



Dondelinger, F., Rogers, S. , Filippone, M. , Cretella, R., Palmer, T.,  
Smith, R., Millar, A., and Husmeier, D. (2012) *Parameter inference in  
mechanistic models of cellular regulation and signalling pathways using  
gradient matching*. In: WCSB2012 - 9th International Workshop on  
Computational Systems Biology, 4-6 Jun 2012, Ulm, Germany.

<http://eprints.gla.ac.uk/65591/>

Deposited on: 20 June 2012

# PARAMETER INFERENCE IN MECHANISTIC MODELS OF CELLULAR REGULATION AND SIGNALLING PATHWAYS USING GRADIENT MATCHING

Frank Dondelinger<sup>1</sup>, Simon Rogers<sup>2</sup>, Maurizio Filippone<sup>2</sup>, Roberta Cretella<sup>3</sup>, Tim M. Palmer<sup>3</sup>, Robert W. Smith<sup>4</sup>, Andrew J. Millar<sup>4</sup> and Dirk Husmeier<sup>1</sup>

<sup>1</sup> School of Mathematics and Statistics, University of Glasgow

<sup>2</sup> School of Computing Science, University of Glasgow

<sup>3</sup> Institute of Cardiovascular and Medical Sciences, University of Glasgow  
University of Glasgow, Glasgow G12 8QQ, Scotland

<sup>4</sup> SynthSys, University of Edinburgh, Edinburgh EH9 3JD

Corresponding author: Dirk.Husmeier@glasgow.ac.uk

## ABSTRACT

A challenging problem in systems biology is parameter inference in mechanistic models of signalling pathways. In the present article, we investigate an approach based on gradient matching and nonparametric Bayesian modelling with Gaussian processes. We evaluate the method on two biological systems, related to the regulation of *PIF4/5* in *Arabidopsis thaliana*, and the JAK/STAT signal transduction pathway.

## 1. INTRODUCTION

A central problem in computational systems biology is the formulation of a consistent mechanistic model description of signalling pathways and molecular processes of cellular regulation. While there have been several attempts at describing these processes at a qualitative level, proper statistical inference is a much more challenging problem. The approach is based on minimising the discrepancy between measured data, e.g. related to the abundance profiles of some molecular components, and their simulated values. This discrepancy is related to some metric, which can be shown to be defined by the assumed noise model in terms of a maximum likelihood approach to inference. For instance, minimising the root mean square deviation between measured and simulated data is equivalent to maximizing the likelihood under the assumption of white additive Gaussian noise. For further details, see [1].

The practical difficulties with this approach are two-fold. First, the likelihood landscape is typically rugged and multimodal, which calls for some form of annealing scheme. Second, each parameter adaptation requires a numerical solution of the differential equations (ODEs), which is computationally expensive and hence limits the number of maximum likelihood (ML) optimization steps or Markov chain Monte Carlo (MCMC) sampling steps that can be carried out at reasonable computational costs.

A potential solution to this problem is the approach of gradient matching. The idea is that rather than aiming to explicitly solve the ODEs, we seek to minimise the discrepancy between the gradients inferred from the slope

of the interpolant and those predicted from the coupled system of ODEs. The former is defined by the regression model and depends on some smoothness or regularization parameters. The latter is defined by the system of coupled ODEs and depends on its parameters, whose determination is the ultimate objective of inference. Earlier approaches pursued a two-step approach, in which first the interpolant was inferred, and then the ODE parameters were inferred by minimizing the discrepancy between the time derivatives predicted from the ODEs and those predicted from the slope of the interpolant [2]. The disadvantage of this approach is that the result of parameter inference critically hinges on the quality of the interpolation scheme, which once completed is kept fixed. A better approach, first suggested in [3], is to allow for some feedback mechanism by which the system of ODEs can act back on the interpolation scheme. For instance, the system of ODEs might only match the slopes of the interpolant for implausible or a priori unlikely parameter configurations, in which case the interpolant should be adjusted. Hence, in order to be viable, the mismatch between the time derivative predicted from the slope of the interpolant and the one obtained from the ODEs should be systematically reduced in an iterative loop, whereby both the ODEs and the smoothness parameters of the regression are adapted simultaneously. A variation of this approach was presented in [4], where the authors employed parallel tempering using the data-smoothing hyperparameter, which improves sampling efficiency.

## 2. METHOD

Our approach is based on non-parametric Bayesian regression with Gaussian processes, following Calderhead et al. in [5]. Their approach draws on the fact that the derivative of a Gaussian process is a Gaussian process again. This renders Gaussian process regression a natural tool for the double objective of nonlinear data interpolation and gradient matching along the lines outlined in the previous section. The result is a hierarchical Bayesian model which allows for parameter inference. We will have a system

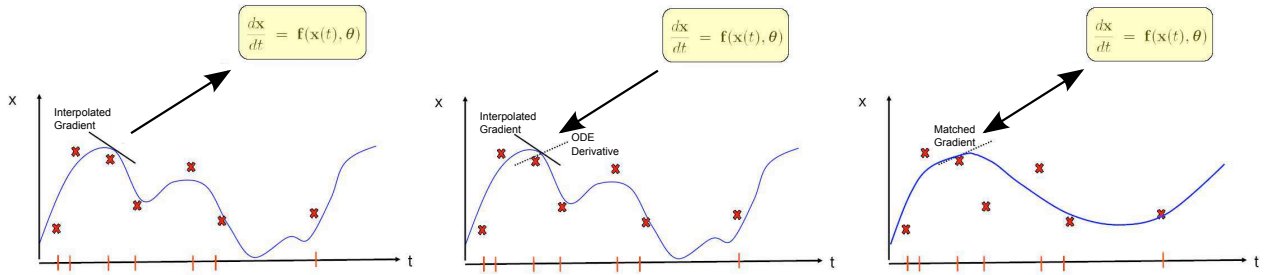


Figure 1. Outline of the approach of gradient matching. Left panel: the slope of the interpolant informs the inference of the ODE parameters  $\theta$ . Centre panel: the derivative  $\frac{dx}{dt}$  predicted based on the updated parameters from the ODEs  $f(x(t), \theta)$  is fed back to the interpolation scheme. Right panel: by iterating the above two steps, the mismatch between the time derivative predicted from the slope of the interpolant and the one obtained from the ODEs is systematically reduced, thereby adapting both the ODE parameters and the smoothness parameters of the regression simultaneously.

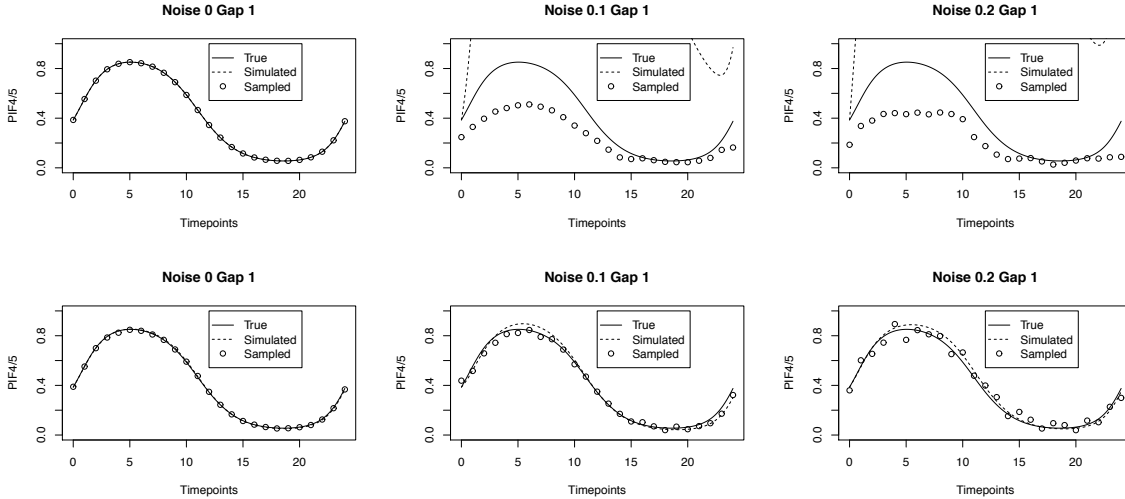


Figure 2. *PIF4/5* expression sampled every hour, with varying observational noise. We show the true (noiseless) expression values, versus expression values simulated from the ODE system using mean sampled parameters. We also show the values of *PIF4/5* sampled from the GP model. Top: Calderhead et al. model [5]. Bottom: Improved model. Note that the dashed line in the top panel is out of scale.

of coupled ODEs, which predict the time derivative; the ODE parameters are adapted so as to minimize the deviation from the time derivatives predicted with the Gaussian process. A novel aspect of our approach, which constitutes an important improvement on the method proposed in [5], is the fact that the smoothness hyper-parameters of the Gaussian process are adapted simultaneously along with the parameters of the ODEs. A mathematical description of this approach is beyond the scope and page limit of this paper and will be presented elsewhere. A schematic representation is depicted in Figure 1.

### 3. DATA

We test our approach on two biological systems; gene regulation in the circadian clock of *Arabidopsis thaliana*, and receptor signal transduction in the JAK/STAT pathway.

#### 3.1. The *PIF4/5* model

We apply our GP parameter inference method to a model for gene regulation of genes *PIF4* and *PIF5* by *TOC1* in

the circadian clock gene regulatory network of *Arabidopsis thaliana*. The overall network is represented by the Locke 2-loop model [6], with fixed parameters set following [7]. Only the parameters involved in regulation of *PIF4* and *PIF5* are inferred. As the expression profiles are very similar, we simplify the model to represent genes *PIF4* and *PIF5* as a combined gene *PIF4/5*. We are interested in the promoter strength  $s$ , the rate constant  $K_d$  and Hill coefficient  $h$  of the regulation by *TOC1*, and the degradation rate  $d$  of the *PIF4/5* mRNA. The regulation is represented by the following ODE:

$$\frac{d[PIF4/5]}{dt} = s \cdot \frac{K_d^h}{K_d^h + [TOC1]} - d \cdot [PIF4/5] \quad (1)$$

where  $[PIF4/5]$  and  $[TOC1]$  represent the concentration of *PIF4/5* and *TOC1*, respectively, and  $t$  represents time.

#### 3.2. The JAK/STAT pathway

We analyse a model for interleukin-6 signalling (IL-6) in vascular endothelial cells. IL-6 binds to a receptor on the

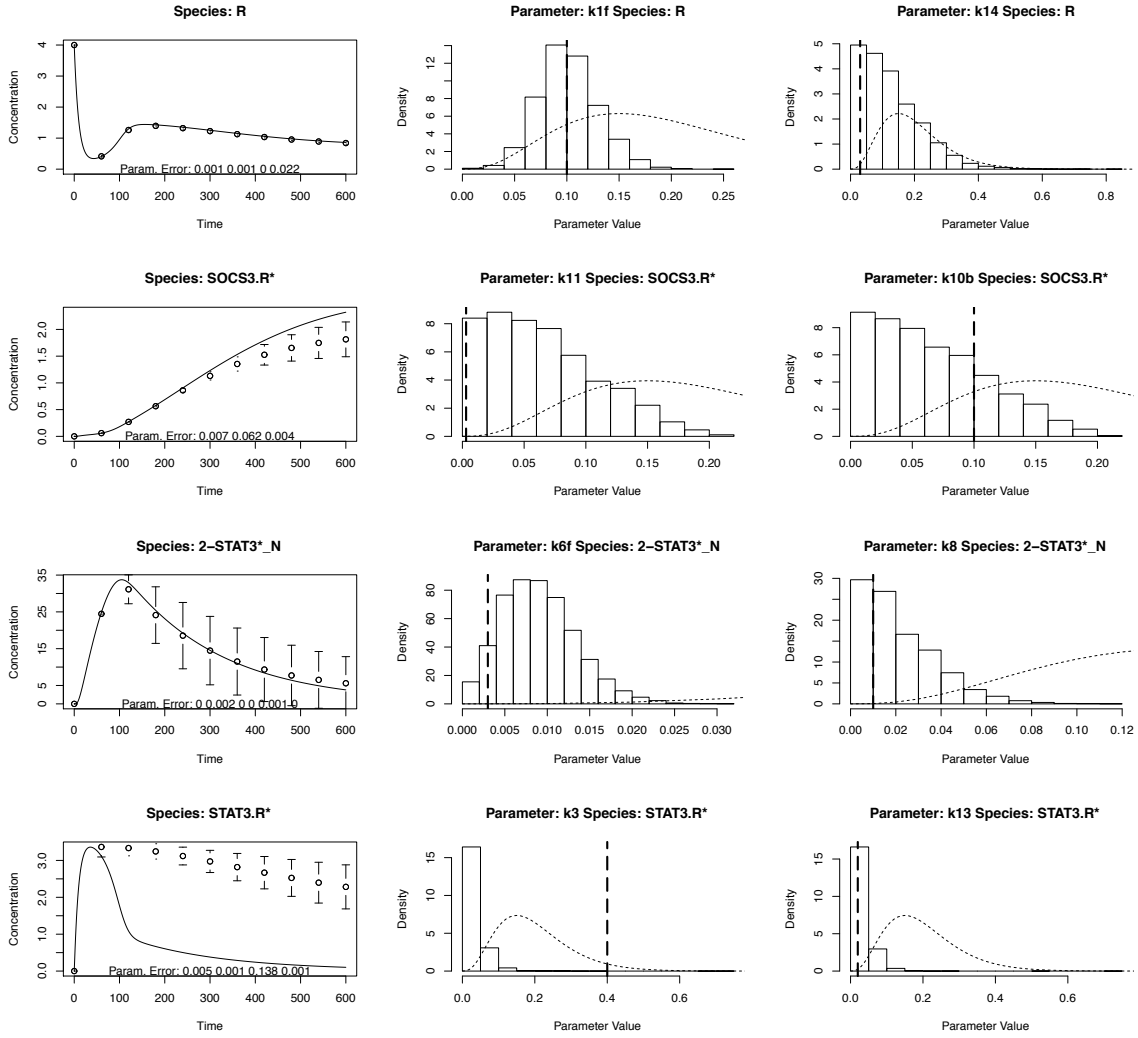


Figure 3. JAK/STAT pathway results. Row 1:  $R$ . Row 2:  $SOCS3.R^*$ . Row 3:  $2STAT3^*_N$ . Row 4:  $STAT3.R^*$ . See (2) for the equations. True (solid line) and inferred (points with error bars) concentrations for each species. The inferred values were obtained by drawing parameter samples from the posterior and forward-simulating the ODE system with these parameters. The error bars show one standard deviation. Histograms show the distribution of sampled parameters, the dashed line represents the gamma prior on the parameters in the hierarchical Bayesian model, and the horizontal line indicates the true value of the parameter.

plasma membrane, activating the JAK/STAT pathway [8]. The receptor is phosphorylated, creating docking sites for signalling molecules like  $STAT3$ .  $STAT3$  binds to the phosphorylated receptor and is phosphorylated itself. Phosphorylated  $STAT3$  molecules are released from the receptor, dimerize and then migrate to the nucleus to trigger mRNA transcription of target proteins like  $SOCS3$ .  $SOCS3$  acts as a feedback mechanism for the signalling pathway: it binds to active receptors to prevent  $STAT3$  activation and to provide a signal termination. The model we consider is a complex system comprising 13 species and 19 parameters. The dynamics of the system are described by mass-action kinetics, with non-linear interactions among species. Under the assumption of full observation of all species, we can decompose the system into 13 subsystems, one per species. This simplifies inference, and allows us to investigate the local identifiability of parameters in this model. For space reasons, we only reproduce the subset of equations consisting of the species in Figure 3 below; the full system will be presented in a future paper.

$$\begin{aligned}
 \frac{d[R]}{dt} &= -k_1^f[R] + k_1^b[R^*] + k_{11}[SOCS3.R^*] + k_{14}[SOCS3.STAT3.R^*] \\
 \frac{d[SOCS3.R^*]}{dt} &= -k_{11}[SOCS3.R^*] + k_{10}^f[SOCS3][R^*] - k_{10}^b[SOCS3.R^*] \\
 \frac{d[2STAT3^*_N]}{dt} &= -k_6^f[2STAT3^*_N][P300] + k_6^b[2STAT3^*_N.P300] + k_5[2STAT3^*_C] - k_7[2STAT3^*_N] + k_8[2STAT3^*_N.P300] \\
 \frac{d[STAT3.R^*]}{dt} &= k_2^f[STAT3][R^*] - k_2^b[STAT3.R^*] - k_{13}[SOCS3][STAT3.R^*] - k_3[STAT3.R^*]
 \end{aligned} \tag{2}$$

Here the square brackets,  $[\cdot]$ , indicate concentrations,

and the  $k_i$  are kinetic parameters.

#### 4. RESULTS

We simulated data from the *PIF4/5* model using the parameters ( $s = 1, K_d = 0.46, h = 2, d = 1$ ), adding observation noise with standard deviation in  $\{0, 0.1, 0.2\}$ . The parameter inference was done by sampling from the posterior using MCMC with the model from [5], as well as using MCMC with the improved model from Section 2. The time period of the measurements is 24 hours, and the interval (gap) between observed points is 1 hour, resulting in 25 datapoints. Figure 2 shows the noiseless concentrations sampled from the GP model, and simulated *PIF4/5* concentrations from the true parameters and from the sampled parameters, for MCMC with the Calderhead et al. model in [5], and MCMC with the improved model.

For the JAK/STAT pathway, we generated simulation data for 600 timepoints with parameter values that gave realistic behaviour for the different species. The data was sampled at intervals of 60 seconds from time zero, making 11 timepoints in total. Due to space restrictions, we only present results for a subset of species. Figure 3 shows the results for the inactive receptor  $R$ , for the  $SOCS3/R^*$  complex (where  $R^*$  is the activated receptor), for the activated 2- $STAT3$  complex in the nucleus and for the  $STAT3/R^*$  complex.

#### 5. DISCUSSION

Our results demonstrate that gradient matching is a promising approach for parameter estimation in ODE systems. We have demonstrated that our improvement on the method described in [5] produces better results in the *PIF4/5* system (Figure 2) in the presence of observation noise.

Our application to the JAK/STAT signalling pathway shows that the approach we have taken is promising, but that some challenges remain. It also allows us to draw some inferences about the properties of this system. For some species, such as  $R$  and  $SOCS3.R^*$  in rows 1 and 2 of Figure 3, we obtain very good predictions for the concentrations of the species, as well as giving a good estimate for some of the parameters, such as  $k_{1f}$  and  $k_{11}$ . However, we can see that there is a lot of uncertainty about the inferred parameters, even though the species concentrations are predicted quite well. This points to a problem with lack of identifiability in parameter space, related to ridges in the likelihood. For example, if the rate limiting chemical kinetics depend on the ratio of two kinetic constants, then the confidence or credible intervals of the individual parameters may be large without that being reflected by the prediction uncertainty, as long as the posterior distribution of the ratio is peaked.

We notice that for species 2- $STAT3^*_N$  (row 3) in Figure 3, the posterior probability mass for the parameters is in the tail of the prior distribution. This implies two things. First, the data are informative with respect to the inference of some parameters. Second, the prior has not been chosen very well for this example, and should be chosen more appropriately. Finally, the bottom row of Figure 3 shows a mismatch between the true and predicted signal of  $STAT3.R^*$ . This points to an intrinsic difficulty with

the gradient matching approach. The mismatch is mainly due to a short transient region around time point 100. In the following time segment the gradient is well matched, whereas the signal itself shows a strong deviation. This indicates that in scenarios of this form, the likelihood landscape for the explicit solution of the ODEs differs systematically from the one obtained with gradient matching.

In conclusion, our gradient matching approach demonstrates good predictions for realistic biological systems, and promises to become a useful tool for parameter estimation in system biology. However, some challenges related to unidentifiable parameters and non-stationary signals still remain.

#### 6. ACKNOWLEDGEMENTS

This work was funded by Bridging-the-Gap EPSRC grant 59229/1. R.W.S. was supported by BBSRC (BB/F59011/1, BB/F005237/1). SynthSys is a Centre for Integrative and Systems Biology partly supported by BBSRC and EPSRC (BB/G019621).

#### 7. REFERENCES

- [1] N. D. Lawrence, M. Girolami, M. Rattray, and G. Sanguinetti, Eds., *Learning and Inference in Computational Systems Biology*, MIT Press, 2010.
- [2] A.A. Poyton, M.S. Varziri, K.B. McAuley, P.J. McLellan, and J.O. Ramsay, "Parameter estimation in continuous-time dynamic models using principal differential analysis," *Computers & chemical engineering*, vol. 30, no. 4, pp. 698–708, 2006.
- [3] J.O. Ramsay, G. Hooker, D. Campbell, and J. Cao, "Parameter estimation for differential equations: a generalized smoothing approach," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 5, pp. 741–796, 2007.
- [4] D. Campbell and R.J. Steele, "Smooth functional tempering for nonlinear differential equation models," *Statistics and Computing*, pp. 1–15, 2012.
- [5] B. Calderhead, M. Girolami, and N.D. Lawrence, "Accelerating Bayesian inference over nonlinear differential equations with Gaussian processes," *Neural Information Processing Systems (NIPS)*, vol. 22, 2008.
- [6] J.C.W. Locke, M.M. Southern, L. Kozma-Bognar, V. Hibberd, P.E. Brown, M.S. Turner, and A.J. Millar, "Extension of a genetic network model by iterative experimentation and mathematical analysis," *Molecular Systems Biology*, vol. 1, pp. (online), 2005.
- [7] A. Pokhilko, S.K. Hodge, K. Stratford, K. Knox, K.D. Edwards, A.W. Thomson, T. Mizuno, and A.J. Millar, "Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock model," *Molecular systems biology*, vol. 6, no. 1, 2010.
- [8] P.C. Heinrich, I. Behrmann, G. Müller-Newen, F. Schaper, and L. Graeve, "Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway," *Biochemical Journal*, vol. 334, no. Pt 2, pp. 297, 1998.