Eberhard O. Voit, PhD
Professor and Georgia Research Alliance Eminent Scholar in Systems Biology
Wallace H Coulter Department of Biomedical Engineering
Georgia Institute of Technology

Model Identification in Computational Systems Biology

Whether a systems biological study attempts to capture the dynamics of a large, complex system or aims to understand the design principles of a rather small system, the key component of the analysis is a mathematical model. For the construction of such a model, it is often quite easy to determine which processes within a system increase or decrease a given variable. By contrast, there are essentially no objective guidelines for determining the optimal mathematical format for each process description. This conundrum raises the question of whether a suitable data analysis can possibly suggest appropriate functions.

Over the past decade, time series data have become available in biology at an increasing rate. The trend is to be welcomed, as these data contain enormous information, which however is implicit and needs to be extracted with computational means. Time series data are particularly beneficial for analyses of metabolic pathway systems, because these are strongly constrained by stoichiometric and other intrinsic features. These constraints effectively bound the space of admissible parameter values that need to be specified in order to translate the pathway system into a computable structure. Even within these bounds, the estimation of suitable parameter values is a major bottleneck in the harvesting of information from time series data. The overriding quality criterion for this step is usually the sum of squared residual errors between data and model. While this criterion is a natural starting point, there are various instances where a good fit alone is not sufficient. I will discuss several examples of such instances. Beyond the estimation of parameter values, the greater challenge lies in the fact that the most appropriate functional forms for describing biological processes are not even known when a system is to be estimated. This structural uncertainty clearly complicates any estimation strategy. I will show that the problem can be ameliorated to some degree if the right types of time series data and computational methods are available.

Voit, E.O.: What if the fit is unfit? Criteria for biological systems estimation beyond residual errors. In: M. Dehmer, F. Emmert-Streib and A. Salvador (Eds.): Applied Statistics for Biological Networks. J. Wiley and Sons, New York, pp. 183-200, 2011.

Goel, G., I-C. Chou, and E.O. Voit: System estimation from metabolic time series data. *Bioinformatics* 24, 2505-2511, 2008.

Iwata, M., F. Shiraishi, and E.O. Voit: Coarse but efficient identification of metabolic pathway systems. *Int. J. Syst. Biol.*4(1), 57-72, 2013.

Voit, E.O.: Characterizability of metabolic pathway systems from time series data. *Math. Biosc.* 246(2):315-25, 2013.