**Developing containerised automated workflows for large scale parallel execution of NetLogo models**

Doug Salt, Gary Polhill, Alessandro Gimona, Jonathan Ball

Information and Computational Sciences Department, The James Hutton Institute, Aberdeen, UK.

**Keywords**: Containers, NetLogo, Workflow, Automation, Reproducibility

It appears that little work has been done on the usage of containers for the large scale parallel execution of NetLogo models. That said, there are generalised frameworks for the running of NetLogo models such as Jansssen et al. (2008) and Janssen et al. (2014), and in addition frameworks for the analysis of data produced by the data produced from from such large scale execution, such as Jin et al. (2017).

Containers are ideally suited to the running of NetLogo models for several reasons. First, the large scale parallel execution of NetLogo models is a classic ‘embarrassingly parallel’ problem. Each NetLogo instance of execution is self-contained and requires inputs and outputs that are independent

from any other instance. Second, containers are an ideal method of providing an execution with a high degree of reliability and specificity. This leads not only to increased reliability, but also improves reproducibility of experiments. Lastly containerisation is ‘platform agnostic’ allowing the usage of two kinds of high performance computing environments, grid and utility computing (Sood et al. 2016) with the possibility of these being triggered automatically on demand from local instances of execution.

We present here such a framework for an automated containerised workflow for running and analysing NetLogo experiments based on file detection. This will use heterogeneous HPC resource including local, grid or utility and any combination thereof. The framework is based on a scheduler, itself containerised which may be deployed locally or remotely, which will allow the monitoring for specific file patterns that constitute a complete experiment. It will invoke a NetLogo container to process the set of files. The output files processed by this container, may in turn provide a trigger for further processing.

**References**:

Janssen, Marco A., Lilian Na'ia Alessa, Michael Barton, Sean Bergin, and Allen Lee. "Towards a community framework for agent-based modelling." Journal of Artificial Societies and Social Simulation 11, no. 2 (2008): 6.

Janssen, Marco A., Allen Lee, and Timothy M. Waring. "Experimental platforms for behavioral experiments on social-ecological systems." Ecology and Society 19, no. 4 (2014).

Jin, Xiongbing, Kirsten Robinson, Allen Lee, J. Gary Polhill, Calvin Pritchard, and Dawn C. Parker. "A prototype cloud-based reproducible data analysis and visualization platform for outputs of agent-based models." Environmental modelling & software 96 (2017): 172-180.

Sood, D., Kour, H. and Kumar, S., 2016. Survey of computing technologies: Distributed, utility, cluster, grid and cloud computing. JNCET, 6(5).