Development of a Quantitative Comparison Tool for Plant Models

Olivier Pieters[1,2], Tom De Swaef[2], Francis wyffels[1]

[1] *IDLab-AIRO -- Ghent University -- imec, Belgium*
[2] *Plant Unit, Research Institute for Agriculture, Fisheries and Food, Belgium*

For correspondence: olivier.pieters@ugent.be

Keywords: model comparison, reservoir computing, machine learning, virtual plant, benchmark

Over the past three decades, plant modelling has become a major research domain. Robust, validated models allow to study the effect of plant treatments on plant development in virtual conditions, providing novel insights and enabling users to optimise field conditions more effectively. Because of this success, a large amount of models has been developed at different scales of detail, from organ level functional-structural plant models (FSPMs) to field-scale crop models. This difference in structure and purpose, offers different optimisations and trade-offs. quantitative-plant.org/model provides an overview of the wide range of available models. However, selecting the right model for a specific purpose is not easy.

Comparing plant models often relies on intricate knowledge of the model, and sometimes experiments. For example, (Feng et al, 2014) compared an empirical crop model with an FSPM. They used detailed knowledge of the models and field trials to compare both models. This significantly obstructs similar studies, due to time and cost requirements involved with learning the inner workings of a model and executing field experiments. To the best of our knowledge, easy to use, quantitative, and model-independent measures are lacking. We propose to build upon (Dambre et al., 2012) and use the reservoir computing (RC) framework to enable fast, quantitative estimates of model performance with a common set of benchmark tasks. The goal of these tasks is to assess the non-linear and memory modelling capabilities of the model through a set of increasingly difficult tasks.

RC is a computational framework related to recurrent neural networks. It is well suited for temporal and sequential data. The input is mapped into a high-dimensional space whose current state is a function of the current input and the previous state of the network. A linear readout of the current network state is used to assess the target task (a benchmark in our case). This structure gives rise to memory due to recursive connections, and non-linearity. The mathematical formulation is included below in the case of discrete time, but is extendable to continuous time. A visual representation is depicted in Figure 1.

$$\boldsymbol{s}_t = f(\mathbf{W}\,\boldsymbol{s}_{t-1}) + \mathbf{A}\,\boldsymbol{x}_t$$
$$\boldsymbol{y}_t = \mathbf{B}\,\boldsymbol{s} + \boldsymbol{b}_0$$

Vectors $\boldsymbol{x}$, $\boldsymbol{s}$ and $\boldsymbol{y}$ represent the input(s), current state and benchmark target(s) of the network at time point t respectively. $\boldsymbol{x}$ and $\boldsymbol{y}$ are reduced to scalars if there is just one input variable and output target, respectively. The output task is computed using a linear combination of network states, represented by **B**. **W** represents the interconnectivity of the network, and how much each state influences the other states in the next time point. $f(\cdot)$ is a non-linear function that is often implemented by means of a tanh or sigmoid that is applied element-wise. Compared to conventional recurrent neural networks, training a reservoir is much simpler. The interconnectivity of the network **W** and input mapping **A** are randomly initialised and kept fixed. Only the linear mapping **B** from the network's state to the output target is optimised. This reduces the optimisation to a linear regression problem, which is well studied.

The RC framework is very general and can be extended to various physical media. For instance, in compliant robots RC is used to outsource part of the control tasks to the physical body (Caluwaerts et al., 2012). Applicability of the RC framework requires two properties: fading memory and non-linearity (Fernando and Sojakka, 2003). These are clearly satisfied by plants, and their modelling counterparts. Models rely heavily on differential equations, whose solutions are generally not linear, and time-dependent.

This analogy between general RC and plant model-based RC is depicted in Figure 1. A wide range of biotic and abiotic factors form the input of the model, while the reservoir is the plant model itself whose state is observed using the dynamic output variables such as dry matter content, photosynthesis and stomatal conductance. These are combined to compute the benchmark tasks. These benchmark tasks are usually mathematical functions that have to be computed that are non-linear time-dependent transformations of the input signals.

Similar to, (Nakajima et al., 2015), the selected set of tasks should assess the non-linear and temporal modelling capabilities of plant models. This provides further insight into the physiological processes involved in learning and anticipating future events within plants.

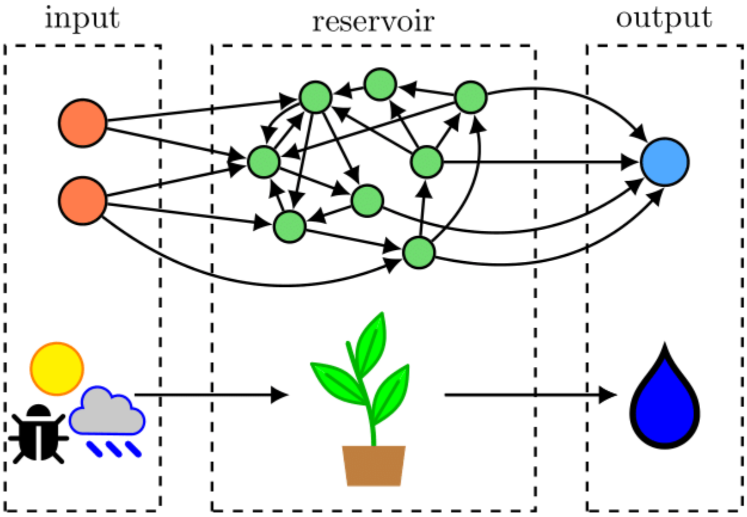

Figure 1 Reservoir computing setup (top) applied to plant models (bottom). The orange, green and blue nodes represent the input ($\boldsymbol{x}$), state ($\boldsymbol{s}$) and output ($\boldsymbol{y}$) of the reservoir respectively. The edges represent the interdependence between nodes (**A**, **B**, and **W**). This is readily applied to the modelling setup we consider here, where the input is the input of the plant model, the reservoir is the plant model itself whose state is observed using the outputs to estimate one of the benchmark tasks.

To conclude, we propose to consider plant models as simulated implementations of reservoirs. From their outputs, a set of user defined benchmark tasks can be calculated. An easy comparison of plant models is then possible by means of these benchmark tasks, providing information about non-linearity, and memory. Since the aim of plant models is the accurate simulation of plants, this can be a stepping-stone for the application of RC to plants, leading to a new framework to explain plant behaviour.

References

Nakajima et al., 2015, *Scientific Reports* 5, 10487, doi: 10.1038/srep10487

Fernando and Sojakka, 2003, *Springer Berlin* 588-597, doi: 10.1007/978-3-540-39432-7_63

Caluwaerts et al., 2012, *Artificial Life* 19:35-36, doi: 10.1162/ARTL_a_00080

Feng et al, 2014, *European Journal of Agronomy* 53:16-17 doi: 10.1016/j.eja.2013.11.005

Dambre et al., 2012, *Scientific Reports* 2:1-7 doi: 10.1038/srep00514