

RESEARCH ARTICLE

# Data-driven reverse engineering of signaling pathways using ensembles of dynamic models

David Henriques<sup>1</sup>, Alejandro F. Villaverde<sup>1,2</sup>, Miguel Rocha<sup>2</sup>, Julio Saez-Rodriguez<sup>3,4\*</sup>, Julio R. Banga<sup>1\*</sup>

**1** Bioprocess Engineering Group, Spanish National Research Council, IIM-CSIC, Vigo, Spain, **2** Centre of Biological Engineering, University of Minho, Braga, Portugal, **3** Joint Research Center for Computational Biomedicine, RWTH-Aachen University, Aachen, Germany, **4** European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, United Kingdom

\* saezrodriguez@combine.rwth-aachen.de (JSR); julio@iim.csic.es (JRB)



**OPEN ACCESS**

**Citation:** Henriques D, Villaverde AF, Rocha M, Saez-Rodriguez J, Banga JR (2017) Data-driven reverse engineering of signaling pathways using ensembles of dynamic models. *PLoS Comput Biol* 13(2): e1005379. doi:10.1371/journal.pcbi.1005379

**Editor:** Kai Tan, University of Pennsylvania, UNITED STATES

**Received:** May 11, 2016

**Accepted:** January 24, 2017

**Published:** February 6, 2017

**Copyright:** © 2017 Henriques et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** Data are from the DREAM challenges and are available at [www.synapse.org/#!/Synapse:syn1720047](http://www.synapse.org/#!/Synapse:syn1720047).

**Funding:** JRB and DH acknowledge funding from the EU FP7 project NICHE (ITN Grant number 289384). JRB acknowledges funding from the Spanish MINECO project SYNBIOfACTORY (grant number DPI2014-55276-C5-2-R). AFV acknowledges funding from the Galician government (Xunta de Galiza) through the I2C postdoctoral fellowship ED481B2014/133-0. The

## Abstract

Despite significant efforts and remarkable progress, the inference of signaling networks from experimental data remains very challenging. The problem is particularly difficult when the objective is to obtain a *dynamic* model capable of predicting the effect of novel perturbations not considered during model training. The problem is ill-posed due to the nonlinear nature of these systems, the fact that only a fraction of the involved proteins and their post-translational modifications can be measured, and limitations on the technologies used for growing cells *in vitro*, perturbing them, and measuring their variations. As a consequence, there is a pervasive lack of identifiability. To overcome these issues, we present a methodology called SELDOM (enSEmble of Dynamic ILogic-based Models), which builds an ensemble of logic-based dynamic models, trains them to experimental data, and combines their individual simulations into an ensemble prediction. It also includes a model reduction step to prune spurious interactions and mitigate overfitting. SELDOM is a data-driven method, in the sense that it does not require any prior knowledge of the system: the interaction networks that act as scaffolds for the dynamic models are inferred from data using mutual information. We have tested SELDOM on a number of experimental and *in silico* signal transduction case-studies, including the recent HPN-DREAM breast cancer challenge. We found that its performance is highly competitive compared to state-of-the-art methods for the purpose of recovering network topology. More importantly, the utility of SELDOM goes beyond basic network inference (i.e. uncovering *static* interaction networks): it builds *dynamic* (based on ordinary differential equation) models, which can be used for mechanistic interpretations and reliable dynamic predictions in new experimental conditions (i.e. not used in the training). For this task, SELDOM's ensemble prediction is not only consistently better than predictions from individual models, but also often outperforms the state of the art represented by the methods used in the HPN-DREAM challenge.

funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Author summary

Signaling pathways play a key role in complex diseases such as cancer, for which the development of novel therapies is a difficult, expensive and laborious task. Computational models that can predict the effect of a new combination of drugs without having to test it experimentally can help in accelerating this process. In particular, network-based dynamic models of these pathways hold promise to both understand and predict the effect of therapeutics. However, their use is currently hampered by limitations in our knowledge of the underlying biochemistry, as well as in the experimental and computational technologies used for calibrating the models. Thus, the results from such models need to be carefully interpreted and used in order to avoid biased predictions. Here we present a procedure that deals with this uncertainty by using experimental data to build an ensemble of dynamic models. The method incorporates steps to reduce overfitting and maximize predictive capability. We find that by combining the outputs of individual models in an ensemble it is possible to obtain a more robust prediction. We report results obtained with this method, which we call SELDOM (enSEmble of Dynamic lOGic-based Models), showing that it improves the predictions previously reported for several challenging problems.

This is a *PLOS Computational Biology* Methods paper

## Introduction

Inferring the molecular circuits of the cell from experimental data is a fundamental question of systems biology. In particular, the identification of signaling and regulatory networks in healthy and diseased human cells is a powerful approach to unravel the mechanisms controlling biological homeostasis and their malfunctioning in diseases, and can lead to the development of novel therapies [1, 2]. Given the complexity of these networks, these problems can only be addressed effectively combining experimental techniques with computational algorithms. Such network inference (or reverse engineering) efforts [3] have been largely developed for gene regulation [4, 5], and to a lesser extent for signal transduction [1]. Extensive work has been published on the inference of molecular circuits, either as static networks—that is, recovering only the topology of interactions—[4–6] or as dynamical systems [7, 8]. It can be beneficial to consider network inference in conjunction with the prediction of data for new conditions, since a precise topology should help in the generation of high quality predictions, and the inability of a model topology to describe a given set of experiments suggests that the model is in some sense wrong or incomplete.

Signal transduction is a highly dynamic process, and the identification and analysis of the underlying systems requires dynamical data of the status of its main players (proteins) upon perturbation with ligands and drugs. These experiments are relatively complex and expensive, and there is a trade-off between coverage and throughput [2] that often makes the problem ill-posed, leading to identifiability issues. The problem of handling parametric and structural uncertainty in dynamic models of biological systems has received great attention in systems biology and biotechnology [9–12]. Inference and identification

methods can be used to find families of dynamic models compatible with the available data, but in general these models will still suffer from lack of identifiability in a certain degree [3].

Ensemble modeling can be used to improve the predictive capabilities of models, helping to overcome the fundamental difficulties associated with lack of structural and/or practical identifiability. The usage of ensemble methods is widespread in fields such as machine learning [13], bioinformatics [14], and weather forecasting, but not so much in computational systems biology, although it has been successfully applied in the context of regulatory [15, 16], metabolic [17, 18], and signaling [19] networks. Although there is no universally agreed explanation of the success of ensemble methods as classifiers in machine learning [20], it has been shown that they can improve generalization accuracy by decreasing variance [21], bias [22] or both [23], and the reasons for this are relatively well understood [13]. A common approach for building an ensemble is to train a number of so-called base learners in a supervised manner, using data re-sampling strategies. An example of the application of such methods in biology can be found in [24], where the inference of gene regulatory networks is formulated as a feature selection problem, and regression is performed using tree-based ensemble methods. This approach was recently extended to accommodate dynamics [25]. Xing et al [26] used ensemble simulations of causal genetic networks to predict genes involved in rheumatoid arthritis. They showed that the use of ensembles allows for quantitative prediction of the effects of perturbation, adding robustness to predictions by accounting for uncertainty in network topology. Besides robustifying predictions, another task in which ensembles can be helpful is in the elucidation of insufficiently characterized circuits. In this context, Kuepfer et al [19] showed how to use ensembles to unravel operating principles in signaling pathways. They created an ensemble of plausible models of the target of rapamycin (TOR) pathway in *S. cerevisiae*, in which the different topologies accounted for uncertainties in network structure: each member of the ensemble extended a core model by including an additional reaction. By clustering the models according to their training errors, they determined the common features shared by those that better reproduced the experimentally observed behaviour. In this way, a new factor was proposed as the key signaling mechanism. Ensembles of dynamic systems have been used for many years in weather forecasting. In that community, sets of simulations with different initial conditions (ensemble modeling) and/or models developed by different groups (multi-model ensemble) are combined to deliver improved forecasts [27, 28]. In the context of metabolism, Lee et al [29] have shown how to use ensembles to assess the robustness of non-native engineered metabolic pathways. Using the ensemble generation method proposed in [18], a sampling scheme is used to generate representative sets of parameters/fluxes vectors, compatible with a known stoichiometric matrix. This approach is based on the fact that this problem is typically underdetermined, *i.e.* there are more reactions/fluxes than metabolites. Thus, model ensembles may be generated by considering all theoretically possible models, or a representative sample of it. The use of an ensemble composed of all models compatible with the data has been applied to gene regulatory [15] and signal transduction networks [30].

If the model structure is unknown, the ensemble generation needs to be completely data-driven. A common approach for inferring network structures from data is to use estimations of information-theoretic measures, such as entropy and mutual information. There is a plethora of methods for inferring static networks, including correlation, Bayesian inference, and information theory metrics [3, 31–33]. We have focused on information-theoretic approaches because of their good properties regarding handling of nonlinear interactions (which are common in signalling pathways) and scalability. For a recent review of

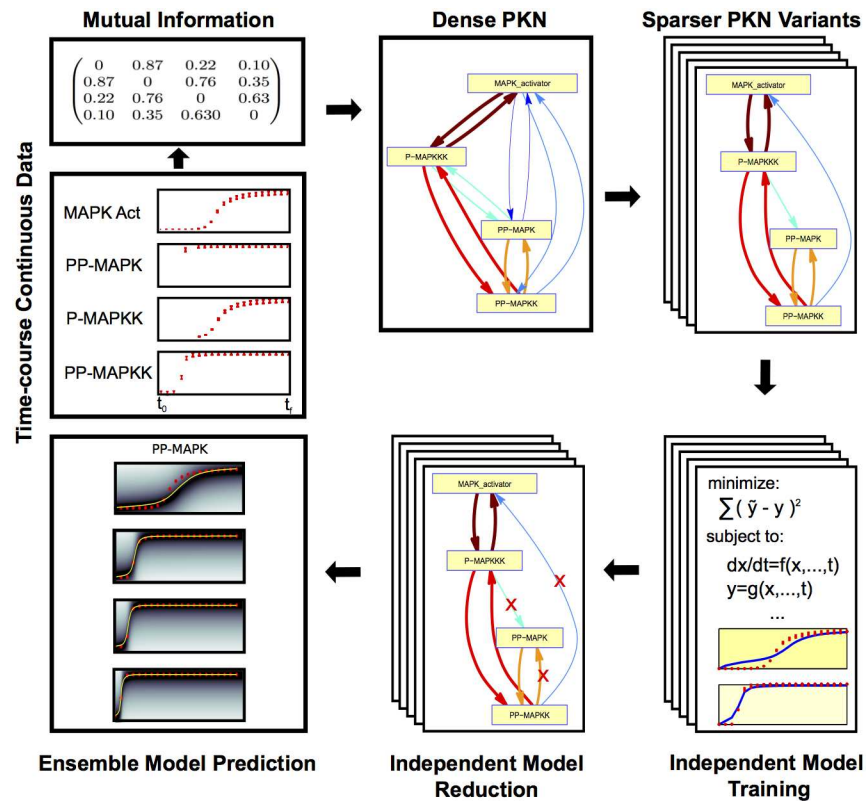
information-theoretic methods, see [6]. The central concept in information theory is entropy, a measure of the uncertainty of a random variable [34]. Mutual information, which can be obtained as a function of the entropies of two variables, measures the amount of information that one random variable provides about another. The mutual information between pairs of variables can be estimated from a data-set, and this can be used to determine the existence of interactions between variables, thus allowing the reverse engineering of network structure. For early examples of this approach, see e.g. the methods reviewed in [35, 36], which covers different modeling formalisms used in gene regulatory network inference (GRN). The use of these techniques is not limited to GRNs; they can be applied to cellular networks in general [37]. Detailed comparisons of some of these methods can be found in several studies [4, 38–40]. Some state-of-the-art information-theoretic methods for network inference are ARACNe [41], and its extensions TD-ARACNE [42] and hARACNe [43], Context Likelihood of Relatedness, (CLR) [44], Minimum Redundancy Networks (MRNET) [45], three-way Mutual Information (MI3) [46], and Mutual Information Distance and Entropy Reduction (MIDER) [47], to name a few. All of them are based on estimating some information-theoretic quantity from the data and applying some criterion for determining the existence of links between pairs of variables. While the details vary from one method to another, it is difficult to single out a clearly “best” method. Instead, it has become clear in recent years that every method has its weaknesses and strengths, and their performance is highly problem-dependent; hence, the best option is often to apply “wisdom of crowds” methods, akin to the ensemble approach described above, as suggested by the results of recent DREAM challenges [48, 49]. In this spirit, recent software tools aim at facilitating the combined use of several methods [50].

Here, we present SELDOM (enSEmble of Dynamic lOgic-based Models), a method developed with the double goal of inferring network topologies, *i.e.* finding the set of causal interactions between a number of biological entities, and of generating high quality predictions about the behaviour of the system under untested experimental perturbations (also known as out-of-sample cross-validation). SELDOM makes no *a priori* assumptions about the model structure, and hence follows a completely data-driven approach to infer networks using mutual information. At the core of SELDOM is the assumption that the information contained in the available data will not be enough to successfully reconstruct a unique network. Instead, it will be generally possible to find many models that provide a reasonable description of the data, each having its own individual bias. Hence SELDOM infers a number of plausible networks, and uses them to generate an ensemble of logic-based dynamic models, which are trained with experimental data and undergo a model reduction procedure in order to mitigate overfitting. Finally, the simulations of the different models are combined into a single ensemble prediction, which is better than the ones produced by individual models.

The remainder of this paper is organised as follows. First, the Methods section provides a step by step description of the procedure followed by SELDOM. Then a number of experimental and *in silico* case studies of signaling pathways of different sizes and complexity are presented. In the Results section the performance of SELDOM is tested on these case studies and benchmarked against other methods. We finish by presenting some conclusions and guidelines for future work.

## Methods

The SELDOM workflow, outlined in Fig 1, combines elements from information theory, ensemble modeling, parametric dynamic model identification, logic-based modeling and



**Fig 1. SELDOM workflow.** The experimental data is used to build an adjacency (a dense DDN) matrix based on the mutual information of all pairs of variables. Through a simple sampling scheme, and limiting the maximum in-degree for each node, a set of more sparse DDNs are generated. Each individual DDN is then used as a scaffold for independent model training and model reduction problems. The resulting models are used to form an ensemble which is able to produce predictions for state trajectories under untested experimental conditions.

doi:10.1371/journal.pcbi.1005379.g001

model reduction. The final objective is to provide high quality predictions of dynamic behavior even for untested experimental conditions. The method starts from time-course continuous experimental data ( $\tilde{y}$ ) and uses data-driven networks (DDNs) as intermediate scaffolds. The workflow can be roughly divided into the following 5 steps:

1. Dense DDN inference using mutual information (MI) from experimental data  $\tilde{y}$ : build an adjacency (dense DDN) matrix based on the mutual information of all pairs of measured variables.
2. Sampling of DDNs: sample  $n_M$  data-driven networks (DDNs) based on the MI.
3. Independent model training: parametric identification of a set of ordinary differential equation models based on the DDNs.
4. Independent model reduction: iterative model reduction procedure of the individual models via a greedy heuristic.
5. Ensemble prediction: build ensemble of models to obtain predictions for state trajectories under untested experimental conditions.

The term network topology is defined here as a directed graph  $G$ . A directed graph (digraph) is a graph where all the edges are directed. The term node or vertex refers to a biological entity such as a protein, protein activity, gene, etc. A directed edge (interaction) starting from node  $v_i$  and pointing to  $v_j$  implies that the behavior of node  $v_i$  interferes with the behavior of node  $v_j$ . In this case,  $v_i$  is said to be adjacent to  $v_j$ . The in-degree of node  $v_i$  ( $deg^-(v_i)$ ) is the number of edges pointing to  $v_i$ . The directed graph  $G$  is composed of the ordered pair  $G(V(G), E(G))$ , where  $V(G)$  is the set of  $n$  vertices and  $E(G)$  is the set of  $m$  edges.

The input to the SELDOM algorithm is an experimental data-set formatted as a MIDAS file [51] and the maximum in-degree ( $deg^-(v_i)$ ) allowed for each node in the networks sampled. The MIDAS file should specify for each experiment the observed signals, the observation times and the treatments/perturbations applied. Two types of perturbations are currently supported: inhibitors and stimuli. These are typical in most experimental studies of signaling pathways, where inhibitors are e.g. small molecules blocking kinase function, and stimuli are upstream ligands (e.g. hormones) whose initial concentration can be manipulated.

### Mutual information

The mutual information  $MI(\tilde{y}_i, \tilde{y}_j)$  between two random variables  $\tilde{y}_i$  and  $\tilde{y}_j$  is a measure of the amount of information that one random variable contains about another. It can also be considered as the reduction in the uncertainty of one variable due to the knowledge of another. It is defined as follows:

$$MI(\tilde{y}_i, \tilde{y}_j) = \sum_{\epsilon=1}^{n_\epsilon} \sum_{s=1}^{n_s^\epsilon} p(\tilde{y}_i^{\epsilon,s}, \tilde{y}_j^{\epsilon,s}) \log \frac{p(\tilde{y}_i^{\epsilon,s}, \tilde{y}_j^{\epsilon,s})}{p(\tilde{y}_i^{\epsilon,s})p(\tilde{y}_j^{\epsilon,s})} \tag{1}$$

where  $\tilde{y}_i$  and  $\tilde{y}_j$  are discrete random vectors with probability mass functions  $p(x)$  and  $p(y)$ , and  $\log$  is usually the logarithm to the base 2, although the natural logarithm may also be used.

Since mutual information is a general measure of dependency between variables, it can be used for inferring interaction networks: the stronger the interaction between two network nodes, the larger their mutual information. If the probability distributions  $p(\tilde{y}_i)$  and  $p(\tilde{y}_j)$  are known,  $MI(\tilde{y}_i, \tilde{y}_j)$  can be derived analytically. In network inference applications, however, this is not possible, so the mutual information must be estimated from data, a task for which several techniques have been developed [52]. In the present work we calculate mutual information using the empirical estimator included in the R package *minet* [53].

### Sampling data-driven networks

Whatever the approach used to estimate the MI, estimation leads to errors, due to factors such as limited measurements or noisy data. Therefore, it is often the case that MI is over-estimated, which results in false positives. Network inference methods usually adopt strategies to detect and discard false positives. For example, ARACNe uses the data processing inequality, which states that, for interactions of the type  $X \rightarrow Y \rightarrow Z$ , it always holds that  $MI(X, Y) \geq MI(X, Z)$ . Thus, by removing the edge with the smallest value of a triplet, ARACNe avoids inferring spurious interactions such as  $X \rightarrow Z$ . However, this in turn may lead to false negatives.

In the present work we are interested in building DDNs that are as dense as possible, in the sense that these should ideally contain all the real interactions, which leads to containing some false positives too (the issue of the false positives will be handled in the independent model reduction step). However, the subsequent dynamic optimization formulation used to train the



models benefits from limiting the number of interactions (i.e. the number of decision variables grows very rapidly with the in-degree).

To find each DDN, we build an adjacency matrix using the array  $MI(\tilde{y}_i, \tilde{y}_j)$ . Each column  $j$  represents the edges starting from  $v_i$  and pointing to  $v_j$ . From this vector we iteratively select as many edges as the maximum in-degree (a pre-defined parameter of the method). In each selection step, an edge is chosen with a probability proportional to  $MI(\tilde{y}_i, \tilde{y}_j)$ . This process is repeated for every node.

### From interaction networks to dynamic models

The DDNs obtained in the previous step represent a set of possible directed interactions. To characterize the dynamics of these interactions we use the multivariate polynomial interpolation technique [54, 55], which transforms discrete models into continuous ones described by ordinary differential equations (ODEs). This interpolation is particularly well suited to represent signaling pathways, since it is able to describe a wide range of behaviours including combinatorial interactions (OR, AND, XOR, etc). Having an ODE-based description allows us to simulate the time courses of the model outputs. By minimizing the difference between those predictions and the experimental data we can obtain a trained model. In the following lines we present the mathematical formulation of the dynamic equations, and in the next subsection we describe the model calibration approach.

Let us represent a Boolean variable by  $x_i \in \{0, 1\}$ . In logic-based ODE models each state variable  $\bar{x}_i$  is a generalization of a Boolean variable, and can have continuous values between zero and one, that is,  $\bar{x}_i \in [0, 1]$ . Every  $i^{th}$  state variable in the model,  $\bar{x}_i$ , represents a species whose behaviour is governed by a set  $\phi_i$  of  $N_i$  variables which act as upstream regulators ( $\bar{x}_{i_1}, \dots, \bar{x}_{i_{N_i}}$ ). Each of the  $N_i$  regulators in  $\phi_i$  is listed by an index  $\phi_{ik}$ . For example, if a variable  $\bar{x}_5$  is regulated by the following subset of  $N_5 = 3$  nodes:  $\phi_5 = \{\bar{x}_3, \bar{x}_7, \bar{x}_8\}$ , we would have three indices  $\phi_{5_1} = 3$ ,  $\phi_{5_2} = 7$ , and  $\phi_{5_3} = 8$ .

For each interaction we describe the nonlinearity that governs the relation between the upstream regulators  $\bar{x}_k$  and the downstream variable  $\bar{x}_i$  using the normalized Hill function  $H_{\phi_{ik}}$ . For each index  $\phi_{ik}$  the normalized Hill function has the form:

$$H_{\phi_{ik}} = \frac{\bar{x}_{\phi_{ik}}^{n_{\phi_{ik}}}}{\bar{x}_{\phi_{ik}}^{n_{\phi_{ik}}} + k_{\phi_{ik}}^{n_{\phi_{ik}}}} \cdot (1 + k_{\phi_{ik}}^{n_{\phi_{ik}}}) \tag{2}$$

We have chosen the normalized Hill function because it is able to represent the switch-like behaviour seen in many molecular interactions [54], as well as other simpler behaviours such as Michaelis-Menten type kinetics. The shape of this curve is defined by the parameters  $n_{\phi_{ik}}$  and  $k_{\phi_{ik}}$ .

Using these Hill functions we write the continuous homologue of the Boolean update function for variable  $\bar{x}_i$  as:

$$\bar{B}_i = \sum_{x_{i_1}=0}^1 \dots \sum_{x_{i_{N_i}}=0}^1 \left[ w_{x_{i_1}, \dots, x_{i_{N_i}}} \cdot \prod_{k=1}^{N_i} (x_{i_k} H_{\phi_{ik}} + [1 - x_{i_k}][1 - H_{\phi_{ik}}]) \right] \tag{3}$$

where the  $w^*$  are parameters that define the model structure. Note that each  $w^*$  has  $N_i$  subindices, that is, as many subindices as regulators (this number varies from one variable to

another). We also remark that  $x_{i_1}, \dots, x_{i_{N_i}}$  are Boolean variables, so the term  $\sum_{x_{i_1}=0}^1 \dots \sum_{x_{i_{N_i}}=0}^1 [\dots]$

represents sums where the  $x_{i^*}$  elements have values zero or one.

**Table 1. Table illustrating the multivariate polynomial interpolation function and model reduction.**

$x_1$	$x_2$	$\bar{B}_i = \dots$	$\bar{B}_i^* = \dots$
0	0	$w_{i,0,0} \cdot (1 - H_{\phi_{i1}}) \cdot (1 - H_{\phi_{i2}}) + \dots$	$w_{i,0,0} \cdot [1 - H_{\phi_{i1}}] \cdot 1 + \dots$
0	1	$w_{i,0,1} \cdot (1 - H_{\phi_{i1}}) \cdot H_{\phi_{i2}} + \dots$	0 + ...
1	0	$w_{i,1,0} \cdot H_{\phi_{i1}} \cdot (1 - H_{\phi_{i2}}) + \dots$	$w_{i,1,0} \cdot H_{\phi_{i1}} \cdot 1 + \dots$
1	1	$w_{i,1,1} \cdot H_{\phi_{i1}} \cdot H_{\phi_{i2}}$	0

The multivariate polynomial interpolation function,  $\bar{B}_i$ , is simplified by setting  $H_{\phi_{ij}}$  to zero, resulting in function  $\bar{B}_i^*$ . In practice this is equivalent to removing the edge  $e_{ij}$ . In the model reduction procedure, the remaining parameters are then estimated starting from the best known solution, and the new (simpler) model is accepted if it is better according to the Akaike information criterion.

doi:10.1371/journal.pcbi.1005379.t001

The time evolution of  $\bar{x}_i$  is then given by:

$$\dot{\bar{x}}_i = (\bar{B}_i - \bar{x}_i) \cdot \frac{1}{\tau_i} \tag{4}$$

where  $\tau_i$  can be seen as the lifetime of species  $x_i$ .

This representation can reproduce several behaviours of interest (see Table 1). For example, if we consider that a variable is controlled by two regulators, an AND type behaviour would be defined by setting  $w_{i,1,1}$  to 1 and the other  $w$ 's ( $w_{i,0,0}$ ,  $w_{i,0,1}$ , and  $w_{i,1,0}$ ) to 0. On the other hand, the OR gate can be represented by setting  $w_{i,1,0}$  and  $w_{i,0,1}$  to 1, and  $w_{i,1,1}$  and  $w_{i,0,0}$  to 0. By linear combinations of these terms it is possible to obtain any of the 16 gates that can be composed of two inputs.

This framework is very general and requires very few assumptions about the system under study. This comes at the cost of a large number of parameters to estimate: in principle, both Hill constants ( $n_{\phi_{ik}}$  and  $k_{\phi_{ik}}$ ) are unknown, as well as the structure parameters  $w^*$  and the lifetime parameter  $\tau_i$ . The following subsection provides a formal definition of the parameter estimation problem.

### Independent model training

For each DDN it is possible to obtain the corresponding dynamic model automatically, as explained in the preceding subsection. Every such model has a number of unknown parameters: for each variable, one or several  $n_*$ ,  $k_*$ ,  $w_*$ , and  $\tau_*$  have to be estimated. To this end we formulate a parameter estimation problem in which the objective function ( $F$ ) is the squared difference between the model predictions ( $y$ ) and the experimental data ( $\bar{y}$ ). The goal is to minimize this cost function for every experiment ( $\epsilon$ ), observed species ( $o$ ) and sampling point ( $s$ ). The model prediction obtained by simulation ( $y$ ) is a discrete data set given by an observation function ( $g$ ) of the model dynamics at time  $t$ . This parameter estimation problem is formally



defined as:

$$\begin{aligned}
 & \underset{n_s, k_s, \tau_s, w_s}{\text{minimize}} & F &= \sum_{\epsilon=1}^{n_\epsilon} \sum_{o=1}^{n_o^\epsilon} \sum_{s=1}^{n_s^{\epsilon,o}} (\tilde{y}_s^{\epsilon,o} - y_s^{\epsilon,o})^2 \\
 & \text{subject to} & & \\
 & & N_i &= \text{deg}^-(v_i) \\
 & & \phi_i &= \{j \mid e_{ij} = 1\}, \quad i = 1, \dots, n, \quad j = 1, \dots, n \\
 & & H_{\phi_{ik}} &= \frac{\bar{x}_{\phi_{ik}}^{n_{\phi_{ik}}}}{\bar{x}_{\phi_{ik}}^{n_{\phi_{ik}}} + k_{\phi_{ik}}^{n_{\phi_{ik}}}} \cdot (1 + k_{\phi_{ik}}^{n_{\phi_{ik}}}) \\
 & & \bar{B}_i &= \sum_{x_{i1}=0}^1 \dots \sum_{x_{iN_i}=0}^1 \left[ w_{x_{i1}, \dots, x_{iN_i}} \cdot \prod_{k=1}^{N_i} (x_{ik} H_{\phi_{ik}} + [1 - x_{ik}][1 - H_{\phi_{ik}}]) \right] \quad (5) \\
 & & \dot{\bar{x}}_i &= (\bar{B}_i - \bar{x}_i) \cdot \frac{1}{\tau_i} \\
 & & \bar{x}_i(t_0) &= \bar{x}_{i0} \\
 & & y &= g(\bar{x}) \\
 & & 0 &\leq w_i \leq 1 \\
 & & \text{LB}_n &\leq n \leq \text{UB}_n \\
 & & \text{LB}_k &\leq k \leq \text{UB}_k \\
 & & \text{LB}_\tau &\leq \tau \leq \text{UB}_\tau,
 \end{aligned}$$

where  $g$  is the output function, which often consists of a subset of the states (if all the states are measured, it is simply  $y = g(\bar{x}) = \bar{x}$ ).

The upper and lower bounds of the parameters (e.g.  $\text{LB}_k$ ) are set to values as wide as possible based on their biochemical meaning and prior knowledge, if existing.

We have recently shown [56] how to train a more constrained version of this problem using mixed-integer nonlinear programming (MINLP). Here, due to its size, the problem is first relaxed into a nonlinear programming (NLP) problem. The corresponding parameter estimation problem is non-convex, so we use the scatter search global optimization method [57] as implemented in the MEIGO toolbox [58].

We note that performing parameter estimation entails repeatedly solving an initial value problem (IVP), which consists of integrating the ODEs from a given initial condition in order to obtain the time course simulation of the model output  $y$ . Several studies that have considered simultaneous network inference and parameter estimation have chosen discretization methods for the solution of the IVP [7, 8]. This has some advantages regarding computational tractability, but forces the  $\dot{x}$  values to be estimated directly from noisy measurements, which is especially challenging when samples are sparse in time. Instead, here we solve the IVP using a state-of-the-art solver for numerical integration of differential equations, CVODE, which is included in the SUNDIALS package [59].

## Independent model reduction

Model reduction is a critical step in SELDOM due to two reasons: (i) we are interested in reducing the network to keep only interactions that are strictly necessary to explain the data (feature selection); (ii) following Occam’s razor principle, it is expected that the ideal model in terms of generalization is the one with just the right level of complexity [60].

Our model reduction procedure is partially inspired by the work of Sunnaker et al [61], where a search tree starting from the most complex model is used to find the complete set of all the simplest models by iteratively deleting parameters. In contrast to [61], we implement a greedy heuristic algorithm. While this method does not offer guarantees of finding the simplest model, it drastically reduces the computational time needed to find the simplest solution. This heuristic helps to maintain diversity in the solutions and ensures that spurious edges are not considered. The iterative model reduction procedure is described in Algorithm 1. At each step (i.e. for each edge), the constraint  $H_{\phi_{ik}}$  is set to zero (see Table 1) and the model is calibrated with a local optimization method, Dynamic Hill Climbing (DHC) [62]. To avoid potential bias caused by the model structure, edges are deleted in a random order.

To decide about the new simplified model we use the Akaike information criterion (AIC), which for the purpose of model comparison is defined as:

$$AIC = 2K + 2n \cdot \ln\left(\frac{F}{n}\right), \tag{6}$$

where  $K$  is the number of active parameters. The theoretical foundations for this simplified version of the AIC can be found in [63].

**Algorithm 1:** Greedy heuristic used to reduce the model. At each step of the model reduction the new (simpler) solution is tested against the previous (more complex) one using the Akaike information criteria (AIC).

**Data:** Time-course continuous data  $\tilde{y}$ , a graph  $G_a(V, E)$  and the optimal parameters  $(n, k, \tau, w)$

**Result:** A simplified graph  $G_a(V, E) *$

**for each**  $e_{\phi_{ik}} \in G_a$  **do**

$$\text{minimize } F = \sum_{\epsilon=1}^{n_\epsilon} \sum_{o=1}^{n_o^\epsilon} \sum_{s=1}^{n_s^{\epsilon,o}} (\tilde{y}_s^{\epsilon,o} - y_s^{\epsilon,o})^2$$

subject to

$$H_{\phi_{ik}} = 0$$

...

**if**  $AIC(n_*, k_*, \tau_*, w_*) < AIC(n, k, \tau, w)$  **then**

$$E_a \leftarrow E_a \setminus e_{\phi_{ik}}$$

$$\{n, k, \tau, w\} \leftarrow \{n_*, k_*, \tau_*, w_*\}$$

**end**

**end**

## Ensemble model prediction

To generate ensemble predictions for the trajectories of state  $x_i$ , SELDOM uses the median value of  $x_i$  across all models for a given experiment  $i_{exp}$  and sampling time  $t_s$ . This is the simplest way to combine a multi-model ensemble projection. More elaborate schemes for optimally combining individual model outputs exist. Gneiting et al. [64] point out that such statistical tools should be used to obtain the full potential of a multi-model ensemble. However, the selection of such weights requires a metric describing the model performance under novel untested conditions (i.e. forecasting), and finding such metric is a non trivial task. For example, in the context of weather forecasting, Tebaldi et al [27] point out that, in the absence of a

metric to quantify model performance for future projections, the usage of simple average is a valid and widely used option that is likely to improve best guess projections due to error cancellation from different models.

## Implementation

SELDOM has been implemented mainly as an R package (R version 2.15), with calls to solvers implemented in C/C++ or Fortran codes using Intel compilers. The SELDOM code is open source and it is distributed as is (with minimal documentation), along with the scripts needed to reproduce all the results and figures.

SELDOM can be installed and run in large heterogeneous clusters and supercomputers. This configuration allows to reduce computation times thanks to the adoption of several parallelization strategies. Model training and reduction are embarrassingly parallel tasks (i.e. they can be performed independently for each individual model). They are automated using shell scripts and a standard queue management system. In addition to this parallelization layer (at the level of individual model training and reduction) there is another parallelization layer at a lower level: for each model, each experiment is simulated in a parallel individual thread using openMP [65], which allows exploiting multi-core processors.

The dynamic optimization problem associated to model training is solved as a master (outer) nonlinear programming problem (NLP) with an inner initial value problem (IVP). The NLPs are solved using the R package MEIGOR [58], with the evaluation of the objective function performed in C code. The solutions of IVPs are obtained by using the CVODE solver [59].

The experimental data is provided using the MIDAS file format, and it is imported and managed using CellNOptR [66].

## Case studies

To assess the performance of SELDOM, we have chosen a number of *in silico* and experimental problems in the reconstruction of signaling networks. Table 2 shows a compact description of some basic properties of these case studies along with a more convenient short name for the purpose of result reporting.

For each case study, two data-sets were derived, one for inference and the second one for performance analysis. We highlight that training and performance assessment data-sets are not just two realizations of the same experimental designs; they were obtained by applying different perturbations, such as different initial conditions or the introduction of inhibitors either experimentally or *in silico*.

**Table 2. An overview of the characteristics of all case studies approached in this work.**

case study	Short name	Reference	Data	$N_{obs}$	$N_{Train}$	$N_{Prediction}$	$deg^-(v_i)$
1a	MAPKp	[67]	<i>in silico</i>	4	10	10	$A = 1, B = 2, C = 3$
1b	MAPKf	[67]	<i>in silico</i>	13	10	10	$A = 3, B = 4, C = 5$
2	SSP	[68]	<i>in silico</i>	13	10	36	$A = 3, B = 4, C = 5$
3	DREAMiS	[69]	<i>in silico</i>	2	20	128	$A = 3, B = 4, C = 5$
4a	DREAMBT20	[70]	Experimental	54	29	8	$A = 3, B = 4, C = 5$
4b	DREAMBT549	[70]	Experimental	52	24	8	$A = 3, B = 4, C = 5$
4c	DREAMMCF7	[70]	Experimental	47	32	8	$A = 3, B = 4, C = 5$
4d	DREAMUACC812	[70]	Experimental	52	32	8	$A = 3, B = 4, C = 5$

The most relevant factors are the number of observed variables, the number of experiments considered for training, the number of experiments considered for prediction and the different maximum in-degrees tested in each case study.

doi:10.1371/journal.pcbi.1005379.t002



**Fig 2. MAPK signaling network.** The model by Huang et al. [67] was used to generate pseudo-experimental data for two sub-problems. The first (MAPKp) partially observed (MAPK-PP, MAPKK-PP and MAPKKK), and the second fully observed MAPKf.

doi:10.1371/journal.pcbi.1005379.g002

The logic-based ODE framework is only able to simulate values between 0 and 1. Therefore, in each case-study, we have normalized the data by dividing every value of a measured protein by the maximum value of that protein found across all conditions and time points in the training data-set. This normalization procedure also facilitates comparison across case-studies.

**Case studies 1a and 1b: MAPK signaling pathway.** Huang et al. [67] developed a model explaining the particular structure of the mitogen-activated protein kinases (MAPKs). This is a highly conserved motif that appears in several signaling cascades (ERK, p38, JNK) [71] composed of 3 kinases. Essentially, Huang et al [67] explain how this arrangement of three kinases sequentially phosphorylated in different sites allows that a graded stimuli is relayed in a ultra-sensitive switch-like manner.

To create this benchmark, the model shown in Fig 2 was used to generate artificial data with no noise. The full system is composed of 12 ODEs. Based in this system, we have derived two case studies, one fully observed (MAPKf) and the second partially observed (MAPKp). The fully observed system is essentially the same as used in [47], while in the partially observed case only one phosphorylation state per kinase was considered (MAPK-PP, MAPKK-PP and MAPKKK).

We highlight that the model representation used in SELDOM is particularly suitable to represent such compact descriptions of signaling mechanisms due to the usage of Hill functions. Additionally, looking at partially observed systems is well in line with experimental practice as state-of-the-art methods for studying signaling pathways are typically targeted to particular states (e.g. phosphorylation) of the proteins (e.g. kinases) involved in the signaling pathways.

Both the data-sets used for training and predictions are composed of 10 different experiments, each with different initial conditions and without added noise. The data used for MAPKp case study is a sub-set of the MAPKf data-set.

**Case study 2: A synthetic signaling pathway.** Resorting to logic-based ODEs, MacNamara et al [68] derived a synthetic model representative of a typical signaling pathway. The

goal was to illustrate the benefits and limitations of different simulations for signaling pathways. This model includes three MAPK systems (p38, ERK and JNK1) and two upstream ligand receptors for EGF and TNF $\alpha$ . Apart from different on/off combinations of EGF and TNF $\alpha$ , the model simulations can be perturbed by inhibiting PI3K and RAF.

The training data-set is composed of 10 experiments with different combinations of ligands (EGF and TNF $\alpha$  on and off) and the inhibitors for RAF and PI3K.

The data-set used to assess performance was generated using the synthetic signaling pathway (SSP) model with the same combinations of EGF and TNF $\alpha$ , but changing the inhibitors. Instead of inhibiting PI3K and RAF, we generate new experiments by considering all other states observed with exception of EGF and TNF $\alpha$ . The final outcome is a validation data-set with 36 experiments.

Both data-sets (training and validation) were partially observed (11 out of 26 variables) and Gaussian noise (with standard deviation  $\sigma = 0.05$  and 0 mean) was added. In this case study the inhibitors are implemented as:

$$\dot{\bar{x}}_{inh,i} = (\bar{B}_i - \bar{x}_i) \cdot \frac{1}{\tau_i} \cdot (1 - inh_i), \quad (7)$$

where  $inh_i$  is chosen as 0.9.

### Case study 3: HPN-DREAM breast cancer network inference, *in silico* sub-challenge.

This is an *in silico* problem developed by the HPN-DREAM consortium. It is a synthetic problem that replicated the reverse phase protein array (RPPA) experimental technique for studying signalling pathways with multiple perturbations as realistically as possible. These perturbations often consist in manipulating ligand concentrations and adding small molecule inhibitors. To achieve this, the authors extended the model from Chen et al. [69], a large dynamic model of ErbB signaling pathways. The model was partially observed (17 variables) and perturbed with a noise model aimed at reproducing the RPPA experimental technique as accurately as possible. In addition to these 17 variables, 3 dummy variables consisting of noise were included to make the challenge even more difficult. All names in the model were replaced by aliases (eg. AB1, AB2, etc).

The training data-set is composed of 20 experiments obtained by considering different combinations of 2 ligands (off, low and high) and 2 small molecule inhibitors. The data-set used for performance assessment is composed of 128 experiments considering the inhibition of the other 15 observed states not considered in the generation of the training set and different combinations of ligand concentrations (off, low and high).

Regarding the implementation of the inhibitors, we followed the same strategy used in the SSP case-study where these are implemented under the assumption that an inhibitor  $inh_i$  of state  $x_i$  directly affects the concentration of  $x_i$ . Such an assumption is based on the challenge design and made following the instructions of the challenge developers.

**Case studies 4a to 4d: HPN-DREAM breast cancer network inference.** One of the richest data-sets of this type was recently made publicly available in the context of the DREAM challenges ([www.dreamchallenges.org](http://www.dreamchallenges.org)). DREAM challenges provide a forum to crowdsource fundamental problems in systems biology and medicine, such as the inference of signaling networks [48, 70], in the form of collaborative competitions. This data-set comprised time-series acquired under eight extracellular stimuli, under four different kinase inhibitors and a control, in four breast cancer cell lines [70].

The HPN-DREAM breast cancer was composed of two sub-challenges. In the experimental sub-challenge the participants were asked to make predictions for 44 observed phosphoproteins, although the complete data-set was larger. As opposed to the *in silico* sub-challenge, the

participants were encouraged to use all the prior knowledge they could use and the experimental protocol along with the real names of the measured quantities, used reagents, inhibitors, etc.

Using different combinations of inhibitors and ligands (on and off), the authors have generated a data-set for several cell-lines. An additional data-set generated with the help of a fourth inhibitor was kept unknown to the participants, who were asked to deliver predictions for several possible inhibitors.

Here, it is assumed that the inhibitors affect mostly the downstream activity of a given kinase. However, it is unknown how it actually influences the kinase concentration or the ability to measure it experimentally. Therefore, the mutual information matrix used to find DDN variants is computed here as:

$$MIM_{inh} = \max(MIM(\tilde{y} \cdot (1 - inh_i)), MIM(\tilde{y})) \tag{8}$$

where  $inh_i$  is a vector of the same size as  $\tilde{y}$ , filled with 0.9 when the inhibition is applied and with 0 otherwise. Regarding the implementation of the dynamic behaviour, this is performed by modifying  $H_{inh,\phi_{ik}}$  of an inhibited species  $x_k$  to:

$$H_{inh,\phi_{ik}} = \frac{\bar{x}^{n_{\phi_{ik}}} \cdot (1 - inh_k)}{\bar{x}^{n_{\phi_{ik}}} \cdot (1 - inh_k) + k^{n_{\phi_{ik}}}} \cdot (1 + k^{n_{\phi_{ik}}}) \tag{9}$$

## Results

### Numerical experiments and method benchmarking

In this section, we describe the numerical experiments carried to show the validity of our ensemble based approach. Besides particular considerations in the data preprocessing or additional constraints added to the dynamic optimization problem which depend on the prior knowledge existent about the case study at hand, SELDOM has two tuning parameters: the ensemble size and the maximum in-degree allowed in the training process. Thus, besides showing how the method performs and illustrating the process we also wanted to show that the method is relatively robust to the choice of these parameters and provide guidelines for the choice of such parameters in future applications.

For each case study we have chosen 3 in-degrees (A, B and C) which are shown in [Table 2](#) and we have chosen a fairly large ensemble size of 100 models.

To assess performance in terms of training and predictive skills of the model, we use the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{\epsilon=1}^{n_\epsilon} \sum_{o=1}^{n_o^\epsilon} \sum_{s=1}^{n_s^{\epsilon,o}} [\tilde{y}_s^{\epsilon,o} - y_s^{\epsilon,o}]^2}{\sum_{\epsilon=1}^{n_\epsilon} \sum_{o=1}^{n_o^\epsilon} n_s^{\epsilon,o}}} \tag{10}$$

To assess performance in terms of network topology inference, we have chosen the area under precision recall (AUPR) curve, where precision (P) and recall are defined as (R):

$$P = \frac{TP}{TP + FP} \tag{11}$$

and

$$R = \frac{TP}{TP + FN}, \tag{12}$$



where TP and FP correspond to the number of true and false positives, respectively and FN corresponds to the number of false negatives.

Other valid metrics exist, such as the Area Under the Receiver Operator Characteristic curve (AUROC). The ROC plots the recall, R, as a function of the false positive rate, FPR, which is defined as

$$FPR = \frac{FP}{FP + TN}. \quad (13)$$

However, it has been argued that ROC curves can paint an excessively optimistic picture of an algorithm's performance [72], because a method can have low precision (i.e. large FP/TP ratio) and still output a seemingly good ROC. Hence we have chosen to use the AUPR measure instead.

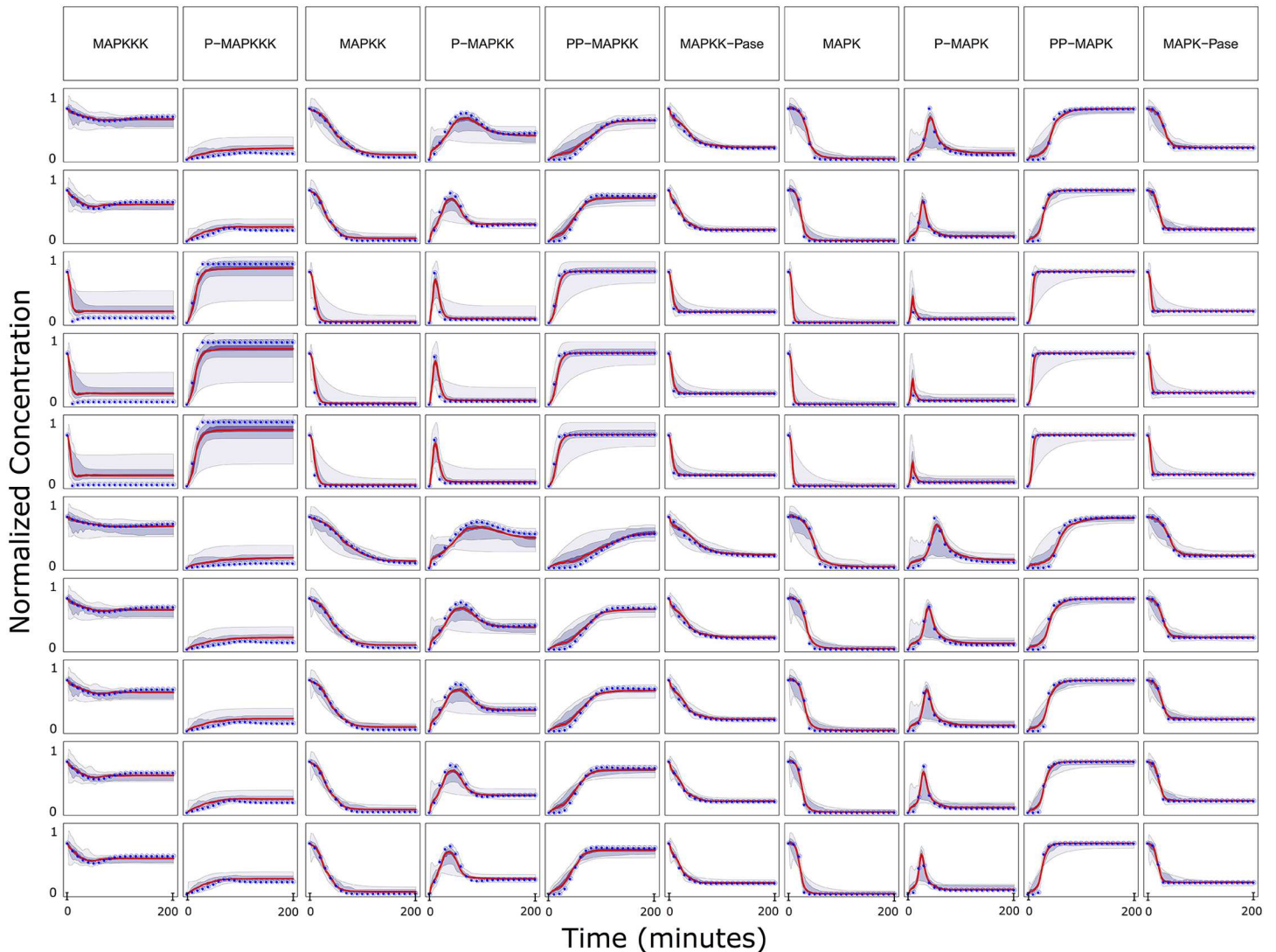
### Predicting trajectories for new experimental perturbations

The training data-sets in each case-study were used to obtain time-course trajectories for untested conditions. The type of trajectories obtained with SELDOM is illustrated with the help of Fig 3 which shows the time-course predictions for different conditions in the case-study 1b (MAPKf). Ensemble trajectories for training and untested conditions are also given for other case-studies in S1 Text and S2 Text, respectively. In most cases the ensemble behaved better than the model with lowest RMSE training value. This effect is particularly evident in the DREAMiS case-study and is reflected in Fig 4. Similar plots are also given for other case studies in S3 Text. In a number of case-studies (DREAMiS, DREAMBT20, DREAMBT549, DREAMMCF7 and DREAMUACC812) there is little correlation between the training RMSE and the prediction RMSE.

In Fig 5, we show the overall picture regarding the predictive skills. Two strategies were considered for the generation of predictions: the best individual model and SELDOM. The RMSE values were normalized by problem and plotted as an heatmap. Additionally, for DREAMiS, DREAMBT20, DREAMBT549, DREAMMCF7 and DREAMUACC812 we added the prediction RMSE values for the top 3 performing participants in the corresponding DREAM sub-challenges (experimental and *in silico*). To compute these RMSE scores we downloaded the participants predictions (available online) and normalized the data by using the maximum value of each measured signal found in the experimental data-set. The greatest gain of using an ensemble approach as shown here is in robustness. The effect of the model reduction was relatively small (yet not neglectable) in terms of RMSE for prediction.

Comparing SELDOM results with those generated during the DREAM challenge, we managed to generate predictions with lower RMSE score than the top 3 participants (Team34, Team8 and Team10) of the *in silico* time-course predictions sub-challenge. Team34 (ranked first) built consensus networks generated by different inference algorithms applied to multiple subsets of the data. To generate the time-course predictions the previously mentioned team used generalized linear models informed by the inferred networks [70]. Regarding the experimental sub-challenge for most cases we obtained similar results to those of the top 3 participants (Team44, Team42 and Team10) with the exception of cell-line DREAMMCF7 where results were slightly better.

The choice of ensemble size parameter affects the predictive skill of the ensemble and the computational resources needed to solve the problem. To verify if this choice was an appropriate one we plotted the average prediction RMSE as a function of the number of models  $n_M$  used to generate the ensemble. The average RMSE was computed by sampling multiple models from the family of models available to compute the trajectories. This is shown in Fig 6 for the

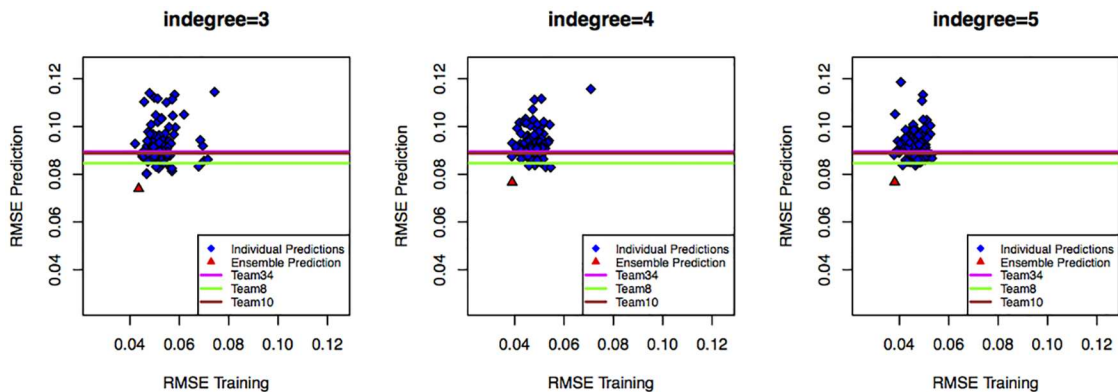


**Fig 3. Time course predictions for case study 1b (MAPKf).** The median in red is surrounded by the predicted non-symmetric 20%,60% and 95% confidence intervals.

doi:10.1371/journal.pcbi.1005379.g003

DREAMiS case-study. Similar curves are given for all case studies as supporting information in [S4 Text](#). With the exception of the combination MAPKp/SELDOM A the outcome for all case-studies is that SELDOM would have done similarly well with a smaller number of models and the prediction RMSE *versus*  $n_M$  always converged asymptotically. The mediocre results from the MAPKp/SELDOM A combination appear to be the result of a poor choice for the maximum in-degree parameter ( $A = 1$ ) which is too small.

**Ensemble for network inference.** To assess the performance of SELDOM for the network topology inference problem, we compared SELDOM with a number of methods implemented in the *minet* package [53]: MRNET [45], MRNETB [73], CLR [44] and ARACNE [41]. This comparison is particularly pertinent in this case as the estimation of the mutual information is done using the same method and parameterization. However, these methods are not designed to recover directed networks. To surmount this limitation, we have introduced the comparison with two other methods for directed networks, TDARACNE [42] and MIDER [47].



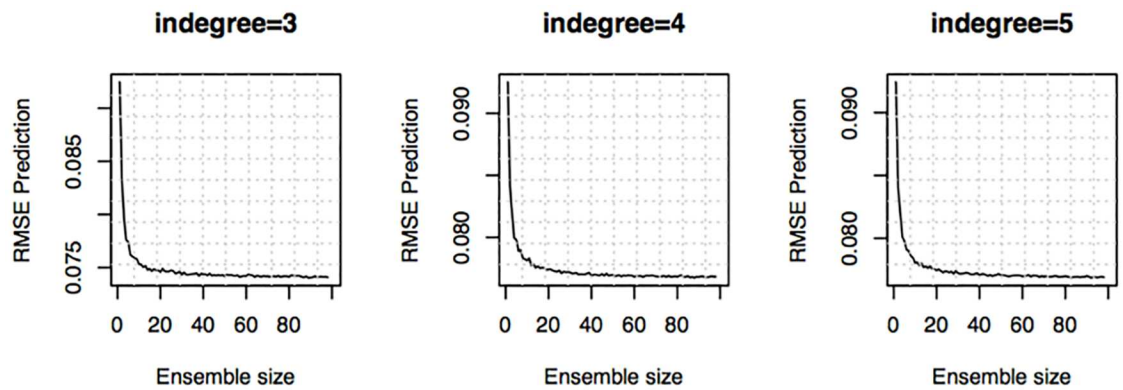
**Fig 4. Relationship between training and prediction RMSE for case study 3 (DREAMiS).** The prediction RMSE is plotted here against the training RMSE for each individual model (blue) and the ensemble (red). RMSE scores for the top 3 performing teams in the HPN-DREAM *in silico* time-course prediction sub-challenge (Team34, Team8 and Team10) are also shown as colored lines.

doi:10.1371/journal.pcbi.1005379.g004

0.075	0.038	0.195	0.089	0.141	0.128	0.144	0.117	best A
0.160	0.041	0.149	0.074	0.084	0.122	0.119	0.107	SELDOM A
0.048	0.029	0.189	0.087	0.114	0.138	0.121	0.149	best B
0.046	0.034	0.161	0.077	0.080	0.121	0.120	0.090	SELDOM B
0.046	0.030	0.262	0.096	0.132	0.159	0.146	0.116	best C
0.042	0.030	0.166	0.077	0.080	0.120	0.114	0.096	SELDOM C
NA	NA	NA	0.089	0.110	0.144	0.140	0.123	Team8
NA	NA	NA	0.089	0.366	0.548	0.115	4.368	Team34
NA	NA	NA	0.095	0.080	0.119	0.117	0.109	Team42
NA	NA	NA	0.085	0.079	0.114	0.110	0.117	Team10
NA	NA	NA	NA	0.080	0.118	0.117	0.109	Team44
mapkp	mapkf	SSpp	DREAMiS	DREAMBT20	DREAMBT549	DREAMUACC812	DREAMMCF7	

**Fig 5. Heatmap with RMSE scores for different methods and case studies.** The color scheme represents RMSE scores normalized by case-study in order to emphasize differences between methods. The color scale moves from green (low RMSE) to blue (high RMSE). The numeric values of the RMSE scores for each method/case-study are also provided in each corresponding cell. SELDOM B and SELDOM C were clearly the most robust strategies doing very well in all problems.

doi:10.1371/journal.pcbi.1005379.g005



**Fig 6. Ensemble predictive skill depending on ensemble size (case study DREAMiS).** This curve was computed by bootstrapping multiple  $n_M$  models from the available 100 models, *i.e.* we sampled multiple realizations of the individual predictions for the same ensemble size and computed the average value. These curves converge asymptotically and show that the chosen ensemble size parameter is adequate. Equivalent predictions could have been obtained with smaller ensemble sizes.

doi:10.1371/journal.pcbi.1005379.g006

In Fig 7, we show the overall results regarding the ability of SELDOM and other network inference methods to reverse engineer the known synthetic networks associated with the models used to generate the data. Comparing with static inference methods, SELDOM behaved consistently well in terms of providing networks with high AUPR score. The sparsest configuration of SELDOM (A) provided the most interesting results.

1.000	0.403	0.377	0.460	SELDOM A
NA	NA	NA	0.464	Team7
0.917	0.323	0.371	0.442	SELDOM C
0.903	0.354	0.325	0.418	SELDOM B
0.500	0.440	0.262	0.453	MRNETB
0.500	0.402	0.256	0.435	ARACNE
0.500	0.441	0.291	0.416	MRNET
0.500	0.452	0.379	0.384	CLR
0.500	0.451	0.251	0.156	Mutual Information
0.597	0.344	0.220	0.210	MIDER
0.189	0.131	0.140	0.428	Correlation
NA	NA	NA	0.319	Team11
NA	NA	NA	0.330	Team18
0.094	0.305	0.211	0.339	TDARACNE
mapkp	mapkf	SSpp	DREAMiS	

**Fig 7. Heatmap with AUPR scores for different methods and case studies.** The color scheme represents AUPR scores normalized by case-study in order to emphasize differences between methods. The color scale moves from red (low AUPR) to yellow (high AUPR). The numeric values of the AUPR scores for each method/case-study are also provided in each corresponding cell. The sparsest version of SELDOM (A) did consistently well in all the case studies. SELDOM B and C did an average job with MAPKf but provided good solutions for all other case-studies. The comparisons are only provided for *in silico* problems with known solution. Additionally, the solution for the top performing teams in the DREAM challenge is only available for DREAMiS.

doi:10.1371/journal.pcbi.1005379.g007



For case study 3 (DREAMiS), we compared the ensemble networks from SELDOM with the networks inferred by the top 3 participants of the *in silico* network inference sub-challenge (Team7, Team11 and Team18). To perform this comparison we downloaded the networks which are available online and computed the AUPR scores. These AUPR scores are shown in Fig 7. SELDOM found networks with AUPR scores that were marginally lower than the top performing participant (Team7) that used GLN [74], a method based on a novel functional  $\chi^2$  test to examine functional dependencies between variables. On the other hand, SELDOM obtained AUPR scores higher than the participants ranked second (Team11) and third (Team18).

Without the independent model reduction step, the results were mediocre regarding the inference of the network topology. The independent model reduction is fundamental for the performance of SELDOM as a method for network inference and the information contained in the dynamics can help discard spurious links. This is illustrated in S5 Text where the mean AUPR score is shown for multiple bootstrapped realizations of the ensemble network with and without applying model reduction. Additionally S6 Text is provided as supporting information where the AUPR curves and AUROC scores are given for all the methods tested in the different case-studies.

We also considered how to validate SELDOM results using experimental data and its capabilities to generate new hypothesis. With this aim, we compared the inferred networks with the current knowledge, as detailed in S7 Text. Briefly, we started by selecting a subset of interactions consistently recovered by SELDOM across all cell-lines. We then searched the literature in a systematic manner to assess the agreement of these predictions with the knowledge currently available in manually curated databases. To this end we used the OmniPath software [75], a comprehensive compendium of literature-curated pathway resources. We found literature support for most (16 out of 20) of the interactions reported with high confidence in all cell-lines. This fact indicates that the predictions made by SELDOM are, generally, consistent with existing biological knowledge. Furthermore, in the cases in which we did not find in OmniPath reports of any interactions, we searched the literature manually and analysed each of the four cases individually. From this analysis we concluded that (i) one of the candidate interactions is a probable true positive (details in S7 Text), (ii) a second one has actually been reported in the literature as an indirect regulator, (iii) a third case is a probable false positive, and (iv) a last case which seems to be a good hypothetical candidate which would require further testing. For this last case we found no literature support for alternative mechanisms. This latter case is an illustrative example of using SELDOM to generate new hypotheses.

## Discussion

In this paper we have presented a novel method for the generation of dynamic predictions in signaling networks. The method (enSEmble of Dynamic lOgic-based Models, SELDOM) handles the indeterminacy of the problem by generating, in a data-driven way, an ensemble of dynamic models combining methods from information theory, global optimization and model reduction. It should be noted that although this method is data-driven (not requiring any prior knowledge of the network), it produces logic-based dynamic models which allow for mechanistic interpretations and are capable of making predictions in different conditions than those used for model calibration.

In this study we focused mostly on the methodological aspects of SELDOM and performed comparisons based on standard metrics. SELDOM can provide cell-specific networks that can be explored to formulate biological hypotheses and guide the design of new experiments to

validate them (as an example, [S7 Text](#) shows the networks obtained with SELDOM for different cell-lines of the HPN-DREAM experimental data sub-challenge).

We applied SELDOM to a number of challenging experimental and *in silico* signal transduction case-studies, including the recent HPN-DREAM breast cancer challenge. Regarding network inference, the ensemble approach did systematically well in all of the *in silico* cases considered in this work. This suggests that exploiting the information contained in the dynamics, as SELDOM does, helps the network inference problem allowing to disregard spurious interactions. We have also shown that, unlike most network inference methods, SELDOM is also capable of making robust dynamic predictions in untested experimental conditions. For this task, SELDOM's ensemble predictions were not only consistently better than the outcomes of individual models, but also often outperformed the state of the art represented by the best performers in the HPN-DREAM challenge. It should be noted that the use of mechanistic dynamic models provides great flexibility regarding the simulation of complex time-varying situations. For example, it is not only possible to simulate inhibitions of several nodes in order to determine which combination produces the desired response, but also to test the outcome of sequential interventions (i.e. taking place at different times during the course of a treatment), which would be impossible to model using statistical approaches that lack mechanistic detail. Another important application is the design of optimal dynamic experiments, i.e. those where the inputs acting as stimuli are designed as time-varying functions ([\[76, 77\]](#)). Furthermore, it is also possible to use the ensemble for extracting biological hypotheses about poorly known parts of a signaling pathway.

The proposed SELDOM pipeline is flexible and can be adapted to any signaling or gene regulation dataset obtained upon perturbation, even if prior knowledge is not available. At the same time, it is also able to incorporate prior knowledge about a problem, for instance in the form of constraints (e.g. the small-molecule inhibitors used in the HPN-DREAM challenge case studies). We have tackled the indeterminacy of the problem by generating a family of solutions, although other strategies, based on data-re-sampling methods and supervised learning (similarly to what has been recently proposed by Huynh-Thu et al. [\[25\]](#)), might work well too. A systematic comparison of ensemble generation methods either based on problem structure or data re-sampling techniques should be considered in further work. Finally, a key point in the usage of ensemble methods is how to combine the models in order to obtain the best prediction possible. In this work we have chosen a simple model averaging framework. If more data become available, more sophisticated methods could be explored.

For the sake of computational reproducibility, we provide the SELDOM code as open source (<http://doi.org/10.5281/zenodo.250558>). This implementation can handle any data file in the MIDAS [\[51\]](#) format. Implementations for all the problems considered are also included in this distribution.

## Supporting information

### **S1 Text. Ensemble time course trajectories for training data.**

(PDF)

### **S2 Text. Ensemble time course trajectories for untested conditions(prediction data).**

(PDF)

### **S3 Text. Relationship between the training and prediction RMSE for individual and ensemble models.** For each case study we show scatter plots with the prediction RMSE as a function of the training RMSE for each individual model and the ensemble.

(PDF)



**S4 Text. Ensemble predictive skill depending on ensemble size for different case-studies.**

The curves shown in the Text were computed by bootstrapping multiple  $n_M$  models from the available 100 models, *i.e.* we sampled multiple realizations of the individual predictions for the same ensemble size and computed the average value. These curves converge asymptotically and show that the chosen ensemble size parameter is adequate. Equivalent predictions could have been obtained with smaller ensemble sizes.

(PDF)

**S5 Text. AUPR score depending on ensemble size for different case-studies with and without applying model reduction.** This curve was computed by bootstrapping multiple  $n_M$  models from the available 100 models, *i.e.* we sampled multiple realizations of the ensemble network for the same ensemble size and computed the average value.

(PDF)

**S6 Text. AUPR curves for different algorithms and case-studies.**

(PDF)

**S7 Text. Networks obtained with SELDOM for different cell-lines of HPN-DREAM experimental data sub-challenge and discussion on the validity of the inferred interactions.**

(PDF)

## Acknowledgments

We thank Dénes Túrei for assistance in the use of the OmniPath software.

## Author contributions

**Conceptualization:** DH JSR JRB.

**Data curation:** DH AFV.

**Formal analysis:** DH AFV JSR JRB.

**Funding acquisition:** JSR JRB.

**Methodology:** DH AFV JSR JRB.

**Project administration:** JSR JRB.

**Resources:** DH AFV JSR JRB.

**Software:** DH AFV.

**Supervision:** AFV MR JSR JRB.

**Validation:** DH AFV.

**Visualization:** DH.

**Writing – original draft:** DH AFV JSR JRB.

**Writing – review & editing:** DH AFV MR JSR JRB.

## References

1. Kholodenko B, Yaffe MB, Kolch W. Computational Approaches for Analyzing Information Flow in Biological Networks. *Sci Signal.* 2012; 5(220):re1. doi: [10.1126/scisignal.2002961](https://doi.org/10.1126/scisignal.2002961) PMID: [22510471](https://pubmed.ncbi.nlm.nih.gov/22510471/)

2. Saez-Rodriguez J, MacNamara A, Cook S. Modeling Signaling Networks to Advance New Cancer Therapies. *Annu Rev Biomed Eng.* 2015; 17(1):143–163. doi: [10.1146/annurev-bioeng-071813-104927](https://doi.org/10.1146/annurev-bioeng-071813-104927) PMID: [26274601](https://pubmed.ncbi.nlm.nih.gov/26274601/)
3. Villaverde AF, Banga JR. Reverse engineering and identification in systems biology: strategies, perspectives and challenges. *J R Soc Interface.* 2014; 11(91):20130505. doi: [10.1098/rsif.2013.0505](https://doi.org/10.1098/rsif.2013.0505) PMID: [24307566](https://pubmed.ncbi.nlm.nih.gov/24307566/)
4. Bansal M, Belcastro V, Ambesi-Impiombato A, di Bernardo D. How to infer gene networks from expression profiles. *Mol Syst Biol.* 2007; 3(1):78. doi: [10.1038/msb4100120](https://doi.org/10.1038/msb4100120) PMID: [17299415](https://pubmed.ncbi.nlm.nih.gov/17299415/)
5. De Smet R, Marchal K. Advantages and limitations of current network inference methods. *Nature Rev Microbiol.* 2010; 8(10):717–729. doi: [10.1038/nrmicro2419](https://doi.org/10.1038/nrmicro2419) PMID: [20805835](https://pubmed.ncbi.nlm.nih.gov/20805835/)
6. Villaverde AF, Ross J, Banga JR. Reverse engineering cellular networks with information theoretic methods. *Cells.* 2013; 2(2):306–329. doi: [10.3390/cells2020306](https://doi.org/10.3390/cells2020306) PMID: [24709703](https://pubmed.ncbi.nlm.nih.gov/24709703/)
7. Chang YH, Gray JW, Tomlin CJ. Exact reconstruction of gene regulatory networks using compressive sensing. *BMC Bioinform.* 2014; 15(1):1–22. doi: [10.1186/s12859-014-0400-4](https://doi.org/10.1186/s12859-014-0400-4) PMID: [25495633](https://pubmed.ncbi.nlm.nih.gov/25495633/)
8. Bonneau R, Reiss DJ, Shannon P, Facciotti M, Hood L, Baliga NS, et al. The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets de novo. *Genome Biol.* 2006; 7(5):1–16. doi: [10.1186/gb-2006-7-5-r36](https://doi.org/10.1186/gb-2006-7-5-r36) PMID: [16686963](https://pubmed.ncbi.nlm.nih.gov/16686963/)
9. Kaltenbach HM, Dimopoulos S, Stelling J. Systems analysis of cellular networks under uncertainty. *FEBS Lett.* 2009; 583(24):3923–3930. doi: [10.1016/j.febslet.2009.10.074](https://doi.org/10.1016/j.febslet.2009.10.074) PMID: [19879267](https://pubmed.ncbi.nlm.nih.gov/19879267/)
10. Schaber J and Liebermeister W and Klipp E. Nested uncertainties in biochemical models. *IET Syst Biol.* 2009; 3(1):1–9. doi: [10.1049/iet-syb:20070042](https://doi.org/10.1049/iet-syb:20070042) PMID: [19154080](https://pubmed.ncbi.nlm.nih.gov/19154080/)
11. Mišković L, Hatzimanikatis V. Modeling of uncertainties in biochemical reactions. *Biotechnol Bioeng.* 2011; 108(2):413–423. doi: [10.1002/bit.22932](https://doi.org/10.1002/bit.22932) PMID: [20830674](https://pubmed.ncbi.nlm.nih.gov/20830674/)
12. Geris L, Gomez-Cabrero D. An Introduction to Uncertainty in the Development of Computational Models of Biological Processes. In: Geris L, Gomez-Cabrero D, editors. *Uncertainty in Biology: A Computational Modeling Approach.* Cham: Springer International Publishing; 2016. p. 3–11. doi: [10.1007/978-3-319-21296-8\\_1](https://doi.org/10.1007/978-3-319-21296-8_1)
13. Dietterich TG. Ensemble Methods in Machine Learning. In: Kittler J and Roli F, editor. *Multiple Classifier Systems: First International Workshop.* Berlin, Heidelberg: Springer; 2000. p. 1–15. doi: [10.1007/3-540-45014-9\\_1](https://doi.org/10.1007/3-540-45014-9_1)
14. Yang P, Yang YH, Zhou BB, Zomaya AY. A Review of Ensemble Methods in Bioinformatics. *Curr Bioinform.* 2010; 5(4):296–308. doi: [10.2174/157489310794072508](https://doi.org/10.2174/157489310794072508)
15. Kauffman S. A proposal for using the ensemble approach to understand genetic regulatory networks. *J Theor Biol.* 2004; 230(4):581–590. doi: [10.1016/j.jtbi.2003.12.017](https://doi.org/10.1016/j.jtbi.2003.12.017) PMID: [15363677](https://pubmed.ncbi.nlm.nih.gov/15363677/)
16. Ud-Dean SMM, Gunawan R. Ensemble Inference and Inferability of Gene Regulatory Networks. *PLoS ONE.* 2014; 9(8):e103812. doi: [10.1371/journal.pone.0103812](https://doi.org/10.1371/journal.pone.0103812) PMID: [25093509](https://pubmed.ncbi.nlm.nih.gov/25093509/)
17. Tan Y, Liao JC. Metabolic ensemble modeling for strain engineers. *Biotechnol J.* 2012; 7(3, SI):343–353. doi: [10.1002/biot.201100186](https://doi.org/10.1002/biot.201100186) PMID: [22021171](https://pubmed.ncbi.nlm.nih.gov/22021171/)
18. Jia G, Stephanopoulos G, Gunawan R. Ensemble kinetic modeling of metabolic networks from dynamic metabolic profiles. *Metabolites.* 2012; 2(4):891–912. doi: [10.3390/metabo2040891](https://doi.org/10.3390/metabo2040891) PMID: [24957767](https://pubmed.ncbi.nlm.nih.gov/24957767/)
19. Kuepfer L, Peter M, Sauer U, Stelling J. Ensemble modeling for analysis of cell signaling dynamics. *Nat Biotechnol.* 2007; 25(9):1001–1006. doi: [10.1038/nbt1330](https://doi.org/10.1038/nbt1330) PMID: [17846631](https://pubmed.ncbi.nlm.nih.gov/17846631/)
20. Re M, Valentini G. Ensemble methods: A review. In: *Advances in Machine Learning and Data Mining for Astronomy.* Chapman & Hall; 2010. p. 563–594. doi: [10.1201/b11822-34](https://doi.org/10.1201/b11822-34)
21. Breiman L. Bagging predictors. *Mach Learn.* 1996; 24(2):123–140. doi: [10.1007/BF00058655](https://doi.org/10.1007/BF00058655)
22. Schapire R, Freund Y, Bartlett P, Lee W. Boosting the margin: A new explanation for the effectiveness of voting methods. *Ann Stat.* 1998; 26(5):1651–1686. doi: [10.1214/aos/1024691352](https://doi.org/10.1214/aos/1024691352)
23. Breiman L. Arcing classifiers. *Ann Stat.* 1998; 26(3):801–824.
24. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. *PLoS ONE.* 2010; 5(9):1–10. doi: [10.1371/journal.pone.0012776](https://doi.org/10.1371/journal.pone.0012776) PMID: [20927193](https://pubmed.ncbi.nlm.nih.gov/20927193/)
25. Huynh-Thu VA, Sanguinetti G. Combining tree-based and dynamical systems for the inference of gene regulatory networks. *Bioinformatics.* 2015; 31(10):1614–1622. doi: [10.1093/bioinformatics/btu863](https://doi.org/10.1093/bioinformatics/btu863) PMID: [25573916](https://pubmed.ncbi.nlm.nih.gov/25573916/)
26. Xing H, McDonagh PD, Bienkowska J, Cashorali T, Runge K, Miller RE, et al. Causal modeling using network ensemble simulations of genetic and gene expression data predicts genes involved in rheumatoid arthritis. *PLoS Comput Biol.* 2011; 7(3):e1001105. doi: [10.1371/journal.pcbi.1001105](https://doi.org/10.1371/journal.pcbi.1001105) PMID: [21423713](https://pubmed.ncbi.nlm.nih.gov/21423713/)

27. Tebaldi C, Knutti R. The use of the multi-model ensemble in probabilistic climate projections. *Phil Trans R Soc A*. 2007; 365(1857):2053–2075. doi: [10.1098/rsta.2007.2076](https://doi.org/10.1098/rsta.2007.2076) PMID: [17569654](https://pubmed.ncbi.nlm.nih.gov/17569654/)
28. Hagedorn R, Doblas-Reyes F, Palmer T. The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus A*. 2005; 57(3):219–233.
29. Lee Y, Rivera JGL, Liao JC. Ensemble Modeling for Robustness Analysis in engineering non-native metabolic pathways. *Metab Eng*. 2014; 25:63–71. doi: [10.1016/j.ymben.2014.06.006](https://doi.org/10.1016/j.ymben.2014.06.006) PMID: [24972370](https://pubmed.ncbi.nlm.nih.gov/24972370/)
30. Guziolowski C, Videla S, Eduati F, Thiele S, Cokelaer T, Siegel A, et al. Exhaustively characterizing feasible logic models of a signaling network using Answer Set Programming. *Bioinformatics*. 2013; 29(18):2320–2326. doi: [10.1093/bioinformatics/btt393](https://doi.org/10.1093/bioinformatics/btt393) PMID: [23853063](https://pubmed.ncbi.nlm.nih.gov/23853063/)
31. Maetschke SR, Madhamshettiwar PB, Davis MJ, Ragan MA. Supervised, semi-supervised and unsupervised inference of gene regulatory networks. *Briefings in bioinformatics*. 2013;p. bbt034. doi: [10.1093/bib/bbt034](https://doi.org/10.1093/bib/bbt034) PMID: [23698722](https://pubmed.ncbi.nlm.nih.gov/23698722/)
32. Banf M, Rhee SY. Computational inference of gene regulatory networks: Approaches, limitations and opportunities. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*. 2016;. doi: [10.1016/j.bbagr.2016.09.003](https://doi.org/10.1016/j.bbagr.2016.09.003) PMID: [27641093](https://pubmed.ncbi.nlm.nih.gov/27641093/)
33. Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP. Causal protein-signaling networks derived from multiparameter single-cell data. *Science*. 2005; 308(5721):523–529. doi: [10.1126/science.1105809](https://doi.org/10.1126/science.1105809) PMID: [15845847](https://pubmed.ncbi.nlm.nih.gov/15845847/)
34. Shannon CE. A Mathematical Theory of Communication. *Bell Syst Tech J*. 1948; 27(3):379–423. doi: [10.1002/j.1538-7305.1948.tb00917.x](https://doi.org/10.1002/j.1538-7305.1948.tb00917.x)
35. De Jong H. Modeling and simulation of genetic regulatory systems: A literature review. *J Comp Biol*. 2002; 9(1):67–103. doi: [10.1089/10665270252833208](https://doi.org/10.1089/10665270252833208) PMID: [11911796](https://pubmed.ncbi.nlm.nih.gov/11911796/)
36. Faria JP, Overbeek R, Xia F, Rocha M, Rocha I, Henry CS. Genome-scale bacterial transcriptional regulatory networks: reconstruction and integrated analysis with metabolic models. *Brief Bioinform*. 2014; 15(4):592–611. doi: [10.1093/bib/bbs071](https://doi.org/10.1093/bib/bbs071) PMID: [23422247](https://pubmed.ncbi.nlm.nih.gov/23422247/)
37. Markowitz F, Spang R. Inferring cellular networks—a review. *BMC Bioinform*. 2007; 8(6):1–17. doi: [10.1186/1471-2105-8-S6-S5](https://doi.org/10.1186/1471-2105-8-S6-S5) PMID: [17903286](https://pubmed.ncbi.nlm.nih.gov/17903286/)
38. Soranzo N, Bianconi G, Altafini C. Comparing association network algorithms for reverse engineering of large-scale gene regulatory networks: Synthetic versus real data. *Bioinformatics*. 2007; 23(13):1640–1647. doi: [10.1093/bioinformatics/btm163](https://doi.org/10.1093/bioinformatics/btm163) PMID: [17485431](https://pubmed.ncbi.nlm.nih.gov/17485431/)
39. Altay G, Emmert-Streib F. Revealing differences in gene network inference algorithms on the network level by ensemble methods. *Bioinformatics*. 2010; 26(14):1738–1744. doi: [10.1093/bioinformatics/btq259](https://doi.org/10.1093/bioinformatics/btq259) PMID: [20501553](https://pubmed.ncbi.nlm.nih.gov/20501553/)
40. Hurley D, Araki H, Tamada Y, Dunmore B, Sanders D, Humphreys S, et al. Gene network inference and visualization tools for biologists: application to new human transcriptome datasets. *Nucleic Acids Res*. 2012; 40(6):2377–2398. doi: [10.1093/nar/gkr902](https://doi.org/10.1093/nar/gkr902) PMID: [22121215](https://pubmed.ncbi.nlm.nih.gov/22121215/)
41. Margolin A, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, et al. ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*. 2006; 7(1):1–15. doi: [10.1186/1471-2105-7-S1-S7](https://doi.org/10.1186/1471-2105-7-S1-S7) PMID: [16723010](https://pubmed.ncbi.nlm.nih.gov/16723010/)
42. Zoppoli P, Morganello S, Ceccarelli M. TimeDelay-ARACNE: Reverse engineering of gene networks from time-course data by an information theoretic approach. *BMC Bioinformatics*. 2010; 11(1):1–15. doi: [10.1186/1471-2105-11-154](https://doi.org/10.1186/1471-2105-11-154) PMID: [20338053](https://pubmed.ncbi.nlm.nih.gov/20338053/)
43. Jang IS, Margolin A, Califano A. hARACNe: improving the accuracy of regulatory model reverse engineering via higher-order data processing inequality tests. *Interface Focus*. 2013; 3(4):20130011. doi: [10.1098/rsfs.2013.0011](https://doi.org/10.1098/rsfs.2013.0011) PMID: [24511376](https://pubmed.ncbi.nlm.nih.gov/24511376/)
44. Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, et al. Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol*. 2007; 5(1):54–66. doi: [10.1371/journal.pbio.0050008](https://doi.org/10.1371/journal.pbio.0050008) PMID: [17214507](https://pubmed.ncbi.nlm.nih.gov/17214507/)
45. Meyer PE, Kontos K, Lafitte F, Bontempi G. Information-theoretic inference of large transcriptional regulatory networks. *EURASIP J Bioinform Syst Biol*. 2007; 2007(1):1–9. doi: [10.1155/2007/79879](https://doi.org/10.1155/2007/79879) PMID: [18354736](https://pubmed.ncbi.nlm.nih.gov/18354736/)
46. Luo W, Hankenson KD, Woolf PJ. Learning transcriptional regulatory networks from high throughput gene expression data using continuous three-way mutual information. *BMC Bioinform*. 2008; 9(1):467. doi: [10.1186/1471-2105-9-467](https://doi.org/10.1186/1471-2105-9-467) PMID: [18980677](https://pubmed.ncbi.nlm.nih.gov/18980677/)
47. Villaverde AF, Ross J, Moran F, Banga JR. MIDER: Network Inference with Mutual Information Distance and Entropy Reduction. *PLoS ONE*. 2014; 9(5):e96732. doi: [10.1371/journal.pone.0096732](https://doi.org/10.1371/journal.pone.0096732) PMID: [24806471](https://pubmed.ncbi.nlm.nih.gov/24806471/)

48. Prill RJ, Saez-Rodriguez J, Alexopoulos LG, Sorger PK, Stolovitzky G. Crowdsourcing Network Inference: The DREAM Predictive Signaling Network Challenge. *Sci Signal*. 2011; 4(189):mr7. doi: [10.1126/scisignal.2002212](https://doi.org/10.1126/scisignal.2002212) PMID: [21900204](https://pubmed.ncbi.nlm.nih.gov/21900204/)
49. Marbach D, Costello JC, Kueffner R, Vega NM, Prill RJ, Camacho DM, et al. Wisdom of crowds for robust gene network inference. *Nat Methods*. 2012; 9(8):796–804. doi: [10.1038/nmeth.2016](https://doi.org/10.1038/nmeth.2016) PMID: [22796662](https://pubmed.ncbi.nlm.nih.gov/22796662/)
50. Hurley DG, Cursons J, Wang YK, Budden DM, Print CG, Crampin EJ. NAIL, a software toolset for inferring, analyzing and visualizing regulatory networks. *Bioinformatics*. 2015; 31(2):277–278. doi: [10.1093/bioinformatics/btu612](https://doi.org/10.1093/bioinformatics/btu612) PMID: [25246431](https://pubmed.ncbi.nlm.nih.gov/25246431/)
51. Saez-Rodriguez J, Goldsipe A, Muhlich J, Alexopoulos LG, Millard B, Lauffenburger DA, et al. Flexible informatics for linking experimental data to mathematical models via DataRail. *Bioinformatics*. 2008; 24(6):840–847. doi: [10.1093/bioinformatics/btn018](https://doi.org/10.1093/bioinformatics/btn018) PMID: [18218655](https://pubmed.ncbi.nlm.nih.gov/18218655/)
52. Steuer R, Kurths J, Daub C, Weise J, Selbig J. The mutual information: Detecting and evaluating dependencies between variables. *Bioinformatics*. 2002; 18(suppl 2):S231–S240. PMID: [12386007](https://pubmed.ncbi.nlm.nih.gov/12386007/)
53. Meyer PE, Lafitte F, Bontempi G. minet: A R/Bioconductor Package for Inferring Large Transcriptional Networks Using Mutual Information. *BMC Bioinform*. 2008; 9(1):461. doi: [10.1186/1471-2105-9-461](https://doi.org/10.1186/1471-2105-9-461) PMID: [18959772](https://pubmed.ncbi.nlm.nih.gov/18959772/)
54. Wittmann DM, Krumsiek J, Saez-Rodriguez J, Lauffenburger DA, Klamt S, Theis FJ. Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. *BMC Syst Biol*. 2009; 3. doi: [10.1186/1752-0509-3-98](https://doi.org/10.1186/1752-0509-3-98) PMID: [19785753](https://pubmed.ncbi.nlm.nih.gov/19785753/)
55. Krumsiek J, Poelsterl S, Wittmann DM, Theis FJ. Odepy—From discrete to continuous models. *BMC Bioinform*. 2009; 3(1):1–21.
56. Henriques D, Rocha M, Saez-Rodriguez J, Banga JR. Reverse engineering of logic-based differential equation models using a mixed-integer dynamic optimization approach. *Bioinformatics*. 2015; 31(18):2999–3007. doi: [10.1093/bioinformatics/btv314](https://doi.org/10.1093/bioinformatics/btv314) PMID: [26002881](https://pubmed.ncbi.nlm.nih.gov/26002881/)
57. Egea JA, Marti R, Banga JR. An evolutionary method for complex-process optimization. *Computers & Operations Research*. 2010; 37(2):315–324.
58. Egea JA, Henriques D, Cokelaer T, Villaverde AF, MacNamara A, Danciu DP, et al. MEIGO: an open-source software suite based on metaheuristics for global optimization in systems biology and bioinformatics. *BMC Bioinform*. 2014; 15(1):1–9. doi: [10.1186/1471-2105-15-136](https://doi.org/10.1186/1471-2105-15-136) PMID: [24885957](https://pubmed.ncbi.nlm.nih.gov/24885957/)
59. Hindmarsh A, Brown P, Grant K, Lee S, Serban R, Shumaker D, et al. SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. *ACM Trans Math Software*. 2005; 31(3):363–396. doi: [10.1145/1089014.1089020](https://doi.org/10.1145/1089014.1089020)
60. Domingos P. The role of occam's razor in knowledge discovery. *Data Min Knowl Discov*. 1999; 3(4):409–425.
61. Sunnaker M, Zamora-Sillero E, Dechant R, Ludwig C, Busetto AG, Wagner A, et al. Automatic Generation of Predictive Dynamic Models Reveals Nuclear Phosphorylation as the Key Msn2 Control Mechanism. *Sci Signal*. 2013; 6(277):ra41. doi: [10.1126/scisignal.2003621](https://doi.org/10.1126/scisignal.2003621) PMID: [23716718](https://pubmed.ncbi.nlm.nih.gov/23716718/)
62. De La Maza M, Yuret D. Dynamic hill climbing. *AI expert*. 1994; 9:26–26.
63. Burnham KP, Anderson DR. Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media; 2003. doi: [10.1007/b97636](https://doi.org/10.1007/b97636)
64. Gneiting T, Raftery AE. Weather forecasting with ensemble methods. *Science*. 2005; 310(5746):248–249. doi: [10.1126/science.1115255](https://doi.org/10.1126/science.1115255) PMID: [16224011](https://pubmed.ncbi.nlm.nih.gov/16224011/)
65. Dagum L, Menon R. OpenMP: An industry standard API for shared-memory programming. *IEEE Comput Sci Eng*. 1998; 5(1):46–55. doi: [10.1109/99.660313](https://doi.org/10.1109/99.660313)
66. Terfve C, Cokelaer T, Henriques D, MacNamara A, Goncalves E, Morris MK, et al. CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC Syst Biol*. 2012; 6(1):1–14. doi: [10.1186/1752-0509-6-133](https://doi.org/10.1186/1752-0509-6-133) PMID: [23079107](https://pubmed.ncbi.nlm.nih.gov/23079107/)
67. Huang C, Ferrell J. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci USA*. 1996; 93(19):10078–10083. doi: [10.1073/pnas.93.19.10078](https://doi.org/10.1073/pnas.93.19.10078) PMID: [8816754](https://pubmed.ncbi.nlm.nih.gov/8816754/)
68. MacNamara A, Henriques D, Saez-Rodriguez J. Modeling signaling networks with different formalisms: A preview. *Methods Mol Biol*. 2013; 1021:89–105. doi: [10.1007/978-1-62703-450-0\\_5](https://doi.org/10.1007/978-1-62703-450-0_5) PMID: [23715981](https://pubmed.ncbi.nlm.nih.gov/23715981/)
69. Chen WW, Schoeberl B, Jasper PJ, Niepel M, Nielsen UB, Lauffenburger DA, et al. Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol Syst Biol*. 2009; 5(1):239. doi: [10.1038/msb.2008.74](https://doi.org/10.1038/msb.2008.74) PMID: [19156131](https://pubmed.ncbi.nlm.nih.gov/19156131/)
70. Hill SM, Heiser LM, Cokelaer T, Unger M, Nesser NK, Carlin DE, et al. Inferring causal molecular networks: empirical assessment through a community-based effort. *Nat Methods*. 2016;. doi: [10.1038/nmeth.3773](https://doi.org/10.1038/nmeth.3773) PMID: [26901648](https://pubmed.ncbi.nlm.nih.gov/26901648/)

71. Johnson G, Lapadat R. Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. *Science*. 2002; 298(5600):1911–1912. doi: [10.1126/science.1072682](https://doi.org/10.1126/science.1072682) PMID: [12471242](https://pubmed.ncbi.nlm.nih.gov/12471242/)
72. Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. In: Proceedings of the 23rd international conference on Machine learning; 2006. p. 233–240.
73. Meyer P, Marbach D, Roy S, Kellis M. Information-Theoretic Inference of Gene Networks Using Backward Elimination. In: BioComp'10, International Conference on Bioinformatics and Computational Biology; 2010. p. 700–705.
74. Song MJ, Lewis CK, Lance ER, Chesler EJ, Yordanova RK, Langston MA, et al. Reconstructing generalized logical networks of transcriptional regulation in mouse brain from temporal gene expression data. *EURASIP Journal on Bioinformatics and Systems Biology*. 2009; 2009(1):1. doi: [10.1155/2009/545176](https://doi.org/10.1155/2009/545176) PMID: [19300527](https://pubmed.ncbi.nlm.nih.gov/19300527/)
75. Túrei D, Korcsmáros T, Saez-Rodriguez J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nature methods*. 2016; 13(12):966–967. doi: [10.1038/nmeth.4077](https://doi.org/10.1038/nmeth.4077) PMID: [27898060](https://pubmed.ncbi.nlm.nih.gov/27898060/)
76. 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(11):e1000558. doi: [10.1371/journal.pcbi.1000558](https://doi.org/10.1371/journal.pcbi.1000558) PMID: [19911077](https://pubmed.ncbi.nlm.nih.gov/19911077/)
77. Banga JR, Balsa-Canto E. Parameter estimation and optimal experimental design. *Essays in biochemistry*. 2008; 45:195–210. doi: [10.1042/BSE0450195](https://doi.org/10.1042/BSE0450195) PMID: [18793133](https://pubmed.ncbi.nlm.nih.gov/18793133/)