ACCURATE INITIAL CONDITION DETERMINATION IN NEXT-GENERATION SYSTEMS BIOLOGY MODELS USING MACHINE LEARNING TECHNIQUES

A. Karakoltizids[1,2], S. Karakitsios[1,2], D. A. Sarigiannis[1,2,3,4]*

[1]HERACLES Research Center on the Exposome and Health, Center for Interdisciplinary Research and Innovation, Aristotle University of Thessaloniki, Thessaloniki, Greece

[2]Environmental Engineering Laboratory, Department of Chemical Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece

[3]University School of Advanced Study IUSS, Pavia, Italy

[4]National Hellenic Research Foundation, Athens, Greece

*(*sarigiannis@auth.gr)*

ABSTRACT

The development of systems biology mathematical models necessitates the determination of numerous parameters, such as turnover numbers, Michaelis-Menten constants, and initial conditions for the developing system of differential equations. Considering the challenges and costs associated with experiments, coupled with ethical concerns regarding animal research, we introduce an in-silico approach based on machine learning and Quantitative Structure Activity Relationship (QSAR) to ascertain metabolite concentrations when experimental data is unavailable. Next Generation Systems Biology (NGSB) models are expansive mathematical models containing hundreds of differential equations. Addressing the challenge of defining initial conditions for the developing system of differential equations, we have devised an AI-based methodology. This approach determines metabolites' concentrations based on the concentrations of neighboring nodes in the pathway considering mass balance and iceberg modeling principles. Currently, the methodology utilizes pathways exclusively from the KEGG database, with plans to integrate other similar databases in future software versions. This methodology has diverse applications in biology, enabling the determination of initial concentrations for both endogenous metabolites and other compounds that are components of a unified network. It ensures that all differential equations in the constructed mathematical model possess an initial condition. Moreover, its utility extends to industrial and pharmaceutical settings, particularly in systems with a high number of freedom degrees, facilitating the determination of concentrations based on reactants and products.

KEYWORDS: Computational Systems Biology, Machine Learning models, QSARs, Mechanistic Models Initialization, Network Modelling

REFERENCES