1004096.

58. Iber D, Fengos G: **Predictive models for cellular signaling networks**. *Methods Mol Biol* 2012, **880**:1-22.

59. Schmidt H, Jirstrand M: **Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology [Internet]**. *Bioinformatics* 2006, **22**:514-515.

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

61. Raue A, Steiert B, Schelker M, Kreutz C, Maiwald T, Hass H, Vanlier J, Tönsing C, Adlung L, Engesser R *et al.*: **Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems**. *Bioinformatics* 2015 http://dx.doi.org/10.1093/bioinformatics/btv405.