Why and how is this evaluation done ?

The objective of the STICS performance evaluation reports are (i) to improve the quality of the STICS model through an iterative process (performance evaluation – anomaly detection – code correction / improvement) and (ii) to give the users a level of confidence in adequacy with the real performances of the model and with its domain of validity.

Through these reports, that will be updated and delivered for each new version of the model, the STICS team is trying to give a fair and synthetic image of the performances of the current version of the model. Simulation results are compared with observations for each species on the whole STICS dataset used by its automatic test and evaluation system (Buis et al, 2016).

This dataset is, for a large part, described in (Coucheney at al., 2015). To perform an evaluation as objective as possible, situations used for the model calibration have been excluded from this dataset as well as observations of soil water and nitrogen used as initial conditions.

All the simulated situations are so far located in France. Model users should be aware that, depending on the species, the soil and climate range and the number of observations of the evaluation dataset may still be scarse, and thus not representative of all situations for which the users may use this model. The dataset will be regularly updated to increase this variability and include situations outside of France.

Next versions of this evaluation exercise should include more species and a more detailed description of the dataset used for each species.

How to read the evaluation reports?

Each species report comprises a brief description of the evaluation dataset (number of observations per output variable, cultivars, …), a comparison between rRMSE obtained with the current version of STICS and with the former one, a global analysis of the performances for the main output variables, and more specific analysis (focusing on the growth dynamic, and/or the yield components and/or the soil water and nitrogen contents) depending on the species and the relative dataset. Graphs, statistical criteria and comments are provided to help readers to better discern the real performances of the model as well as its domain of validity.

The description of the graphs, the definition of the evaluated output variables and of the performance criteria used in the reports are given here-after.

All graphs and statistical criteria have been performed using the CroPlotR R package (Vezy et al, 2021).

Graphs

Residuals are the differences between the observed value and simulated values. In the evaluation reports, they are plotted in function of the observed values.

Performance criteria

All the performance criteria used in the evaluation reports are defined in the CroPlotR documentation (see https://sticsrpacks.github.io/CroPlotR/reference/predictor_assessment.html).

Output variables

The definitions and units of the output variables mentionned in the evaluation reports can be found in the file output.csv contained in the config folder of the JavaStics distribution.

References

Buis, S., Coucheney, E., Launay, M., Lecharpentier, P., Mary, B., Ripoche, D., Beaudoin, N., Ruget, F., Garcia de Cortazar-Atauri, I., Justes, E., Constantin, J., Andrianarisoa, K.S., le Bas, C., Léonard, J. Multicriteria evaluation of the STICS soil-crop model and implementation of an automated evaluation system. International Crop Modelling Symposium ICropM, 15-17 March 2016, Berlin.

Coucheney, E., Buis, S., Launay, M., Constantin, J., Mary, B., de Cortázar-Atauri, I.G., Ripoche, D., Beaudoin, N., Ruget, F., Andrianarisoa, K.S., 2015. Accuracy, robustness and behavior of the STICS soil–crop model for plant, water and nitrogen outputs: Evaluation over a wide range of agro-environmental conditions in France. Environmental Modelling & Software 64, 177-190.

Vezy, R., Buis, S., Lecharpentier, P., Giner, M. (2021). CroPlotR: A Package to Analyse Crop Model Simulations Outputs with Plots and Statistics. https://github.com/SticsRPacks/CroPlotR, https://doi.org/10.5281/zenodo.4442330.)