Unraveling the dynamics in genetic networks from gene expression data

J. Omony[1], L.H. de Graaff[2], G. van Straten[1] and A.J.B. van Boxtel[1]
[1]System and Control group, Wageningen University,
P.O. Box 17, 6700 AA Wageningen, The Netherlands.
[2]Laboratory of Microbiology, Wageningen University,
e-mail: jimmy.omony@wur.nl

Abstract. Unraveling complex biological networks holds the key to understanding numerous biological phenomena such as disease traits, plant and animal growth and cancer development e.t.c. Many proposed network reconstruction approaches still have low accuracy values and do not capture detailed dynamics of a given underlying network structures.

In this study we propose a method to unravel genetic networks for a regulon system using the differential equations approach. Because gene expression data is subject to considerable amount of noise, stochastic modeling approaches (which are sensitive to noise) are often preferred to the deterministic modeling approach. The choice of identification approach depends on the data attributes and quality. A perturbed system with multiple target genes was considered and network identification performed. The performance of the identification criteria was the assessed. The Euler derivatives and polynomial smoothing curves were used to approximate the transcription rates and linear regression models were fitted to the data. Statistically significant transcription factors were selected during the identification process.

Using smoothing did not indicate a clear-cut difference from the results from non-interpolated data with respect to the identification performance (fig. 1). This is because of the high system sensitivity to external perturbations. The reliability of the regression techniques were limited by network size and noisy data. Estimation of transcription rates yields generates more noise which in turn leads to increased chances of wrong identification. An increase in network size and data noise levels leads to the regression assumptions being compromised, hence, poorer identifications.
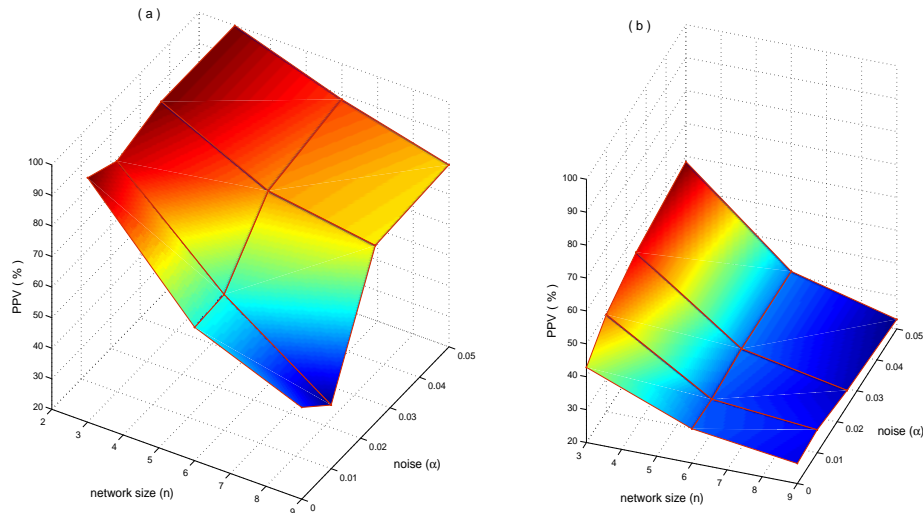

Figure 1: (a) Variation of positive predictive value (PPV) with relative noise α and network size *n* for non interpolated data using Euler's transcription rate approximation. (b) PPV versus noise and network size but with polynomial transcription rate approximation.