A COMPREHENSIVE AI-BASED APPROACH TO PARAMETERIZATION OF NEXT-GENERATION SYSTEMS BIOLOGY MODELS

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

Next-generation systems biology models (NGSB models) are constructed on a large scale, incorporating hundreds of differential equations with millions of interactions and metabolic transformations. These models strive to closely approximate real-life scenarios by encompassing all endogenous metabolites, genes, and enzymes involved in a biological process, as well as those influenced by disruptors or other biological events. Developing such a mathematical model demands extensive expertise and accurate estimation of numerous kinetic parameters, including turnover numbers and Michaelis–Menten constants. Given that many naturally occurring turnover numbers remain unquantified from wet lab experiments, there is a need for either additional experiments or a robust methodology for their estimation. To address this, a dataset sourced from publicly available databases, BRENDA and SABIO RK, was introduced. Subsequently, training and test sets were created based on limited information from these databases. Deep Neural Network (DNN) models were then trained and finely optimized to estimate turnover numbers. The preprocessed integrated dataset includes essential components for each enzyme and reaction, incorporating enzyme sequences, molecular fingerprints, and other chemical attributes. This approach led to the development of a well-optimized model with the selection of appropriate parameters. The model can predict a turnover with an R^2 of 0.67. An online version of these models in the R-Shine format is anticipated to be available soon, providing users with the capability to implement their own applications.

KEYWORDS: Computational Systems Biology, Deep Learning models, QSARs, Michaelis-Menten Kinetics

REFERENCES