From qualitative representation to executable equations – Applying GarpN to a surface-limited cellulose deconstruction model

Marco Kragten[1], **Kamal Kansou**[2], **Flavien Alonzo**[2], **Jan Wielemaker**[3] and **Bert Bredeweg**[1,4]

[1]Faculty of Education, Amsterdam University of Applied Sciences, Amsterdam, The Netherlands
[2]INRAE, Nantes, France
[3]Faculty of Sciences, Vrije Universiteit Amsterdam, The Netherlands
[4]Informatics Institute, Faculty of Science, University of Amsterdam, Amsterdam, The Netherlands
{m.kragten, b.bredeweg}@hva.nl, {kamal.kansou, flavien.alonzo}@inrae.fr, j.wielemaker@vu.nl

Abstract

This paper reports on a case-study in which GarpN was applied to a scientific model. GarpN is being developed to support integration of qualitative and quantitative modelling by translating qualitative representations into mathematical equations and simulations. In the present case, we constructed a qualitative representation of the enzymatic degradation of cellulose. GarpN was then used to generate a quantitative model. We compared the generated equations with the reference model based on the original implementation. The findings indicate that GarpN can support qualitative-to-quantitative translation, while also making explicit where additional modelling decisions and settings are required.

1 Introduction

Qualitative [1] and quantitative [2,3] modelling offer complementary ways of representing and reasoning about dynamic systems. Quantitative models typically describe system behaviour using numerical information, allowing precise simulation and prediction. Qualitative models, in contrast, make conceptual notions explicit and represent systems in terms of entities, quantities, causal relations, quantity spaces, and possible states and transitions.

In previous work [4,5], we focus on secondary education and analysed how the qualitative vocabulary used in DynaLearn [6] relates to mathematical equations and numerical simulation. We presented GarpN, a software tool that translates qualitative representations into executable equations, runs numerical simulations, and matches quantitative behaviour to qualitative states.

In this paper we report on applying GarpN to the scientific model on surface-limited cellulose deconstruction described by Maurer et al. [7]. We use this model as a benchmark for investigating whether a qualitative representation can be translated into an executable quantitative model that is structurally equivalent to the existing reference implementation.

The content of this paper is therefore as follows. Section 2 summarizes the current status of GarpN. Section 3 describes the scientific model and Section 4 the qualitative model derived from that. Section 5 presents the translation details, while Section 6 describes the currently open issues. Sections 7 and 8 conclude the paper.

2 Background: GarpN translation steps

GarpN [4,5] is being developed to integrate qualitative and quantitative modelling. Currently, the GarpN procedure consists of three steps:

1. Translation of a qualitative representation into mathematical equations.
2. Numerical simulation of the generated equations.
3. Matching between quantitative behaviour and qualitative states.

In DynaLearn, a qualitative representation describes a system using entities and quantities. Quantities may have quantity spaces, derivatives, correspondences, causal relationships of type influence or proportionality, and calculi [6,8]. In GarpN, these ingredients are translated into mathematical equations. For example, a positive influence from one quantity to another can be translated into a differential equation, while proportional relationships are translated into value relationships or derivative relationships depending on the selected translation mode.

The translation results in a script with four sections: (*i*) equations, (*ii*) constants, (*iii*) initial values, and (*iv*) time-related settings. This is important because a qualitative representation typically does not contain all the numerical information required for simulation. In qualitative modelling, non-zero values are often only specified ordinally, for example as ‘positive’ or ‘negative’. Likewise, causal relations specify the direction of influence but not necessarily the numerical value, unit conversion, or the precise functional form. Therefore, translating a qualitative representation into a quantitative model requires additional numerical decisions.

The current case-study investigates these translation issues in the context of a scientific model whose reference equations are known.

3 Enzymatic cellulose deconstruction

Cellulose is the most abundant complex biopolymers and the main constituent of plants. Its deconstruction by specific enzymes, such as cellulase, is at the source of many biological and industrial processes. Maurer et al. [7] describe a Langmuir-Michaelis-Menten type model for the adsorption of cellulase onto a cellulose surface and the subsequent hydrolysis. The model distinguishes between free enzyme in solution, enzyme adsorbed onto the surface, enzyme complexed with the substrate, and product formation. Cellulase is in excess and the surface has a limited capacity, so adsorption is controlled by the free surface.

For the present comparison, we used Maurer's model implementation in Python and adapted from [9]:

State variables:

E_s = surface enzyme
E_c = complexed enzyme
P = product
E_f = free enzyme

Available surface:

$$S_{available} = R_{max} - (E_s + E_c)$$

Reference model:

$$dE_s/dt = k_a E_f S_{available} - k_d E_s - k_{-1} R_s E_s + k_{-1} E_c$$
$$dE_c/dt = k_{-1} R_s E_s - k_{-1} E_c$$
$$dP/dt = k_{-2} E_c$$
$$dE_f/dt = -k_a E_f S_{available} + k_d E_s$$

Corresponding flows:

$$r_{ads} = k_a E_f S_{available}$$
$$r_{des} = k_d E_s$$
$$r_{comp} = k_{-1} R_s E_s$$
$$r_{off} = k_{-1} E_c$$
$$r_{cat} = k_{-2} E_c$$

These equations formed the reference against which the GarpN-generated equations are compared.

4 The qualitative representation

We constructed a qualitative representation of the phenomenon described by Maurer et al. [7] adapted from previous work [9] (Fig. 1 and Table 1). The aim was to represent the conceptual structure of the model rather than to manually reproduce the mathematical equations. The representation includes the relevant quantities for Substrate, Active enzyme, Surface enzyme, Free enzyme, and Product.

The construction process focused on identifying the quantities that function as concentrations, the quantities representing flows (or rates), and the causal dependencies between those quantities. Attention was given to the surface limitation mechanism, because adsorption depends not only on free enzyme but also on available surface.

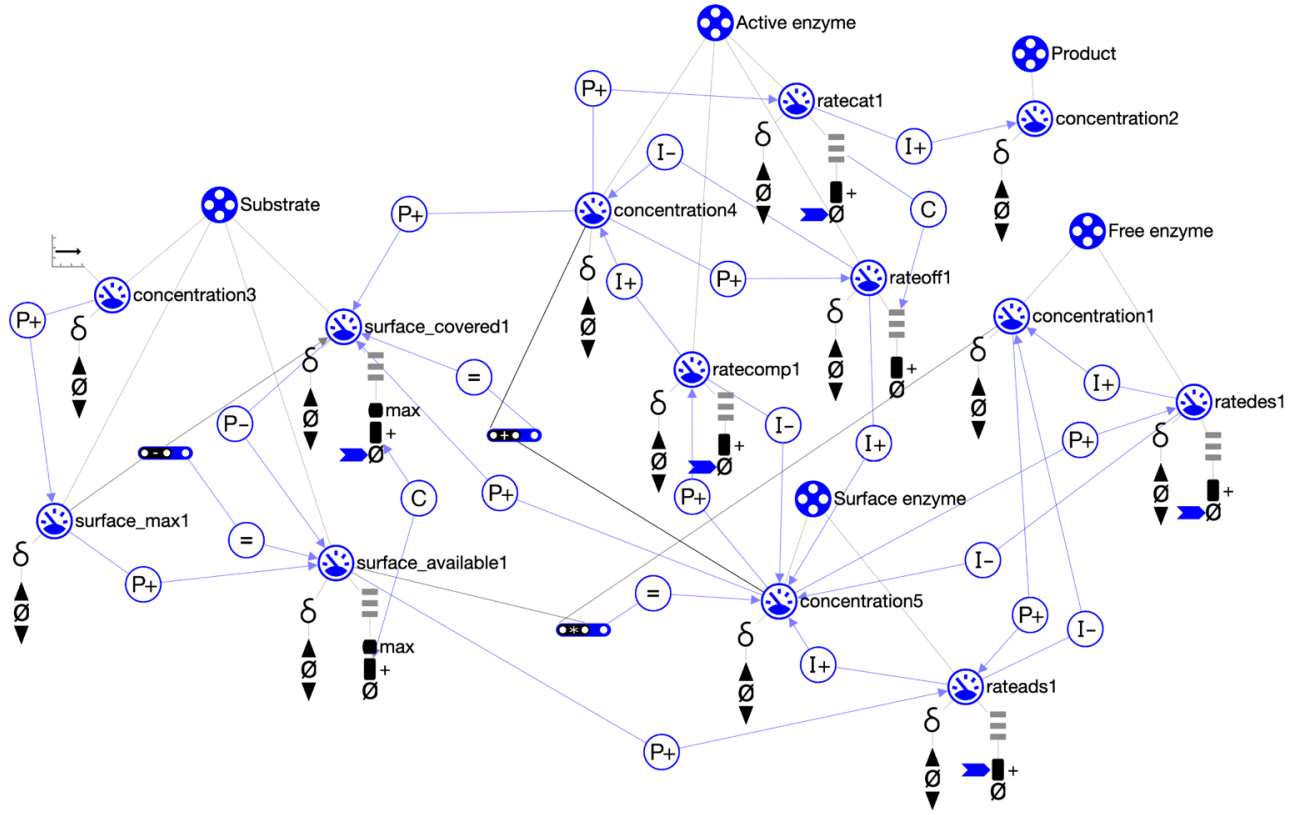

Figure 1. Qualitative representation of the surface-limited cellulose deconstruction model implemented in DynaLearn.

Table 1. Simulation results for the model shown in Fig. 1. S refers to State, <v, ∂> refers to value and derivative (change) respectively, and u refers to unknown value. The simulation generates two behaviour paths: $S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4$ and $S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_5$.

Entity	Quantity	S_1	S_2	S_3	S_4	S_5
Substrate	concentration3	<u, 0>	<u, 0>	<u, 0>	<u, 0>	<u, 0>
	surface_max1	<u, 0>	<u, 0>	<u, 0>	<u, 0>	<u, 0>
	surface_covered1	<0, +>	<+, +>	<+, +>	<max, 0>	<max, 0>
	surface_available1	<max, ->	<+, ->	<+, ->	<0, 0>	<0, 0>
Active enzyme	concentration4	<u, 0>	<u, +>	<u, +>	<u, 0>	<u, 0>
	ratecomp1	<0, +>	<+, +>	<+, +>	<+, 0>	<+, 0>
	ratecat1	<0, 0>	<0, +>	<+, +>	<+, 0>	<+, 0>
	rateoff1	<0, 0>	<0, +>	<+, +>	<+, 0>	<+, 0>
Surface enzyme	concentration5	<u, +>	<u, +>	<u, +>	<u, 0>	<u, 0>
	rateads1	<+, ->	<+, ->	<+, ->	<+, ->	<+, 0>
Free enzyme	concentration1	<u, ->	<u, ->	<u, ->	<u, ->	<u, 0>
	ratedes1	<0, +>	<+, +>	<+, +>	<+, 0>	<+, 0>
Product	concentration2	<u, 0>	<u, 0>	<u, +>	<u, +>	<u, +>

5 Translation by GarpN and initial settings

The qualitative representation was translated into a quantitative simulation script using GarpN. The generated script contained equations, constants, initial values, and time settings. The resulting equations were then inspected and compared to the reference model.

For readability, we rewrote the generated quantities as follows:

E_f = concentration1(Free enzyme)
P = concentration2 (Product)
S = concentration3 (Substrate)
E_c = concentration4 (Active enzyme)
E_s = concentration5 (Surface enzyme)

GarpN generated equations (original shown in Fig 2):

$$S = c16 + c17\,\Delta t$$
$$r_{des} = c7\,E_s$$
$$S_{max} = c1\,S$$
$$S_{covered} = E_c + E_s$$
$$S_{available} = S_{max} - S_{covered}$$
$$r_{ads} = c4\,E_f * c5\,S_{available}$$
$$E_f(t + \Delta t) = E_f(t) + (c14\,r_{des} - c15\,r_{ads})\,\Delta t$$
$$r_{comp} = c6\,E_s$$
$$r_{off} = c3\,E_c$$
$$E_s(t + \Delta t) = E_s(t) + (r_{ads} - r_{des} - r_{comp} + r_{off})\,\Delta t$$
$$E_c(t + \Delta t) = E_c(t) + (c8\,r_{comp} - c9\,r_{off})\,\Delta t$$
$$r_{cat} = c2\,E_c$$
$$P(t + \Delta t) = P(t) + r_{cat}\,\Delta t$$

Initial values:

$$E_f(0) = 3$$
$$E_s(0) = 0$$
$$E_c(0) = 0$$
$$P(0) = 0$$

Table 2 shows the settings of the constants.

Table 2. Values (V) assigned to constants (C).

C	V	Interpretation
c1	1	Scaling from substrate to maximum surface
c2	4.2	Catalysis rate
c3	3	Decomplexation/off rate
c4	1.7	Adsorption rate component
c5	1	Scaling of available surface in adsorption
c6	0.16	Complexation rate
c7	0.14	Desorption rate
c8	1	Scaling of complexation in active enzyme update
c9	1	Scaling of off-rate in active enzyme update
c14	1	Scaling of desorption in free enzyme update
c15	1	Scaling of adsorption in free enzyme update
c16	2.9	Substrate/surface-capacity value
c17	0	Rate of change of substrate/surface capacity

The time step was $\Delta t = 0.001$ and the simulation was run for 1000 iterations, corresponding to t_end = 1. GarpN used Euler integration. The Python reference model used adaptive RK45 integration. Therefore, the comparison reported here focuses primarily on equation-level and structural equivalence.

6 Results

The first step in the comparison is to verify that the generated quantities during the translation correspond to the state variables in the reference implementation (E_r, P, S, E_c and E_s, but also $S_{covered}$ and $S_{available}$).

More importantly, the next step is to compare each generated flow (process) with its counterpart in the reference model. Table 3 shows the details of this mapping. Indeed, with the selected values, the generated model matches the reference model: c4*c5 = 1.7, c7 = 0.14, c6 = 0.16, c3 = 3, c2 = 4.2, and c1*c16 = 2.9. Thus, after specification of the constants, the GarpN-generated equations reproduce the structure of the reference model.

6.1 Addition versus multiplication

Although the final equations are structurally equivalent, the comparison revealed several points where equivalence depends on explicit (human) choices and settings. One specific issue concerns the occasional need for multiplication instead of addition in the case of multiple proportional dependencies.

Currently, multiple proportional dependencies influencing a target quantity are represented as additions by GrapN. This is thus also the case for *ratesads1* (of Surface enzyme) (r_{ads}) which has positive proportional relationships with *surface_available1* (of Substrate) and *concentration1* (Free enzyme) (E_r). Hence, $r_{ads} = c4\ E_f + c5\ S_{available}$ is generated. However, in the reference model these details are represented as: $r_{ads} = k_a\ E_f\ S_{available}$ (or: $r_{ads} = c4\ E_f * c5\ S_{available}$). By using addition (instead of multiplication), the model strictly speaking is not equivalent and in fact significantly different. An additive form implies, among others, that the process could still occur if one of the two influencing factors is absent. This changes the mechanism and the resulting dynamics. It shows that qualitative-to-quantitative translation sometimes requires more than identifying the direction of causality. For some processes the functional form of the involved relationship is essential.

6.2 Surface capacity must remain constant

In the reference model, the maximum surface capacity is a constant:

$$R_{max} = 2.9$$

The GarpN equations express maximum surface through:

$$S = c16 + c17\ \Delta t$$
$$S_{max} = c1\ S$$

With c16 = 2.9, c17 = 0, and c1 = 1, this gives $S_{max} = 2.9$. Thus, the GarpN model becomes equivalent to the reference model. However, this equivalence depends on setting c17 = 0. If c17 is not zero, surface capacity would change over time, resulting in a different model. This point illustrates that quantities that are conceptually constant in a scientific model may still require explicit treatment in the generated simulation script. Since qualitative representations do not always distinguish constants from slowly changing or externally determined quantities, the user must inspect and set these values carefully.

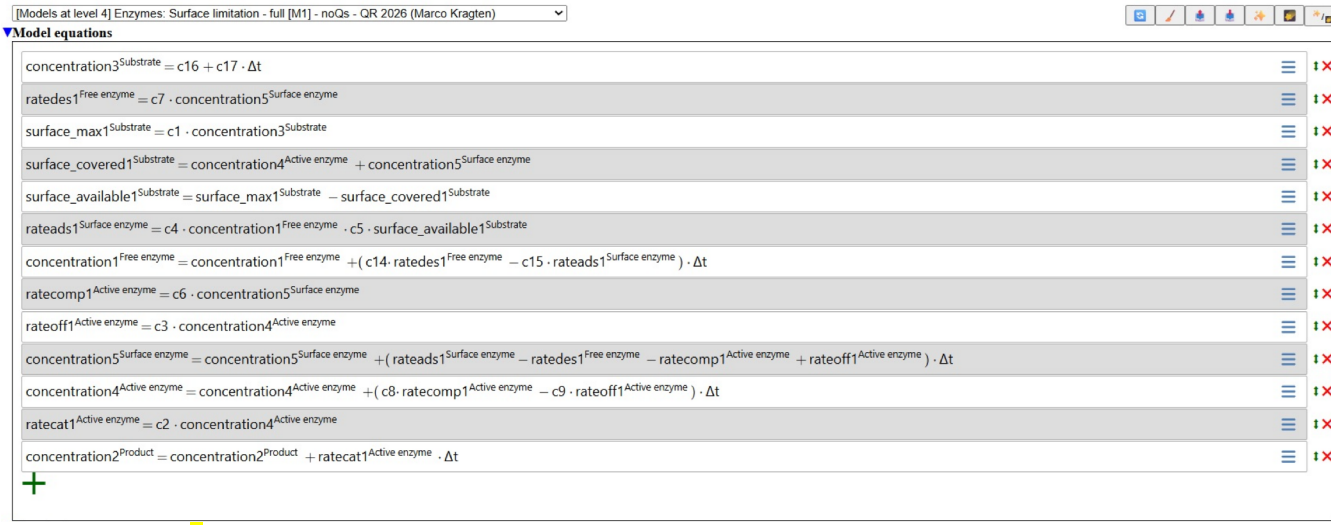

Figure 2. Screenshot of Model equations as generated by GarpN based on the qualitative model shown in Fig. 1.

Table 3. Process mapping with reference model.

Process	Reference model	GarpN-generated equations	Mapping
Adsorption	$r_{ads} = k_a\ E_f\ S_{available}$	$r_{ads} = c4\ E_f * c5\ S_{available}$	Equivalent, if $c4*c5 = k_a$
Desorption	$r_{des} = k_d\ E_s$	$r_{des} = c7\ E_s$	Equivalent, if $c7 = k_d$
Complexation	$r_{comp} = k_{-1}\ R_s\ E_s$	$r_{comp} = c6\ E_s$	Equivalent, if $c6 = k_{-1}\ R_s$
Decomplexation	$r_{off} = k_{-1}\ E_c$	$r_{off} = c3\ E_c$	Equivalent, if $c3 = k_{-1}$
Catalysis	$r_{cat} = k_{-2}\ E_c$	$r_{cat} = c2\ E_c$	Equivalent, if $c2 = k_{-2}$
Surface coverage	$E_s + E_c$	$E_s + E_c$	Equivalent
Surface availability	$R_{max} - (E_s + E_c)$	$S_{max} - S_{covered}$	Equivalent, if $S_{max} = R_{max}$

6.3 Numerical simulation

Fig 3. shows the numerical simulation results of the model generated by GarpN, using a fixed-step Euler integration:

$$X(t + \Delta t) = X(t) + f(X,t)\,\Delta t$$

with $\Delta t = 0.001$ and 1000 iterations. The reference Python implementation uses adaptive RK45 integration. Therefore, even when the equations are structurally equivalent, the numerical trajectories may differ because of the integration method.

This difference is not a conceptual mismatch. If the GarpN equations are structurally equivalent and the time step is sufficiently small, Euler simulations should approximate the reference solution.

The present settings result in a simulation interval of $t = 0$ to $t = 1$. The current Python reference implementation was run over $t = 0$ to $t = 100$. Therefore, a direct numerical comparison requires aligning the time intervals, either by running the Python model over $t = 0$ to $t = 1$, or by increasing the GarpN number of iterations.

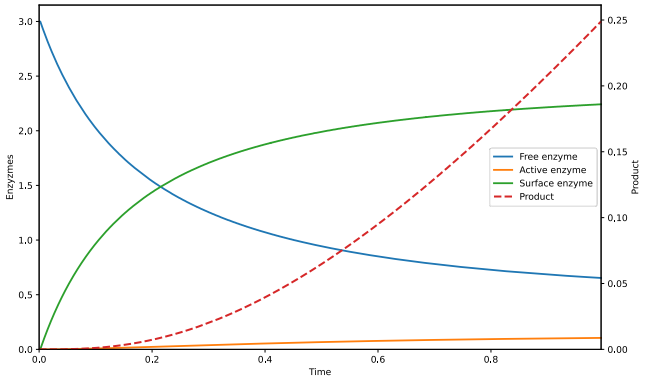

Fig. 3. Numerical simulation of the model generated by GarpN.

7 Discussion

The case-study presented in this paper shows that GarpN can generate equations that reproduce the structure of a mechanistic scientific model. The generated equations corresponded closely to the reference Langmuir-Michaelis-Menten surface-limitation model. The main flows in the reference model were all represented in the GarpN output, including surface coverage and surface availability.

However, the case also shows that equivalence is not fully automatic. The translation from a qualitative representation into a quantitative model requires additional numerical and semantic information. Three points were particularly informative. First, adsorption had to be represented as a multiplicative relation between free enzyme and available surface. This is not a minor numerical detail but a central mechanistic property of the model. A qualitative model can show that both quantities are relevant, but the quantitative implementation must specify how they combine.

Second, surface capacity had to be made constant by setting $c17 = 0$. Without this setting, the generated model would allow surface capacity to change over time. This reveals a general issue for qualitative-to-quantitative translation: constants in a quantitative model may still appear as quantities in the qualitative representation and therefore require explicit numerical treatment.

Third, surface enzyme had to be treated as a dynamic stock. The final GarpN model did this correctly, but the comparison made clear why this distinction matters. A stock changes through inflows and outflows; an auxiliary quantity is recalculated directly from other variables. Treating a stock as an auxiliary relation would fundamentally change the model.

These findings suggest that GarpN should not be understood simply as an automatic equation generator. Rather, it functions as a formal translation environment that makes modelling decisions explicit. The generated equations provide a starting point, but model equivalence requires inspection of functional forms and numerical settings.

For educational use, this may be a strength rather than a weakness. By comparing the qualitative representation, the generated equations, and the numerical simulation, learners can be supported in understanding what is required to move from a conceptual causal structure to an executable numerical simulation.

For research and development, the case suggests several improvements to GarpN: clearer support for multiplicative process terms, stronger distinction between stocks and auxiliary quantities, explicit treatment of constants, and automated checks that compare generated equations with expected stock-and-flow patterns.

8 Conclusion

This paper presented a case-study in which GarpN was applied to address a scientific model. The case shows that a qualitative representation can be translated into executable equations that are structurally equivalent to the quantitative reference model.

At the same time, the case revealed that equivalence depends on explicit modelling decisions. The translation from a qualitative representation to a quantitative model is not merely a syntactic conversion. It is a modelling process in which the qualitative structure provides an important foundation, while additional quantitative refinements are also required.

The next step is to complete the numerical comparison by running GarpN and the Python reference model over the same time interval and comparing their trajectories. More broadly, this case provides input for further development of GarpN as a tool for integrated qualitative and quantitative modelling.

Acknowledgments

The research presented here is co-funded by the Dutch funding agency Regieorgaan SIA (https://regieorgaan-sia.nl/) via projects Pluk de data! (RAAK.PUB09.059) and Synergie. (RAAK.PUB12.031). We would like to thank co-workers for their constructive feedback and support.

9 References

[1] Forbus, K. D. (2019). *Qualitative representations: How people reason and learn about the continuous world*. MIT Press.

[2] Winsberg, E. & Alvarado, R. (2026). Computer Simulations in Science. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/simulations-science/

[3] Soofastaei, A. (Ed.). (2023). Numerical Simulation - Advanced Techniques for Science and Engineering. *IntechOpen*. https://doi.org/10.5772/intechopen.112097

[4] Kragten, M., & Bredeweg, B. (2024). Qualitative and quantitative modelling of dynamic systems: how do they relate? *37th International Workshop on Qualitative Reasoning*, Santiago de Compostela, Spain.

[5] Kragten, M., Wielemaker, J., & Bredeweg, B. (2025). Integrated Qualitative and Quantitative Modelling in GarpN. *38th International Workshop on Qualitative Reasoning*, Montreal, Canada.

[6] Bredeweg, B., Liem, J., Beek, W., Linnebank, F., Gracia, J., Lozano, E., Wißner, M., Bühling, R., Salles, P., Noble, R. A., Zitek, A., Borisova, P., & Mioduser, D. (2013). DynaLearn: An intelligent learning environment for learning conceptual knowledge. *AI Magazine*, 34(4), 46-65.

[7] Maurer, S. A., Bedbrook, C. N., & Radke, C. J. (2012) Cellulase Adsorption and Reactivity on a Cellulose Surface from Flow Ellipsometry. *Industrial & Engineering Chemistry Research*, 51(35), 11389-11400.

[8] Bredeweg, B., Linnebank, F., Bouwer, A., & Liem, J. (2009). Garp3: Workbench for qualitative modelling and simulation. *Ecological Informatics*, 4(5-6), 263-281.

[9] Kansou, K., Rémond, C., Paës, G., Bonnin, E., Tayeb, J., & Bredeweg, B. (2017). Testing scientific models using Qualitative Reasoning: Application to cellulose hydrolysis. *Scientific Reports,* 7, Article 14122. https://doi.org/10.1038/s41598-017-14281-4