AgGlob: Workflow for simulation of agronomic models at a global scale

Helene Raynal, Christine Le Bas, Bertuzzi Patrick, Eric Cahuzac, Eric Casellas, Julie Constantin, Pomeon Thomas, Benoit Toutain

AgGlob: Workflow for simulation of agronomic models at a global scale

Raynal Hélène[1] (helene.raynal@inra.fr), Le Bas Christine[2], Bertuzzi Patrick[3], Cahuzac Eric[4], Casellas Eric[1], Constantin Julie[5], Pomeon Thomas[4], Toutain Benoît[6]

[1] *UR 875 MIAT, MIA, INRA, Castanet Tolosan, France*; [2] *US 1106 INFOSOL, EA, INRA, Orléans, France*; [3] *US 1116 Agroclim, EA, INRA, Avignon, France*; [4] *US 0685 ODR, SAE2, INRA, Castanet Tolosan, France*; [5] *UMR 1248 AGIR, EA, INRA, Castanet Tolosan, FRance*; [6] *INFOSOL, EA, INRA, Orléans, France*

Keywords: Crop modeling, Computational modeling, parallel computing, Global simulation.

Introduction

Simulation of biophysical models over large areas is used in different contexts linked to global agronomy issues (Müller et al., 2017). It is useful for analyzing crop performances at a regional scale (Shelia et al., 2019), for estimating vulnerability of crop production to climate change (Elliot et al., 2014, Montella et al., 2015)... One of approaches is Global Gridded Biophysical Models (GGBMs). It consists to use a crop model developed at field scale and to run it on different sites in order to take into account the heterogeneity of soils, climates and farming practices over the area. The sites are organized according to a spatial grid, with a fine resolution (some km^{-2}). It is possible to run these massive simulations thanks to the development of clusters. It is also possible because more and more data are available to characterize soil, climate and farming practices at fine resolution. Nevertheless, many difficulties remain. They concerned i) the coordination of the actors involved in the process of production of results, ii) the heterogeneity of data formats that makes tricky to reuse them iii) the design and the realization of the campaign of simulations, iv) the validation of simulation results by automated tests, v) the reproducibility of results and traceability, vi) methods and visualization tools suitable to the mass of results to analyze. To overcome these problems, we propose the AgGlob framework, based on a workflow developed on a Galaxy platform instance. (https://galaxyproject.org/).

Materials and Methods

A workflow consists in a sequence of treatments where each step is dependent on occurrence of the previous step. The first treatments concern the access to data stored in external databases and their processing in order to make them compatible to the crop model. For this step, we have developed basic bricks of the workflow.

- *"Climate data" tool*: access to datasets composed of daily observations of temperature, radiation, rain and PET. These data are provided under conditions by SICLIMA database, (INRA climate series provider). The SAFRAN grid is used (Meteo France standard with a resolution of 8km). Queries and post-processing are automated.
- *"Soil data" tool*: access to BDGSF (French Soil Geographic Database) maintained by Infosol (INRA provider). An SQL query requests the data. It puts the data into the format expected by the crop model.
- *"Land use" and "Farming practices" tool:* All the information concerning soil land use in France and farming practices come from the national surveys: "French Land Parcel identification system" and "Enquêtes pratiques agricoles". The information are stored in ODR database (INRA provider) and aggregated at the grid scale. The objective is to have for each cell of the grid, the most representative i) soils, ii) rotations and iv)farming practices. All these layers of information are combined based on the conceptual work done in INRA study "Evaluation Française des Ecosystèmes et des Services Ecosystémiques" . The results is a table where each line corresponds to a point to simulate with all the information required for simulation (crop rotation, sowing date …) on a concise form.

The second step of the workflow consists in preparing the campaign of simulations with the bricks:

- "Simulation Campaign" tool: The previous table is transformed into a text file. The user can download and modify it, in order to design a new simulation campaign (scenario). The tool includes algorithms

for testing the validity. Then, this text file is sent to the parallelization service of RECORD simulation platform (Bergez et al., 2014), embedded in the tool.
- “Crop simulation” tool: The model used is STICS encapsulated in RECORD (Bergez et al., 2014). It runs the campaign simulation on the cluster.
The third step concerns the post-processing of simulation results, with automated checks of simulation results, and the production of indicators.

Results and Discussion
AgGlob is available on an INRA Galaxy instance. It is connected to a distant cluster (Meso@LR) where the simulation jobs are run. It includes a formalization of GGBMs campaign simulation, that we consider as a standard reusable in other projects. A campaign of simulations generates large amount of results. Some consistency checks have been integrated in order to help the user in detecting problems. The workflow can be plugged on other Galaxy instances.

Conclusion
AgGlob is an interesting framework for GGBMs simulation. It helps in coordinating the actors involved, because the different steps of processing are clearly identified and formalized. It also offers a solution for the integration and aggregation of data necessary for simulation, by using automated processing algorithm. It is enough generic to be easily extend to other crop models and to other data. It is also an implementation of the FAIR principles in the domain of GGBMs work, therefore it enhances the reproducibility and traceability of results.

References:

- Shelia V, Hansen J, Sharda V, Porter C, Aggarwal P, Wilkerson C, Hoogenboom G (2019) A multi-scale and multi-model gridded framework for forecasting crop production, risk analysis, and climate change impact studies. Environ Model Softw, 115: 144-154
- Müller C, Elliott J, Chryssanthacopoulos J, Arneth A, Balkovic J, Ciais P, Deryng D, Folberth C, Glotter M, Hoek S, Iizumi T, Izaurralde R C, Jones C, Khabarov N, Lawrence P, Liu W, Olin S, Pugh T A M, Ray, D K, Reddy, A, Rosenzweig, C, Ruane, A C, Sakurai, G, Schmid, E, Skalsky, R, Song, C X, Wang, X, de Wit, A, and Yang, H (2017) Global gridded crop model evaluation: benchmarking, skills,
- Bergez JE, Raynal H, Launay M, Beaudoin N, Casellas E, Caubel J, Chabrier P, Coucheney E, Dury J, Garcia de Cortazar-Atauri I, Justes E, Mary B, Ripoche R, Ruget F (2014) Evolution of the STICS crop model to tackle new environmental issues: New formalisms and integration in the modelling and simulation platform RECORD. Environ Model Softw, 62: 370-384
- Elliott J, Kelly D, Chryssanthacopoulos J, Glotter M, Jhunjhnuwala K, Best N, Wilde M, Foster I (2014) The parallel system for integrating impact models and sectors (pSIMS). Environ Model Softw, 62 : 509-516,
- Montella R, Kelly D, Xiong W, Brizius A, Elliott J, Madduri R, Maheshwari K, Porter C, Vilter P, Wilde M, Zhang M, Foster I (2015) FACE-IT: A science gateway for food security research. Concurrency Computat: Pract Exper, doi: 101002/cpe3540