

Simulation tool for statistical application and comparison of multi-response multivariate estimators

Raju Rimal^{a,*}, Trygve Almøy^a, Solve Sæbø^b

^a*Faculty of Chemistry and Bioinformatics, Norwegian University of Life Sciences, Ås, Norway*

^b*Professor, Norwegian University of Life Sciences, Ås, Norway*

Abstract

While data science is battling to extract information from the enormous explosion of data, many estimators and algorithms are being developed for better prediction. Researchers and data scientists often introduce new methods and evaluate them based on various aspects of data. However, studies on the impact of/on a model with multiple response variables are limited. This study compares some newly-developed (envelope) and well-established (PLS, PCR) prediction methods based on real data and simulated data specifically designed by varying properties such as multicollinearity, the correlation between multiple responses and position of relevant principal components of predictors. This study aims to give some insight into these methods and help the researcher to understand and use them in further studies.

Keywords: model-comparison, multi-response, simrel

*Corresponding Author

Email addresses: `raju.rimal@nmbu.no` (Raju Rimal), `trygve.almoy@nmbu.no` (Trygve Almøy), `solve.sabo@nmbu.no` (Solve Sæbø)