Neutral Network Models for Biochemical Systems

Marco Vilela [1,2], Susana Vinga[3,7], Marco A. Grivet Mattoso Maia[3], Eberhard O. Voit[4], Jonas S. Almeida[5,1]

1: Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Rua da Quinta Grande 6, Apartado 127, 2780-156 Oeiras, Portugal
2: Chemistry Department, Texas A&M University, College Station, TX
3: Instituto de Engenharia de Sistemas e Computadores: Investigação e Desenvolvimento (INESC-ID), R. Alves Redol 9, 1000-029 Lisboa, Portugal
4: Pontifícia Universidade Católica do Rio de Janeiro, Centro Técnico-Científico, Centro de Estudo em Telecomunicações. Rua Marquês de São Vicente 225, Gávea. 22453-900 - Rio de Janeiro, RJ - Brasil
5: Integrative BioSystems Institute and Dept. Biomedical Engineering, Georgia Institute of Technology and Emory University, 313 Ferst Drive, Atlanta, GA 30332, USA
6: Dept. Bioinformatics and Computational Biology, University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA.
7: Faculdade Ciências Médicas - Universidade Nova de Lisboa (FCM-UNL), C Mártires Pátria 130, 1169-056 Lisboa, Portugal

mvilela@mathbiol.org

Overview

Background

The major difficulty in modeling biological systems from multivariate time series is the identification of parameter sets that endow a model with dynamical behaviors sufficiently similar to the experimental data. Directly related to this parameter estimation issue is the task of identifying the structure and regulation of ill-characterized systems.

Results

We propose a method for the identification of admissible parameter sets of canonical S-systems from biological time series. The method is based on a Monte Carlo process that is combined with an improved version of our previous parameter optimization algorithm [3]. The method maps the parameter space into the network space, which characterizes the connectivity among components, by creating an ensemble of decoupled S-system models that imitate the dynamical behavior of the time series with sufficient accuracy. The concept of sloppiness is revisited in the context of these S-system models with an exploration not only of different parameter sets that produce similar dynamical behaviors but also different network topologies that yield dynamical similarity.

Conclusions

The proposed parameter estimation methodology was applied to actual time series data from the glycolytic pathway of the bacterium *Lactococcus lactis* [2] and led to ensembles of models with different network topologies. In parallel, the parameter optimization algorithm was applied to the same dynamical data upon imposing a pre-specified network topology derived from prior biological knowledge, and the results from both strategies were compared. The results suggest that the proposed method may serve as a powerful exploration tool for testing hypotheses and the design of new experiments.

Methods

$$S_i = \alpha_i \prod_{j=1}^{n} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n} X_j^{h_{ij}} \qquad \text{S-system equations [1]}$$

$$\mathcal{L} = \begin{bmatrix} \ln(X_1(t_1)) & \cdots & \ln(X_j(t_1)) & \cdots & \ln(X_n(t_1)) \\ \ln(X_1(t_2)) & \cdots & \ln(X_j(t_2)) & \cdots & \ln(X_n(t_2)) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \ln(X_1(t_M)) & \cdots & \ln(X_j(t_M)) & \cdots & \ln(X_n(t_M)) \end{bmatrix} \qquad \psi_i = \begin{bmatrix} \psi_{i1} \\ \psi_{i2} \\ \psi_{i3} \\ \cdots \\ \psi_{iM} \end{bmatrix}$$

ES is Eigenspace of W

$$\ln\left(\alpha_i \prod_{j=1}^{n} X_j^{g_{ij}}\right) = ES\psi \qquad \ln\left(\beta_i \prod_{j=1}^{n} X_j^{h_{ij}}\right) = ES\delta$$

$$S_i = \alpha_i \prod_{j=1}^{n} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n} X_j^{h_{ij}}$$

$$S_i = e^{ES\psi} - e^{ES\delta}$$

$$F = \log\left[\left(S_i - \hat{S}_i(\delta)\right)^T \left(S_i - \hat{S}_i(\delta)\right)\right]$$

$$\frac{\partial \psi}{\partial \delta} = ES^{+}\left[\left(\left(S_i + e^{ES\delta}\right)^{-1} \circ e^{ES\delta}\right) \delta ES\right]$$

$$\frac{\partial F}{\partial \delta} = 2\left(-e^{ES\psi} \circ \left(ES\frac{\partial \psi}{\partial \delta}\right) + e^{ES\delta} \circ ES\right)^T \left(S_i - \hat{S}_i(\delta)\right) \qquad \text{Levenberg-Marquardt algorithm}$$

Results

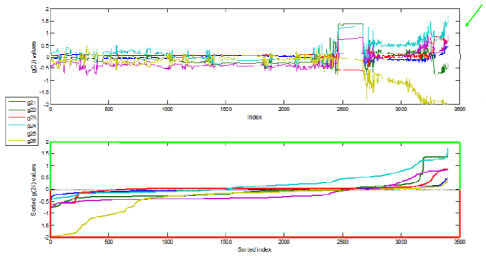


Figure 1 – Kinetic orders estimated for the G6P (X2) production term with potential inclusion of all system variables. A) The two indices refer to the two species considered by the interaction; for instance, g26 indicates the effect of X6 (acetate) on G6P production. B) The same parameter sets as shown in A, but ordered individually by magnitude, showing the possible variation admissible each parameter. The light green y-axis represents the region of parameter space with possible activation interaction; analogously, the light red y-axis represents the region of possible inhibition interaction.

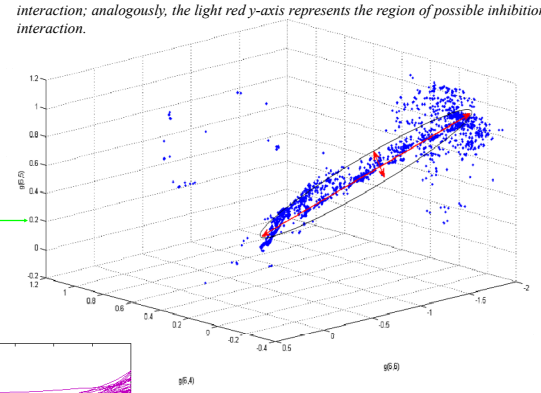
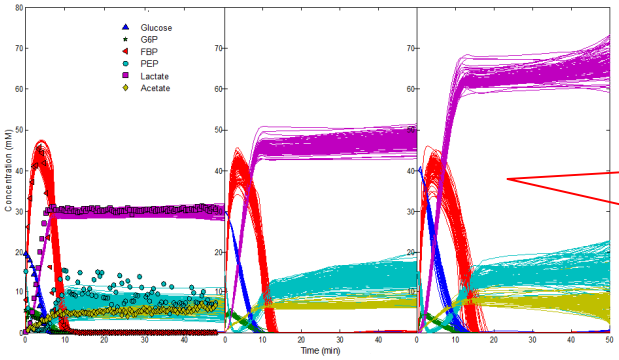


Figure 2 – Quantification of sloppiness for one of the equations of the Lactococcus model. The figure shows the 3-D projection of the ellipsoid that represents the region of the parameter space that produces similar dynamical behaviors. The arrows show the direction of the sloppy and stiff directions in the 3-D projection, corresponding respectively to eigenvectors with small and large eigenvalues of the Hessian matrix of the cost function.



Good generalization capability

New ensemble of models. *Systems integrated with different initial concentrations for glucose substrate (20, 30 and 40 mM)*

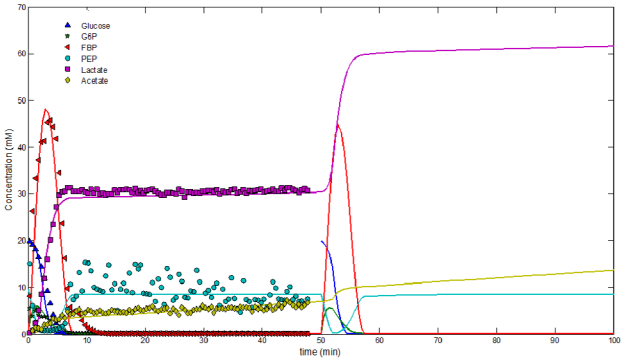


Figure 5 – Glucose double pulse simulation. A second 20 mM glucose pulse was supplied to the system after 50 min, resulting in the further accumulation of lactate and acetate.

Acknowledgments

The authors acknowledge partial support by project DynaMo (PTDC/EEA-ACR/69530/2006; S. Vinga, PI) from the Portuguese Science Foundation (FCT) and the INESC-ID. This research was supported by the National Science Foundation's Division of Biological Infrastructure (DBI) and Directorate for Computer & Information Science & Engineering (CISE). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsoring institutions.

References

[1] Voit EO: **Computational analysis of biochemical systems : a practical guide for biochemists and molecular biologists**. Cambridge ; New York: Cambridge University Press; 2000
[2] Ramos A, Neves AR, Santos H: **Metabolism of lactic acid bacteria studied by nuclear magnetic resonance**. *Antonie Van Leeuwenhoek* 2002, **82**(1-4):249-261
[3] Vilela M, Chou IC, Vinga S, Vasconcelos AT, Voit EO, Almeida JS: **Parameter optimization in S-system models**. *BMC Syst Biol* 2008, **2**:35
[4] Ciliberti S, Martin OC, Wagner A: **Robustness can evolve gradually in complex regulatory gene networks with varying topology**. *PLoS Comput Biol* 2007, **3**(2):e15