

Reservoir Computing With in silico Plants

Maxime Cannoodt

Student number: 01702828

Supervisors: Prof. dr. ir. Francis wyffels, Dr. ir. Michiel Stock

Councillors: Dr. ir. Olivier Pieters, Dr. Tom De Swaef (Instituut voor Landbouw-, Visserij- en Voedingsonderzoek)

Master's dissertation submitted in order to obtain the academic degree of
Master of Science in Computer Science Engineering

Academic year 2021-2022

Reservoir Computing With in silico Plants

Maxime Cannoodt

Student number: 01702828

Supervisors: Prof. dr. ir. Francis wyffels, Dr. ir. Michiel Stock

Councillors: Dr. ir. Olivier Pieters, Dr. Tom De Swaef (Instituut voor Landbouw-, Visserij- en Voedingsonderzoek)

Master's dissertation submitted in order to obtain the academic degree of
Master of Science in Computer Science Engineering

Academic year 2021-2022

This master's dissertation is part of an exam. Any comments formulated by the assessment committee during the oral presentation of the master's dissertation are not included in this text.

The author gives permission to make this master dissertation available for consultation and to copy parts of this master dissertation for personal use. In all cases of other use, the copyright terms have to be respected, in particular with regard to the obligation to state explicitly the source when quoting results from this master dissertation.

Maxime Cannoodt
May 26, 2022

Acknowledgments

I would like to express gratitude towards the people without whom this thesis would not have been possible.

First, I want to thank my counselor and advisor Olivier Pieters. The feedback he provided on the countless iterations of this book has been invaluable to the end product. I wish to thank my promotor Francis wyffels for continuously challenging my understanding of my research results and coming up with new ideas to try when I felt like I reached a dead end. I want to thank both of them especially for introducing me to the topic of reservoir computing. It was certainly an unconventional pick from the list of dissertation topics, but every time I saw my peers struggle to stay interested in their subjects, I was glad I chose the unbeaten path.

I also want to thank Tom De Swaef for filling the gaps in my biological knowledge and helping me understand concepts in computer plant modeling. Rami Albasha deserves a special mention for answering my questions about the inner workings of their grapevine simulation. And thanks, Francis and One-Planet Research Center, for inviting me to their research meeting. Visiting WUR's phenotyping facilities and listening to their active projects helped me understand better how plant reservoir computing fits in the landscape of agricultural technology.

Finally, I want to show gratitude to the people who have accompanied me during this arduous journey. Thank you to Jone Hoong and Tom for the social work sessions on Wednesdays and Thursdays at the iGent tower. I admit I might have distracted you two more than I should have, but the weekly social interaction helped me stay sane as we slowly worked towards our finished thesis books. My boyfriend Felix deserves a massive thank you for his patience when I worked late into the night and for listening to me vent my frustrations when progress was slow, or results were not as expected. Last but not least, I want to thank my family for their patience when I was more distant and absent during this academic year. I have always struggled to find the right balance between school and family, but now we can finally celebrate the end!

This dissertation topic challenged me to change my belief about what computation means. After seven months of research, I hope that I can explain the concepts of plant reservoir computing to you in an accessible way. I believe that we are only at the beginning of the road for plant reservoir computing. If the knowledge and technology can scale, I find the opportunities in agricultural tech immensely exciting.

Abstract

We present a framework to evaluate the computational capabilities of plant physiological processes in plant models for physical reservoir computing (PRC). PRC is a computing paradigm that enables the use of physical substrates for computation. Recently, Pieters (2022) proposed and successfully demonstrated the use of plants in a PRC context. We expand upon this work *in silico* using plant models for grapevine and wheat. The use of computer simulations gave us the freedom to observe the plant directly, without the need for sensing technology and laboratory setups. In our experiments, we investigated the use of leaf temperature, transpiration rate, photosynthesis rate, light absorption, water flow, and hydraulic pressure as readout traits. Each of these traits displayed reservoir dynamics and correlated well with biologically relevant tasks. Surprisingly, our framework highlighted unusual behavior in mechanistic plant models, which the authors did not yet report. We propose that the plant-modeling community can adapt the methods in this work to verify the behavior of plant models. We believe that this is just the starting point for plant reservoir computing. If knowledge and technology can scale, there will be countless opportunities in agricultural technology.

Keywords: Physical Reservoir Computing (PRC), Unconventional Computing, Plant Physiology, functional-structural plant models (FSPMs).

Reservoir Computing With *in silico* Plants

Maxime Cannoodt

Supervisors: Prof. dr. ir. Francis wyffels, Dr. ir. Michiel Stock
 Counselors: Dr. ir. Olivier Pieters, Dr. ir. Tom De Swaef (ILVO)

Abstract—We present a framework to evaluate the computational capabilities of plant physiological processes in plant models for physical reservoir computing (PRC). PRC is a computing paradigm that enables the use of physical substrates for computation. Recently, the use of live plants in a PRC context was demonstrated. We expand upon this work *in silico* using plant models for grapevine and wheat. In our experiments, we investigated the use of leaf temperature, transpiration rate, photosynthesis rate, light absorption, water flow, and hydraulic pressure as readout traits. Each of these traits displayed reservoir dynamics and correlated well with biologically relevant tasks. Surprisingly, our framework highlighted unusual behavior in mechanistic plant models, which the authors did not yet report. We propose that the plant-modeling community can adapt the methods in this work to verify the behavior of plant models. We believe that this is just the starting point for plant reservoir computing. If knowledge and technology can scale, there will be countless opportunities in agricultural technology.

Keywords—Physical Reservoir Computing (PRC), Unconventional Computing, Plant Physiology, functional-structural plant models (FSPMs).

I. INTRODUCTION

RESERVOIR computing (RC) is a holistic machine learning technique commonly used with time series data. RC is based on the emergent computational properties of nonlinear dynamical systems, the so-called reservoir. The key to understanding reservoir computing is that all non-linearity and memory required to solve a problem are already present in the (observable) reservoir dynamics. These dynamics are combined using a simple linear readout function to obtain the output value for a certain task. RC was first introduced as an alternative to trained recurrent neural networks and as a more biologically plausible computing model [1]–[3]. Later it was observed that two key features of said novel ideas: combining non-linearity and memory, can be identified in many physical systems as well, and there exists a rich literature on the use of physical media as computational resources [4], giving rise the physical reservoir computing (PRC). Here, the reservoir state is partially inferred from sensor data observing the dynamical system from different spatial positions [5], as illustrated in Figure 1.

Plants can also be interpreted as a computational resource. The plant’s environment is its input, and its health, growth, and behavior result from processing this information. Interpreting plants as information-processing organisms is not a new idea. A large body of research indicates that plants observe their environment, have memory, and are capable of learning complex patterns [9]. Experiments by Adamatzky et al. have shown

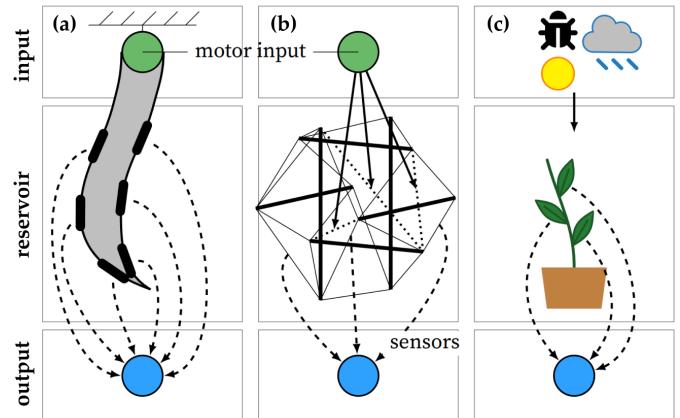


Fig. 1: Illustration of different physical reservoir computing implementations. (a) Bioinspired silicone arm with integrated deformation sensors [6]. (b) Tensegrity robot made of beams and springs [7]. Force sensors measure the dynamics of the spring-mass system. (c) The plant physiological reservoir proposed in this work. Illustration from Pieters (2022)⁸ with permission from the author.

limited yet promising success in exploiting plant intelligence to implement various functions and algorithms [10]. Their work serves as experimental evidence that plants can compute highly nonlinear functions. Most recently, Pieters demonstrated the first successful application of reservoir computing with plants [8]. His research shows that *Fragaria × ananassa* (strawberry) is capable of performing several environmental, eco-physiological, and computational tasks, paving the way to apply reservoir computing more broadly to a variety of crops.

The work by Pieters has shown that reservoir computing with plants is possible. Now, a more thorough examination of the reservoir dynamics of plant physiological processes is warranted. In this work, we further investigated the reservoir dynamics present in plants. We conducted our research with computer simulations of plants. Working with simulations provided the freedom to observe the plant’s physiological processes without the need for sensing technology or laboratory setups. It gives us a first intuition into which processes show promising reservoir dynamics, which informs future research with *in vivo* specimens.

II. MATERIALS AND METHODS

A. Experimental Design

To perform useful computations, a reservoir must show input separation with regard to the relevant tasks and fading

memory of past environmental conditions [6]. We investigated linear task separation and memory capacity in several observable eco-physiological processes by performing relevant environmental and eco-physiological reservoir tasks. We also verified whether these physiological processes displayed the fading memory property, an essential precondition for performing deterministic computations. We propose an experimental framework for quantitatively measuring these properties in plant physiological processes. The results from these experiments highlight which physiological reservoirs to pursue in future RC research.

1) Input Separation: The reservoir readout is a strictly linear function of the observed reservoir state. So, to perform nonlinear computation, the reservoir must transform its inputs such that the desired task becomes a linear function of the reservoir observations. To measure the input separation with regard to biologically relevant tasks, we used three categories of targets, previously established by Pieters: (i) predicting current environmental conditions, (ii) predicting the plant's eco-physiological performance, and (iii) computational benchmarks. Predicting environmental inputs gives us a first indication of the plant's coupling with the environment. Next, eco-physiological tasks indicate whether the observed reservoir is effective at solving biological problems. Finally, computational benchmarks can measure the reservoir's degree of nonlinearity and memory. We used three benchmarks: a delay line to measure memory capacity, a polynomial expansion to measure nonlinearity, and the Nonlinear Autoregressive Moving Average (NARMA) benchmark, a common benchmark for RC [6]. The NARMA system is defined as:

$$y(t+1) = \alpha y(t) + \beta y(t) \left[\sum_{i=0}^{n-1} y(t-i) \right] + \gamma x(t-n+1)x(t) + \delta \quad (1)$$

where n tunes the nonlinearity and memory of the task and the parameters α , β , γ , and δ are set to 0.3, 0.05, 1.5 and 0.1 [6]. Each computational task uses an environmental input as base signal.

We measured the prediction accuracy of each physiological reservoir for a given task using the normalized mean square error (NMSE) metric:

$$\text{NMSE} = \frac{1}{N} \sum_{t=1}^N \frac{(y[t] - \hat{y}[t])^2}{\text{var}(y)} \quad (2)$$

A lower score means a more accurate prediction. The NMSE has several advantages over regular mean square error [8]. Because it is normalized, the results can be compared between reservoir-task pairings, including between plants. It is also easy to interpret: a perfect predictor scores 0.0, while predicting the signal mean for every step yields a score of 1.0.

2) Fading Memory: To perform reliable computations, a reservoir must display the same behavior whenever it is subjected to the same input sequence [11]. In practice, this means past inputs should have a fading influence over the reservoir's present state [12]. The period over which this influence fades forms a trade-off between short-term accuracy and long-term memory. To test fading memory in physiological

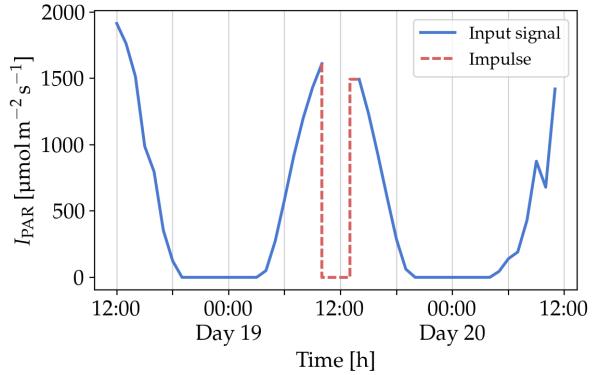


Fig. 2: Applying an impulse of 3 h and a value of 0 $\mu\text{mol m}^{-2} \text{s}^{-1}$ to the incident PAR of CN-Wheat.

processes, we designed an experiment in which an impulse is applied to the plant for a brief period, as illustrated in Figure 2. We then compared the trajectory of the physiological reservoir with that of a control experiment that did not receive the impulse. The impulse can be applied to any of the environmental inputs. The amplitude and duration of the impulse should remain within realistic boundaries so that an actual plant would not suffer permanent damage that alters the reservoir dynamics.

To quantify the divergence $\delta(t)$ between two reservoir trajectories at a given time step, we used a modified NMSE metric:

$$\delta(t) = \frac{1}{p} \sum_{i=1}^p \frac{(\mathbf{X}_i^C(t) - \mathbf{X}_i^E(t))^2}{\text{var}(\mathbf{X}_i^C)} \quad (3)$$

where p is the size of the observed reservoir, \mathbf{X}^C the reservoir state in the control experiment, and \mathbf{X}^E the reservoir state during the impulse experiment. A reservoir affected by the impulse will show a peak in divergence once the impulse is applied. If the physiological process displays fading memory, the divergence should go down to pre-impulse levels in a finite amount of time after the artificial stimulus is removed.

B. Selected Plant Models

We used functional-structural plant models (FSPMs) to test the experimental framework. FSPMs represent the plant as a 3D structure of interconnected elements. Each structural element simulates physiological processes such as photosynthesis, transpiration, and water and carbon flow [13]. This type of model has become the state-of-the-art in plant modeling in the past decade [14].

We present results from two FSPMs: HydroShoot [15] and CN-Wheat [16]. HydroShoot is an FSPM of common grapevine. It was published by Albasha et al. as a study of gas exchange in large plant canopies under water deficit conditions and to compare the effectiveness of different canopy shapes. CN-Wheat, on the other hand, is a mechanistic model of common wheat developed by Barillot et al. The model was made to simulate the effect of fertilization regimes on carbon (C) and nitrogen (N) metabolism in wheat culms during the

post-flowering stages of development. Table I compares the model specifications for the discussed models.

In addition to input separation and fading memory, the PRC framework assumes the reservoir dynamics are unchanging over time. Plants are not static reservoirs in the general case because they go through several growth phases that drastically change their structure and behavior. Because of this, we specifically selected plant models with a static plant structure or a dynamic structure in a stage of post-vegetative growth. For *in vivo* scenarios, stationarity can be approximated by investigating reservoir dynamics within a time window for which the growth rate is negligible [8].

C. Considered Physiological Reservoirs

Many physiological processes inside the plant can be considered a potential reservoir. We limited ourselves to processes that are observable using existing, non-destructive sensing technologies. The observable physiological processes for HydroShoot and CN-Wheat are presented in Table II. We had a preference for processes that are observable using image-based sensors because contactless sensors do not influence the reservoir dynamics [8]. Also, contactless sensing technologies may one day be the key to scaling plant RC to larger arrays of plants. Finally, we did not consider processes that violate fading memory.

D. Reservoir Readout Model

For the reservoir readout function we used a simple linear regression model with a bias term:

$$\hat{y}[t] = w_0 + \sum_{i=1}^p w_i x_i[t] \quad (4)$$

Where \hat{y} is the predicted value, x_i the reservoir measurements, and w_i the model parameters. To train the model parameters, we used a gradient descent algorithm to minimize the ridge regularization loss:

$$L_{\text{ridge}} = \frac{1}{N} \sum_{i=0}^{N-1} (y[i] - \hat{y}[i])^2 + \lambda \sum_{i=1}^p |w_i|^2 \quad (5)$$

where λ is a regularization parameter that prevents overfitting. To find the optimal λ , we used a parameter sweep with logarithmic spacing and a leave-one-out cross-validation scheme. We dedicated 50% of the experimental data for training and cross-validation and the other 50% for testing the final model on unseen data.

TABLE I: Comparison of HydroShoot and CN-Wheat: simulation details.

	HydroShoot	CN-Wheat
Simulation step	1 h	1 h
Simulation duration	7 days	50 days
Model structure	static (idealized reservoir)	Post-vegetative growth
Model size	$\approx 10^3$ elements	15 elements

III. RESULTS AND DISCUSSION

A. Environmental and Physiological Tasks

We considered air temperature (T_{air}), relative humidity (RH) and incident photosynthetically active radiation (PAR) for the environmental targets. For the physiological targets we used transpiration rate (E), absorbed PAR (Φ_{PAR}) and net photosynthesis rate (A_n). The physiological target A_n was only available for HydroShoot, while both models offer it as a physiological reservoir.

Figure 4a showcases the performance of each physiological reservoir as boxplots. For the HydroShoot model, we observe that each reservoir performs roughly the same for a given target. In CN-Wheat, the difference between the reservoirs is more outspoken. For example, RH correlates relatively well with E in CN-Wheat, where none of the processes in HydroShoot correlate well with RH. We must also note that CN-Wheat significantly outperforms HydroShoot in most tasks.

We study the time series prediction accuracy more closely in Figure 4c. In this figure, we see that the E reservoir captures the broad dynamics of Φ_{PAR} for both HydroShoot and CN-Wheat. However, HydroShoot cannot reproduce finer details, underpredicting the highest targets and overpredicting the lowest. Moreover, HydroShoot's predictions sometimes lag behind the target (Figure 4c, samples A and C), and other times lead the target (Figure 4c, sample D). In contrast, CN-Wheat's transpiration rate captures finer details much better.

B. Computational Benchmarks

We based the computational benchmarks on two different input signals (T_{air} and PAR). For each benchmark, we compared the reservoir's advantage (or disadvantage) over a linear model of the input signal as we increased the benchmark's difficulty. In what follows, we discuss the results for the surface temperature (T_s) and transpiration rate (E) reservoirs specifically.

Figure 4b shows the results for the delay line benchmark. Both reservoirs display memory capacity for this task. The memory for predicting the T_{air} input can be explained by thermal inertia; the difference in heat capacity between air and the plant's water mass results in a slowed reaction to changes in ambient temperature. The readout function can exploit this inertia to learn about past inputs. This is only a first-order interpretation, as it does not consider e.g. heat lost in evaporation. Similarly, thermal inertia can explain the results for incident solar radiation.

Next, Figure 4d presents the accuracy predicting a polynomial expansion. A polynomial relationship exists between air temperature and respiration rate, as the prediction error reaches a minimum for a degree of five. For the other input-reservoir pairings, we see a similar performance characteristic as the linear model plus a bias penalty induced by the noisy input-reservoir relation.

Finally, we consider the NARMA benchmark. In Figure 4f we increased the n parameter from 2 up to 24. For the benchmark based on air temperature, only surface temperature T_s

TABLE II: Comparison of HydroShoot and CN-Wheat: physiological processes observable in each structural element. Not all are shown in the table; we only present the processes that are experimentally observable and that do not violate fading memory. Check marks indicate whether the physiological process is modeled by the FSPM.

Symbol	Description	Unit	HydroShoot	CN-Wheat
A_n	Net photosynthesis rate	$\mu\text{mol m}^{-2} \text{s}^{-1}$	✓	✓
T_s	Surface temperature	$^{\circ}\text{C}$	✓	✓
g_s	Stomatal conductance	$\text{mol m}^{-2} \text{s}^{-1}$	✓	✓
E	Transpiration rate	$\text{mol H}_2\text{O m}^{-2} \text{s}^{-1}$	✓	✓
F	Water flow across segment	kg s^{-1}	✓	
Ψ	Mean water potential of segment	MPa	✓	

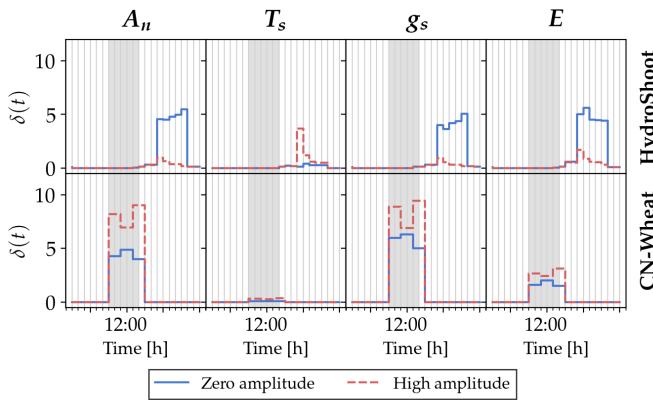


Fig. 3: Predicting a time-delayed signal of T_{air} and PAR using surface temperature (T_s) and transpiration rate (E). The reservoirs' performance is compared to a linear model that uses the benchmark's input as the only feature. The error bars show the scores' standard deviation.

outperformed the linear model. The same thermal inertia theory from before can explain the correlation with this NARMA benchmark. Both T_s and transpiration rate E outperform the linear model for the benchmark based on incident PAR, particularly for n values 8, 12, and 18. The task becomes easier for n equal to 18 and 24. This is because of the averaging effect of Equation 1 as the summation window extends over (nearly) the entire diurnal period of 24 h. Figure 4e shows the time series prediction for the NARMA8 task based on incident PAR, as predicted by surface temperature T_s .

C. Fading Memory

Figure 3 show the divergence $\delta(t)$ (Equation 3) in the physiological reservoirs caused by impulses of 5 h. We applied the impulses to incident PAR with zero and high amplitudes. We used a value of approx. two times the natural maximum signal strength for the high amplitude.

For HydroShoot, the data shows that the reservoir divergence caused by the impulse lasts for the same duration as the impulse itself. However, the reservoir reacts at an 8 h delay. Investigating the experimental data revealed a continuously shifting misalignment between the diurnal cycle of the input signal and the oscillations in the physiological dynamics.

In an extreme case, the peak in transpiration rate lagged nearly a half-period behind the incident PAR. In addition, The data indicates that the reservoir dynamics diverge slightly during the daytime, even between plants subjected to identical conditions. This indicates that the observed physiological processes display somewhat chaotic dynamics during daytime hours. The chaotic behavior may be a result of numerical instabilities in the simulation. However, there exist reports of chaotic behavior in plant physiological responses to light in the literature [17]. The shifting delay in reservoir response is the most likely explanation for the delayed response to the artificial stimulus. We must also attribute the slight disturbance that precedes the peak in divergence to the natural chaotic behavior. After accounting for these two factors, no lasting memory effects are visible at the simulation scale of 1 h; if there are any transient dynamics, they must occur over a much shorter time frame.

The results for CN-Wheat reveal that the simulations responsible for our physiological reservoir update at a 2 h interval, as described in the model's follow-up paper [18]. Looking past the misalignment caused by the 2 h simulation step, we see that the reservoir divergence disappears immediately when the real input signal resumes. The most likely explanation is that the transient behavior lasts less than two hours, such that the coarse simulation steps mask any memory effects.

These initial results imply that each considered reservoir has fading memory of incident PAR that lasts less than two hours. Regardless, we can conclude that all reservoirs show stationary behavior w.r.t. this abiotic input.

D. Potential Weaknesses

We used the same measurements for the reservoir as are used by the model to compute the physiological regression tasks. This introduced positive bias into the correlation between the reservoir and the physiological targets. This bias is especially pronounced in the accuracy of CN-Wheat for predicting total transpiration rate and absorbed PAR; CN-Wheat achieves near-perfect accuracy for these tasks. To reduce this bias, we could have introduced some signal noise to the observations to make them less precise. However, it is difficult to determine what amount of noise is appropriate in this case.

Another weakness is the trust in the accuracy of the FSPMs. Plant models are generally not verified as one hundred percent accurate copies of their real-world counterparts. As we described in Section III-C, the HydroShoot model displayed unexpected behavior that was not reported in the original paper. To emphasize that our results characterize the behavior of plant models, we used the names of the FSPMs instead of the plant genus.

Finally, the lack of time resolution in the plant simulations leaves us with a weak take-away about transient memory dynamics in the studied plant species. Though we can conclude that the considered physiological reservoirs show stationary behavior at the hour scale, we cannot state the exact duration during which past inputs influence the present reservoir dynamics.

E. Applications

If plants can behave like reservoir computers, RC may become a versatile addition to the agricultural R&D toolbox. For example, new insights could be gained into eco-physiological processes by observing and measuring the relationship between environmental signals (reservoir input) and plant behavior (reservoir dynamics). We suspect there are many possibilities for discovering new phenomena with this holistic data-driven approach, especially on the time scale between seconds and minutes. In addition, reservoir computing can have applications in phenotyping. The RC framework may identify phenotypes that are better at integrating their environmental inputs into an adaptive response. Especially rapid responses are currently underexplored in phenotyping research, even though they play an essential role in plant life [19], [20]. Even further down the road, RC can find applications in autonomous greenhouses. RC can measure the plant's response to the conditions set by the grower or an autonomous control system in real-time. A closed-circuit control loop could even be designed to let the plant itself controls the delivery of light, water, and nutrients.

IV. CONCLUSION

Regression benchmarks show promising results for plant RC with several plant physiological processes. In addition, the analysis of fading memory revealed chaotic behavior in physiological processes and unexpected behavior in a plant model that was previously unreported.

To select physiological reservoirs for future research, we recommend cross-referencing our results of reservoir performance with available sensing technologies. For example, surface temperature, transpiration rate, and stomatal conductance form a good combination of reservoir features because they can be observed using thermal cameras [8].

The study of using plants as a substrate for reservoir computing is still in its infancy. There are several yet unexplored topics to extend the foundation of plant RC. We currently only have data on RC with three species: strawberry plants in prior work and now *in silico* versions of grapevine and wheat. Another area that requires attention is how RC can be reconciled with plant growth.

We believe that this is only the beginning of the road for plant reservoir computing. If the knowledge and technology can scale, the opportunities in agricultural tech will be numerous.

REFERENCES

- [1] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks - with an erratum note," 2002, [Online]. Available: <https://www.ai.rug.nl/minds/uploads/EchoStatesTechRep.pdf>.
- [2] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Computation*, vol. 14, no. 11, pp. 2531–2560, Nov. 1, 2002, ISSN: 0899-7667, 1530-888X, DOI: 10.1162/089976602760407955.
- [3] J. Steil, "Backpropagation-decorrelation: Online recurrent learning with $O(n)$ complexity," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, vol. 2, Budapest, Hungary: IEEE, 2004, pp. 843–848, ISBN: 978-0-7803-8359-3, DOI: 10.1109/IJCNN.2004.1380039.
- [4] G. Tanaka, T. Yamane, J. B. Héroux, *et al.*, "Recent advances in physical reservoir computing: A review," *Neural Networks*, vol. 115, pp. 100–123, Jul. 2019, ISSN: 08936080, DOI: 10.1016/j.neunet.2019.03.005.
- [5] J. Burms, K. Caluwaerts, and J. Dambre, "Reward-modulated hebbian plasticity as leverage for partially embodied control in compliant robotics," *Frontiers in Neurorobotics*, vol. 9, Aug. 17, 2015, ISSN: 1662-5218, DOI: 10.3389/fnbot.2015.00009.
- [6] K. Nakajima, H. Hauser, T. Li, and R. Pfeifer, "Information processing via physical soft body," *Scientific Reports*, vol. 5, no. 1, p. 10487, Sep. 2015, ISSN: 2045-2322, DOI: 10.1038/srep10487.
- [7] K. Caluwaerts, M. D'Haene, D. Verstraeten, and B. Schrauwen, "Locomotion without a brain: Physical reservoir computing in tensegrity structures," *Artificial Life*, vol. 19, no. 1, pp. 35–66, Jan. 2013, ISSN: 1064-5462, 1530-9185, DOI: 10.1162/ARTL_a_00080.
- [8] O. Pieters, "Reservoir computing with plants," Ph.D. dissertation, Ghent University. Faculty of Engineering and Architecture, 2022, 235 pp., [Online]. Available: <http://hdl.handle.net/1854/LU-8739713>.
- [9] S. Mancuso, *The revolutionary genius of plants: a new understanding of plant intelligence and behavior*, First Atria books hardcover edition. New York, NY: Atria Books, an imprint of Simon & Schuster, Inc, 2018, 225 pp., OCLC: on1048600309, ISBN: 978-1-5011-8785-8.
- [10] A. Adamatzky, S. Harding, V. Erokhin, *et al.*, "Computers from plants we never made: Speculations," in *Inspired by Nature*, S. Stepney and A. Adamatzky, Eds., vol. 28, Series Title: Emergence, Complexity and Computation, Cham: Springer International Publishing, 2018, pp. 357–387, ISBN: 978-3-319-67996-9 978-3-319-67997-6, DOI: 10.1007/978-3-319-67997-6_17.
- [11] K. Nakajima, "Physical reservoir computing—an introductory perspective," *Japanese Journal of Applied Physics*, vol. 59, no. 6, p. 060501, Jun. 1, 2020, ISSN: 0021-4922, 1347-4065, DOI: 10.35848/1347-4065/ab8d4f.
- [12] M. Lukoševičius, H. Jaeger, and B. Schrauwen, "Reservoir computing trends," *KI - Künstliche Intelligenz*, vol. 26, no. 4, pp. 365–371, Nov. 2012, ISSN: 0933-1875, 1610-1987, DOI: 10.1007/s13218-012-0204-5.
- [13] J. R. Coussement, T. De Saepe, P. Lootens, and K. Steppen, "Turgor-driven plant growth applied in a soybean functional-structural plant model," *Annals of Botany*, vol. 126, no. 4, pp. 729–744, Sep. 14, 2020, ISSN: 0305-7364, 1095-8290, DOI: 10.1093/aob/mcaa076.
- [14] G. Louarn and Y. Song, "Two decades of functional-structural plant modelling: Now addressing fundamental questions in systems biology and predictive ecology," *Annals of Botany*, vol. 126, no. 4, pp. 501–509, Sep. 14, 2020, ISSN: 0305-7364, 1095-8290, DOI: 10.1093/aob/mcaa143.
- [15] R. Alasha, C. Fournier, C. Pradal, *et al.*, "HydroShoot: A functional-structural plant model for simulating hydraulic structure, gas and energy exchange dynamics of complex plant canopies under water deficit—application to grapevine (*Vitis vinifera*)," in *silico Plants*, vol. 1, no. 1, diz007, Jan. 1, 2019, ISSN: 2517-5025, DOI: 10.1093/insilicoplants/diz007.
- [16] R. Barillot, C. Chambon, and B. Andrieu, "CN-wheat: a functional-structural model of carbon and nitrogen metabolism in wheat culms after anthesis. II. model evaluation," *Annals of Botany*, vol. 118, no. 5, pp. 1015–1031, Oct. 2016, ISSN: 0305-7364, 1095-8290, DOI: 10.1093/aob/mcw144.
- [17] S. Shabala, R. Delbourgo, and I. Newman, "Observations of bifurcation and chaos in plant physiological responses to light," *Functional Plant Biology*, vol. 24, no. 1, p. 91, 1997, ISSN: 1445-4408, DOI: 10.1071/PP96075.
- [18] R. Barillot, C. Chambon, and B. Andrieu, "CN-wheat, a functional-structural model of carbon and nitrogen metabolism in wheat culms after anthesis. i. model description," *Annals of Botany*, vol. 118, no. 5, pp. 997–1013, Oct. 2016, ISSN: 0305-7364, 1095-8290, DOI: 10.1093/aob/mcw143.
- [19] J.-J. Alarcon and M. Malone, "Substantial hydraulic signals are triggered by leaf-biting insects in tomato," *Journal of Experimental Botany*, vol. 45, no. 7, pp. 953–957, 1994, ISSN: 0022-0957, 1460-2431, DOI: 10.1093/jxb/45.7.953.
- [20] T. De Saepe, K. Vermeulen, N. Vergote, *et al.*, "Plant sensors help to understand tipburn in lettuce," *Acta Horticultae*, no. 1099, pp. 63–70, Sep. 2015, ISSN: 0567-7572, 2406-6168, DOI: 10.17660/ActaHortic.2015.1099.3.

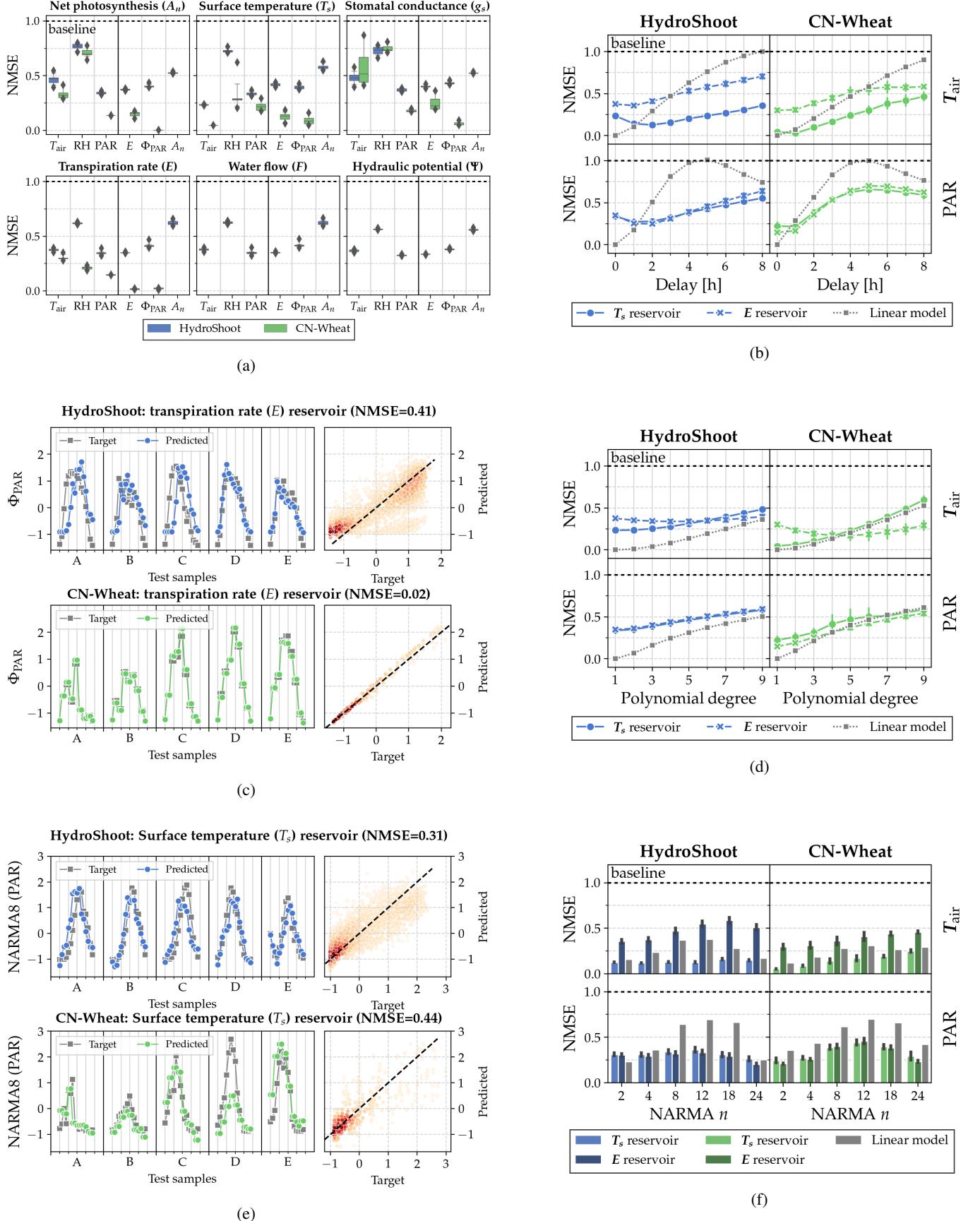


Fig. 4: Results for input, physiological, and computational tasks. Each figure shows the distribution of test scores for 16 random reservoir samplings. (a): Reservoir performance for input and physiological regression tasks. The plots show prediction accuracy for a suite of regression tasks for each of the physiological reservoirs from Table II. The regression target of net photosynthesis rate A_n was only available for the HydroShoot model. The box shows the quartiles of the distribution; the whiskers indicate the 5th and 95th percentiles. (b, d, f): Predicting a delayed signal, polynomial expansion, and NARMA task based on T_{air} and PAR using surface temperature (T_s) and transpiration rate (E). The reservoirs' performance is compared to a linear model that uses the benchmark's input as the only feature. The error bars show the scores' standard deviation; (c, e): Time series prediction of absorbed PAR and NARMA8 using E and T_s , respectively. The left plot shows predictions for five days randomly sampled from the test set. The right shows correlation between target and predicted values. Darker colors indicate a greater density of points.

Contents

Acknowledgments	i
Abstract	ii
Extended Abstract	ii
Contents	ix
List of Figures	xii
List of Tables	xiii
Acronyms	xiii
I Background	1
1 Introduction	2
1.1 Ongoing Efforts in Agricultural Research	2
1.2 Reservoir Computing	3
1.3 Research Outline	4
1.4 Chapters overview	5
2 Physical Reservoir Computing	6
2.1 Reservoir Computing	6
2.1.1 Echo State Network	7
2.1.2 Applications	8
2.2 Physical Reservoir Computing	9
2.2.1 Generalized Mathematical Model	10
2.3 The Echo State Property	11
2.4 Time Scale of Reservoir Dynamics	12
2.4.1 Sampling of a Continuous-Time Reservoir	13
2.5 Evaluating Reservoir Performance	13
2.5.1 Input Separability	14
2.5.2 Fading Memory	14

Contents

2.5.3	Computational Benchmarks	15
2.6	Simulation of Physical Reservoirs	17
2.7	Summary	17
3	Functional-structural Plant Models	18
3.1	Selecting Plant Models for RC Experiments	18
3.2	Functional-structural Plant Models (FSPMs)	19
3.3	Summary	21
II	Methodology	22
4	Experimental Design	23
4.1	Reservoir Dynamics in Plants	23
4.1.1	Research Objectives	24
4.2	Framework For Measuring Plant Reservoir Dynamics	24
4.2.1	Input Separation	24
4.2.2	Fading Memory	25
4.3	Considered Reservoirs	26
4.4	Summary	26
5	Selected Plant Models	29
5.1	Selection Criteria	29
5.2	Related Work	30
5.3	Selected Models	31
5.3.1	HydroShoot	31
5.3.2	CN-Wheat	32
5.3.3	Simulation Details	33
5.3.4	Environmental Inputs	34
5.3.5	Regression Targets	34
5.3.6	Available Reservoirs	34
5.4	Summary	34
6	Implementation Details	36
6.1	Preparing the Simulations	36
6.2	Observing The Plant Model	36
6.3	Data Preprocessing	37
6.4	Train and Test Data Split	37
6.5	Model Validation	37
6.6	Implementing the Impulse Experiment	38
6.7	Summary	39

Contents

III Results	40
7 Results And Discussion	41
7.1 Input Separation	41
7.1.1 Environmental Inputs and Physiological Tasks	41
7.1.2 Computational Benchmarks	44
7.1.3 Comparison with Previous Research	45
7.2 Fading Memory	48
7.2.1 CN-Wheat	48
7.2.2 HydroShoot	48
7.3 Weaknesses in Our Methods and Potential Improvements	51
7.4 Summary	53
8 Conclusion	54
8.1 Implications of the Results	54
8.2 Improving Mechanistic Plant Modeling	54
8.3 Future Work	54
Appendices	57
A Meteorological Data	57
B HydroShoot: Additional Information	59
B.1 Simulation Details	59
B.2 Reservoir Composition	59
B.3 Regression Dataset	59
B.4 Cross-Validation Strategy	59
B.5 Setup Impulse Experiment	60
C CN-Wheat: Additional Information	61
C.1 Simulation Details	61
C.2 Reservoir Composition	61
C.3 Regression Dataset	61
C.4 Cross-Validation Strategy	61
C.5 Setup Impulse Experiment	62
Bibliography	62

List of Figures

2.1	A reservoir computing model with a RNN as the reservoir, commonly referred to as an echo state network (ESN).	7
2.2	Illustration of different reservoir computing implementations, both theoretical and physical versions.	10
2.3	Demonstration of the Echo State Property and illustration of linear input separation. . . .	12
2.4	Memory capacity for ESNs with different leak rates (from Verstraeten (2009)).	13
2.5	Two methods for testing the fading memory property of a reservoir, demonstrated using a ESN.	16
3.1	Interactions at micro- and macroscopic scales present in functional-structural plant models (from Louarn et al. (2020)).	19
3.2	Two examples of functional-structural plant models.	20
5.1	FSPMs selected for use in experiments in this work.	33
6.1	Illustration of the data that is discarded as a preprocessing step and how the remaining data is split into a train and test set.	37
6.2	Illustration of leave-one-group-out cross-validation strategy.	38
6.3	Adding an artificial impulse to an environmental input.	39
7.1	Reservoir performance for input and physiological regression tasks.	42
7.2	Time series prediction of input PAR and absorbed PAR using transpiration rate (E). . . .	43
7.3	Predicting a time-delayed signal of T_{air} and PAR using surface temperature (T_s) and transpiration rate (E).	44
7.4	Predicting a polynomial expansion of T_{air} and PAR using surface temperature (T_s) and transpiration rate (E).	46
7.5	Predicting the NARMA benchmark based on T_{air} and PAR using surface temperature (T_s) and transpiration rate (E).	46
7.6	Time series prediction of the NARMA benchmark using surface temperature (T_s). . . .	47
7.7	CN-Wheat: Reservoir divergence $\delta(t)$ (Equation 4.3) after applying an impulse to PAR of various lengths and amplitudes.	49
7.8	HydroShoot: Reservoir divergence $\delta(t)$ after applying an impulse to PAR of various lengths and amplitudes.	50
7.9	Two figures highlighting unstable behavior observed in the HydroShoot FSPM.	52
A.1	Hourly meteorological inputs used in CN-Wheat simulations.	57

A.2	Hourly meteorological inputs used in HydroShoot simulations.	58
B.1	Data grouping method used to combine data from overlapping simulation runs of HydroShoot.	60
C.1	Some structural elements of CN-Wheat stop functioning after a specific amount of days. .	62
C.2	Data grouping method used to combine data from three simulations of CN-Wheat.	62

List of Tables

2.1	Examples of subjects in RC applications.	8
2.2	Applications and related benchmark tasks of RC.	9
4.1	Imaging techniques for various plant-physiological processes.	27
4.2	Non-imaging techniques for various plant-physiological processes.	28
5.1	FSPMs we explored for use in RC experiments. Models highlighted in bold are used for the final results presented in this work.	31
5.2	Comparison of HydroShoot and CN-Wheat: simulation details.	33
5.3	Comparison of HydroShoot and CN-Wheat: environmental models.	35
5.4	Comparison of HydroShoot and CN-Wheat: available regression tasks.	35
5.5	Comparison of HydroShoot and CN-Wheat: physiological processes observable in each structural element.	35

Acronyms

BPDC Backpropagation-Decorrelation. 7

BPTT Backpropagation-Through-Time. 6

CSM Crop Surface Model. 18

ESN Echo State Network. xii, 6, 13, 14, 16

ESP Echo State Property. 11, 12

FSPM Functional-Structural Plant Model. xii, xiii, 19, 20, 24, 31, 33, 34, 36, 41, 48, 52, 53

LSM Liquid State Network. 7

MSE Mean Squared Error. 15, 25

NARMA Nonlinear Autoregressive Moving Average. 15, 44, 45

NMSE Normalized Mean Square Error. 25, 38

PAR Photosynthetically active radiation. 32, 35, 39, 45, 48, 53, 60, 62

PRC Physical Reservoir Computing. 12, 17, 23, 25, 29

RC Reservoir Computing. 3, 4, 6, 15, 17, 18, 23, 24, 26, 29–31, 34, 35, 53

RNN Recurrent Neural Network. 6

Part I

Background

Chapter 1

Introduction

New challenges confront the field of agricultural research every year. The biggest challenge is, of course, climate change (Bisoffi, 2019). Changing climate conditions threaten many common crops, including staple crops such as corn, wheat, rice, sunflower, and cotton (Hatfield et al., 2014). Other areas of concern include food security, rising population, declining land productivity, and urban food strategy (European Commission, 2009, 2020).

1.1 Ongoing Efforts in Agricultural Research

A significant amount of research is dedicated to discovering and breeding crop varieties that are well-adapted to the novel climate, water, and soil conditions caused by global warming. Phenotyping plays a vital role in this. In agricultural research, phenotyping refers to identifying species varieties (phenotypes) with positive traits from a large pool of genetic variants (genotypes) (Walter, Liebisch, et al., 2015). For example, suppose researchers attempt to breed crops more resilient to saltwater. In that case, the researchers will measure the saltwater resistance of each specimen in their pool of genetic variants and selectively breed the phenotypes with the best resistive properties. However, the phenotype of an organism is the outcome of complex interactions between genetic and environmental factors, resulting in a high-dimensional problem space (Walter, Liebisch, et al., 2015). Because of this, current *in vivo* research is limited to measuring the effect of one or two environmental variables on a plant, even though the plant's response integrates many more variables that are unaccounted for (Pieters, 2022). Novel approaches should aim to more accurately model the interaction between the phenotype and the whole environment.

Another active line of research is autonomous greenhouse farming. Crop growth in a greenhouse setting is affected by many factors, including light, temperature, humidity, CO₂, and water and nutrient delivery (Hemming, 2022). The greenhouse farmer must finely control these factors to balance crop yield with major cost drivers like water and energy intensity. In addition, the grower must control pests and diseases to have a successful harvest. A farmer who wishes to optimize their operating efficiency must manage all these interlinked aspects, which requires an intricate understanding of the crop species from the grower. The need for this deep, crop-specific domain knowledge forms a bottleneck for rapidly expanding the number of greenhouses in operation or scaling the number of different crops grown in a single facility. Autonomous farming aims to assist the farmer through (partial) automation of the decision-making process and delivery system while minimizing resource use. This goal can be achieved using advanced sensor technology combined with AI techniques such as computer vision and machine learning (Hemming, 2022). Autonomous greenhouses may one day answer the need for food security and food delivery in areas with high urban density and decrease the strain on rural farmlands to feed the rapidly growing urban population.

On a more fundamental level, research must be conducted to better understand the eco-physiological processes that drive plant growth and health. A better understanding of the inner workings of plants will help accelerate the development of new resistant varieties. It can also help reveal the soundest indicators of plant well-being to incorporate into the control systems of autonomous greenhouses. Yet, it is insufficient to comprehend the physical interactions between the plant and its abiotic environment in isolation. The plant responds to the sum of its environment, resulting in emergent behavior that may be unaccounted for in theoretical findings and lab experiments (Poorter et al., 2016).

Mechanistic models are state-of-the-art for modeling the interactions between the plant and its abiotic environment. These models simulate the plant at the lowest level, commonly implementing physiological processes at the cell organ level such as photosynthesis, transpiration, and water and carbon flow (Coussement et al., 2020). It then becomes possible to study emergent behaviors that result from these base processes. By subjecting mechanistic models to novel environmental conditions, they can simulate the plant's response to various scenarios, such as new soil types or drought. This use case makes mechanistic models of particular interest in understanding macroscopic plant behavior and phenotyping. This type of model has shown very encouraging results in the past decade (Louarn et al., 2020).

However, there are some drawbacks to mechanistic models. First, this model is computationally expensive as it relies on solving many differential equations for every time step. Because of this, the simulated time scale is frequently limited to hours or days. The computational cost makes them unsuited for modeling interactions that occur at the second or minute scale, even though those play a significant role in the behavior of plants (Pieters, 2022). Second, the development of mechanistic models is closely coupled to the theoretical understanding of a species's physiology. As a result, the model is only as accurate as the current understanding of the plant's low-level processes. Finally, developing a mechanistic model is a significant undertaking. A large portion of the work cannot be generalized to other species, which forms a significant bottleneck in the development speed of plant models. In conclusion, though mechanistic models have become increasingly relevant in the literature, there is still room for alternative approaches that can be applied more broadly and are faster to deploy.

1.2 Reservoir Computing

Reservoir Computing (RC) is a holistic machine learning technique commonly used with time series data. Instead of relying on resource-intensive and data-hungry deep learning techniques, RC works by exploiting the computational properties of nonlinear dynamical systems, which are called reservoirs in this context. The key to understanding reservoir computing is that all non-linearity and memory required to solve a problem are already present in the reservoir dynamics. Only a simple linear regression model is needed to transform the information in the reservoir into an interpretable output value. RC was first introduced as a computer model (Jaeger, 2002a; Maass et al., 2002; Steil, 2004). However, non-linear dynamical behavior can be observed in many physical systems as well, and there exists a rich literature on the use of physical media as computational resources (Tanaka et al., 2019).

Plants can also be interpreted as a computational resource. The plant's environment is its input, and its health, growth, and behavior are a result of processing this information. Interpreting plants as intelligent information-processing organisms is not a new idea. A large body of research indicates that plants are aware of their environment, have memory, and are capable of learning complex patterns

(Mancuso, 2018). Experiments by Adamatzky et al. (2018) have shown limited yet promising success in exploiting plant intelligence to implement various functions and algorithms. Though our goal is not to implement general-purpose computing using plants, their work serves as experimental evidence that plants can indeed compute highly nonlinear functions. Most recently, Pieters et al. (2021) have demonstrated the first successful application of reservoir computing with plants. Their research shows that *Fragaria × ananassa* (strawberry) is capable of performing several environmental, eco-physiological, and computational tasks, paving the way to apply reservoir computing more broadly to a variety of crops.

If plants do indeed behave like computing reservoirs, RC may become a versatile addition to the agricultural research and development toolbox. For example, observing and measuring the relationship between environmental signals (reservoir input) and plant behavior (reservoir dynamics), we could gain insights into eco-physiological processes not yet uncovered or understood. We suspect there are still considerable opportunities for discovering new phenomena with this holistic data-driven approach, especially on the time scale between seconds and minutes. In addition, reservoir computing can have applications in phenotyping. The RC framework may be able to identify phenotypes that are better or faster at integrating their environmental inputs into an adaptive response. Especially rapid responses are currently underexplored in phenotyping research, even though they play an essential role in plant life (Alarcon et al., 1994; De Swaef, Vermeulen, et al., 2015a). Even further down the road, RC can be applied in autonomous greenhouses to observe the plant’s real-time response to the conditions set by the grower or autonomous control system. A closed-circuit control loop can be designed that integrates the real-time behavior of the plant, where the plant itself controls the delivery of light, water, and nutrients.

1.3 Research Outline

The work by Pieters has shown that reservoir computing with plants is possible. Now, a more thorough examination of the reservoir characteristics of plant-physiological processes is warranted. In this work, we further investigated the reservoir dynamics present in plants. We investigated several observable eco-physiological processes’ linear task separation and memory capacity by performing relevant environmental and eco-physiological reservoir tasks. These experiments identified a set of promising physiological reservoirs for future RC research. We also verified whether these physiological processes displayed the fading memory property, an essential precondition for performing deterministic computations.

We conduct our research with *in silico* plants. Working with simulations provides the freedom to observe the plant’s physiological processes without the need for sensing technology or laboratory setups. It gives us a first intuition into which processes show promising reservoir dynamics, which informs future research with *in vivo* specimens. RC can also contribute to FSPM development by identifying flaws in physiological models that previously went unnoticed.

1.4 Chapters overview

To understand the research presented in this book, Chapter 2 starts with a brief primer on the principles of reservoir computing and the use of physical media for computation. In Chapter 3, we give a brief introduction to the research field of plant modeling and simulation. Then, Chapter 4 elaborates on the research goals set for this work and explains the experimental design. Chapters 5 and 6 delve deeper into our methods by explaining the selection process for computer plant models and the experiment implementation details, respectively. Next, Chapter 7 reports our experimental data and discusses those results. Finally, 8 concludes the book with a summary of our key takeaways and a discussion of future work.

Chapter 2

Physical Reservoir Computing

2.1 Reservoir Computing

Reservoir computing (RC) is a machine learning method proposed in the early 2000s as an alternative to training recurrent neural networks (RNN) (Jaeger, 2002a). Training weights in recurrent networks requires computationally expensive and highly linear training algorithms such as backpropagation-through-time (BPTT). Moreover, BPTT often suffers from numerical instability and often yields sub-optimal solutions (Bengio et al., 1994).

In the RC framework, the untrained RNN (called the reservoir) is interpreted as a black box dynamical system. Input signals are transformed by the nonlinear activations of the neurons and are integrated with the past state of the reservoir through recurrent connections (Jaeger et al., 2004). As a result, the reservoir displays both nonlinear behavior and memory capacity. We can then exploit these properties to solve a variety of regression tasks.

Figure 2.1 shows a diagram of a general RC model using an RNN as the reservoir. To accomplish regression tasks, an RC model consists of two parts: the nonlinear reservoir that transforms environmental inputs and a linear readout function that observes the reservoir to perform a given task. The linear readout function observes the state of the reservoir and applies a linear regression model to determine the final output. When there is no feedback from the readout function back to the reservoir, it is possible to train multiple readouts that perform different tasks based on the same reservoir dynamics.

The key to understanding reservoir computing is that all non-linearity and memory required to solve a given task are already present in the reservoir dynamics. The readout function then learns to distill the information present in the observed reservoir state to solve a given task. In this way, RC is similar to the kernel method from classical machine learning, where the reservoir acts as a nonlinear expansion and temporal convolution of the input signals (Hermans et al., 2012). However, kernel techniques do not account for time explicitly, whereas reservoir computing inherently occurs in the time domain. This strong connection to the time domain makes RC particularly useful for solving tasks with time series data and real-time applications.

In its original form, the neurons inside the reservoir are randomly connected and are generally not trained. However, there is literature that proposes task-independent tuning schemes that improve the stability of the reservoir (Lukoševičius et al., 2009; Norton et al., 2006; Tanaka et al., 2020). Reservoir computing models are notably more data-efficient and easier to train because there is no need to train the recurrent part of the network using complex and data-hungry algorithms such as BPTT.

The literature has proposed several implementations of reservoirs. The echo state network (ESN), as proposed by Jaeger (2002a), is a purely mathematical implementation where neuron activations are determined as a nonlinear function of the weighted sum of incoming connections. ESNs operate in discrete

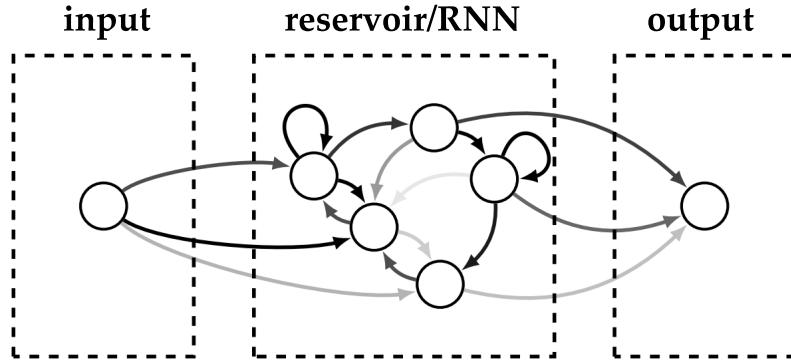


Figure 2.1: A reservoir computing model with a RNN as the reservoir, commonly referred to as an echo state network (ESN). Figure reproduced from Pieters (2022) with permission from the author.

time only and are ideal for software implementations. Subsection 2.1.1 gives a detailed mathematical description of the echo state network. Independently, Maass et al. (2002) proposed the liquid state machine (LSM), an idea that originated from computational neuroscience. In this model, the recurrent network is formed by spiking neurons that can operate in both discrete and continuous time. The neurons receive spike trains (discrete events in time) as input and generate an output sequence of spikes accordingly. A third approach proposed by Steil (2004) is called Backpropagation-Decorrelation (BPDC). It is a learning algorithm for online training of the readout weights.

2.1.1 Echo State Network

The Echo State Network (ESN) is a mathematical framework that implements the core concepts of reservoir computing (Jaeger, 2002b). In this framework, the reservoir is a discrete-time RNN using a sigmoidal nonlinearity (such as the tanh function). The activation weights of the reservoir are initialized randomly and remain fixed. The linear readout function takes the reservoir state as input and is fit to the desired target signal.

The reservoir state in time step t is calculated from the input signal \mathbf{u} using the following update rules (Lukoševičius et al., 2012),

$$\tilde{\mathbf{x}}[t] = \tanh(\mathbf{W}^{\text{in}}[1; \mathbf{u}[t]] + \mathbf{W}\mathbf{x}[t-1]) \quad (2.1)$$

$$\mathbf{x}[t] = (1 - \alpha)\mathbf{x}[t-1] + \alpha\tilde{\mathbf{x}}[t] \quad (2.2)$$

where \mathbf{x} signifies the reservoir neurons' activations, $\tilde{\mathbf{x}}$ the activation update, and α a scalar leaking rate. The readout function \mathbf{y} is fulfilled by a simple linear regression of the reservoir state and an optional direct feed-in of the input signal.

$$\mathbf{y}[t] = \mathbf{W}^{\text{out}}[1; \mathbf{u}[t]; \mathbf{x}[t]] \quad (2.3)$$

The weights \mathbf{W}^{out} of the readout function can be trained using a variety of learning methods. A com-

mon offline approach for fitting the function output to a target signal is ridge regression (Lukoševičius et al., 2012):

$$\mathbf{W}^{\text{out}} = \mathbf{Y}_{\text{target}} \mathbf{X}^T (\mathbf{X} \mathbf{X}^T + \lambda^2 \mathbf{I})^{-1} \quad (2.4)$$

where λ is a regularization parameter that prevents overfitting. For $\lambda = 0$, the model attempts to find weights that fit the training data exactly. The results is a high training accuracy, but low accuracy on unseen data. Large model weights are penalized for $\lambda > 0$, resulting in smaller weights and a more generalized model. However, if λ is too large, the model parameters shrink too much, and the model loses expressiveness. The optimal value for λ should be selected using validation data separate from training and test data. Other regression techniques are also possible. For example, Burms et al. (2015) have demonstrated online training of the readout function using reward-modulated Hebbian plasticity rules.

2.1.2 Applications

Reservoir computing has enjoyed success in a multitude of research domains because of its simple computational model, the potential for real-time processing, and broad applicability to problems in the physical world (Tanaka et al., 2019). Early applications of RC include time series prediction tasks such as stock price forecasting and weather forecasting (Nakajima, 2020). Table 2.1 shows a comprehensive summary of the domains to which RC has been applied as compiled by Tanaka et al. (2019). Similarly, Table 2.2 gives a selection of works that demonstrate the application of reservoir computing.

Table 2.1: Examples of subjects in RC applications. This table originally appeared in Tanaka et al. (2019) and has been reused under the CC BY 4.0 license.

Category	Examples
Biomedical	EEG, fMRI, ECG, EMG, heart rates, biomarkers, BMI, eye movement, mammogram, lung images.
Visual	Images, videos.
Audio	Speech, sounds, music, bird calls.
Machinery	Vehicles, robots, sensors, motors, compressors, controllers, actuators.
Engineering	Power plants, power lines, renewable energy, engines, fuel cells, batteries, gas flows, diesel oil, coal mines, hydraulic excavators, steam generators, roller mills, footbridges, air conditioners.
Communication	Radio waves, telephone calls, Internet traffic.
Environmental	Wind power and speed, ozone concentration, PM2.5, wastewater, rainfall, seismicity.
Security	Cryptography.
Financial	Stock price, stock index, exchange rate.
Social	Language, grammar, syntax, smart phone.

Table 2.2: Applications and related benchmark tasks of RC. This table originally appeared in Tanaka et al. (2019) and has been reused under the CC BY 4.0 license.

Applications	Benchmark tasks
Pattern classification	Spoken digit recognition (Verstraeten, Schrauwen, Stroobandt and Van Campenhout, 2005), Waveform classification (Paquot et al., 2012), Human action recognition (Soh & Demiris, 2012), Handwritten digit image recognition (Jalalvand, Van Wallendael, & Van de Walle, 2015)
Time series forecasting	Chaotic time series prediction (Jaeger, 2001), NARMA time series prediction (Jaeger, 2003)
Pattern generation	Sine-wave generation (Jaeger, 2002), Limit cycle generation (Hauser, Ijspeert, Füchslin, & Maass, 2012)
Adaptive filtering and control	Channel equalization (Jaeger & Haas, 2004)
System approximation	Temporal XOR task (Bertschinger & Natschläger, 2004), Temporal parity task (Bertschinger & Natschläger, 2004)
Short-term memory	Memory capacity (Jaeger, 2002)

2.2 Physical Reservoir Computing

The RC model easily extends to the physical world by interpreting physical systems that exhibit non-linear dynamics and memory effects as a reservoir substrate (Nakajima, 2020). In other words, this Physical Reservoir Computing (PRC) framework aims to exploit dynamics found in the natural world to perform computational tasks. Figure 2.2 shows a mapping of the RNN-based RC architecture to two physical applications from the literature and the plant RC case discussed in this thesis. In the physical systems (Figures 2.2b, c, and d), it is impossible to observe the complete state of the reservoir as in the RNN-based reservoir. Instead, the reservoir state is inferred from sensor data observing the dynamical system in real-time (Burms et al., 2015; Nakajima et al., 2015).

PRC is a field of interdisciplinary research, implementing reservoirs using a large variety of materials (Tanaka et al., 2019). Examples of substrates being researched are analog electrical circuits (Soriano et al., 2015; Zhao et al., 2016), FPGAs (Antonik, 2018), memristive circuits (Walter, Röhrbein, et al., 2015), photonic circuits (Van der Sande et al., 2017), spintronics (Torrejon et al., 2017), and soft body robotics (Caluwaerts, 2014; Caluwaerts et al., 2011; Nakajima et al., 2013). Biological reservoirs are closely investigated as well, with a particular interest in neurobiology (Hafizovic et al., 2007; Hinaut et al., 2013) and ecological dynamics (Didovsky et al., 2015; Jones et al., 2007; Ushio et al., 2021). Most recently, Pieters et al. (2021) demonstrated promising results in the use of *in vivo* plants as reservoir substrate.

A common misconception about PRC research is that the goal is to achieve general-purpose computation with unconventional reservoirs. However, the aim is not to outperform silicon computers at carrying out generic calculations. Instead, in most cases, the goal is to bypass the need for precise modeling of a physical system by observing how the system responds to inputs rather than attempting to predict how the inputs influence the state of the system (Nakajima, 2020). This method is particularly applicable in real-time control scenarios such as robotics, where real-time integration of sensor data is

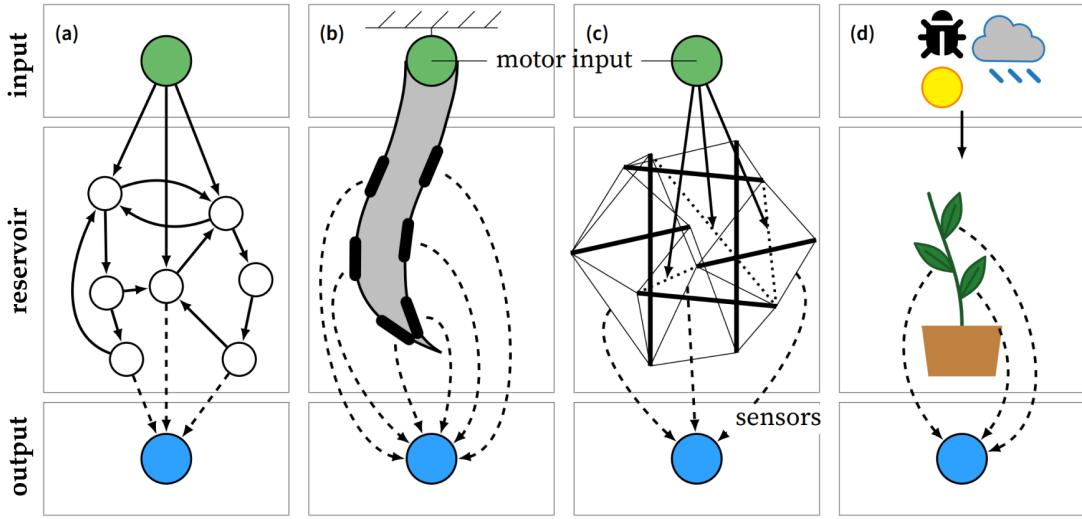


Figure 2.2: Illustration of different reservoir computing implementations, both theoretical and physical versions. Each of the subfigures depicts a different type of implementation: (a) RNN-based version; (b) soft body built of silicone (Nakajima et al., 2015); (c) compliant robot made of a tensegrity structure (Caluwaerts et al., 2013); and (d) plant as reservoir. Figure and caption reused from Pieters (2022) with permission from the author.

more feasible than power-hungry physics simulations (see e.g. Burms et al., 2015; Caluwaerts, 2014), or for the prediction of chaotic systems where powerful computers are not able to predict the system better than low-powered reservoir models (Gauthier et al., 2021).

2.2.1 Generalized Mathematical Model

The mathematical model of the Echo State Network (section 2.1.1) can be generalized to describe any reservoir computer, physical or otherwise (Nakajima, 2020). Doing so will be helpful to determine what defines the RC system's computational capacity in the following sections. This section establishes a mathematical vocabulary for describing any PRC model in discrete-time notation.

The reservoir update function can be described as a function $f(\cdot, \cdot)$ that maps the previous reservoir state and the current input to a new reservoir state,

$$\mathbf{x}[t] = f(\mathbf{x}[t-1], \mathbf{u}[t]) \quad (2.5)$$

where \mathbf{x} describes the entire reservoir state at step t and \mathbf{u} represents all external factors that influence the dynamics of the reservoir substrate. Assuming the reservoir dynamics result in emergent memory effects, it is possible to reformulate the reservoir state \mathbf{x} as a function of the past inputs it has seen:

$$\mathbf{x}[t] = \phi(\mathbf{u}[t], \mathbf{u}[t-1], \mathbf{u}[t-2], \dots) \quad (2.6)$$

This function ϕ is called the input echo function and is a key characteristic of each particular reservoir

substrate. The target task is defined as any arbitrary function T of the past time steps, thus requiring memory to solve:

$$\mathbf{y}[t] = T(\mathbf{u}[t], \mathbf{u}[t-1], \mathbf{u}[t-2], \dots) \quad (2.7)$$

Finally, the linear readout function ψ is described as a learned approximation of T , based on the reservoir state:

$$\begin{aligned} \psi(\mathbf{x}[t]) &= \psi(\phi(\mathbf{u}[t], \mathbf{u}[t-1], \mathbf{u}[t-2], \dots)) \\ &\approx T(\mathbf{u}[t], \mathbf{u}[t-1], \mathbf{u}[t-2], \dots) = \mathbf{y}[t] \end{aligned} \quad (2.8)$$

It is important to note that the input echo function ϕ is surjective for virtually all reservoirs. In other words, a unique sequence of inputs does not necessarily result in a sequence of unique reservoir states. Not only is this natural, as no reservoir can have infinite memory of past events, Section 2.3 shows that this summarizing property of ϕ is vital to derive any meaningful computation from reservoir computers.

2.3 The Echo State Property

Strictly speaking, for a reservoir to possess reproducible computational properties, it must show a deterministic relation between its input sequence and its state (Nakajima, 2020). Put plainly, applying the same input sequence should reliably produce the same sequence of output states.

This precondition is particularly strict on physical systems because it requires precise control over the initial conditions of the physical reservoir (Equation 2.5). This requirement is virtually unattainable as it is impossible to account for all physical factors of an environment that influence the reservoir before the start of the experiment. Fortunately, a lesser but sufficient condition exists to perform useful computation with a reservoir.

The echo state property (ESP) (also referred to as the fading memory property) defines a less strict precondition under which it is possible to derive computations from a dynamical system. Informally, the ESP states that a sufficient condition for reservoir computing is for the input echo function ϕ (Equation 2.6) to have fading memory of past inputs (Lukoševičius et al., 2012). If this assumption is valid, then the trajectories of the reservoir state will converge when subjected to the same input sequence, regardless of the initial conditions. A side-effect of the echo state property is that it limits the reservoir's memory of past events, thus limiting the available integration window for memory-bound tasks.

Figure 2.3a demonstrates the principle using an ESN. It shows the evolution of a single node in the network for two runs. At step 0, we initialized the reservoir state with a different random state vector. After step 0, we subject the reservoir to the same input sequence for both scenarios. While the initial trajectory depends on the node's starting value, the neuron synchronizes to the input sequence entirely by step 75. The node's evolution no longer depends on the initial value from that point onwards.

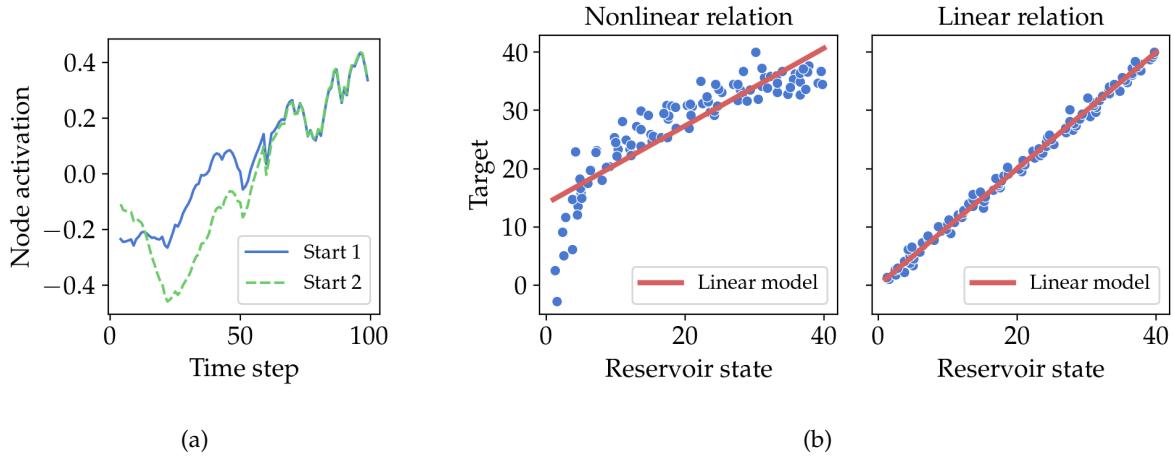


Figure 2.3: Demonstration of the Echo State Property and illustration of linear input separation. (a) Demonstration of the Echo State Property in an ESN with 100 nodes, leak rate 0.16 and spectral radius 1.25. The ESN is subjected to the same input sequence, with two different initial state vectors. The plot shows the activation value of the same node through time. (b) Example of linear input separation between a target task and the reservoir state. The left plot shows a nonlinear relation (in this case, logarithmic). The right plot shows a linear relation. The red line shows a linear regression model fitted to the data.

Fading memory can be recognized in physical dynamical systems as well. Oftentimes it is the result of a dampening force such as electrical resistance or drag. For example, Nakajima et al. (2015) implemented PRC using a silicone octopus arm submerged in water. In their particular case, the fluid resistance of the water tank exerts a dampening effect on the silicone arm, resulting in a fading memory.

The understanding behind this property is that the response of many dynamical systems eventually synchronizes to the input signal, suppressing chaotic behavior (Nakajima, 2020). However, the mathematical understanding of the ESP and its relation to dynamical systems is still a topic of active research (see e.g. Lu et al., 2018; Yildiz et al., 2012).

2.4 Time Scale of Reservoir Dynamics

As Section 2.1 already stated, reservoir computing occurs inherently in the time domain. That means we must carefully consider the time scale at which the reservoir dynamics occur. The time scale of reservoir dynamics can vary significantly. For example, imagine a physical reservoir implemented as a resistive-capacitive electrical network. The time it takes for the network to reach a resting state can be changed by increasing or decreasing the resistance and capacitance of the elements present in the network. We can tune the dynamics' time scale to range from milliseconds to minutes or hours.

A faster time scale also means the reservoir returns to its resting state faster, decreasing the system's memory capacity. Figure 2.4 from Verstraeten (2009) illustrates the effect of time scale on memory capacity. The plot shows the memory capacity for ESNs with twenty nodes and different leak rates. The smaller the leak rate, the wider but flatter the memory curve becomes. That means that networks with a lower leak rate show better long-term memory at the cost of precision in memory of more recent inputs.

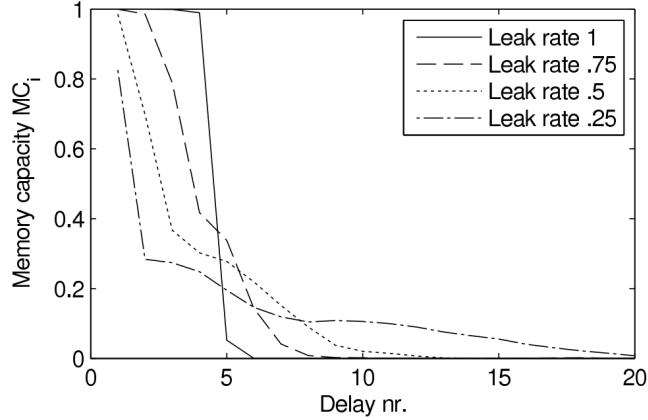


Figure 2.4: Memory capacity for ESNs with different leak rates (see Subsection 2.1.1). Networks with a smaller leak rate show better long-term memory at the cost of decreased accuracy in short-term memory. This figure originally appeared in Verstraeten (2009).

The takeaway here is that the time scale of the reservoir dynamics plays a vital role in its ability to solve a given task. A reservoir highly synchronized to its inputs will display faster dynamics and thus a shorter but greater precision integration window for solving tasks. Conversely, a reservoir with higher memory capacity will typically exhibit slower dynamics and thus a longer integration window, at the expense of short-term precision.

2.4.1 Sampling of a Continuous-Time Reservoir

While the dynamics of a physical reservoir operate in continuous time, most applications implement the readout model digitally using a conventional computer and sensors that discretely sample the input signals and reservoir state. According to the Shannon-Nyquist sampling theorem, to sample a continuous signal without losing information we must sample at twice the rate of the largest frequency component present in that signal (Shannon, 1949). Hence, determining the right time scale at which the desired reservoir dynamics occur plays a prominent role in PRC system design. Undersampling critical signals in the reservoir will result in information loss by introducing aliasing in the sampled signal. In conclusion, effectively sampling the reservoir state presents an additional optimization problem when implementing RC with physical reservoirs.

2.5 Evaluating Reservoir Performance

Before attempting to solve a specific computational task, it is important to map out what the physical reservoir can do by benchmarking its computational properties. To have computational capacity, a reservoir should show input separability and fading memory (Nakajima et al., 2015). Evaluating these properties is in fact a way of characterizing the echo response function of the system as described in Section 2.2.1.

2.5.1 Input Separability

The readout function is a strictly linear model. To solve nonlinear tasks, the reservoir must transform its inputs such that the task becomes a linear function of the reservoir state. In other words, there must be a linear relation between the reservoir's observable state and the target task. Figure 2.3b shows an example of linear input separation in a reservoir.

However, the nonlinear transformation applied by a finite reservoir cannot solve any arbitrary task. In addition, the presence of notable recurrent effects in the reservoir may degrade precision in recalling recent inputs (Section 2.4 explores this phenomenon further). When the input transformation applied by the reservoir results in an observable state from which a readout function can compute the desired task efficiently, we speak of good input separability. Good input separability is a function of both nonlinearity and memory, and the optimal mix of both properties varies depending on the task.

One way to evaluate the input separation of a reservoir relative to a given task is to train a reservoir readout and a control readout model (Ushio et al., 2021). The former fits to the reservoir state, while the latter fits to the environmental signals directly. If the expanded state of the reservoir contains helpful input separation for the task, the reservoir readout will perform better than the control model.

2.5.2 Fading Memory

Section 2.3 already stated the importance of fading memory in physical reservoirs. However, this property is difficult to measure directly. Nonetheless, we will outline two ideas for measuring this property in physical systems.

The first method attempts to measure the warmup time of the reservoir. The idea is simple: initialize the same reservoir with different starting conditions, then measure the time it takes for the reservoir state to synchronize to the input. If the physical system displays the fading memory property, then the trajectories of the reservoirs will eventually synchronize, even if initialized with different starting conditions. Figure 2.5a demonstrates this method on an ESN with different leak rates. In this experiment, An ESN is subjected to identical input series with two different initial reservoir states x . The rate at which the reservoir syncs to the input signal depends on the time scale of the reservoir dynamics dictated by the leak rate.

In the second method, we subject two reservoirs to identical input signals. One reservoir will serve as the test subject, while the other acts as a control. After the two systems have synchronized to the input, we briefly apply an environmental impulse to the test reservoir, after which the original input signal resumes. We then measure the time for the test reservoir to resynchronize with the control reservoir. This impulse can look different depending on the type of reservoir. For example, we can temporarily remove the reservoir from its environment such that it no longer receives any environmental stimuli. Another possibility is to apply an unexpected and artificial stimulus for a short time. Regardless of how we realize the impulse, we must apply it long enough to have a noticeable impact on the trajectory of the reservoir dynamics. Otherwise, it is impossible to measure the time until the reservoir synchronizes to the regular inputs. Figure 2.5c demonstrates this method on an ESN. This time, all input values between time steps 60 and 80 are set to a fixed value of ten times the standard deviation of the input signal. After step 80, the reservoir gradually returns to its original behavior, at a rate proportional to its leak rate.

Both methods rely on measuring the divergence in behavior of a reservoir when subjected to different conditions. To quantitatively compare the difference between reservoir states at time t , we can use a distance metric such as Mean Squared Error (MSE) on its observable state X :

$$\text{MSE}(t) = \frac{1}{N} \sum_{i=1}^N \left(X_i^{(1)}(t) - X_i^{(2)}(t) \right)^2 \quad (2.9)$$

where N is the size of the reservoir's observable state vector.

Note that in both experiments, the time it takes for the reservoirs to synchronize also informs about its recurrent dynamics and memory capacity.

2.5.3 Computational Benchmarks

To better compare results between studies of different reservoirs, some standard benchmarks appear in the literature. The most common RC benchmark is the nonlinear autoregressive moving average (NARMA). We can tune the nonlinearity and memory of the NARMA task with the n parameter. The NARMA system is defined as:

$$y(t+1) = \alpha y(t) + \beta y(t) \left[\sum_{i=0}^{n-1} y(t-i) \right] + \gamma x(t-n+1)x(t) + \delta \quad (2.10)$$

where the parameters α , β , γ , and δ are set to 0.3, 0.05, 1.5 and 0.1 (Nakajima et al., 2015).

For some reservoirs, it is necessary to tune the NARMA system's memory to the correct time scale to get stable results (Pieters, 2022). For this the adapted formula can be used:

$$y(t+1) = \alpha y(t) + \beta y(t) \left[\sum_{i=0}^{n-1} y(t-\tau i) \right] + \gamma x(t-\tau n+1)x(t) + \delta \quad (2.11)$$

where τ is tuned to the desired time scale of the NARMA system's memory.

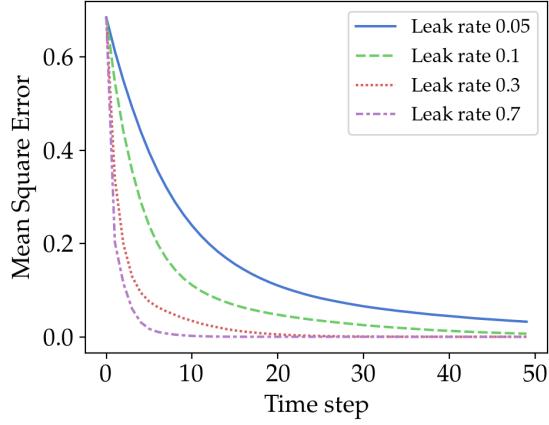
We can use other simple benchmarks to evaluate the memory and nonlinear properties. For example, predicting the past inputs of the reservoir for increasing delay tests the memory of the system:

$$y(t) = x(t-\tau) \quad (2.12)$$

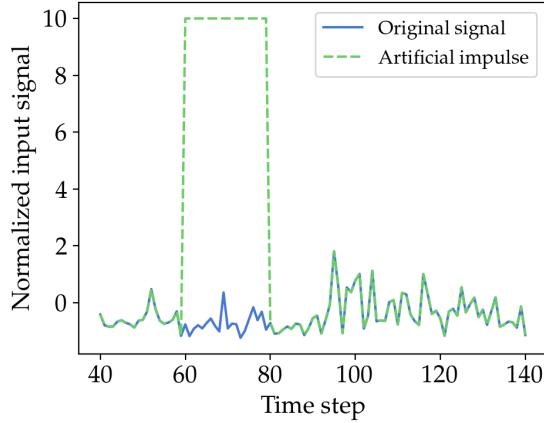
where greater values of τ require more memory to solve ($\tau > 0$). To test reservoir nonlinearity, a simple exponentiation of the input signal can be used as the prediction target:

$$y(t) = (x(t))^n \quad (2.13)$$

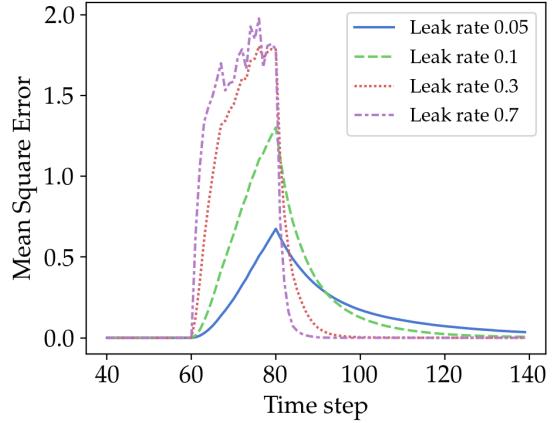
where n is greater or less than 1.



(a) First method: divergence between reservoirs



(b) Second method: applying an artificial impulse



(c) Second method: divergence between reservoirs

Figure 2.5: These figures show experimental data from testing the fading memory property of an ESN as described in Subsection 2.5.2. Both experiments use a reservoir of 100 nodes with randomly initialized weights. For stability, we normalized the weights to have a spectral radius below 1.25 (Caluwaerts, 2014). As input data, we used the California Housing regression dataset (Kelley Pace et al., 1997) which can be found in the sklearn Python package. (a) The divergence in reservoir state between two runs of the same reservoir with identical inputs and different starting state vectors. The starting states are sampled from a uniform distribution between -1 and 1. The disparity between the reservoirs is measured by the MSE between their respective state vectors (Equation 2.9). (b) An artificial impulse with an amplitude of ten times the standard deviation of the original signal is applied between time steps 60 and 80. (c) The divergence in reservoir state between two runs of the same reservoir using identical starting conditions. One run receives an artificial impulse between time step 60 and 80 on all input dimensions, as depicted by Subfigure b.

Finally, to test a reservoir's ability to predict the trajectory of chaotic systems, Ushio et al. (2021) have proposed a test using strange attractors. In this test, a time series of a Lorenz attractor acts as the input to the reservoir. The reservoir then has to predict near-future values of this input time series. The test evaluates how well the reservoir can synchronize with a chaotic system, which is a common application for RC (Tanaka et al., 2019).

2.6 Simulation of Physical Reservoirs

If the behavior of the ϕ -function is understood completely, it is possible to create a computer model that simulates the reservoir dynamics (Nakajima, 2020). Using a simulated reservoir can have various advantages in research. (i) It is possible to iterate faster on designing synthetic physical reservoirs, such as FPGAs or photonic circuits, by validating designs in software before constructing the physical reservoir. (ii) Some reservoirs operate at long time scales (Ushio et al., 2021). Simulation can accelerate the research cycle by running experiments faster than in real-time. (iii) In simulation, it is possible to operate in a highly controlled environment, which makes it possible to solve a constrained version of a problem first. For example, it is possible to dictate identical initial conditions in every experiment run or enforce static reservoir dynamics where the physical counterpart may experience some malleability caused by the environment.

2.7 Summary

This chapter showed how physical reservoir computing outsources computation to physical dynamical systems. This works thanks to the emergent computational qualities present in nonlinear dynamical behavior. A simple linear readout function is used to transform the reservoir's state into a useful output. In computer models, we can observe the reservoir state directly. In physical implementations, we observe the reservoir using sensor data.

To perform useful computations, the reservoir must perform the suitable nonlinear transformations and convolutions to its inputs such that the problem becomes linear. Another condition for PRC is a fading memory of past inputs; Fading memory is essential to performing deterministic computations.

To test whether a physical material is suitable for reservoir computing, we can design a suite of regression targets to quantify the reservoir's correlation with relevant computational tasks. In addition, we proposed two types of experiments for testing the fading memory condition: one measures the dependence on initial conditions, and the other measures the stationarity of the reservoir dynamics in time.

Chapter 3

Functional-structural Plant Models

This chapter looks at current developments in the field of *in silico* plants, intending to design RC experiments with simulated organisms. Research with *in vivo* plants is difficult and time-consuming as it involves meticulous experimental design, long observation times, and of course keeping the plants healthy during the experiment(Pieters et al., 2021). For these reasons, it is of great interest to explore computer simulations of plants that can accelerate research in reservoir computing applications.

3.1 Selecting Plant Models for RC Experiments

There exist many different types of plant models. To get an idea of the various model categories one can visit *Quantitative Plant*^{*}, a website that aggregates open-source plant models from the literature. Listed models range from crop surface models (CSM) that estimate biomass and crop yield from aerial imagery (Tilly et al., 2014) to sub-plant models that simulate gas exchange between the canopy and the environment (Sinoquet et al., 2001).

To select plant models for use in RC experiments, they must meet certain criteria. Firstly, the model must track the state of a single plant over time, as the scope of the research presented limits the interpretation of the reservoir to a single organism. This rules out the use of e.g. crop surface models because they only inform about large populations of plants; in these models information on the level of an individual organism is lost (although it would certainly be interesting to apply reservoir computing principles to the ecological dynamics of a crop field).

Secondly, the plant's state should be observable from multiple fixed points of the plant's anatomy. These observed points must be distinct structural elements of the plant that simulate the physiological processes of a real specimen and must communicate with one another as a dynamical system. We set this requirement because an RC system using a plant model as reservoir is intended to represent the equivalent setup in the real world. This requirement entails a preference for mechanistic models that make few prior assumptions about plant behavior, leaving macroscopic behavior to emerge naturally from microscopic interactions at the cellular level (Gauthier et al., 2020). Ideally, the observable parameters of the plant are also empirically measurable *in vivo* to enable the research results from this work to be better transferable to a real-world setting.

Finally, the time resolution of the simulation should be sufficiently fine to observe dynamics at multiple time scales as well as avoid signal aliasing (see Section 2.4).

^{*}<https://www.quantitative-plant.org/model>. See Lobet (2017) and Lobet et al. (2013) for more details.

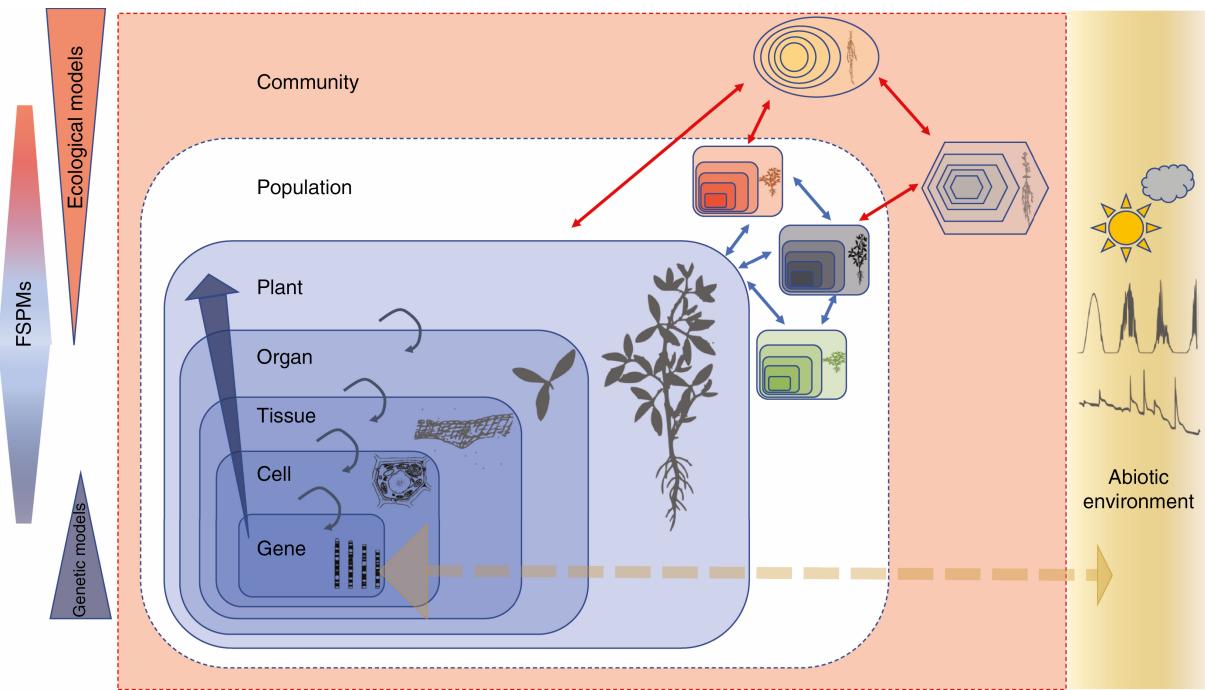


Figure 3.1: Interactions at micro- and macroscopic scales present in functional-structural plant models. FSPMs typically implement a subset of the scales depicted in the figure. The interactions between structural elements at one level govern the macroscopic behavior at the higher level. All levels interact with the environment. In some models, the behavior of the lowest level units is governed by variable genetic properties. This figure is reused from Louarn et al. (2020) with permission from the publisher.

3.2 Functional-structural Plant Models (FSPMs)

A promising category of models that can meet the stated criteria is that of functional-structural plant models (FSPM) (Godin et al., 2005; Louarn et al., 2020). FSPM model the plant as a collection of interconnected units that represent a small structural part of the plant in space. These units interact with the environment and with each other to simulate eco-physiological processes (Coussement et al., 2020). The structural elements of the simulated plant form a hierarchy of interactions at different scales. Together they govern the macroscopic behavior of the organism (Louarn et al., 2020). Most FSPMs simulate interactions across a range of scales, from cell organelles to whole plant organs and even populations of multiple organisms. A schematic representation of the different scales of interactions is shown on Figure 3.1.

Functional-structural plant models owe their success to more than two decades of interdisciplinary collaboration between biology, biophysics, ecology, and computer science (Louarn et al., 2020). At present, FSPMs have been applied in developmental biology (Cieslak et al., 2016; Gauthier et al., 2020; Schneider et al., 2019), integrative and systems biology (Chang et al., 2019; Millar et al., 2019; Zhu et al., 2016) and applied plant sciences (Albasha et al., 2019; Forschungszentrum Jülich, Germany et al., 2021; Renton et al., 2017). The accelerated development in this area of research is enabled by common frameworks that standardize the description of e.g. plant architecture (Boudon et al., 2012; Godin et al., 1998) and eco-physiological interactions (Fournier et al., 2016; Zhou et al., 2020). Popular platforms for functional plant modelling workflows are OpenAlea (Pradal et al., 2008, 2015), GroIMP (Kniemeyer et al., 2007)



Figure 3.2: Two examples of functional-structural plant models. Figures originally appeared in their respective cited works. (a) Soybean FSPM (Coussement et al., 2020). Figure reused with permission from the publisher (Copyright 2018 Oxford University Press). (a) V-Mango (Boudon et al., 2020). Figure reused under the CC BY 4.0 license.

and GreenLab (Reffye et al., 2021).

The precise set of simulated physiological processes varies from model to model. Commonly modeled factors include hydraulic pressure, water and carbon flow, and photosynthesis (Albasha et al., 2019; Coussement et al., 2020; Zhou et al., 2020). The complexity of the environmental simulation also differs between models, ranging from manual application of input values in root nodes (Zhou et al., 2020) to ray-traced application of lighting conditions, including self-shading (Coussement et al., 2020). Further still, some FSPMs model plant growth over the span of an entire growth season (Coussement et al., 2020; Lecarpentier et al., 2019) where others assume a static plant geometry with predefined limits for simulation duration and model accuracy (Albasha et al., 2019).

Figure 3.2 highlights a small selection of models from the literature. First, Soybean FSPM (Coussement et al., 2020) is a highly detailed mechanistic plant growth model. Soybean FSPM was developed to understand better a plant's macroscopic behavior in relation to water. The model simulates vegetative growth driven by hydrostatic pressure inside the plant cells and incorporates a ray-traced lighting model with dynamic lighting and shading conditions (Figure 3.2a). V-Mango (Boudon et al., 2020) models the vegetative and reproductive development of mango trees to enable *in silico* experimentation with better cultivation practices. The model allows researchers to understand why certain characteristics of mango trees make them more susceptible to pests and to study the influence of the three-dimensional tree architecture on fruit growth. While this is far from an exhaustive list of what has been accomplished with FSPMs, it should give the reader an idea of what is possible with this first-principles approach to plant modeling.

An important caveat is that these models typically do not simulate a live plant as an accurate digital twin of an *in vivo* plant. Instead, they are commonly developed with a specific research question in mind and model only the physiological processes that are deemed correlated to the studied phenomenon (Louarn et al., 2020). Hence, it is warranted to perform a careful comparison between the conditions and assumptions of the original study and those of a new context before reusing the model for tasks other than its original purpose. Note that digital twin initiatives are also an active line of research. For

example, researchers at WUR have been developing a digital twin system for tomato plants (NPEC, 2020).

3.3 Summary

This chapter introduced the research field of functional-structural plant modeling, a type of model that has become the state-of-the-art in plant modeling in the past years. FSPMs represent the plant as a 3D structure of interconnected elements, where each structural element simulates eco-physiological processes such as photosynthesis, transpiration, and water and carbon flow. FSPMs are ideal for reservoir computing research because they model the interaction in space and time between the plant's structure and the abiotic environment. Therefore, if plant physiology displays properties of nonlinear dynamical systems, this will be captured by a sufficiently detailed FSPM.

Part II

Methodology

Chapter 4

Experimental Design

Chapter 2 established the necessary background on RC and explained how physical media can function as a computational resource. Then, Chapter 3 introduced the rich research field of functional-structural plant modeling. Now, we present how we will identify reservoir dynamics in plants and how we can quantitatively measure the reservoir performance, using simulated plants as the experiment subject.

4.1 Reservoir Dynamics in Plants

To recap Chapter 2: to solve a computational problem with a physical medium, it must possess two fundamental properties. The first is linear input separation relative to the target signal. A reservoir can achieve this by performing a nonlinear expansion of its environmental inputs. To perform tasks that require memory, the reservoir must also possess recurrent dynamics; the reservoir integrates current inputs with its memory of the past. The second requirement is a fading memory of past inputs. Without a fading influence of past inputs, computations carried out by the reservoir cannot consistently be reproduced. The question remains: do plants display these properties?

First, we consider linear input separation. To make a target a linear function of the reservoir state, the reservoir must perform some nonlinear transformation of its inputs. An example of this in plants is how grass responds to soil temperature. The growth rate of grass shows a roughly parabolical relationship with soil temperature, reaching an optimum around roughly 15-20°C, after which growth declines due to heat stress (Hurtado-Uria et al., 2013). This shows that a plant's behavior can show a nonlinear relation with its environment.

Next, we consider the fading memory property. Clearly, this rule does not hold for the general case; a plant temporarily subjected to extreme cold or heat will die and no longer process any information. Favorable environmental conditions will influence the plant's growth rate, permanently altering reservoir dynamics as a function of past inputs. Nonetheless, we might be able to identify operating ranges for specific eco-physiological processes and time scales for which the fading memory property holds up approximately.

In addition to the properties above, the PRC framework assumes the reservoir dynamics are unchanging over time. Naturally, changing the reservoir's constitution also changes its nonlinear characteristics and memory capacity. Plants are not stationary reservoirs in the general case because they go through several growth phases that change their structure and behavior drastically. But here, too, we can approximate stationarity by investigating reservoir dynamics within a time window for which the growth rate is negligible.

In conclusion, there is sufficient evidence to suggest that plants can display computational properties for eco-physiologically relevant tasks. However, plant RC research is still in its infancy. It remains

unexplored under which conditions the RC framework applies to plant behavior and what subset of eco-physiological processes show reservoir-like dynamics.

4.1.1 Research Objectives

Pieters (2022) has already demonstrated that plants can function as a medium for reservoir computing. This work aims to expand on the foundation for plant RC. We aim to answer the following questions: From the subset of plant physiological processes that are empirically observable, which processes show promising input separation for eco-physiological tasks and computational benchmarks? And do these properties display the fading memory principle at the relevant time scale? To answer these questions, we propose a framework for quantitatively measuring the reservoir properties of the plant's physiological functions. This framework enables direct comparison of the reservoir performance of different physiological processes and across multiple plants.

We used FSPMs to test this framework. Using simulated plant has several advantages. It gives us the freedom to observe the plant directly, without the need for sensing technology and laboratory setups. Consequently, it speeds up the research process and gives us a first intuition into suitable plant reservoirs. This can help bootstrap future *in vivo* research. In addition, we believe that RC can contribute to the plant modeling community by identifying potential weaknesses in FSPMs. The reason is that RC can be used for analyzing short-term as well as large scale variations; whereas current modeling efforts focuses only on long term behavior, even though research shows that rapid physiological responses play an essential role in plant life (Alarcon et al., 1994; De Swaef, De Schepper, et al., 2015).

4.2 Framework For Measuring Plant Reservoir Dynamics

In this section, we propose our framework for quantitatively measuring input separation and fading memory properties of eco-physiological processes in plants. With these methods we will compare the capabilities of different processes. The results will highlight which observable plant properties to pursue in future RC research.

4.2.1 Input Separation

First, we measure the linear input separation of the reservoir w.r.t. various target tasks. We propose a collection of regression targets based on the environment and physiological state of the plant. We choose these tasks to test the reservoir's capacity to solve biological problems. Classification tasks are not considered because they are less relevant from an eco-physiological standpoint; most processes are real-valued and occur in continuous time. We fit a simple linear readout model to the reservoir observations to test the prediction accuracy. We then compare the predicted series against the ground truth data to compute an accuracy metric.

To paint a complete picture of the reservoir, we use a mix of target tasks. We propose three categories of targets, previously established by Pieters (2022): (i) environmental inputs, (ii) eco-physiological state, and (iii) computational benchmarks. Predicting environmental inputs gives us a first indication of the

plant's coupling with the environment. For example, we can use ambient temperature, incident solar radiation or relative humidity as a prediction target. Next, eco-physiological tasks indicate whether the observed reservoir is effective for solving biological problems. Targets in this category can be, e.g. transpiration rate or photosynthesis rate on the organism level. Finally, computational benchmarks measure a reservoir's degree of nonlinearity and memory. We propose three targets: a delay line to measure memory capacity (Equation 2.12), a polynomial expansion to measure nonlinearity (Equation 2.13), and the NARMA benchmark to measure both properties (Equation 2.11). Each artificial target uses an environmental series as input. These targets are not practical tasks for a plant, but the results can help contextualize plants in the broader PRC literature.

We fit a linear readout model for every reservoir-target combination. For the readout model we propose a simple linear regression model with a bias term:

$$\hat{y}[t] = \mathbf{W}^{\text{out}} [1; \mathbf{X}[t]] \quad (4.1)$$

To isolate the computational capacity of the reservoir from the natural correlation between the target and the environment, the environmental inputs are not fed into the readout model; the prediction relies on the reservoir only. Model parameters \mathbf{W}^{out} are fitted to training data using ridge regression (Equation 2.4). The regularization λ parameter is tuned for each reservoir-target pairing. To find the optimal λ , we propose a parameter sweep with logarithmic spacing, validated using cross-validation on the training data.

We evaluate the reservoir performance on a set of held-out test data. Prediction accuracy is measured using the Normalized Mean Square Error (NMSE) metric:

$$\text{NMSE} = \frac{1}{N} \sum_{t=1}^N \frac{(y[t] - \hat{y}[t])^2}{\text{var}(y)} \quad (4.2)$$

A lower score means a more accurate prediction. The NMSE has several advantages over regular MSE (Pieters, 2022). Because it is normalized, the results can be compared across reservoir-target pairings, including between plants. It is also easy to interpret: a perfect predictor scores 0.0, while predicting the signal mean for every step yields a score of 1.0.

4.2.2 Fading Memory

To test fading memory, we propose the impulse experiment described in Section 2.5.2 and illustrated in Figures 2.5b and 2.5c. The impulse can be applied to any of the environmental inputs. The amplitude and duration of the impulse should remain within realistic boundaries so that an actual plant would not suffer permanent damage that alters the reservoir dynamics.

To quantitatively measure the divergence $\delta(t)$ between two reservoir trajectories at a given time step, we propose a modified NMSE metric:

$$\delta(t) = \frac{1}{N} \sum_{i=1}^N \frac{(\mathbf{X}_i^C(t) - \mathbf{X}_i^E(t))^2}{\text{var}(\mathbf{X}_i^C)} \quad (4.3)$$

Where N is the size of the observed reservoir, \mathbf{X}^C the reservoir state in the control experiment, and \mathbf{X}^E the reservoir state during the impulse experiment. We can use a line plot of the divergence during and after the impulse to inspect the fading memory property of the reservoir. A reservoir affected by the impulse should show a peak in divergence once the impulse is applied. If the physiological process displays fading memory, the divergence should go down to pre-impulse levels in a finite amount of time after the artificial stimulus is removed. Note that a substantial divergence between reservoirs subjected to the same inputs also indicates the echo state property may not hold.

4.3 Considered Reservoirs

Many physiological processes inside the plant can be considered as a reservoir. To ensure the results from this work can form the foundation for future research, we direct our attention towards those processes which are observable using existing sensing technologies. In his doctoral dissertation, Pieters (2022) compiled comprehensive tables of physiological traits that are observable using imaging techniques (Table 4.1) and contact-based techniques (Table 4.2). We can use these tables to choose which processes to look for when selecting plant models.

This work shows a preference for reservoirs observable by contactless imaging systems. In the future, contactless systems will play a key role in scaling plant RC to large arrays of plants because they do not influence reservoir dynamics and can observe multiple plants at once. However, imaging technologies for plant RC also have some drawbacks. They are more susceptible to noise induced by ambient factors than contact-based sensors. Dynamic ambient factors must also be accounted for when calibrating the measurements. Image-based systems tend to have a lower resolution, capturing less subtle variations than contact sensors (Pieters, 2022).

4.4 Summary

This chapter proposed a framework for quantitatively measuring reservoir characteristics of plant-physiological behavior. In the first test, we use a suite of regression targets and a standard readout model to measure the capacity to perform nonlinear tasks. The second test investigates the fading memory property in an experiment where a brief artificial stimulus is applied to the plant's environment. For the considered reservoirs, we prefer observable physiological traits using contactless sensing technologies. Still, the selection of reservoirs shown in the results will depend on what is available in the selected plant models.

Table 4.1: Imaging techniques for various plant-physiological processes. From Pieters (2022): “Depending on sensor type, different target traits can be investigated. Moreover, the timescale of the physiological process is also indicated. When measuring on this timescale, we can observe noticeable variation in that particular trait. It ranges days (d), hours (h) to minutes (min) and seconds (s).” Table reused from Pieters (2022) with permission from the author.

sensor	target trait	timescale	reference
thermal camera	leaf temperature T_{leaf}	s→min	Costa et al. (2013)
	stomatal conductance ($\leftarrow T_{\text{leaf}}$)	s→min	Jones (1999)
	transpiration ($\leftarrow T_{\text{leaf}}$)	s→min	Maes et al. (2012)
	drought stress ($\leftarrow T_{\text{leaf}}$)	d	De Swaef et al. (2021)
	diseases and pathogens ($\leftarrow T_{\text{leaf}}$)	h→d	Costa et al. (2013)
depth camera/LiDAR	heat stress ($\leftarrow T_{\text{leaf}}$)	min→h	Janka et al. (2013)
	3D plant architecture	h→d	Busemeyer et al. (2013), Friedli et al. (2016)
	leaf area growth	h→d	Walter et al. (2007)
	plant height	h→d	Borra-Serrano et al. (2019)
	plant biomass	h→d	Golzarian et al. (2011)
RGB camera	plant size and architecture	h→d	Lootens et al. (2016)
	drought stress (reflectance and biomass)	d	Mazis et al. (2020)
	diseases and pathogens (reflectance)	h→d	Chaele et al. (2007)
	leaf photosynthesis	min→h	Baker (2008)
	stomatal conductance	s→min	Nejad et al. (2005)
chlorophyll fluorescence imager	chilling stress	min→h	Devacht et al. (2011)
	drought stress	d	Meyer et al. (1999)
	diseases and pathogens	min→h	Berger et al. (2007)
	heat stress	h	Briantais et al. (1996)
	salt stress	h→d	Moradi et al. (2007)
hyperspectral camera	nutrient deficiency	d	Ciompi et al. (1996)
	leaf pigments	d	Blackburn (2007)
	canopy density	d	Carlson et al. (1997)
	photosynthesis (PRI)	min→h	Inoue et al. (2008)
	water content	h→d	Whetton et al. (2018)
	drought stress	d	Berger et al. (2010)
	diseases and pathogens	h→d	Bock et al. (2010) and Van De Vijver et al. (2020)
	nutrient deficiency	d	Zhao et al. (2005)

Table 4.2: Non-imaging techniques for various plant-physiological processes. From Pieters (2022): “Depending on sensor type, different target traits can be investigated. Moreover, the timescale is also indicated in which measurements with noticeable variation can be recorded, ranging from days (d), hours (h) to minutes (min) and seconds (s).” Table reused from Pieters (2022) with permission from the author.

sensor	target trait	timescale	reference
weighing scale	transpiration	s→min	Wallach et al. (2010)
	plant biomass	d	Van leperen et al. (1994)
	drought stress (biomass)	d	Wallach et al. (2010)
LVDT	leaf length	s→min	Barillot et al. (2020)
	stem diameter	s→min	De Swaef, De Schepper, et al. (2015)
leaf clip	leaf thickness	s→min	De Swaef, Vermeulen, et al. (2015b)
	stem water potential	s→min	Vandegehuchte et al. (2014)
Zim-probe	leaf turgor pressure	s→min	Rodriguez-Dominguez, et al. (2019)
	formation of cavitations	s→min	De Baerdemaeker et al. (2019)
acoustic	water uptake	s→min	Ehosioke et al. (2020)
	root biomass	s→min	Ehosioke et al. (2020)
sap flow	water uptake	min	Hanssens et al. (2015)
	transport rate	min	

Chapter 5

Selected Plant Models

For the reasons stated in Chapter 4, we conduct our experiments with *in silico* plants. This chapter starts with a set of formal requirements that plant simulations must meet for use in physical reservoir computing. Then follows an overview of the FSPMs from the literature we have examined for use in this research. Finally, we introduce the final selection of models and compare their traits and characteristics.

5.1 Selection Criteria

Section 3.1 already formulated the general requirements of a plant model for use in the PRC framework. This section now formalizes the search criteria from the point of view of the experiments we wish to conduct. The selected models must meet the following specifications:

1. **Modeled as a dynamical system.** To observe reservoir dynamics, there must be a dynamical relationship between the structural elements of the plant.
2. **Local observability of eco-physiological processes.** To create a readout of the reservoir, standard eco-physiological functions, such as surface temperature, transpiration rate, and stomatal conductance, must be observable at each of the discrete elements of the plant.
3. **Sufficiently small simulation steps.** Small simulation steps can reveal dynamics that occur at shorter time scales. Ideally, we can observe dynamics at the second to hour scale.
4. **(Pseudo-)static reservoir.** Reservoir Computing assumes the reservoir dynamics are unchanging through time. To satisfy this condition, we limit our search to plant models 1 Research on the effect of plant growth on reservoir dynamics is out of scope for this research.
5. **Practically observable physiological processes.** Section 4.3 explained that there is a preference for models that simulate physiological processes which we can observe using existing (contactless) sensing technologies.

The first and second requirements are met by narrowing the search to functional-structural plant models, as was established in Chapter 3. However, we must evaluate the third, fourth, and fifth conditions on a case-by-case basis. For the fifth condition, Tables 4.1 and 4.2 can be used to determine which models to consider.

Some additional requirements arise from a practical aspect. First, the plant simulation must be able to run for a large number of time steps. This is needed to build a sufficiently large dataset to train readout models on. Also, observing fading memory effects requires a sufficient length of simulated steps. Second, the model software must be easily adapted to extract the evolution of the reservoir state during simulation and store it in a time series dataset.

5.2 Related Work

There are diverse models in the literature, but many do not meet all our demands. There are a few common shortcomings. For one, many FSPMs only model a limited set of physiological processes (Louarn et al., 2020). These simulations are too narrow to gain convincing evidence of reservoir dynamics present in real-world plants. Other published works show sophisticated architectures and solid real-world validations but do not make their simulations open source (Prieto et al., 2020).

Table 5.1 shows a selection of plant models we examined more closely. We briefly investigated V-Mango (Boudon et al., 2020) and WALTer (Lecarpentier et al., 2019). Ultimately, we disregarded these models because they model growth with time steps of one day. This step size is too large to explore reservoir dynamics at the time scales we want to observe (1 s-1 h).

PlaNet-Maize (Lobet et al., 2014) and Soybean FSPM (Coussement et al., 2020) are growth models, but the step size of one hour warrants taking a deeper look; when the growth rate is very slow compared to the time scale of the target task, the reservoir can be regarded as approximately static, as Pieters (2022) demonstrated with *in vivo* strawberry plants. Unfortunately, we could not further explore this direction due to practical reasons. PlaNet-Maize is open-source, but the code repository contains no documentation on getting started or making modifications. The soybean model is closed-source but was made available to us by co-author T. De Swaef. Regrettably, we met many difficulties setting up the simulation environment, and the GroIMP platform made it unintuitive to implement adaptations as an untrained user. Because of these reasons, combined with the limited time scope of this project, we did not explore these models further.

HydroShoot (Albasha et al., 2019) and a similar, unnamed model by Prieto et al. (2020) are both FSPMs of grapevine; they have a static geometry and model a diverse set of physiological functions involved in gas exchange. As far as we could find, Prieto et al. did not publish the code for their model. In contrast, HydroShoot is open source and very well documented. It was straightforward to set up using the Conda package manager for Python. The use of well-documented libraries from the OpenAlea platform made it easy to track the plant’s state or adjust its environment.

Finally, WheatFspm (Gauthier et al., 2020) is another growth model. WheatFspm builds upon the prior work CN-Wheat (Barillot et al., 2016b), a post-vegetative growth model of the same species. Because we prefer (pseudo-)static models, we used CN-Wheat for our experiments instead. The simulation code provided in the source repository already extracts all physiological time series, making CN-Wheat a very convenient model for our purposes.

Both HydroShoot and CN-Wheat have a good overlap in the physiological processes and meteorological inputs they simulate. These shared qualities enable a closer comparison of how the models behave as reservoirs when solving the same regression tasks. The following section introduces the two selected models in more detail and compares them from a RC perspective.

Table 5.1: FSPMs we explored for use in RC experiments. Models highlighted in bold are used for the final results presented in this work.

Model	Crop	Growth	Time Step	Implementation	Open Source
PlaNet-Maize (Lobet et al., 2014)	Maize	Yes	1 h	Java	Yes
CN-Wheat (Barillot et al., 2016b)	Common wheat	No	1 h	Python (OpenAlea)	Yes
HydroShoot (Albasha et al., 2019)	Common grapevine	No	1 h	Python (OpenAlea)	Yes
WALTer (Lecarpentier et al., 2019)	Common wheat	Yes	1 d	Python (OpenAlea)	Yes
Soybean FSPM (Coussement et al., 2020)	Soybean	Yes	1 h	Java (GroIMP)	No
V-Mango (Boudon et al., 2020)	Mango tree	Yes	1 d	Python (OpenAlea), R	Yes
WheatFspm (Gauthier et al., 2020)	Common wheat	Yes	1 h	Python (OpenAlea)	Yes
Prieto et al. (2020) (unnamed)	Common grapevine	No	1 h	Python (OpenAlea)	No

5.3 Selected Models

5.3.1 HydroShoot

HydroShoot is an FSPM of *Vitis vinifera*, otherwise known as the common grapevine. It was published by Albasha et al. (2019) as a study of the gas exchange of large plant canopies under water deficit conditions and to compare the effectiveness of different canopy shapes. It is written in the Python programming language and is built on the OpenAlea platform.

HydroShoot consists of three main modules, simulating the water potential across structural segments, the energy budget of individual leaves, and transpiration and carbon assimilation of individual leaves. Hydraulic pressure gradients across elements induce water flow, creating dynamical behavior between the elements. The authors experimentally validated that gas exchange, leaf hydraulic pressure, and temperature are adequately simulated during daytime hours. However, the model is prone to underestimate leaf temperature during nighttime hours. Nonetheless, on the surface, HydroShoot appears to be a realistic simulation of plant dynamics and thus a good candidate for RC research.

The model uses a static, predefined structure. The geometry is hand-modeled, based on measurements taken from real grapevine specimens. Figure 5.1a shows an example of one of the canopy configurations modeled for HydroShoot. Because HydroShoot does not simulate plant growth or tissue death, we should expect it to behave as an idealized static reservoir. The entire structure consists of hundreds of elements, about three hundred of which are photosynthesizing leaves, depending on the 3D model.

HydroShoot uses hourly meteorological inputs of solar incident shortwave irradiance (PAR), air temperature (T_{air}), relative humidity (RH), wind speed (u), CO₂ concentration, and atmospheric pressure (P_a). However, in the simulations for the 2019 publication, the latter two are fixed to values of 400 ppm and 101.3 kPa respectively.

The simulations in the original publication ran for four days with hourly time steps. In correspondence with the authors, R. Albasha informed us that the model has been used in simulations up to one month in length, using 2 h steps during the day and a 12 h step for nighttime. However, extending the simulation time requires modeling recurrent soil water potential over a longer time than is provided in the original simulation, which is out of scope for this work. As a compromise, we ran the model in runs of 7 days with a simplified model of constant recurrent soil water potential (see Appendix B.1).

5.3.2 CN-Wheat

CN-Wheat is a mechanistic model of *Triticum aestivum*, more widely known as common wheat. CN-Wheat was developed by Barillot et al. (2016b) and is also built on the OpenAlea platform. The model was created to simulate carbon (C) and nitrogen (N) metabolism in wheat culms in the post-flowering stages of development.

The model simulates a rich set of eco-physiological processes that drive the C and N lifecycle inside the plant. This includes photosynthesis and transpiration, synthesis of CN metabolites, metabolite transport in phloem tissue, and interaction between metabolites and the root system. The network of phloem tissue is simplified as a shared pool of C and N that facilitates communication between the functional-structural elements. This simplified transport model limits flow dynamics to the gradients between the elements and this shared pool of resources. It reduces the potential for more intricate dynamics within the network, such as the local influence of neighboring elements over each other's behavior.

Figure 5.1b depicts the entire structural architecture of the CN-Wheat crop. The modest structure consists of less than twenty explicitly observable organs. From this set, only 10-12 organs track eco-physiological functions that are of interest for reservoir computing; processes such as carbon accumulation are linear and trivially show no fading memory properties. Because the model assumes a post-flowering growth stage, C and N are only expended to grow grains and roots, while other elements are fixed in mass. Eco-physiological dynamics may still be affected by a sub-model simulating tissue death, affecting the total green area of the specimen.

CN-Wheat uses the following hourly meteorological inputs: incident photosynthetically active radiation (PAR), air temperature (T_{air}), relative humidity (RH), air CO₂ concentration, and wind speed (u). The model also considers the relative distribution of PAR among the photosynthetic organs. In the used simulations, CO₂ concentration is fixed to 360 ppm for all time steps.

In our experiments, we adopted the NEMA (Nitrogen Economy Model within plant Architecture) sim-

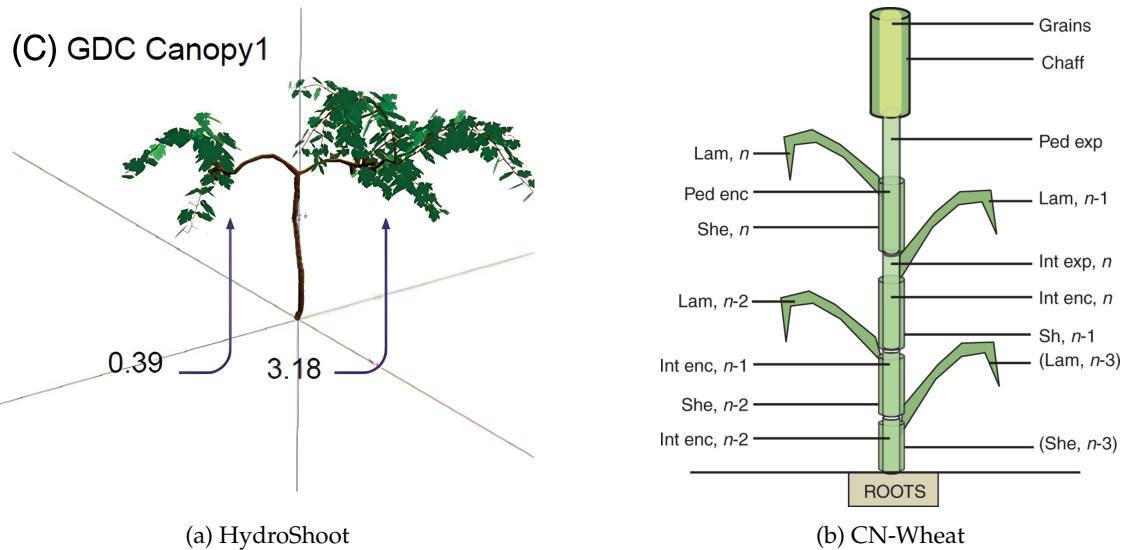


Figure 5.1: FSPMs selected for use in experiments in this work. (a) A grapevine specimen from HydroShoot with the canopy trained in the “Geneva Double Curtain” configuration. This figure is reused from Albasha et al. (2019) under the CC BY 4.0 license. (b) The architecture of the CN-Wheat model. This figure is reused from Barillot et al. (2016b) with permission from the publisher (Copyright 2016 Oxford University Press).

ulation configurations used in the original 2016 publication. The simulations examine the behavior of the plant under three different fertilization regimes. In this configuration, the model runs for up to 50 simulated days at a time scale of hourly steps.

5.3.3 Simulation Details

Table 5.2 compares the simulation details of both FSPMs. HydroShoot’s simulation length is relatively short, so the simulation must be run multiple times using different meteorological inputs to generate a large enough dataset. Because of HydroShoot’s large size, we will limit our observations to a small selection of leaf elements. While CN-Wheat only has fifteen elements, we must use an even smaller subset as the reservoir because some physiological targets are a linear function of the elements.

Table 5.2: Comparison of HydroShoot and CN-Wheat: simulation details.

	HydroShoot	CN-Wheat
Simulation step	1 h	1 h
Simulation duration	7 days	50 days
Model structure	static (idealized reservoir)	Post-vegetative growth
Model size	$\approx 10^3$ elements	15 elements

5.3.4 Environmental Inputs

The environmental inputs of both models are listed in Table 5.3. Significant abiotic factors are present in both models: air temperature, relative air humidity, wind speed and incident sunlight. CO₂ concentration and atmospheric pressure are model inputs but use constant values during simulation. The available wind speed data for CN-Wheat varies daily instead of hourly, making it unsuited as a regression target or input to computational benchmarks.

5.3.5 Regression Targets

Table 5.4 lists the physiological regression targets that can be used for each model. To create more overlap between the two models, it is possible to calculate some targets not present in CN-Wheat from the observed structural elements.

5.3.6 Available Reservoirs

The observable physiological processes for HydroShoot and CN-Wheat are presented in Table 5.5. We limited the table to processes that would be experimentally observable, as listed in Tables 4.1 and 4.2. Both models track the photosynthesis rate, surface temperature, transpiration rate, and stomatal conductance. HydroShoot has some additional hydraulic processes available.

5.4 Summary

First, we formalized the selection criteria to choose which plant models are suited for our research. We based these criteria on considerations from an RC, a plant-physiological, and a practical perspective. Next, we held a brief discussion of each FSPM we considered from the literature. Then, we made a final selection of two models: HydroShoot and CN-Wheat. We introduced the background of each model and gave its implementation details. Finally, we summarized their properties relevant to our experiments: the environmental inputs, the available regression targets, and the observations that can be used as a reservoir.

Table 5.3: Comparison of HydroShoot and CN-Wheat: environmental models.

Symbol	Description	Unit	HydroShoot	CN-Wheat
T_{air}	Air temperature	°C	Hourly	Hourly
RH	Relative humidity	%	Hourly	Hourly
u	Wind speed	m s ⁻¹	Hourly	Daily
PAR	Photosynthetically active radiation	W m ⁻² μmol _{photons} m ⁻² s ⁻¹	Hourly /	/ Hourly
CO ₂	CO ₂ concentration	μmol _{CO₂} mol ⁻¹	400 ppm	360 ppm
P_a	Atmospheric pressure	kPa	101.3 kPa	/

Table 5.4: Comparison of HydroShoot and CN-Wheat: available regression tasks. Check marks indicate whether the target is computed by the FSPM. Targets annotated with * are not calculated in the simulation, but can be computed from other simulation outputs.

Symbol	Description	Unit	HydroShoot	CN-Wheat
E	Transpiration rate rate	g h ⁻¹	✓	✓
A_n	Net photosynthesis rate	μmol s ⁻¹	✓	
T_{leaf}	Mean leaf temperature	°C	✓	*
Φ_{PAR}	Absorbed PAR	W m ⁻² μmol s ⁻¹	✓	*
R_{shoot}	Shoot respiration	μmol _C s ⁻¹		✓
R_{roots}	Root respiration	μmol _C s ⁻¹		✓

Table 5.5: Comparison of HydroShoot and CN-Wheat: physiological processes observable in each structural element. Not all are shown in the table; we only present the processes that are experimentally observable and that do not violate RC assumptions in an obvious way. Check marks indicate whether the physiological process is modeled by the FSPM.

Symbol	Description	Unit	HydroShoot	CN-Wheat
A_n	Net photosynthesis rate	μmol m ⁻² s ⁻¹	✓	✓
T_s	Surface temperature	°C	✓	✓
g_s	Stomatal conductance	mol m ⁻² s ⁻¹	✓	✓
E	Transpiration rate	mol _{H₂O} m ⁻² s ⁻¹	✓	✓
F	Water flow across segment	kg s ⁻¹	✓	
Ψ	Mean water potential of segment	MPa		✓

Chapter 6

Implementation Details

Chapter 4 established the experimental framework for this work. Then, Chapter 5 detailed the two FSPMs we selected for our experiments. In this chapter, we explain how we implemented the experiments. We discuss the adaptations that we made to the simulations, how we handled data for the regression tasks, and how we implemented the impulse experiment.

6.1 Preparing the Simulations

We used the same environmental inputs as the models' original publications for all experiments because the authors only verified the models for these climate conditions. The complete hourly data for HydroShoot and CN-Wheat can be found in Appendix A. HydroShoot required small adaptations to the soil water model because empirical soil water data was only available for four consecutive dates. We ran CN-Wheat simulations with no modifications. Details for reproducing the exact simulation setup for both models, as well as descriptions of model modifications we made can be found in Appendices B.1 and C.1.

6.2 Observing The Plant Model

We extracted time series for all physiological processes listed in Table 5.5 from each structural element of the FSPM. The reservoirs are then constructed from this data. This work investigates homogeneous reservoirs only; we only considered the spatial variations of a single process, not the dynamics between two or more processes. However, some preliminary work with heterogeneous reservoirs showed promising results, which can be explored in further research.

It is impossible to get exact measurements or observe the entire plant at once in a real-world scenario. To make our results more representative for follow-up research, we decided only to observe a fraction of the plant's structure to limit the observability of the reservoir. We used a reservoir sample size of 32 for HydroShoot and 7 for CN-Wheat. This corresponds with approximately 10% and 70% of the total structure, respectively. We chose 32 for HydroShoot to limit the number of readout parameters and because we observed that the reservoir performance showed negligible improvements for larger sample sizes. We only had ten structural elements to observe the chosen physiological processes in CN-Wheat. We selected a sample size of 7 to avoid the total observability of the reservoir. Observing the entire reservoir may introduce bias as the plant model uses the same observations to compute the physiological tasks. Appendices B.2 and C.2 further elaborate on how the reservoir was constructed for HydroShoot and CN-Wheat, respectively.

6.3 Data Preprocessing

First, the raw simulation outputs are combined into a single dataset. Appendices B.3 and C.3 explain how we did this for each model specifically. Once we have a single dataset, some preprocessing steps are applied. Next, we discarded data from the first four days of each simulation run to avoid transient dynamics present during model warm-up. We also discarded 8 hours of nighttime from each day, between 22:00 and 4:00. By discarding these samples from the dataset, we focus the model's capacity on predicting the more detailed and intricate daytime dynamics instead of concentrating on the broader day/night dynamic. Figure 6.1 illustrates the data that is discarded.

Finally, we rescaled both the target and the reservoir series to have zero mean and unit variance. When we scaled each dimension independently, we noticed a drop in model performance. We hypothesize that this is due to lost information about the relative difference between observed elements. Because the observations of the reservoir elements are not independent of each other, we instead rescaled the reservoir using the mean and variance of the reservoir as a whole.

6.4 Train and Test Data Split

To avoid bias in the model scores, we require a set of unseen test data to validate the performance of the final model. Therefore the optimal model parameters should be trained using only a subset of the total dataset. We used a train-to-test ratio of 0.5 for all our regression experiments. Figure 6.1 shows how we divided the data samples into train and test sets. Time series data naturally shows a high degree of temporal correlation. To reduce the surface area for information leaking between train and test sets, we divided the data into blocks of four days. The previously mentioned process of discarding the nighttime samples further limits the information leakage between adjacent blocks of train and test data.

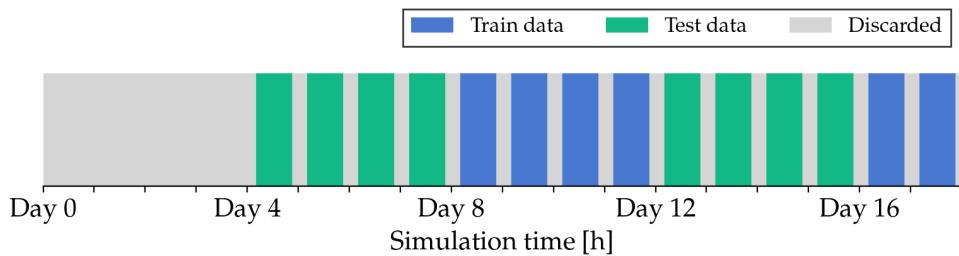


Figure 6.1: Illustration of the data that is discarded as a preprocessing step and how the remaining data is split into a train and test set.

6.5 Model Validation

The model parameters for each reservoir-target pairing are fitted using ridge regression (Equation 2.4). To find the optimal regularization parameter λ , we performed a parameter sweep from 0.0001 to 100 with fifty logarithmically spaced values. We validated the choice of λ using the cross-validation method.

In this scheme, the training data is split into “folds” of roughly equal size. The model is trained using data from all but one fold. The final fold is used to validate the model, using the NMSE metric. We repeat this until we have used all folds as validation data. The mean NMSE across all validation folds suggests the model performance for that value of λ . We pick the value of λ that yields the best mean validation score. Once the optimal λ is selected, we train a final model using all available training data.

Figure 6.2 shows how we created our cross-validation folds. We group the folds per calendar day to avoid putting data from the same simulation day in both the train and validation sets. Otherwise, we risk introducing bias in the validation score thanks to the high temporal correlation between samples of the same input day. We used the leave-one-group-out strategy, which means we used only one day of data for validation in each iteration. See Appendices B.4 and C.4 for plant information specific to the HydroShoot and CN-Wheat datasets.

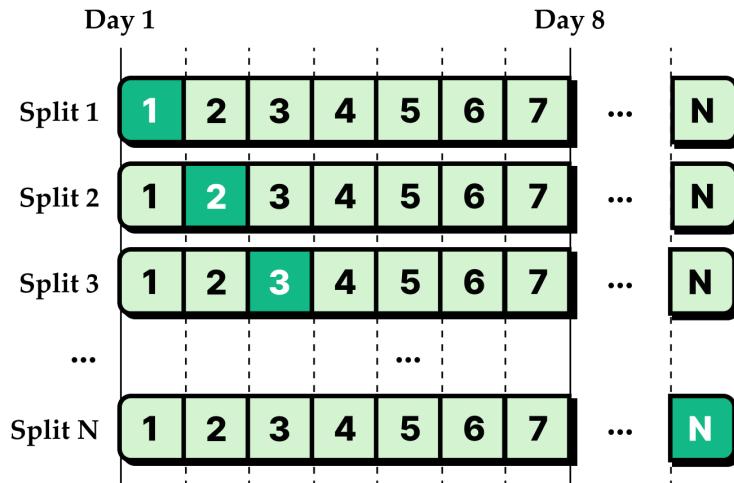


Figure 6.2: Leave-one-group-out cross-validation strategy. Each block represents samples generated from the same calendar day of meteorological inputs. In each split, the highlighted data is used as validation data, the other folds are used for training.

6.6 Implementing the Impulse Experiment

To implement the impulse experiment, we modified the meteorological input data before running the simulations. Figure 6.3 demonstrates how the impulse is added to an environmental signal. The stimulus was applied at least four days after the simulation started to account for transient warmup dynamics in the plant model. We chose four days because this is the same as the warmup period we assumed in Section 6.3. The model is allowed to run at least five more days after the impulse to capture the fading memory effect completely. The impulse is always centered on midday and is applied instantaneously. Appendices B.5 and C.5 disclose the exact simulation configuration for HydroShoot and CN-Wheat respectively.

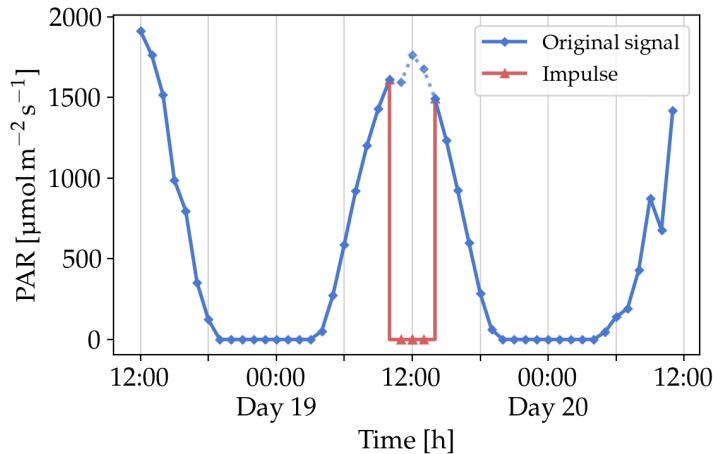


Figure 6.3: Adding an artificial impulse to an environmental input. In this figure we applied an impulse with a width of three hours and a value of $0 \mu\text{mol m}^{-2} \text{s}^{-1}$ to the incident PAR of CN-Wheat.

6.7 Summary

This chapter disclosed the implementation details required to reproduce the experimental results presented in this work. We discussed how the reservoir dataset is constructed, what data preprocessing is applied, how a test dataset is created, and we how validated the readout model parameters. We also documented how we implemented the impulse experiment. We reported additional information on how HydroShoot and CN-Wheat were used in Appendices B and C.

Part III

Results

Chapter 7

Results And Discussion

In this chapter we apply the framework established in Chapter 4 to the FSPMs we selected in Chapter 5. With this framework, we tested the reservoir characteristics of each of the physiological processes in Table 5.5. We first study the linear input separation for physiologically relevant regression tasks. Then, we examine the fading memory property of the considered physiological processes. Finally, we point out some weaknesses in our methods and how they could have been mitigated.

7.1 Input Separation

We used several regression tasks to test the linear input separation in the physiological reservoirs. These tasks fall into three categories: (i) predicting current environmental conditions, (ii) predicting the plant's eco-physiological performance, and (iii) computational benchmarks. For the reservoir readout, we used a random sample of the plant structure, with a size of 32 and 7 for HydroShoot and CN-Wheat, respectively. We considered 16 different random reservoir samples for each plant model to determine the performance variability caused by the selected observations.

7.1.1 Environmental Inputs and Physiological Tasks

Let us first consider categories (i) and (ii). For the environmental targets, we considered air temperature (T_{air}), relative humidity (RH) and incident photosynthetically active radiation (PAR). We chose these targets because of their physiological relevance and to compare the results with those already obtained for strawberry plants in Pieters (2022). For the physiological targets we used transpiration rate (E), absorbed PAR (Φ_{PAR}) and net photosynthesis rate (A_n). A_n was only available in HydroShoot.

Figure 7.1 showcases the performance of each physiological reservoir as boxplots. For the HydroShoot model, we observe that each reservoir performs roughly the same for a given target. In CN-Wheat, the difference between the reservoirs is more outspoken. For example, RH correlates relatively well with E in CN-Wheat, where none of the processes in HydroShoot correlate well with RH. We must also note that CN-Wheat significantly outperforms HydroShoot in most tasks.

We study the time series prediction accuracy more closely in Figure 7.2. In this figure, we see that the E reservoir captures the broad dynamics of PAR and Φ_{PAR} for both HydroShoot and CN-Wheat. However, HydroShoot cannot reproduce finer details, underpredicting the highest values and overpredicting the lowest values. Moreover, the model's predictions sometimes lag behind the target (Figure 7.2a, samples A and C), other times the predictions lead the target (Figure 7.2a, sample D). In contrast, CN-Wheat's transpiration rate captures finer details much better.

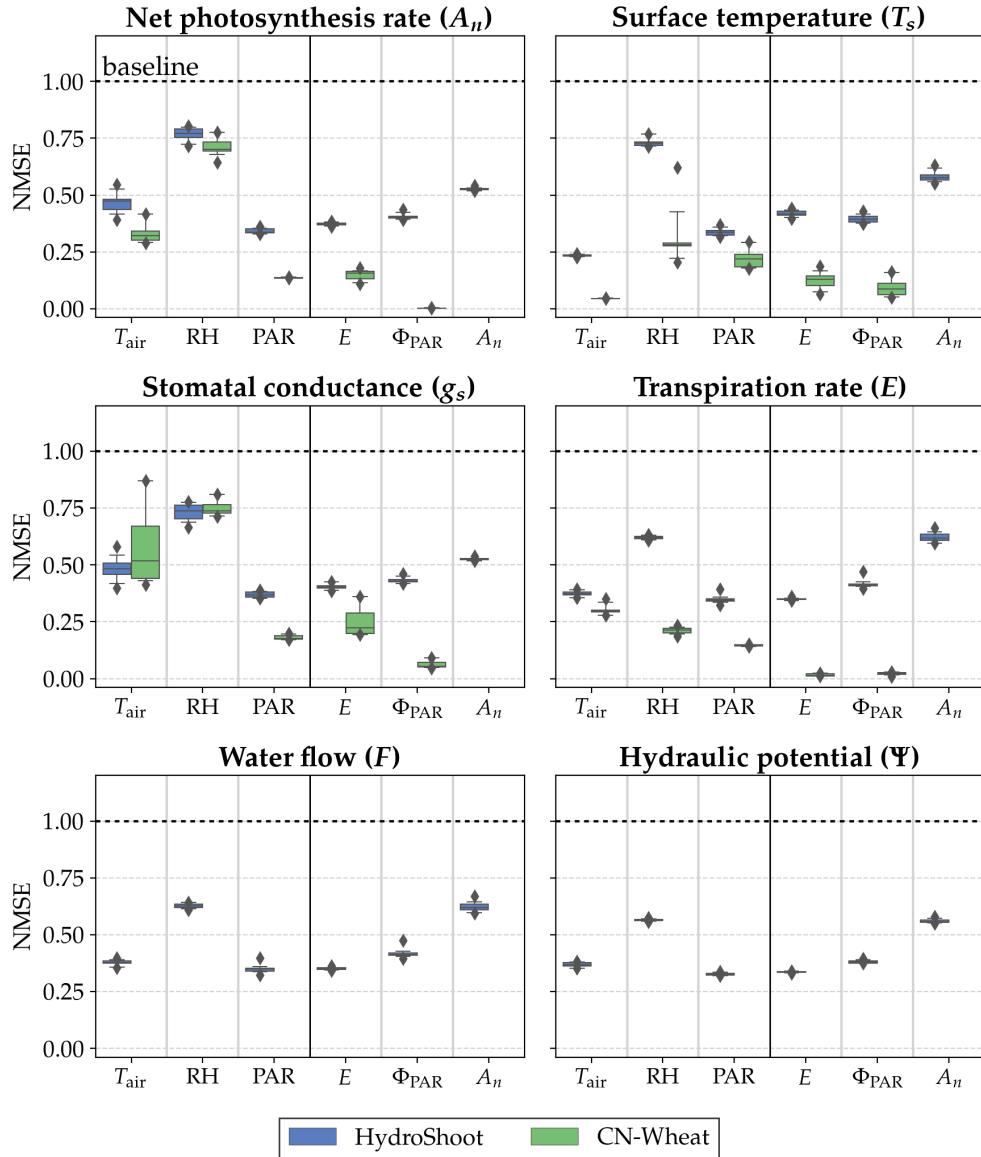


Figure 7.1: Reservoir performance for input and physiological regression tasks. The different plots show the prediction accuracy (y axis) for a suite of regression tasks (x axis) for each of the physiological reservoirs from Table 5.5. The regression target of net photosynthesis rate A_n was only available for the HydroShoot model. Each boxplot shows the distribution of test across 16 random reservoir samplings. The box shows the quartiles of the distribution; the whiskers indicate the 5th and 95th percentiles. Outliers are shown as a diamond shape.

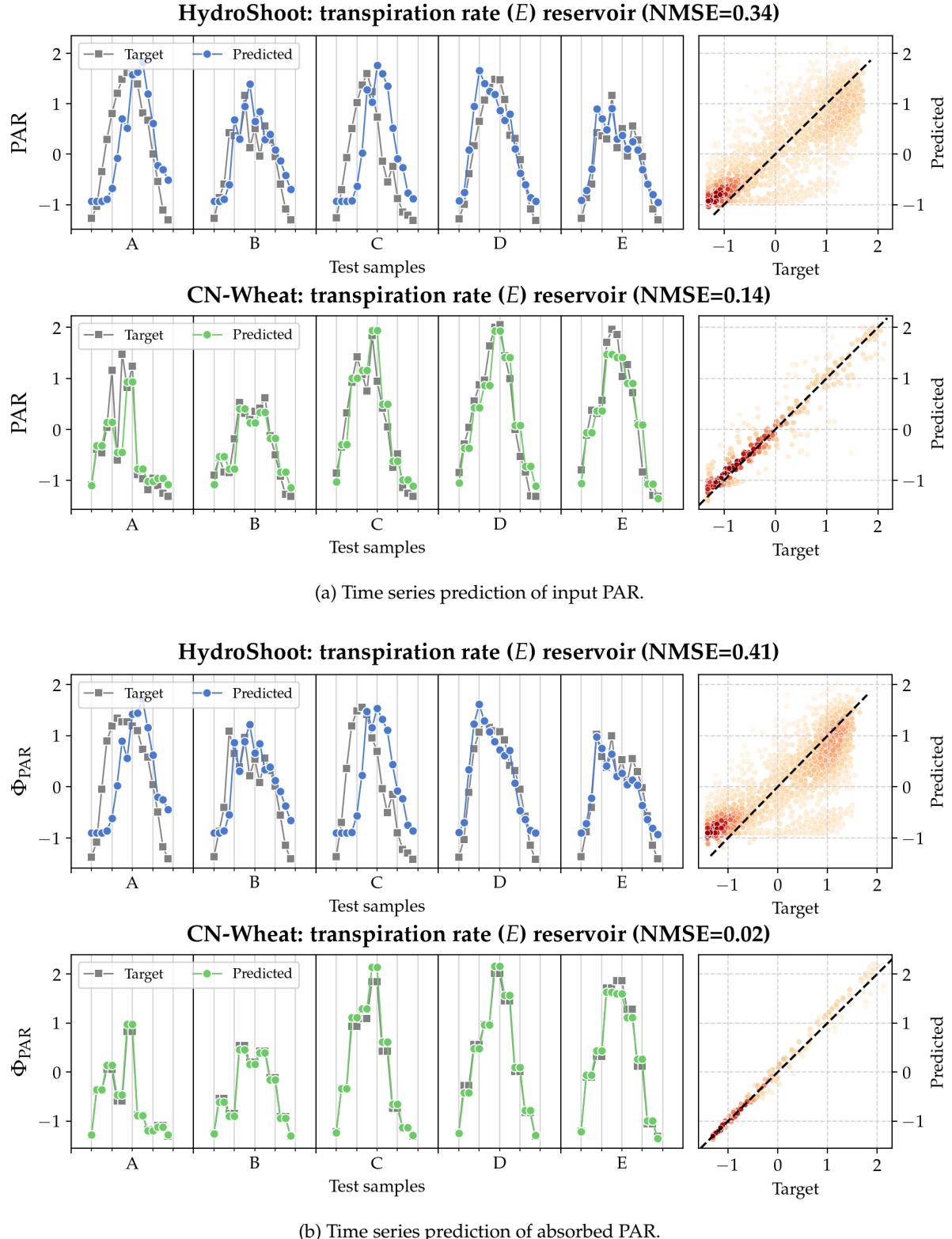


Figure 7.2: Time series prediction of input PAR and absorbed PAR using transpiration rate (E). The left plot shows time series prediction for five random samples from the test set (see Section 6.4), the right shows the correlation between the target and predicted values. Darker colors indicate a greater density of points.

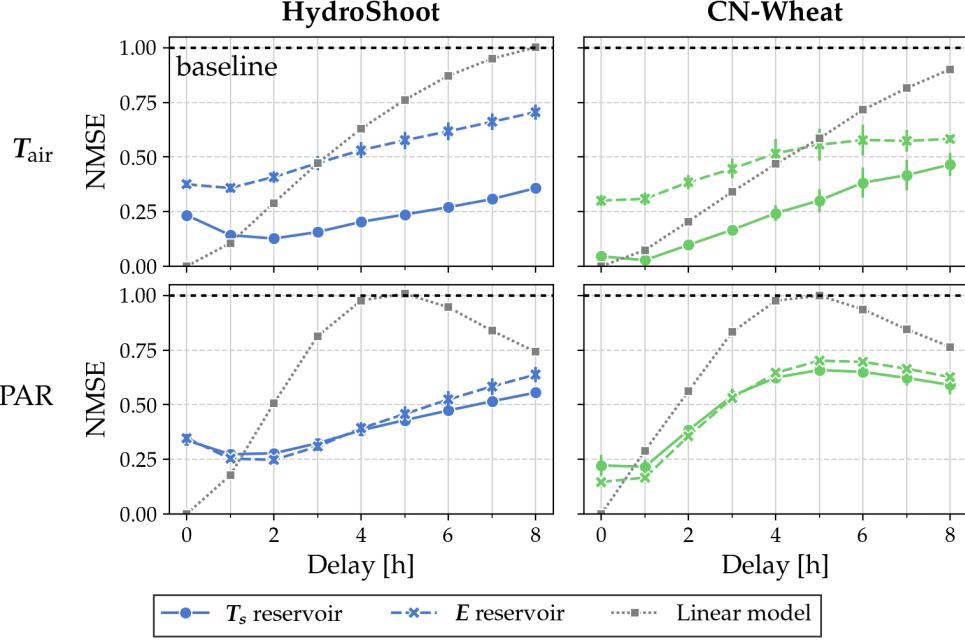


Figure 7.3: Predicting a time-delayed signal of T_{air} and PAR using surface temperature (T_s) and transpiration rate (E). In each plot, we compare the reservoirs' performance with a linear model that uses the benchmark's input as the only feature. The error bars show the scores' standard deviation.

7.1.2 Computational Benchmarks

Next, we discuss the results for the computational benchmarks. We present results for the delay line, polynomial expansion and NARMA benchmarks based on two different input signals (T_{air} and PAR). For each benchmark, we compared the reservoir's advantage (or disadvantage) over a linear model of the input signal as we increased the benchmark's difficulty. In what follows, we discuss the results for the surface temperature (T_s) and transpiration rate (E) reservoirs specifically.

Figure 7.3 shows the results for the delay line benchmark. Both reservoirs display memory capacity for solving this particular task. For the T_{air} input, one explanation for the observed memory is thermal inertia. During the diurnal cycle, heat transfers between the air and the plant's water mass, which has a higher heat capacity than air. This difference in heat capacity results in a slowed reaction to changes in ambient temperature. The readout function can exploit this inertia to learn about past inputs. Note that this is only a first-order interpretation, as it, for example, does not account for heat lost in evaporation. A similar hypothesis can explain the results for incident PAR, as solar radiation affects the plant's temperature. Lastly, we note that the transpiration rate shows similar performance characteristics as the surface temperature, indicating a possible relationship between the two processes.

Next, Figure 7.4 presents the accuracy predicting a polynomial expansion. Interestingly, we observe a possibly polynomial relationship between the air temperature and the respiration rate, as the prediction error reaches a minimum for a degree of five. For the other input-reservoir pairings, we see a similar performance characteristic as the linear model of the input signal, plus a bias penalty induced by the noisy relation between the reservoir and the input.

Finally, we consider the NARMA benchmark. In Figure 7.5 we increased the n parameter from 2 up to 24. This means that, for $n = 24$, the NARMA system integrates inputs from up to 24 hours in the past. For the benchmark based on air temperature, only the T_s reservoir outperformed the linear model. The same thermal inertia theory from the delay line task can explain this reservoir's correlation with this task. Both surface temperature T_s and transpiration rate E outperform the linear model for the benchmark based on incident PAR, particularly for n values 8, 12, and 18. The two reservoirs display similar performance for this task, further hinting at a relationship between the two. From $n = 18$ onwards, the performance of both the baseline model and the reservoirs improve over smaller values of n . This can be explained as follows. The input of the NARMA formula (Equation 2.10) has a period of twenty-four steps. Setting $n = 24$ means integrating over an entire day-night cycle in the formula. This summation term then shows low variance as the input temperature and incident PAR have relatively constant cycles in our experiments (see the input data in Appendix A). The result is a smoothed, low-variance NARMA signal, which becomes easier to predict even with a linear model. Figure 7.6 shows the time series prediction for the NARMA8 task based on incident PAR, as predicted by surface temperature T_s .

7.1.3 Comparison with Previous Research

Pieters (2022) performed similar experiments using real-world strawberry plants. His experiments used leaf thickness measurements of seven leaves as the observed reservoir, sampled every second. Because leaf thickness strongly correlates with transpiration rate E (Giuliani et al., 2013), we compare our results for the E reservoir with the results reported using leaf thickness measurements. We must emphasize that we are comparing prediction at different time scales; our models ran hourly, whereas Pieters' experiment was conducted at the second scale.

The NMSE values we obtained for the input and physiological tasks are in line with what is reported by Pieters, with HydroShoot performing a bit worse and CN-Wheat a bit better than what was reported for the strawberry plant. HydroShoot performed notably worse at predicting E and A_n than the strawberry plant. However, in terms of reservoir observability, Pieters' experiment aligns closer with our reservoir for CN-Wheat than for HydroShoot.

Pieters used incident PAR as a base for the computational benchmarks, so we compare our results for that input. We obtained significantly better results for the polynomial tasks. The strawberry plant's prediction accuracy capped out at an NMSE of nearly 1.0 for only a degree of 6, whereas our *in silico* plants achieve NMSE scores under 0.6 for up to a degree of 9. However, our baseline linear model also performed better than Pieters' control experiment, suggesting there may be a difference in methodology that makes the tasks complicated to compare. The data for the delay line task can be compared for delays of 1-5 h. Here, the results obtained in our experiments are similar to those observed for the strawberry plant. The results for the NARMA benchmark are more difficult to compare because The NARMA task used in Pieters' work operated at the minute scale, using the alternative formulation from Equation 2.11 with $\tau = 60$. Because of the hourly time step in our plant models, we designed our NARMA task at the hour scale (i.e. with $\tau = 1$). We obtained significantly better results with our version of the benchmark than were reported for the minute scale task; The strawberry plant did not achieve NMSE values lower than 0.5 for any NARMA task. As we explained in the previous section, this disparity may be because the NARMA task at the hour scale results in a more smoothed signal.

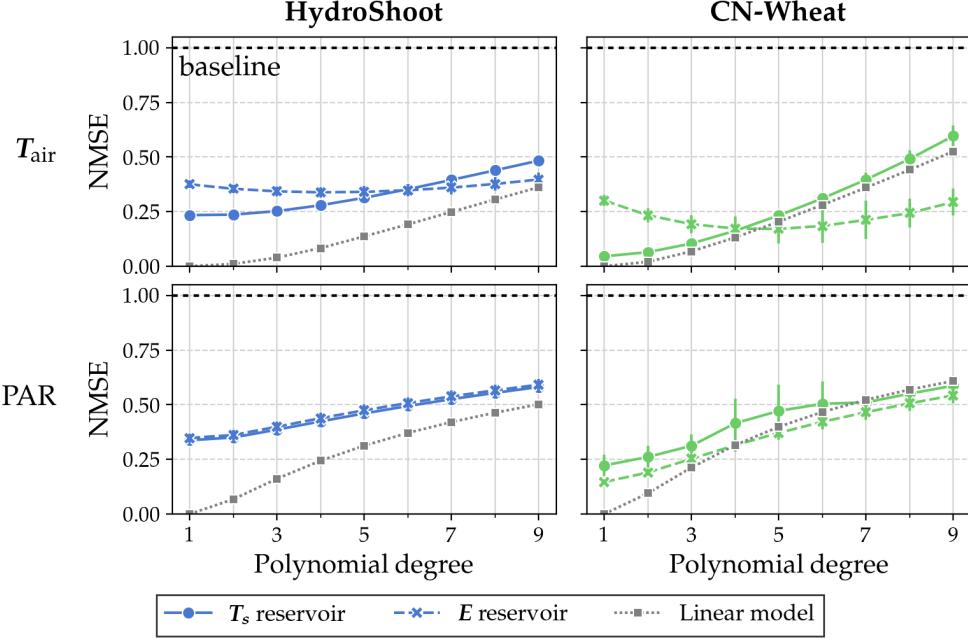


Figure 7.4: Predicting a polynomial expansion of T_{air} and PAR using surface temperature (T_s) and transpiration rate (E). In each plot, we compare the reservoirs' performance with a linear model that uses the benchmark's input as the only feature. The error bars show the scores' standard deviation.

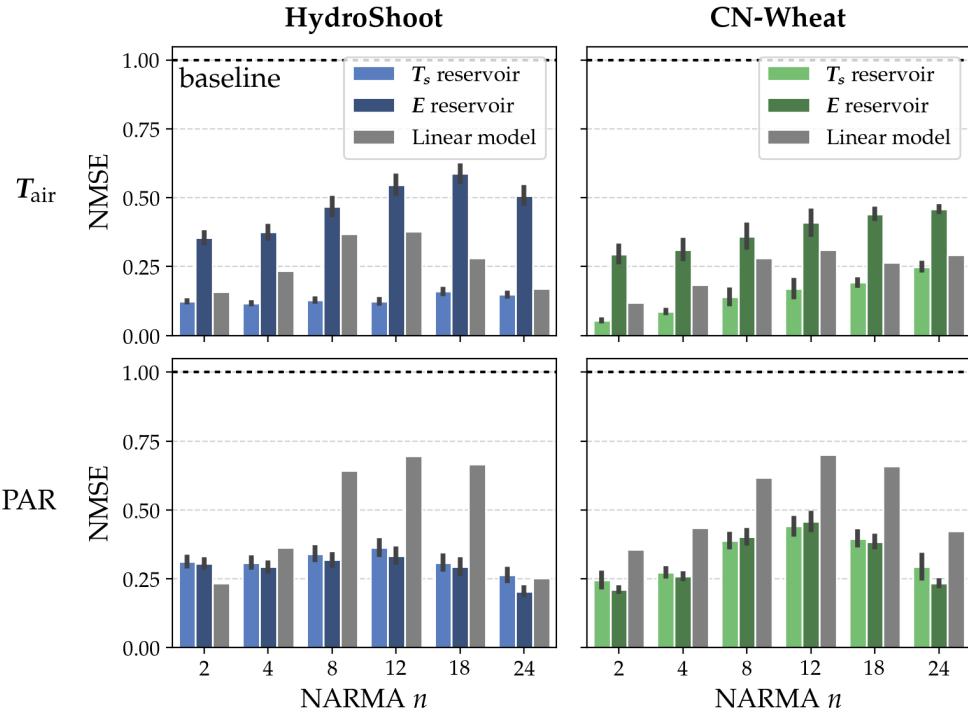


Figure 7.5: Predicting the NARMA benchmark based on T_{air} and PAR using surface temperature (T_s) and transpiration rate (E). In each plot, we compare the reservoirs' performance with a linear model that uses the benchmark's input as the only feature. The error bars show the scores' standard deviation.

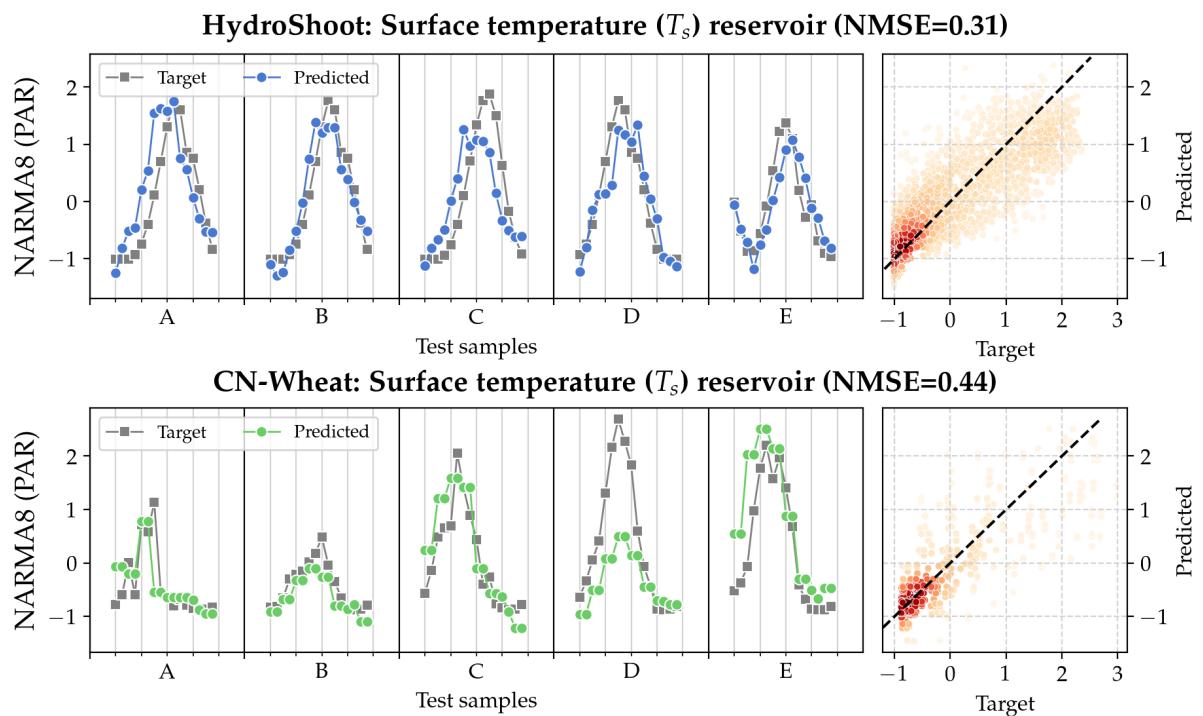


Figure 7.6: Time series prediction of the NARMA benchmark using surface temperature (T_s). The left plot shows time series prediction for five random samples from the test set (see Section 6.4), the right shows the correlation between the target and predicted values. Darker colors indicate a greater density of points.

7.2 Fading Memory

To test the fading memory principle, we applied an impulse of varying length and amplitude to the incident PAR of each plant model. The experiment revealed unexpected properties in the FSPMs. Figures 7.7 and 7.8 show the divergence $\delta(t)$ (Equation 4.3) in each physiological process caused by the impulse for CN-Wheat and HydroShoot respectively.

7.2.1 CN-Wheat

Let us first study the effects of the artificial stimulus on CN-Wheat (Figure 7.7). It is immediately apparent that the reservoir's divergence changes at intervals of two hours instead of the hourly steps of the inputs. Indeed, Barillot et al. (2016a) confirms that some submodels of CN-Wheat run at two-hour intervals to limit computational cost. The figure indicates that the affected submodels control each of our physiological reservoirs. On the one hand, this is unfortunate because, for our use case, the practical simulation scale is 2 h instead of 1 h. On the other hand, even at this coarser simulation scale, the model was able to perform tasks well at the 1 h scale in Section 7.1. Perhaps, the results could be even more promising when simulated with an hourly time step.

Because we chose the impulse lengths to be odd, we see some misalignment between the impulse steps and the steps that show disturbance in the reservoir. Looking past the misalignment caused by the 2 h simulation interval, we see that the reservoir divergence immediately disappears when the real input signal resumes. The most likely explanation is that the transient behavior lasts less than two hours, such that the coarse simulation steps mask any memory effects. An alternative explanation is that the FSPM's state is a memoryless function of the immediate environmental conditions. We think this is unlikely because the plant model should be dynamic by construction. We also reject the explanation that the impulse is too short to impact the reservoir because we see the same behavior resulting from a 1 h or 9 h stimulus. Either way, we can safely conclude that all reservoirs in the CN-Wheat model show stationary behavior w.r.t. the PAR input as their dynamics resumed as if the impulse did not occur.

At first glance, these results for CN-Wheat seem to contradict the behavior we saw in Figure 7.3 for the delay line task. We saw that T_s and E can predict past incident PAR better than a linear baseline model (though the scores are overall still quite bad). An alternative explanation for this performance is that the 2 h simulation steps add memory of the past hour to the reservoir for every second time step. Figure 7.3 supports this theory: the reservoirs display near-identical prediction accuracy for delays of 0 h and 1 h before sharply increasing for a delay of 2 h.

7.2.2 HydroShoot

Next, we discuss the outcome of the impulse experiment for the HydroShoot FSPM. Figure 7.8 shows that the disturbance in the reservoir caused by the impulse in incident PAR lasts for the same length as the applied stimulus. However, the reservoir reacts at an 8 h delay. We first investigate this delayed reaction before drawing further conclusions from this data.

Figure 7.9a shows the divergence metric for the entire simulation duration and compares various experimental setups. Note that each experiment had identical initial conditions and experienced identical

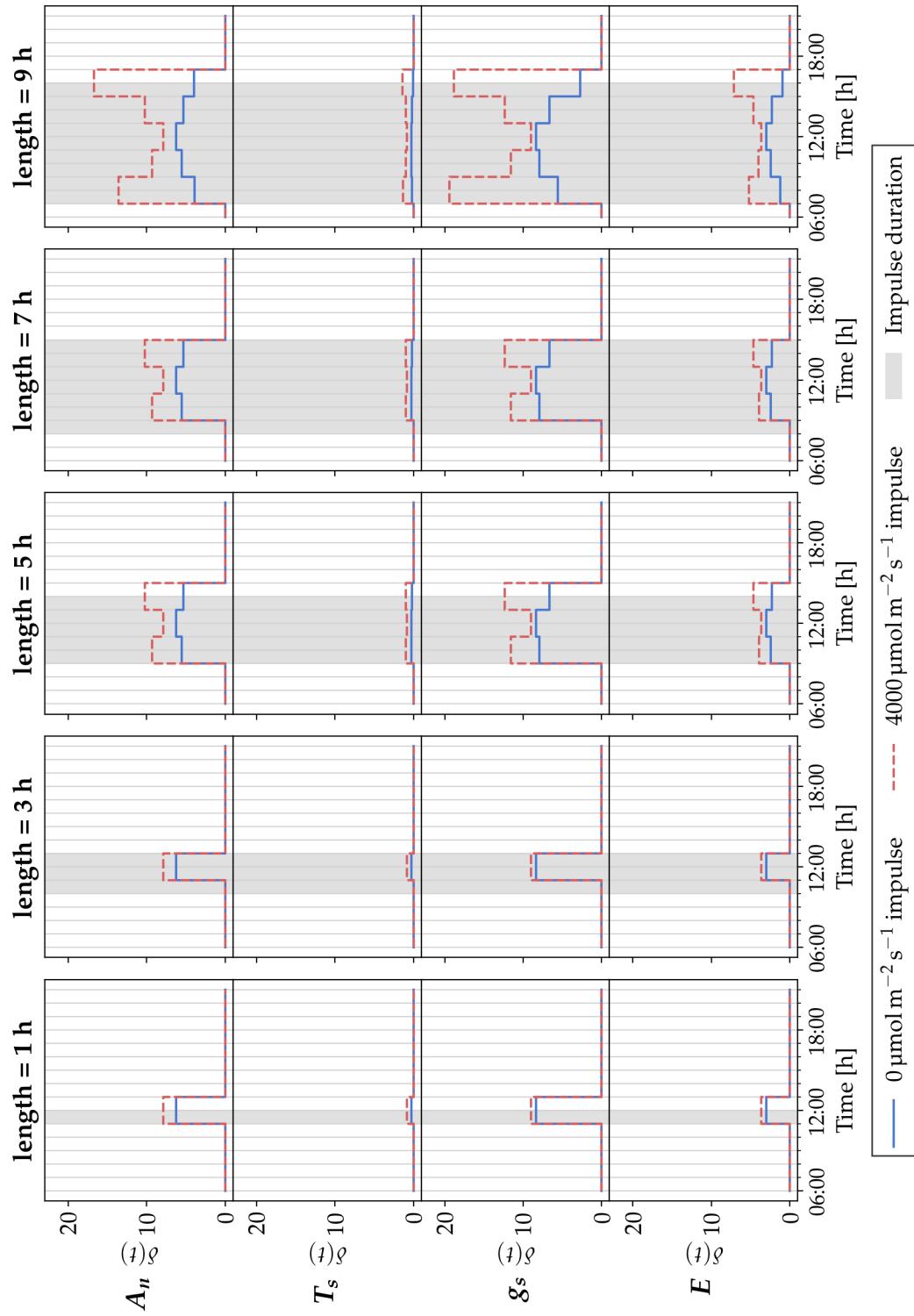


Figure 7.7: CN-Wheat: Reservoir divergence $\delta(t)$ (Equation 4.3) after applying an impulse to PAR of various lengths and amplitudes.

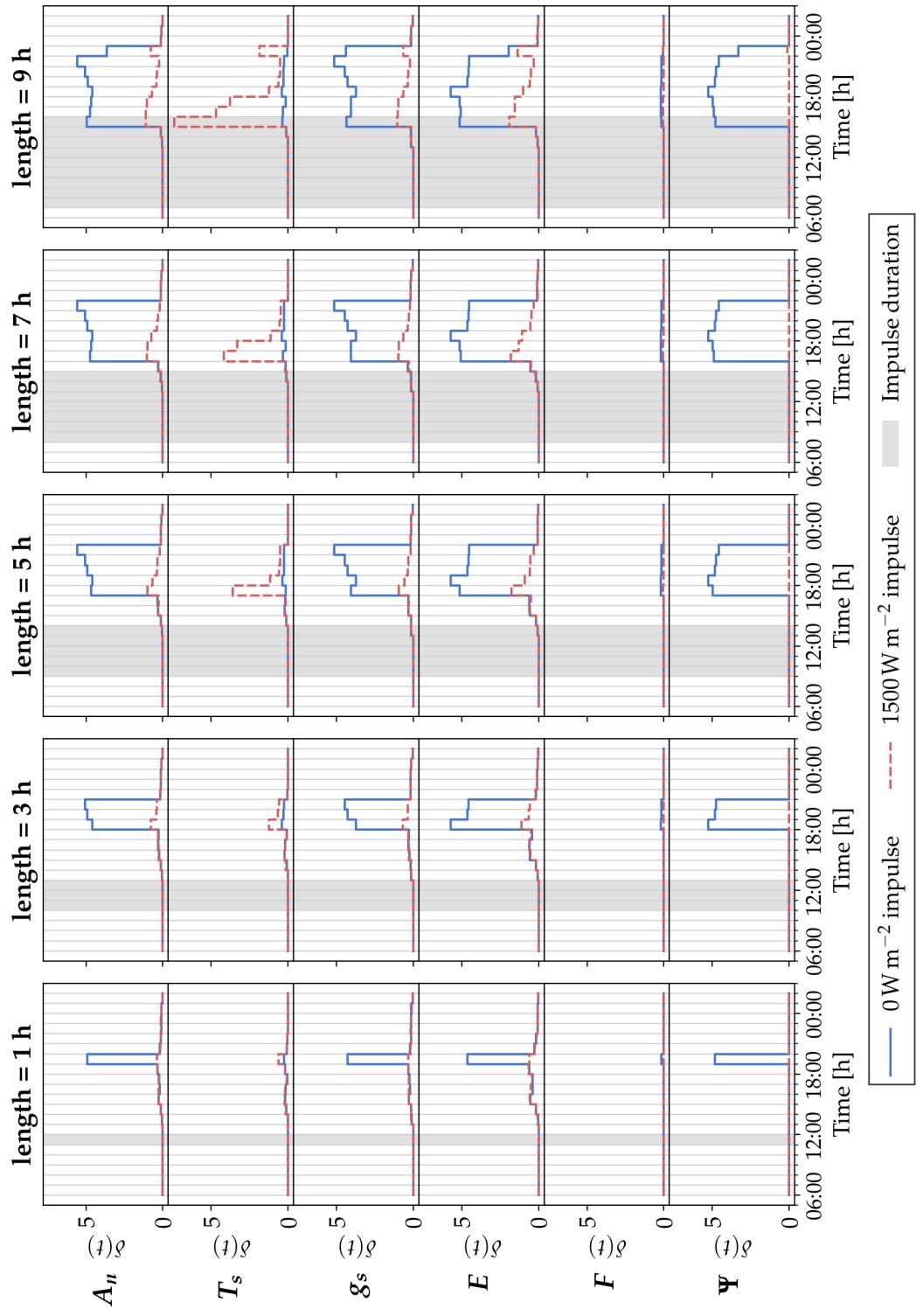


Figure 7.8: HydroShoot: Reservoir divergence $\delta(t)$ after applying an impulse to PAR of various lengths and amplitudes.

environmental conditions up to the start of the impulse on day 5. The plot shows that the reservoir dynamics diverge slightly during daytime even under identical conditions. This indicates that the observed physiological processes display somewhat chaotic dynamics during daytime hours. Keep in mind that this cyclical divergence is quite low compared to the disturbance caused by the impulse ($\delta(t)$) of 0.4 vs. 3.5 in Figure 7.9a). We did not observe cyclical divergence in CN-Wheat. The chaotic behavior may be a result of numerical instabilities in the simulation. However, previous reports of chaotic behavior in plant physiological processes exist in the literature. For example, Shabala et al. (1997) reported chaotic dynamics in plant physiological responses to light, which aligns with our observations. Therefore, the chaos observed in HydroShoot may be a more realistic model than the deterministic behavior of CN-Wheat.

Another anomaly in the HydroShoot simulation is shown in Figure 7.9b. The figure shows a superposition of the input PAR and the mean transpiration rate of the observed reservoir elements. The data reveals a continuously shifting misalignment between the diurnal cycle of the input signal and the oscillations in the physiological dynamics. In an extreme example, the peak in transpiration rate lags nearly a half-period behind the incident PAR on day 4. Comparing this data to Figure 7.9a reveals that the cyclical peaks in reservoir divergence we discussed earlier are aligned in time with these time-shifted reservoir responses. We observed this behavior for any simulation starting day. We were unable to identify the root cause for these dynamics, and the CN-Wheat model did not exhibit similar behavior. This anomaly may also explain the leading and lagging predictions we observed in Figure 7.2.

Continuing the analysis of the experimental results: the shifting delay in reservoir response is the most likely explanation for the delayed response to the artificial stimulus. Since we know of the chaotic response to incident light, we must attribute the slight disturbance that precedes the delayed peak in reservoir divergence to the natural chaotic behavior and not the impulse. Because the peak in disturbance has the same duration as the impulse and accounting for the assumed delay in reaction of 8 h in the reservoir's response, we cautiously propose the same conclusion as for CN-Wheat. That is, no lasting memory effects are visible at the simulation scale of 1 h; if there are any transient dynamics, they must occur over a much shorter time frame.

7.3 Weaknesses in Our Methods and Potential Improvements

We can identify some potential weaknesses that may have influenced our results. For one, we used the same measurements for the reservoir as are used by the model to compute the physiological regression tasks. This introduces positive bias into the correlation between the reservoir and the physiological targets. This bias is especially pronounced in the accuracy of CN-Wheat for predicting total transpiration rate and absorbed PAR; the readout model has a distinct advantage because it could directly observe 70% of the plant's internals. To reduce this bias, we could have introduced some signal noise to the observations to make them less precise. However, it is difficult to determine what amount of noise is appropriate in this case.

Another weakness is the trust we placed in the accuracy of the used FSPMs. The authors of these plant models have only validated their accuracy for the use cases presented in their publications. Plant models are generally not verified as one hundred percent accurate copies of their real-world counterparts. As we observed in Section 7.2.2, the HydroShoot model showed some unexpected behavior that was not

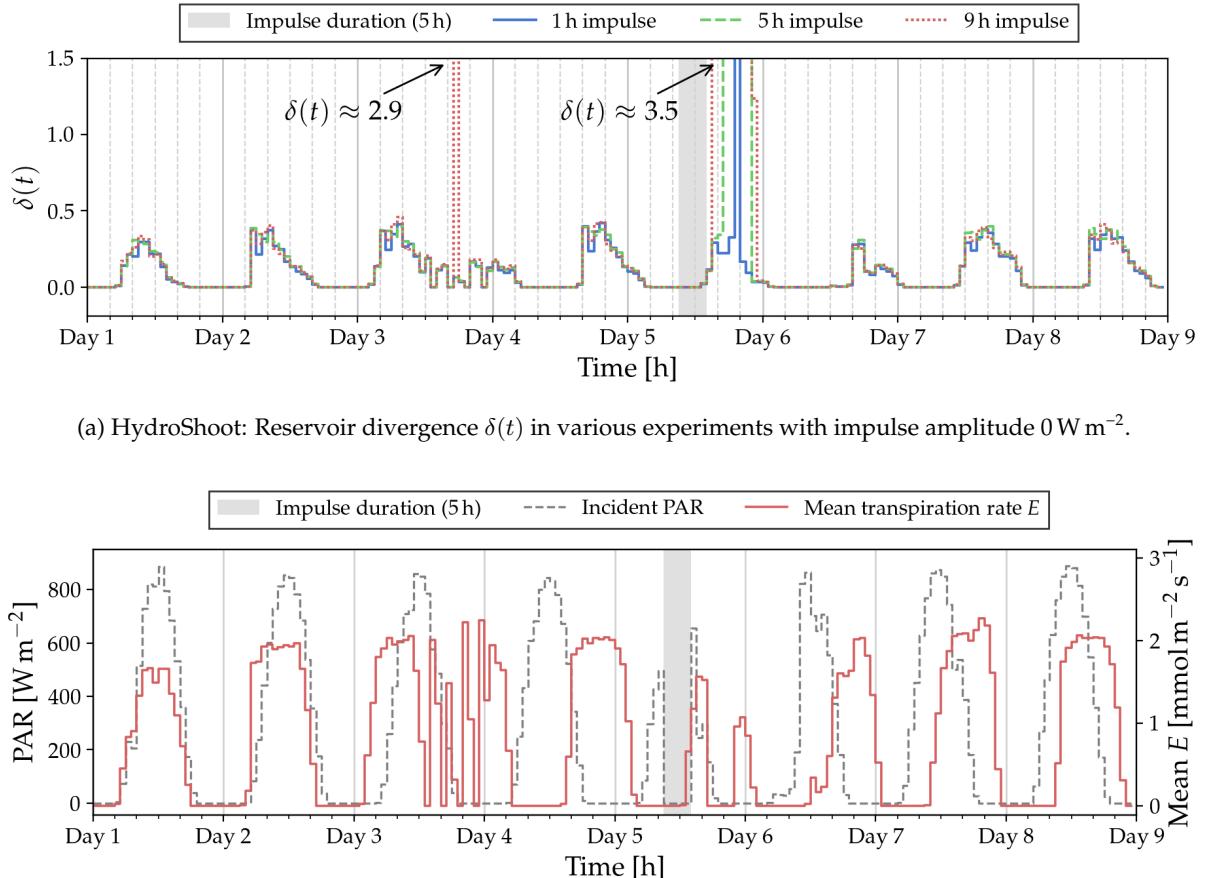


Figure 7.9: Two figures highlighting unstable behavior observed in the HydroShoot FSPM. (a) Reservoir divergence $\delta(t)$ in various experiments with impulse amplitude 0 W m^{-2} . Under all circumstances, a cyclical pattern of reservoir divergence emerges. (b) Input data and modeled transpiration rate E from an impulse experiment with length 5 h. The superposition of the data reveals a shifting misalignment between the diurnal pattern of incident PAR and the consequent transpiration rate.

reported in the original paper. To emphasize that our results characterize the behavior of plant models and not live plants, we used the names of the FSPMs throughout this chapter instead of the plant genus.

Finally, the lack of time resolution in the plant simulations leaves us with a weak take-away about transient memory dynamics in the studied plant species. Though we can conclude that the considered physiological reservoirs show stationary behavior at the hour scale, we cannot state the exact period for which past inputs influence the present reservoir dynamics.

7.4 Summary

In this chapter, we demonstrated the use of eco-physiological processes in computer plant models for reservoir computing. Our results for regression tasks highlighted several promising processes that showed good predictive correlation with biologically relevant tasks. The results can be used in future plant RC research to decide which physiological processes to pursue. The impulse experiment was less conclusive. However, the initial results seem to imply that each considered reservoir has fading memory that lasts less than two hours, with regards to incident PAR. The difficulties in using the selected FSPMs for this experiment highlighted the need for plant models with a better time resolution than 1 h. Finally, through our analysis of reservoir dynamics, we observed chaotic characteristics of eco-physiological processes, as well as unexpected behavior in a plant model that was previously unreported. Because the application of RC was able to reveal these behaviors in plant models so readily, we propose that RC can play a more prominent role not only in plant computation but also in the development and verification of FSPMs.

Chapter 8

Conclusion

The previous chapter reported the quantitative results of our experiments along with our interpretation of the data and a discussion of their implications. This chapter offers broader conclusions to the experiments we conducted. Then, we briefly recap the weaknesses we found in mechanistic plant models and how they can be improved. Finally, we propose a few directions for future research and potential applications of plant RC.

8.1 Implications of the Results

Our experiments showed promising results for plant RC with several plant physiological processes. To pick reservoirs for future research, we recommend cross-referencing our results of reservoir performance with available sensing technologies to select the physiological reservoir with the highest potential. For example, surface temperature, transpiration rate, and stomatal conductance form a good combination of reservoir features because they can all be observed using thermal cameras (see Table 4.1).

8.2 Improving Mechanistic Plant Modeling

Our research with FSPMs also highlighted some blind spots in mechanistic plant modeling. Most importantly, we see a need for developing mechanistic models that operate at a much smaller time resolution, i.e. minute- or second-scale. Another area that deserves further attention is modeling chaotic behavior in plant physiological processes. Understanding chaos in reservoirs and studying how input signals can amplify or suppress chaotic behavior is paramount to the successful application of physical reservoir computing (Nakajima, 2020). The formulas and methods in this book can be adapted into a toolset for validating the behavior of plant models. For example, Equation 4.3 can be used to measure the chaos present in the plant model quantitatively.

8.3 Future Work

The study of using plants as a substrate for reservoir computing is still in its infancy. There are several yet unexplored topics to extend the foundation of plant RC. For example, we currently only have data on RC with three species: strawberry plants from Pieters (2022), and now *in silico* versions of grapevine and wheat. Further research must establish a broader basis for predicting the reservoir properties of a species, e.g. by studying the role of the size of the plant. Another area that requires attention is how RC can be reconciled with plant growth. For example, we can use online learning approaches such as

Chapter 8. Conclusion

Hebbian learning or methods inspired by neuromorphic systems to fit readout functions that update as the plant's physiological behavior changes.

Looking towards applications, there are numerous exciting paths to explore. For example, plant RC could detect abiotic stresses such as heat or drought stress. Reservoir computing could also be applied in phenotyping to characterize the plant's physiological response to various environments. Even more ambitious, plant RC could be used to design a closed-loop system that optimizes greenhouse conditions for vegetative growth and crop yield. We could use reward-modulated Hebbian learning techniques (Burms et al., 2015) to create a self-learning control system that adapts to the needs of the plant as it grows. Here we can recommend the continued use of CN-Wheat (Barillot et al., 2016b) and WheatFspm (Gauthier et al., 2020) for an *in silico* proof-of-concept. These FSPMs model vegetative and grain growth, respectively, and showed stable behavior in our experiments.

We believe that the results from this work form good additional evidence that plants can be used as a substrate for physical reservoir computing. Our results can serve as a guide for future research in plant RC, both *in vivo* and *in silico*. This is just the starting point for plant reservoir computing. If knowledge and technology can scale, there will be countless opportunities in agricultural technology.

Appendices

Appendix A

Meteorological Data

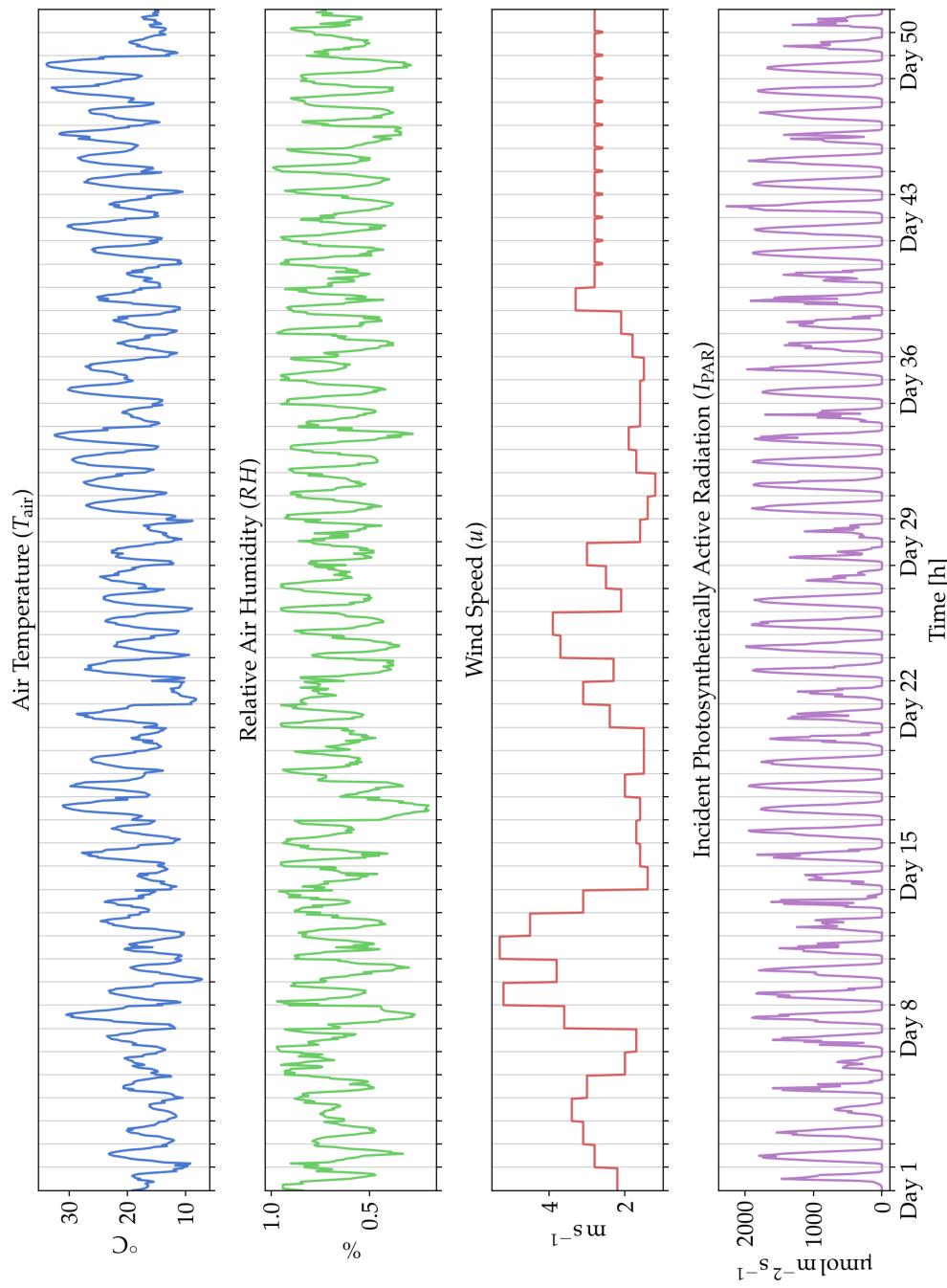


Figure A.1: Hourly meteorological inputs used in CN-Wheat simulations. The data was recorded in Clermont-Ferrand in central France. The data is a reconstruction of various dates in 2005. See Barillot et al. (2016a) for more details.

Appendix A. Meteorological Data

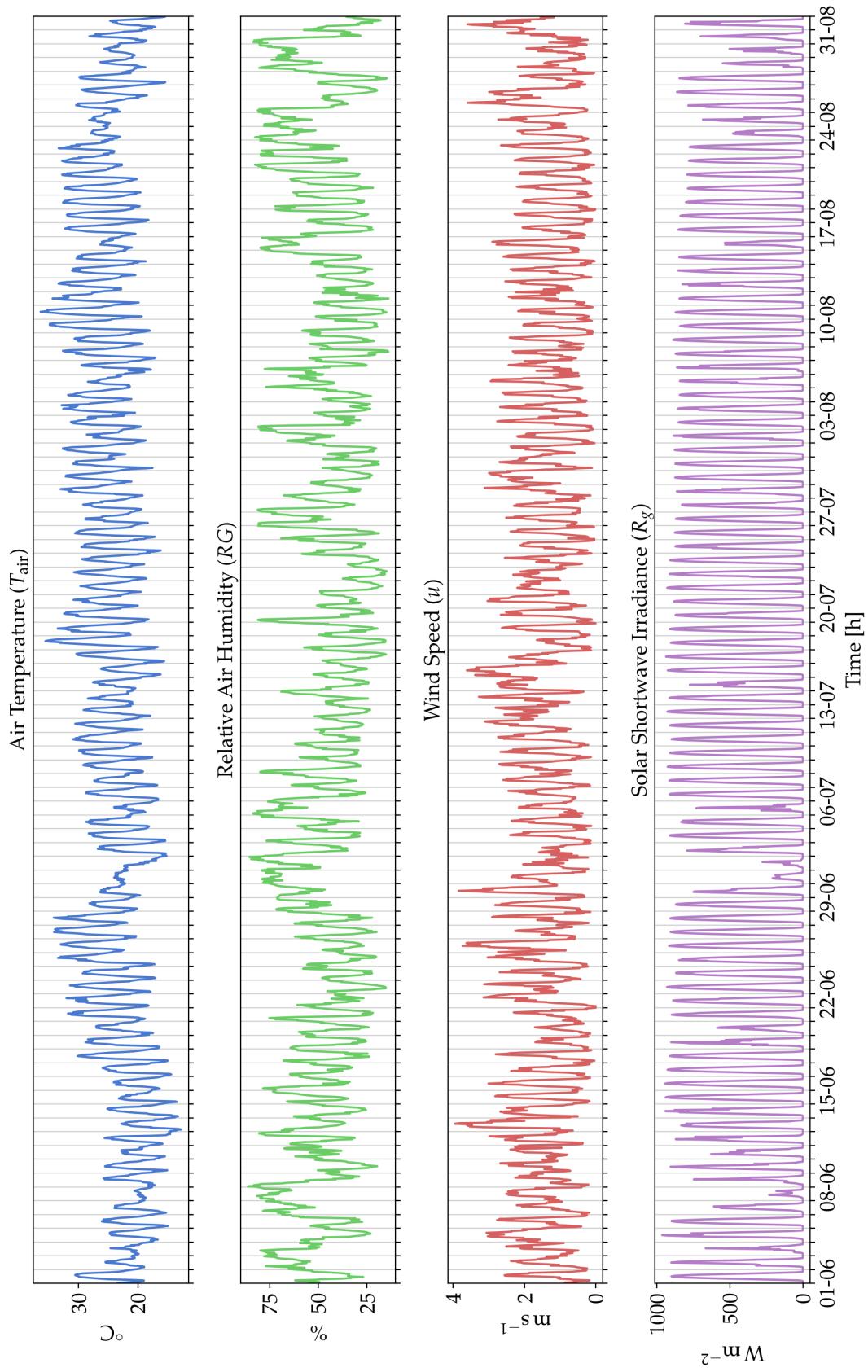


Figure A.2: Hourly meteorological inputs used in HydroShoot simulations. The data originates from Montpellier, southern France and was recorded in 2012. See Albasha et al. (2019) for more details.

Appendix B

HydroShoot: Additional Information

B.1 Simulation Details

We used the “GDC Canopy 1” configuration from the public repository* in no water deficit mode. We adapted the simulation duration to run longer than the four days in the original publication. The model’s first author, R. Albasha, communicated to us that the model had been used for simulation durations of up to a month. However, the hydraulic model uses the soil water potential Ψ_{soil} at the start of each day as input data. Adding a realistic soil water model was out of scope for this work. We settled on a modest extension to seven days for regression experiments and ten days for the impulse experiment. For this, we used a simple soil water model: we set Ψ_{soil} input to a constant value of -0.40 MPa at the start of each simulation day.

B.2 Reservoir Composition

We limited the reservoir to leaf elements only. Leaves are the least invasive to observe when transferring the experimental setup to an *in vivo* scenario. Besides, most processes from Table 5.5 occur only in leaves. Because the model structure includes 360 leaves, we randomly selected a subset of 32. The same random subset was used for all experiments.

B.3 Regression Dataset

We built the dataset from simulations using three months of input data (see Figure A.2). To collect this data, we ran more than eighty seven-day simulations, each started with an offset of one calendar day. Figure B.1 illustrates how the overlapping simulation data was grouped for train, validation and test splitting.

B.4 Cross-Validation Strategy

Initially, we used the leave-one-group-out strategy as outlined in Section 6.5. However, we noticed that the regularization parameter λ has a smooth and relatively flat optimization landscape for this particular dataset. Because of this, we shifted to using a group-K-fold strategy with five folds. The difference in model performance is negligible, but the speedup in optimization time is considerable.

*The simulation code can be found at <https://github.com/openalea/hydroshoot>

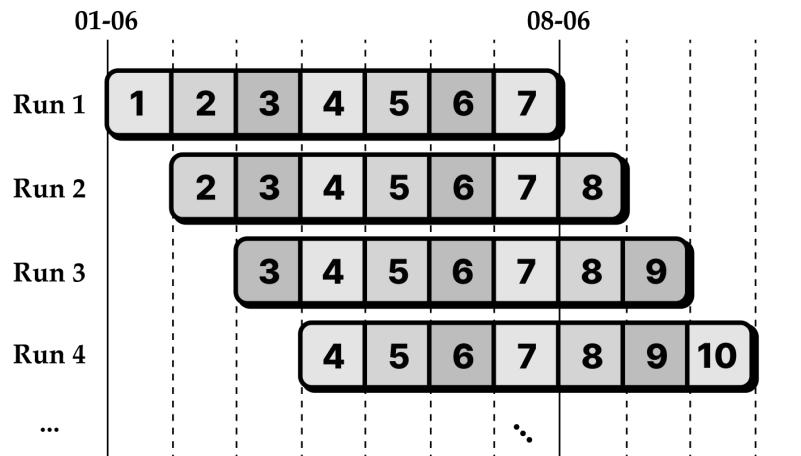


Figure B.1: Data grouping method used to combine data from overlapping simulation runs of HydroShoot. Data with the same group number must stay together in a train, validation, or test set.

B.5 Setup Impulse Experiment

For the impulse experiments, we used the same simulation configuration as described in B.1. We ran the simulations for ten days, starting at 2012-08-01. We applied the impulse on midday of the fifth day. The model then ran for five more days. We applied the stimulus to the incident solar irradiance (PAR). We used values of 0 and 1500 W m^{-2} . The latter corresponds to $\approx 1.6 \times$ the maximum amplitude of the natural signal.

Appendix C

CN-Wheat: Additional Information

C.1 Simulation Details

We used the post-vegetative simulations used for the “NEMA” experiment from Barillot et al. (2016a)*. We used the simulations for the H0, H3 and H15 fertilization regimes. The simulations were run as-is for their entire fifty-day duration.

C.2 Reservoir Composition

In CN-Wheat, photosynthesis occurs in more than just the leaf elements. Because the structural model is already small at just fifteen elements, we allowed all shoot elements to be part of the reservoir. This left ten elements that track eco-physiological processes of interest (namely surface temperature T_s , transpiration rate E , stomatal conductance g_s , and net photosynthesis rate A_n). However, we still selected a random sample of seven elements so that the readout model does not have total observability over the internal state of the plant model.

C.3 Regression Dataset

We noticed in initial experiments that the reservoir dynamics have a sudden change at a specific time step in each simulation configuration (see Figure C.1). The time step is different for each of the fertilization regimes. Because this regime change likely violates the assumption of a pseudo-static reservoir, we trimmed the output data to only include data from before the regime change. After trimming, 29, 35, and 37 days of data from the H0, H3 and H15 simulations remain.

Because the same plant is used in all three scenarios, we combined the data into a single dataset; this gives us three different samples per time step. Model scores trained on the combined dataset were comparable to or slightly lower than models trained using data from a single scenario. But combining the datasets resulted in a lower variance in scores. Figure C.2 illustrates how the simulation data was grouped for train, validation and test splitting.

C.4 Cross-Validation Strategy

We used the leave-one-group-out strategy from Section 6.5 for finetuning the regularization parameter.

*All simulations are available at <https://github.com/openalea-incubator/WheatFspm>

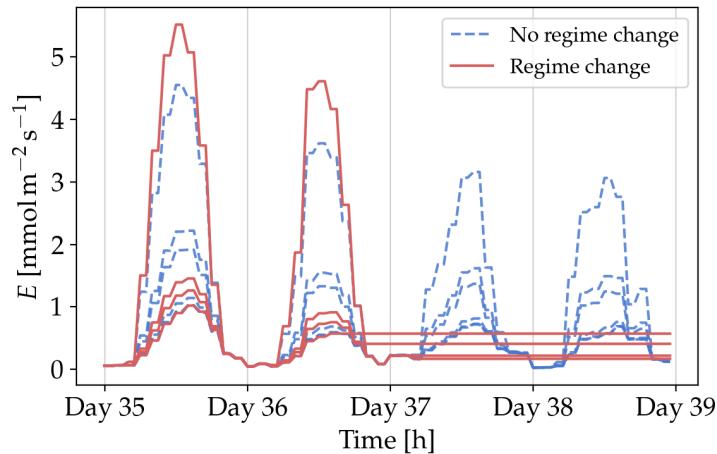


Figure C.1: Some structural elements of CN-Wheat stop functioning after a specific amount of days. Depicted on the figure is the transpiration rate of each shoot element in the H3 simulation. The red elements change their behavior after day 37.

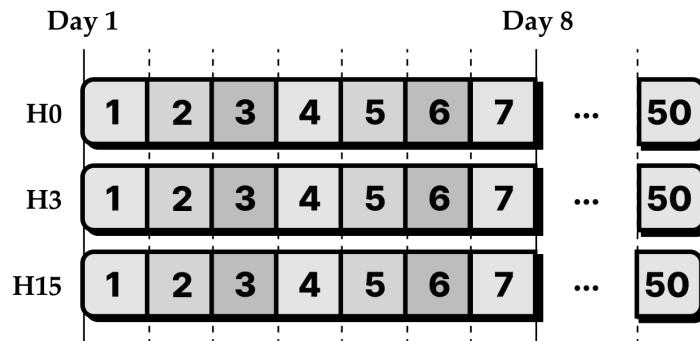


Figure C.2: Data grouping method used to combine data from three simulations of CN-Wheat. Data with the same group number must stay together in a train, validation, or test set.

C.5 Setup Impulse Experiment

The H3 simulation was used for all impulse experiments. We applied the stimulus on day 19 of the meteorological inputs from Figure A.1. Each simulation ran its entire fifty-day course. The impulse was applied to incident photosynthetically active radiation (PAR). We used amplitudes of 0 and 4000 $\mu\text{mol m}^{-2} \text{s}^{-1}$. The latter corresponds to approximately twice the maximal naturally occurring signal strength.

Bibliography

- Adamatzky, Andrew et al. (2018). "Computers from Plants We Never Made: Speculations". In: *Inspired by Nature*. Ed. by Susan Stepney et al. Vol. 28. Series Title: Emergence, Complexity and Computation. Cham: Springer International Publishing, pp. 357–387. ISBN: 978-3-319-67996-9 978-3-319-67997-6. DOI: 10.1007/978-3-319-67997-6_17.
- Alarcon, J-J. et al. (1994). "Substantial hydraulic signals are triggered by leaf-biting insects in tomato". en. In: *Journal of Experimental Botany* 45.7, pp. 953–957. ISSN: 0022-0957, 1460-2431. DOI: 10.1093/jxb/45.7.953.
- Albasha, R et al. (Jan. 2019). "HydroShoot: a functional-structural plant model for simulating hydraulic structure, gas and energy exchange dynamics of complex plant canopies under water deficit—application to grapevine (*Vitis vinifera*)". en. In: *in silico Plants* 1.1, diz007. ISSN: 2517-5025. DOI: 10.1093/insilicoplants/diz007.
- Antonik, Piotr (2018). *Application of FPGA to real-time machine learning*. New York, NY: Springer Berlin Heidelberg. ISBN: 978-3-319-91052-9.
- Baker, Neil R. (2008). "Chlorophyll Fluorescence: A Probe of Photosynthesis In Vivo". In: *Annual Review of Plant Biology* 59.1, pp. 89–113. DOI: 10.1146/annurev.arplant.59.032607.092759. pmid: 18444897.
- Barillot, Romain et al. (Oct. 2016a). "CN-Wheat, a functional-structural model of carbon and nitrogen metabolism in wheat culms after anthesis. I. Model description". en. In: *Annals of Botany* 118.5, pp. 997–1013. ISSN: 0305-7364, 1095-8290. DOI: 10.1093/aob/mcw143.
- (Oct. 2016b). "CN-Wheat, a functional-structural model of carbon and nitrogen metabolism in wheat culms after anthesis. II. Model evaluation". en. In: *Annals of Botany* 118.5, pp. 1015–1031. ISSN: 0305-7364, 1095-8290. DOI: 10.1093/aob/mcw144.
- Barillot, Romain et al. (Dec. 16, 2020). "Leaf Elongation Response to Blue Light Is Mediated by Stomatal-Induced Variations in Plant Transpiration in *Festuca Arundinacea*". In: *Journal of Experimental Botany* (eraa585). ISSN: 0022-0957. DOI: 10.1093/jxb/eraa585.
- Bengio, Y. et al. (Mar. 1994). "Learning long-term dependencies with gradient descent is difficult". In: *IEEE Transactions on Neural Networks* 5.2, pp. 157–166. ISSN: 1045-9227, 1941-0093. DOI: 10.1109/72.279181.
- Berger, Bettina et al. (Aug. 1, 2010). "High-Throughput Shoot Imaging to Study Drought Responses". In: *Journal of Experimental Botany* 61.13, pp. 3519–3528. ISSN: 0022-0957. DOI: 10.1093/jxb/erq201.
- Berger, Susanne et al. (Dec. 1, 2007). "Plant Physiology Meets Phytopathology: Plant Primary Metabolism and Plant–Pathogen Interactions". In: *Journal of Experimental Botany* 58.15-16, pp. 4019–4026. ISSN: 0022-0957. DOI: 10.1093/jxb/erm298.
- Bisoffi, Stefano (2019). *A meta-analysis of recent foresight documents in support of the 5th SCAR Foresight Exercise*. Standing Committee on Agricultural Research (SCAR). URL: <https://scar-europe.org/images/FORESIGHT/CASA-Study-Meta-Analysis-Foresight-SUB.pdf> (visited on 04/15/2022).
- Blackburn, George Alan (Mar. 1, 2007). "Hyperspectral Remote Sensing of Plant Pigments". In: *Journal of Experimental Botany* 58.4, pp. 855–867. ISSN: 0022-0957. DOI: 10.1093/jxb/erl123.
- Bock, C. H. et al. (Mar. 10, 2010). "Plant Disease Severity Estimated Visually, by Digital Photography and Image Analysis, and by Hyperspectral Imaging". In: *Critical Reviews in Plant Sciences* 29.2, pp. 59–107. ISSN: 0735-2689. DOI: 10.1080/07352681003617285.

Bibliography

- Borra-Serrano, Irene et al. (2019). "Canopy Height Measurements and Non-Destructive Biomass Estimation of Lolium Perenne Swards Using UAV Imagery". In: *Grass and Forage Science* 74.3, pp. 356–369. ISSN: 1365-2494. DOI: 10.1111/gfs.12439.
- Boudon, Frédéric et al. (2012). "L-Py: An L-System Simulation Framework for Modeling Plant Architecture Development Based on a Dynamic Language". In: *Frontiers in Plant Science* 3. ISSN: 1664-462X. DOI: 10.3389/fpls.2012.00076.
- Boudon, Frédéric et al. (Sept. 2020). "V-Mango: a functional-structural model of mango tree growth, development and fruit production". en. In: *Annals of Botany* 126.4, pp. 745–763. ISSN: 0305-7364, 1095-8290. DOI: 10.1093/aob/mcaa089.
- Briantais, Jean-Marie et al. (May 1, 1996). "Heat Stress Induces in Leaves an Increase of the Minimum Level of Chlorophyll Fluorescence, Fo: A Time-Resolved Analysis". In: *Photosynthesis Research* 48.1, pp. 189–196. ISSN: 1573-5079. DOI: 10.1007/BF00041008.
- Burms, Jeroen et al. (Aug. 2015). "Reward-Modulated Hebbian Plasticity as Leverage for Partially Embodied Control in Compliant Robotics". en. In: *Frontiers in Neurorobotics* 9. ISSN: 1662-5218. DOI: 10.3389/fnbot.2015.00009.
- Busemeyer, Lucas et al. (Mar. 2013). "BreedVision — A Multi-Sensor Platform for Non-Destructive Field-Based Phenotyping in Plant Breeding". In: *Sensors* 13.3, pp. 2830–2847. DOI: 10.3390/s130302830.
- Caluwaerts, K. et al. (Jan. 2013). "Locomotion Without a Brain: Physical Reservoir Computing in Tensegrity Structures". en. In: *Artificial Life* 19.1, pp. 35–66. ISSN: 1064-5462, 1530-9185. DOI: 10.1162/ARTL_a_00080.
- Caluwaerts, Ken (2014). *Design and computational aspects of compliant tensegrity robots*.
- Caluwaerts, Ken et al. (2011). "The body as a reservoir: locomotion and sensing with linear feedback". In: *Conference proceedings : 2nd international conference on morphological computation*. URL: <https://biblio.ugent.be/publication/1203118>.
- Carlson, Toby N. et al. (Dec. 1, 1997). "On the Relation between NDVI, Fractional Vegetation Cover, and Leaf Area Index". In: *Remote Sensing of Environment* 62.3, pp. 241–252. ISSN: 0034-4257. DOI: 10.1016/S0034-4257(97)00104-1.
- Chaerle, Laury et al. (Mar. 1, 2007). "Monitoring and Screening Plant Populations with Combined Thermal and Chlorophyll Fluorescence Imaging". In: *Journal of Experimental Botany* 58.4, pp. 773–784. ISSN: 0022-0957. DOI: 10.1093/jxb/erl257.
- Chang, Tian-Gen et al. (Jan. 2019). "Systems models, phenomics and genomics: three pillars for developing high-yielding photosynthetically efficient crops". en. In: *in silico Plants* 1.1. ISSN: 2517-5025. DOI: 10.1093/insilicoplants/diy003.
- Cieslak, Mikolaj et al. (Nov. 2016). "Integrating Physiology and Architecture in Models of Fruit Expansion". In: *Frontiers in Plant Science* 7. ISSN: 1664-462X. DOI: 10.3389/fpls.2016.01739.
- Ciompi, Stefania et al. (Aug. 16, 1996). "The Effect of Nitrogen Deficiency on Leaf Gas Exchange and Chlorophyll Fluorescence Parameters in Sunflower". In: *Plant Science* 118.2, pp. 177–184. ISSN: 0168-9452. DOI: 10.1016/0168-9452(96)04442-1.
- Costa, J. Miguel et al. (Oct. 1, 2013). "Thermography to Explore Plant–Environment Interactions". In: *Journal of Experimental Botany* 64.13, pp. 3937–3949. ISSN: 0022-0957. DOI: 10.1093/jxb/ert029.
- Coussement, Jonas R et al. (Sept. 2020). "Turgor-driven plant growth applied in a soybean functional-structural plant model". en. In: *Annals of Botany* 126.4, pp. 729–744. ISSN: 0305-7364, 1095-8290. DOI: 10.1093/aob/mcaa076.
- De Baerdemaeker, Niels J. F. et al. (Dec. 17, 2019). "X-Ray Microtomography and Linear Discriminant Analysis Enable Detection of Embolism-Related Acoustic Emissions". In: *Plant Methods* 15.1, p. 153. ISSN: 1746-4811. DOI: 10.1186/s13007-019-0543-4.

Bibliography

- De Swaef, T., K. Vermeulen, et al. (Sept. 2015a). "Plant sensors help to understand tipburn in lettuce". In: *Acta Horticulturae* 1099, pp. 63–70. ISSN: 0567-7572, 2406-6168. DOI: 10.17660/ActaHortic.2015.1099.3.
- De Swaef, Tom, Veerle De Schepper, et al. (Oct. 1, 2015). "Stem Diameter Variations as a Versatile Research Tool in Ecophysiology". In: *Tree Physiology* 35.10, pp. 1047–1061. ISSN: 0829-318X. DOI: 10.1093/treephys/tpv080.
- De Swaef, Tom, Kristof Vermeulen, et al. (2015b). "Plant Sensors Help to Understand Tipburn in Lettuce". In: *Acta Horticulturae*. 2nd International Symposium on Horticulture in Europe (SHE). Vol. 1099. International Society for Horticultural Science (ISHS), pp. 63–70. ISBN: 978-94-6261-096-5. DOI: <http://dx.doi.org/10.17660/ActaHortic.2015.1099.3>.
- De Swaef, Tom et al. (Jan. 2021). "Applying RGB- and Thermal-Based Vegetation Indices from UAVs for High-Throughput Field Phenotyping of Drought Tolerance in Forage Grasses". In: *Remote Sensing* 13.1, p. 147. DOI: 10.3390/rs13010147.
- Devacht, S. et al. (June 1, 2011). "Evaluation of Cold Stress of Young Industrial Chicory (*Cichorium Intybus L.*) Plants by Chlorophyll a Fluorescence Imaging. I. Light Induction Curve". In: *Photosynthetica* 49.2, pp. 161–171. ISSN: 03003604, 15739058. DOI: 10.1007/s11099-011-0015-1.
- Didovskyk, Andriy et al. (Jan. 2015). "Distributed Classifier Based on Genetically Engineered Bacterial Cell Cultures". en. In: *ACS Synthetic Biology* 4.1, pp. 72–82. ISSN: 2161-5063, 2161-5063. DOI: 10.1021/sb500235p.
- Ehosioke, Solomon et al. (2020). "Sensing the Electrical Properties of Roots: A Review". In: *Vadose Zone Journal* 19.1, e20082. ISSN: 1539-1663. DOI: 10.1002/vzj2.20082.
- European Commission, Directorate General for Research and Innovation (2009). *New challenges for agricultural research : climate change, food security, rural development, agricultural knowledge system. 2nd foresight exercise*. eng. LU: Publications Office. URL: <https://data.europa.eu/doi/10.2777/6185> (visited on 04/15/2022).
- (2020). *Resilience and transformation: report of the 5th SCAR Foresight exercise expert group : natural resources and food systems : transitions towards a 'safe and just' operating space*. eng. LU: Publications Office. URL: <https://data.europa.eu/doi/10.2777/717705> (visited on 04/15/2022).
- Forschungszentrum Jülich, Germany et al. (Jan. 2021). "Advances in root architectural modeling". In: *Burleigh Dodds Series in Agricultural Science*. Burleigh Dodds Science Publishing, pp. 3–32. ISBN: 978-1-78676-360-0. DOI: 10.19103/AS.2020.0075.02.
- Fournier, Christian et al. (2016). *Caribu (OpenAlea)*. URL: <https://github.com/openalea-incubator/caribu>.
- Friedli, Michael et al. (Jan. 29, 2016). "Terrestrial 3D Laser Scanning to Track the Increase in Canopy Height of Both Monocot and Dicot Crop Species under Field Conditions". In: *Plant Methods* 12.1, p. 9. ISSN: 1746-4811. DOI: 10.1186/s13007-016-0109-7.
- Gauthier, Daniel J. et al. (Dec. 2021). "Next generation reservoir computing". en. In: *Nature Communications* 12.1, p. 5564. ISSN: 2041-1723. DOI: 10.1038/s41467-021-25801-2.
- Gauthier, Marion et al. (Sept. 2020). "A functional structural model of grass development based on metabolic regulation and coordination rules". en. In: *Journal of Experimental Botany* 71.18. Ed. by Robert Hancock, pp. 5454–5468. ISSN: 0022-0957, 1460-2431. DOI: 10.1093/jxb/eraa276.
- Giuliani, R. et al. (July 2013). "Coordination of Leaf Photosynthesis, Transpiration, and Structural Traits in Rice and Wild Relatives (Genus *Oryza*)". en. In: *PLANT PHYSIOLOGY* 162.3, pp. 1632–1651. ISSN: 0032-0889, 1532-2548. DOI: 10.1104/pp.113.217497.
- Godin, C. et al. (Mar. 1998). "A Multiscale Model of Plant Topological Structures". en. In: *Journal of Theoretical Biology* 191.1, pp. 1–46. ISSN: 00225193. DOI: 10.1006/jtbi.1997.0561.

Bibliography

- Godin, Christophe et al. (June 2005). "Functional-structural plant modelling". en. In: *New Phytologist* 166.3, pp. 705–708. ISSN: 0028-646X, 1469-8137. DOI: 10.1111/j.1469-8137.2005.01445.x.
- Golzarian, Mahmood R. et al. (Feb. 1, 2011). "Accurate Inference of Shoot Biomass from High-Throughput Images of Cereal Plants". In: *Plant Methods* 7.1, p. 2. ISSN: 1746-4811. DOI: 10.1186/1746-4811-7-2.
- Hafizovic, S. et al. (Aug. 2007). "A CMOS-based microelectrode array for interaction with neuronal cultures". en. In: *Journal of Neuroscience Methods* 164.1, pp. 93–106. ISSN: 01650270. DOI: 10.1016/j.jneumeth.2007.04.006.
- Hanssens, Jochen et al. (2015). "High Light Decreases Xylem Contribution to Fruit Growth in Tomato". In: *Plant, Cell & Environment* 38.3, pp. 487–498. ISSN: 1365-3040. DOI: 10.1111/pce.12411.
- Hatfield, J. et al. (2014). *Ch. 6: Agriculture. Climate Change Impacts in the United States: The Third National Climate Assessment*. Tech. rep. U.S. Global Change Research Program. DOI: 10.7930/J02Z13FR.
- Hemming, Silke (2022). *Why autonomous greenhouses?* URL: <https://www.wur.nl/en/project/Why-autonomous-greenhouses.htm> (visited on 04/15/2022).
- Hermans, Michiel et al. (Jan. 2012). "Recurrent Kernel Machines: Computing with Infinite Echo State Networks". en. In: *Neural Computation* 24.1, pp. 104–133. ISSN: 0899-7667, 1530-888X. DOI: 10.1162/NECO_a_00200.
- Hinaut, Xavier et al. (Feb. 2013). "Real-Time Parallel Processing of Grammatical Structure in the Fronto-Striatal System: A Recurrent Network Simulation Study Using Reservoir Computing". en. In: *PLoS ONE* 8.2. Ed. by Thomas Boraud, e52946. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0052946.
- Hurtado-Uria, Cristina et al. (Nov. 2013). "Relationships between meteorological data and grass growth over time in the south of Ireland". en. In: *Irish Geography* 46.3, pp. 175–201. ISSN: 0075-0778, 1939-4055. DOI: 10.1080/00750778.2013.865364.
- Inoue, Y. et al. (Jan. 15, 2008). "Normalized Difference Spectral Indices for Estimating Photosynthetic Efficiency and Capacity at a Canopy Scale Derived from Hyperspectral and CO₂ Flux Measurements in Rice". In: *Remote Sensing of Environment* 112.1, pp. 156–172. ISSN: 0034-4257. DOI: 10.1016/j.rse.2007.04.011.
- Jaeger, Herbert (2002a). "The "echo state" approach to analysing and training recurrent neural networks - with an Erratum note". In: URL: <https://www.ai.rug.nl/minds/uploads/EchoStatesTechRep.pdf>.
- (2002b). *Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach*. Vol. 5. Issue: 01. GMD-Forschungszentrum Informationstechnik Bonn.
- Jaeger, Herbert et al. (Apr. 2004). "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication". en. In: *Science* 304.5667, pp. 78–80. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1091277.
- Janka, Eshetu et al. (June 1, 2013). "High Temperature Stress Monitoring and Detection Using Chlorophyll a Fluorescence and Infrared Thermography in Chrysanthemum (*Dendranthema Grandiflora*)". In: *Plant Physiology and Biochemistry* 67, pp. 87–94. ISSN: 0981-9428. DOI: 10.1016/j.plaphy.2013.02.025.
- Jones, Ben et al. (Apr. 2007). "Is there a Liquid State Machine in the Bacterium Escherichia Coli?" In: *2007 IEEE Symposium on Artificial Life*. Honolulu, HI, USA: IEEE, pp. 187–191. ISBN: 978-1-4244-0701-9. DOI: 10.1109/ALIFE.2007.367795.
- Jones, H. G. (1999). "Use of Thermography for Quantitative Studies of Spatial and Temporal Variation of Stomatal Conductance over Leaf Surfaces". In: *Plant, Cell & Environment* 22.9, pp. 1043–1055. ISSN: 1365-3040. DOI: 10.1046/j.1365-3040.1999.00468.x.
- Kelley Pace, R. et al. (May 1997). "Sparse spatial autoregressions". en. In: *Statistics & Probability Letters* 33.3, pp. 291–297. ISSN: 01677152. DOI: 10.1016/S0167-7152(96)00140-X.

Bibliography

- Kniemeyer, O. et al. (2007). "Groimp as a Platform for Functional-Structural Modelling of Plants". en. In: *Functional-Structural Plant Modelling in Crop Production*. Ed. by J. Vos et al. Dordrecht: Springer Netherlands, pp. 43–52. ISBN: 978-1-4020-6032-8 978-1-4020-6034-2. DOI: 10.1007/1-4020-6034-3_4.
- Lecarpentier, Christophe et al. (June 2019). "WALTer: a three-dimensional wheat model to study competition for light through the prediction of tillering dynamics". en. In: *Annals of Botany* 123.6, pp. 961–975. ISSN: 0305-7364, 1095-8290. DOI: 10.1093/aob/mcy226.
- Lobet, Guillaume (July 2017). "Image Analysis in Plant Sciences: Publish Then Perish". en. In: *Trends in Plant Science* 22.7, pp. 559–566. ISSN: 13601385. DOI: 10.1016/j.tplants.2017.05.002.
- Lobet, Guillaume et al. (2013). "An online database for plant image analysis software tools". en. In: *Plant Methods* 9.1, p. 38. ISSN: 1746-4811. DOI: 10.1186/1746-4811-9-38.
- Lobet, Guillaume et al. (Oct. 2014). "A modeling approach to determine the importance of dynamic regulation of plant hydraulic conductivities on the water uptake dynamics in the soil-plant-atmosphere system". en. In: *Ecological Modelling* 290, pp. 65–75. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2013.11.025.
- Lootens, Peter et al. (June 10, 2016). "High-Throughput Phenotyping of Lateral Expansion and Regrowth of Spaced Lolium Perenne Plants Using on-Field Image Analysis". In: *Plant Methods* 12.1, p. 32. ISSN: 1746-4811. DOI: 10.1186/s13007-016-0132-8.
- Louarn, Gaëtan et al. (Sept. 2020). "Two decades of functional-structural plant modelling: now addressing fundamental questions in systems biology and predictive ecology". en. In: *Annals of Botany* 126.4, pp. 501–509. ISSN: 0305-7364, 1095-8290. DOI: 10.1093/aob/mcaa143.
- Lu, Zhixin et al. (June 2018). "Attractor reconstruction by machine learning". en. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.6, p. 061104. ISSN: 1054-1500, 1089-7682. DOI: 10.1063/1.5039508.
- Lukoševičius, Mantas et al. (Aug. 2009). "Reservoir computing approaches to recurrent neural network training". en. In: *Computer Science Review* 3.3, pp. 127–149. ISSN: 15740137. DOI: 10.1016/j.cosrev.2009.03.005.
- Lukoševičius, Mantas et al. (Nov. 2012). "Reservoir Computing Trends". en. In: *KI - Künstliche Intelligenz* 26.4, pp. 365–371. ISSN: 0933-1875, 1610-1987. DOI: 10.1007/s13218-012-0204-5.
- Maass, Wolfgang et al. (Nov. 2002). "Real-Time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations". en. In: *Neural Computation* 14.11, pp. 2531–2560. ISSN: 0899-7667, 1530-888X. DOI: 10.1162/089976602760407955.
- Maes, W. H. et al. (Aug. 1, 2012). "Estimating Evapotranspiration and Drought Stress with Ground-Based Thermal Remote Sensing in Agriculture: A Review". In: *Journal of Experimental Botany* 63.13, pp. 4671–4712. ISSN: 0022-0957. DOI: 10.1093/jxb/ers165.
- Mancuso, Stefano (2018). *The revolutionary genius of plants: a new understanding of plant intelligence and behavior*. eng. First Atria books hardcover edition. OCLC: on1048600309. New York, NY: Atria Books, an imprint of Simon & Schuster, Inc. ISBN: 978-1-5011-8785-8.
- Mazis, Anastasios et al. (June 1, 2020). "Application of High-Throughput Plant Phenotyping for Assessing Biophysical Traits and Drought Response in Two Oak Species under Controlled Environment". In: *Forest Ecology and Management* 465, p. 118101. ISSN: 0378-1127. DOI: 10.1016/j.foreco.2020.118101.
- Meyer, Sylvie et al. (Nov. 1, 1999). "Heterogeneous Inhibition of Photosynthesis over the Leaf Surface of Rosa Rubiginosa L. during Water Stress and Abscisic Acid Treatment: Induction of a Metabolic Component by Limitation of CO₂ Diffusion". In: *Planta* 210.1, pp. 126–131. ISSN: 1432-2048. DOI: 10.1007/s004250050661.

Bibliography

- Millar, Andrew J et al. (Apr. 2019). "Practical steps to digital organism models, from laboratory model species to 'Crops in silico'". en. In: *Journal of Experimental Botany* 70.9, pp. 2403–2418. ISSN: 0022-0957, 1460-2431. DOI: 10.1093/jxb/ery435.
- Moradi, Foad et al. (June 1, 2007). "Responses of Photosynthesis, Chlorophyll Fluorescence and ROS-Scavenging Systems to Salt Stress During Seedling and Reproductive Stages in Rice". In: *Annals of Botany* 99.6, pp. 1161–1173. ISSN: 0305-7364. DOI: 10.1093/aob/mcm052.
- Nakajima, Kohei (June 2020). "Physical reservoir computing—an introductory perspective". en. In: *Japanese Journal of Applied Physics* 59.6, p. 060501. ISSN: 0021-4922, 1347-4065. DOI: 10.35848/1347-4065/ab8d4f.
- Nakajima, Kohei et al. (2013). "A soft body as a reservoir: case studies in a dynamic model of octopus-inspired soft robotic arm". In: *Frontiers in Computational Neuroscience* 7. ISSN: 1662-5188. DOI: 10.3389/fncom.2013.00091.
- Nakajima, Kohei et al. (Sept. 2015). "Information processing via physical soft body". en. In: *Scientific Reports* 5.1, p. 10487. ISSN: 2045-2322. DOI: 10.1038/srep10487.
- Nejad, Abdolhossein Rezaei et al. (2005). "Stomatal Response Characteristics of Tradescantia Virginiana Grown at High Relative Air Humidity". In: *Physiologia Plantarum* 125.3, pp. 324–332. ISSN: 1399-3054. DOI: 10.1111/j.1399-3054.2005.00567.x.
- Norton, D. et al. (2006). "Preparing More Effective Liquid State Machines Using Hebbian Learning". In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. Vancouver, BC, Canada: IEEE, pp. 4243–4248. ISBN: 978-0-7803-9490-2. DOI: 10.1109/IJCNN.2006.246996.
- NPEC (Sept. 2020). *WUR is working on Digital Twins for tomatoes, food and farming*. English. URL: <https://www.npec.nl/news/wur-is-working-on-digital-twins-for-tomatoes-food-and-farming/> (visited on 05/25/2022).
- Pieters, Olivier (2022). "Reservoir computing with plants". en. PhD thesis. Ghent University. Faculty of Engineering and Architecture. URL: <http://hdl.handle.net/1854/LU-8739713>.
- Pieters, Olivier et al. (2021). "Plants as a Substrate for Physical Reservoir Computing". en.
- Poorter, Hendrik et al. (Dec. 2016). "Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field". en. In: *New Phytologist* 212.4, pp. 838–855. ISSN: 0028-646X, 1469-8137. DOI: 10.1111/nph.14243.
- Pradal, Christophe et al. (2008). "OpenAlea: a visual programming and component-based software platform for plant modelling". en. In: *Functional Plant Biology* 35.10, p. 751. ISSN: 1445-4408. DOI: 10.1071/FP08084.
- Pradal, Christophe et al. (June 2015). "OpenAlea: scientific workflows combining data analysis and simulation". en. In: *Proceedings of the 27th International Conference on Scientific and Statistical Database Management*. La Jolla California: ACM, pp. 1–6. ISBN: 978-1-4503-3709-0. DOI: 10.1145/2791347.2791365.
- Prieto, Jorge A et al. (Sept. 2020). "A functional-structural plant model that simulates whole- canopy gas exchange of grapevine plants (*Vitis vinifera* L.) under different training systems". en. In: *Annals of Botany* 126.4, pp. 647–660. ISSN: 0305-7364, 1095-8290. DOI: 10.1093/aob/mcz203.
- Reffye, Philippe de et al. (Feb. 2021). "Two decades of research with the GreenLab model in agronomy". en. In: *Annals of Botany* 127.3, pp. 281–295. ISSN: 0305-7364, 1095-8290. DOI: 10.1093/aob/mcaa172.
- Renton, Michael et al. (May 2017). "Modelling crop-weed competition: Why, what, how and what lies ahead?". en. In: *Crop Protection* 95, pp. 101–108. ISSN: 02612194. DOI: 10.1016/j.cropro.2016.09.003.
- Rodriguez-Dominguez, C. M. et al. (July 15, 2019). "Sensitivity of Olive Leaf Turgor to Air Vapour Pressure Deficit Correlates with Diurnal Maximum Stomatal Conductance". In: *Agricultural and Forest Meteorology* 272–273, pp. 156–165. ISSN: 0168-1923. DOI: 10.1016/j.agrformet.2019.04.006.

Bibliography

- Schneider, Anne et al. (Oct. 2019). "Light Regulation of Axillary Bud Outgrowth Along Plant Axes: An Overview of the Roles of Sugars and Hormones". In: *Frontiers in Plant Science* 10, p. 1296. ISSN: 1664-462X. DOI: 10.3389/fpls.2019.01296.
- Shabala, Sergey et al. (1997). "Observations of Bifurcation and Chaos in Plant Physiological Responses to Light". en. In: *Functional Plant Biology* 24.1, p. 91. ISSN: 1445-4408. DOI: 10.1071/PP96075.
- Shannon, C.E. (Jan. 1949). "Communication in the Presence of Noise". In: *Proceedings of the IRE* 37.1, pp. 10–21. ISSN: 0096-8390. DOI: 10.1109/JRPROC.1949.232969.
- Sinoquet, H. et al. (Apr. 2001). "RATP: a model for simulating the spatial distribution of radiation absorption, transpiration and photosynthesis within canopies: application to an isolated tree crown: 3D model of radiation, photosynthesis and transpiration". en. In: *Plant, Cell & Environment* 24.4, pp. 395–406. ISSN: 01407791. DOI: 10.1046/j.1365-3040.2001.00694.x.
- Soriano, Miguel C. et al. (June 2015). "Minimal approach to neuro-inspired information processing". In: *Frontiers in Computational Neuroscience* 9. ISSN: 1662-5188. DOI: 10.3389/fncom.2015.00068.
- Steil, J.J. (2004). "Backpropagation-decorrelation: online recurrent learning with O(N) complexity". In: *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*. Vol. 2. Budapest, Hungary: IEEE, pp. 843–848. ISBN: 978-0-7803-8359-3. DOI: 10.1109/IJCNN.2004.1380039.
- Tanaka, Gouhei et al. (July 2019). "Recent advances in physical reservoir computing: A review". en. In: *Neural Networks* 115, pp. 100–123. ISSN: 08936080. DOI: 10.1016/j.neunet.2019.03.005.
- Tanaka, Takuma et al. (July 2020). "Effect of recurrent infomax on the information processing capability of input-driven recurrent neural networks". en. In: *Neuroscience Research* 156, pp. 225–233. ISSN: 01680102. DOI: 10.1016/j.neures.2020.02.001.
- Tilly, Nora et al. (Mar. 2014). "Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice". en. In: *Journal of Applied Remote Sensing* 8.1, p. 083671. ISSN: 1931-3195. DOI: 10.1117/1.JRS.8.083671.
- Torrejon, Jacob et al. (July 2017). "Neuromorphic computing with nanoscale spintronic oscillators". en. In: *Nature* 547.7664, pp. 428–431. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature23011.
- Ushio, Masayuki et al. (Sept. 2021). *Computational capability of ecological dynamics*. en. preprint. Ecology. DOI: 10.1101/2021.09.15.460556.
- Van De Vijver, Ruben et al. (Jan. 1, 2020). "In-Field Detection of Alternaria Solani in Potato Crops Using Hyperspectral Imaging". In: *Computers and Electronics in Agriculture* 168, p. 105106. ISSN: 0168-1699. DOI: 10.1016/j.compag.2019.105106.
- Van der Sande, Guy et al. (May 2017). "Advances in photonic reservoir computing". In: *Nanophotonics* 6.3, pp. 561–576. ISSN: 2192-8614, 2192-8606. DOI: 10.1515/nanoph-2016-0132.
- Van leperen, W. et al. (Jan. 1, 1994). "A New Method to Measure Plant Water Uptake and Transpiration Simultaneously". In: *Journal of Experimental Botany* 45.1, pp. 51–60. ISSN: 0022-0957. DOI: 10.1093/jxb/45.1.51.
- Vandegehuchte, Maurits W. et al. (Sept. 1, 2014). "Modelling Reveals Endogenous Osmotic Adaptation of Storage Tissue Water Potential as an Important Driver Determining Different Stem Diameter Variation Patterns in the Mangrove Species Avicennia Marina and Rhizophora Stylosa". In: *Annals of Botany* 114.4, pp. 667–676. ISSN: 0305-7364. DOI: 10.1093/aob/mct311.
- Verstraeten, David (2009). "Towards generic Reservoir Computing: time scales and novel reservoirs". In: *Reservoir Computing: computation with dynamical systems*, pp. 73–178.
- Wallach, Rony et al. (July 1, 2010). "Development of Synchronized, Autonomous, and Self-Regulated Oscillations in Transpiration Rate of a Whole Tomato Plant under Water Stress". In: *Journal of Experimental Botany* 61.12, pp. 3439–3449. ISSN: 0022-0957. DOI: 10.1093/jxb/erq168.

Bibliography

- Walter, Achim, Frank Liebisch, et al. (2015). "Plant phenotyping: from bean weighing to image analysis". en. In: *Plant Methods* 11.1, p. 14. ISSN: 1746-4811. DOI: 10.1186/s13007-015-0056-8.
- Walter, Achim et al. (2007). "Dynamics of Seedling Growth Acclimation towards Altered Light Conditions Can Be Quantified via GROWSCREEN: A Setup and Procedure Designed for Rapid Optical Phenotyping of Different Plant Species". In: *New Phytologist* 174.2, pp. 447–455. ISSN: 1469-8137. DOI: 10.1111/j.1469-8137.2007.02002.x.
- Walter, Florian, Florian Röhrbein, et al. (Dec. 2015). "Neuromorphic implementations of neurobiological learning algorithms for spiking neural networks". en. In: *Neural Networks* 72, pp. 152–167. ISSN: 08936080. DOI: 10.1016/j.neunet.2015.07.004.
- Whetton, Rebecca L. et al. (Feb. 1, 2018). "Hyperspectral Measurements of Yellow Rust and Fusarium Head Blight in Cereal Crops: Part 1: Laboratory Study". In: *Biosystems Engineering* 166, pp. 101–115. ISSN: 1537-5110. DOI: 10.1016/j.biosystemseng.2017.11.008.
- Yildiz, Izzet B. et al. (Nov. 2012). "Re-visiting the echo state property". en. In: *Neural Networks* 35, pp. 1–9. ISSN: 08936080. DOI: 10.1016/j.neunet.2012.07.005.
- Zhao, Chenyuan et al. (Sept. 2016). "Novel Spike based Reservoir Node Design with High Performance Spike Delay Loop". en. In: *Proceedings of the 3rd ACM International Conference on Nanoscale Computing and Communication*. New York NY USA: ACM, pp. 1–5. ISBN: 978-1-4503-4061-8. DOI: 10.1145/2967446.2967447.
- Zhao, Duli et al. (May 1, 2005). "Nitrogen Deficiency Effects on Plant Growth, Leaf Photosynthesis, and Hyperspectral Reflectance Properties of Sorghum". In: *European Journal of Agronomy* 22.4, pp. 391–403. ISSN: 1161-0301. DOI: 10.1016/j.eja.2004.06.005.
- Zhou, Xiao-Ran et al. (Jan. 2020). "CPlantBox, a whole-plant modelling framework for the simulation of water- and carbon-related processes". en. In: *in silico Plants* 2.1, diaa001. ISSN: 2517-5025. DOI: 10.1093/insilicoplants/diaa001.
- Zhu, Xin-Guang et al. (May 2016). "Plants *in silico* : why, why now and what?-an integrative platform for plant systems biology research: Plants *in silico*". en. In: *Plant, Cell & Environment* 39.5, pp. 1049–1057. ISSN: 01407791. DOI: 10.1111/pce.12673.

Reservoir Computing With in silico Plants

Maxime Cannoodt

Student number: 01702828

Supervisors: Prof. dr. ir. Francis wyffels, Dr. ir. Michiel Stock

Councillors: Dr. ir. Olivier Pieters, Dr. Tom De Swaef (Instituut voor Landbouw-, Visserij- en Voedingsonderzoek)

Master's dissertation submitted in order to obtain the academic degree of
Master of Science in Computer Science Engineering

Academic year 2021-2022