DOI 10.18699/BGRSSB-2018-110

Development of nonlinear regression models of flowering time control by climatic factors in soybean and chickpea

K. Kozlov[1]*, L. Novikova[1, 2], I. Seferova[2], S. Nuzhdin[1, 3], M. Samsonova[1]
[1] *Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*
[2] *Federal Research Center the N. I. Vavilov All-Russian Institute of Plant Genetic Resources, St. Petersburg, Russia*
[3] *University of Southern California, Los Angeles, CA, USA*
** e-mail: kozlov_kn@spbstu.ru*

Key words: soybean, chickpea, flowering time, nonlinear regression, grammatical evolution

Motivation and Aim: Climate impact on agriculture is getting stronger due to combined effects of rising average temperatures, reduced water supply in dry regions and more frequent extreme weather conditions. Agriculture must intensify, become more sustainable, and possess greater resilience to pests and climate. Crucial to this effort are predictive models that connect agricultural traits to climatic factors.
Methods and Algorithms: The analytic representation of non-linear dependence of a agronomic traits on climatic factors is build using a formal approach called "Grammatical Evolution". This technique constructs N functions from "words" of length L according to the rules of a defined grammar. In this grammar a word represents expression that may encode predictor (X) or operation on expressions. The model is further build as a linear combination of N functions with LASSO approach [1]. The set of predictors is determined by minimization of approximation error with Differential evolution (DEEP). The model quality was assessed with determination coefficient R^2. Each model is characterized by the vector of fixed length that contains predictor indices and coefficients, thus the approach allows us to quantitatively compare the model structures.
Results: The method was applied to predict the flowering time in two datasets, namely 379 samples of 9 soybean accessions of different origin phenotyped at Pushkin VIR station in 1999–2013 [2] ($R^2 = 0.60$) and chickpea VIR landraces from Turkey and Ephiopia, phenotyped in Syria ($R^2 = 0.47$). The model vectors build for Turkey and Ephiopia chickpea varieties were uncorrelated (Pearson $r = 0.13$, $p = 0.4 > 0.05$) and their difference was statistically significant according to Mann-Whitney test ($p = 0.04 < 0.05$).
Conclusion: The flowering time models developed in this work are more accurate in comparison with earlier models [2] but possess more complex structure and contain hyperbolic functions. The results obtained with chickpea dataset showed the ability to quantitative compare varieties using statistical testing of difference in the model structures.
Acknowledgements: Supported by the Federal Targeted Program (Agreement No. 14.575.21.0136). Calculations were performed in Supercomputer Center of Peter the Great St. Petersburg Polytechnic University.

References

1. Kozlov et al. (2018) A mathematical model of the impact of climatic factors on soybean development. Biophysics. 63(1):175-176.
2. Seferova I., Novikova L. (2015) Climatic factors that impact development of fast ripening accessions of soybean in Nothwest Russia. Applied Botany, Genetics and Selection. 176:88-97.