Computers and Electronics in Agriculture 218 (2024) 108664

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Reusability check-based refinement of a biophysical fishpond model

P. Sharma [a], G. Gyalog [a], L. Berzi-Nagy [b], F. Tóth [a], Z. Nagy [a], B. Halasi-Kovács [a], D.L. Fazekas [a], D. Mezőszentgyörgyi [c], B. Csukas [d], M. Varga [c,*]

[a] Research Center for Fisheries and Aquaculture, Institute of Aquaculture and Environmental Safety, Hungarian University of Agriculture and Life Sciences, 35 Anna liget, 5540, Szarvas, Hungary
[b] Agroloop Hungary Ltd., 7/A, Illatos út, 1097, Budapest
[c] Institute of Animal Sciences, Hungarian University of Agriculture and Life Sciences, 40 Guba, 7400, Kaposvar, Hungary
[d] Hungarian University of Agriculture and Life Sciences, 1 Pater K, 2100, Godollo, Hungary

ARTICLE INFO

Keywords:
Reusability
Biophysical model
Reduced / extended model
Pond aquaculture
Scaling-up

ABSTRACT

Given the increasing importance of data- and model-driven design and control in food production systems, this paper addresses the need to improve the reproducibility, replicability, and reusability of datasets, models, and modeling frameworks. While sensor data and machine learning-based control and operation of agricultural and aquacultural systems face reproducibility and replicability challenges, reusability is becoming critical for computational model-based design and planning of complex processes. This study evaluates the reusability of an existing pond aquaculture model and outlines a systematic, stepwise approach for reusability enhancements. The suggested methodology starts with an established reference model of a typical production fishpond and improves its reusability through pilot-scale experiments, covering key aspects of pond farming technologies. The reference model is subjected to stepwise reusability checks using measured data from the respective pilot units, progressing from simpler (reduced) to more complex (extended) cases. Each step concludes with the necessary parameter or sub-model refinements, which remain unchanged in subsequent steps. The refined model is validated with the measured data from other pilot experiments. This process can be repeated until satisfactory results are obtained. The resulting model is then tested to scale up a production pond model using limited case-specific input data. In addition, a hypothetically modified scenario is studied to address discrepancies between measured and simulated data.

1. Introduction

Nowadays, driven by the increasing amount of measured data and advanced methods in the surging number of publications, attention has shifted to the reproducibility, replicability, and reusability of the accumulated knowledge (Plesser, 2018; European Commission, 2020). Reproducibility pertains to the availability of the complete datasets, methodological descriptions, and computational codes. Replicability involves testing these resources for another dataset to solve the same problem by someone else, elsewhere. Reuse encompasses not only the repeated use of data, methods, and models but also leveraging research findings beyond the original research context. Within the original topic of research, reusability means the application of data, models, and tools for the solution of simplified or extended problems, as well as for scaling-up the results. In a broader context reusability can be extended to other scientific fields or disciplines (e.g., for innovation, knowledge transfer, transdisciplinary research), marking the final stage of the research life cycle.

In the context of agro-environmental systems, computational tools with improved data and model reusability, are crucial to ensure a reliable food supply under changing environmental conditions (Liu et al., 2002; de Kok et al., 2015; Varga et al., 2020). In this context, Palma et al. (2016) introduced an open-access platform for agricultural data management, underpinned by a novel data model that integrates and extends different standards, including those related to GIS. The CGIAR Ontologies Community of Practice also initiated efforts to improve knowledge representation and ontology development for effective use of multidisciplinary data (Arnaud et al., 2020). In marine science, Schaap et al. (2022) showed the importance of data reuse through their comprehensive review of marine-related data infrastructures. Holzworth et al. (2010) highlighted the growing awareness in the modeling community of the need for model sharing, despite the paradox that redeveloping

* Corresponding author.
E-mail address: varga.monika@uni-mate.hu (M. Varga).

https://doi.org/10.1016/j.compag.2024.108664
Received 23 January 2023; Received in revised form 16 January 2024; Accepted 19 January 2024
Available online 2 February 2024

models is often easier than adapting existing models to new contexts. They shared insights from their work on the APSIM farm system modeling framework to address this challenge. Stella and colleagues (2015) also provided an example of model reuse by repurposing a sugarcane model for a giant reed, highlighting the evolving landscape of reusability in agricultural research.

Barriers to model reuse are multifaceted, primarily arising from the complex nature of agro-environmental models, resulting in a time-consuming process for understanding, modifying, and re-implementing them (Holzworth et al.,2010). Most of the models often have limited extensibility to specific cases, lacking standardization within the modeling framework. Hence, the challenge of model reuse is exacerbated as sub-models designed for one framework are not easily transferable to others (Donatelli et al., 2012). Extending biophysical models also demands additional initial data sets and parameters, making the process resource-intensive.

Today, rapid advances in sensors and machine learning provide good solutions for local control and operation problems, however, these methods have limited ability to capture complex environmental interactions (Lokers et al., 2016). Databases and data-driven models often lack completeness and consistency (Talari et al., 2022), while the knowledge remains fragmented rather than integrated across the farm-to-fork chain (Repar et al., 2023; Rosen et al., 2022). While machine learning effectively processes big data for short-term decision-making, it fails to explain causal relationships in complex systems (Durden et al., 2017). Consequently, the importance of first principles-based ('a priori', mechanistic, biophysical) models is crucial (Chary et al., 2022; Cohen et al., 2022).

In the sphere of agricultural and aquacultural systems, effective model-based design and planning necessitate the utilization of previously validated mechanistic, biophysical models, allowing for the modification of parameters to address a range of scenarios, including various species and technologies, different environmental conditions, scaling-up, etc.

The freshwater fishpond sector occupies a central position within the European aquaculture industry. Consequently, a significant demand exists for reusable models aimed at augmenting productivity, efficiency, and quality while reducing costs in these systems (Varga et al., 2020). In Europe, approximately 38 % of the freshwater aquaculture production is derived from fishponds located in Central and Eastern European countries, with Cyprinids constituting the majority of production (comprising almost 80 %, Gyalog et al., 2022). Fishpond production practices in many countries, however, continue to be 'extensive' or 'semi-intensive', yielding between 500 and 2500 kg/ha/year (Horváth et al., 2002). These complex systems function as open ecological systems, where natural and technological processes are in synergy and cannot be isolated (Aquaculture Advisory Council, 2021). Many fish farms have transformed into multifunctional units, integrating into the social, ecological and economic dimensions of fish farming (Palásti et al., 2020; Popp et al., 2019). Ecological intensification is suggested to attain the sustainable goals within fishpond production systems by merging artificial interventions with ecosystem services (Dong et al., 2022). Hence, it is crucial to consider a broad range of allochthonous and autochthonous factors that impact pond aquaculture. Specifically, understanding the effects of managerial interventions — such as feeding, manuring, stocking density, and pond level control — on complex fishpond ecosystem processes is vital for achieving effective production (Varga et al., 2020).

Several models have been developed since the 1970 s to understand fishpond performance (Ogawa & Mitsch, 1979; Prinsloo & Schoonbee, 1984; Svirezhev et al., 1984; Hagiwara & Mitsch, 1994; Jorgensen, 1995; Kochba et al., 1994). Bolte et al. (2000) constructed a decision-support tool based on databases, knowledge-based components, and models of the pond ecosystem, and various decision-support features. Emphasizing the reusability of data, models and code, the authors created a decision support tool, the POND model. Recent advancements in information technology have enabled the creation of interactive, software-based systems, that comprehend complex interactions in the fishpond environments, resulting in the reusability of process algorithms and specific model units. For example, utilizing mass-balanced models like Ecopath (Christensen et al., 2008), Aubin et al., 2021, drew inferences about the fish diet in common carp polyculture through a combination of experiments and trophic web interactions modeling. However, achieving a more comprehensive understanding of the complex relations and interacting mechanisms among various elements in the pond environment remains a challenge (Mathisen et al., 2016). To develop a precise pond management model, a deeper comprehension and consideration of detailed sub-groups within the food web, such as phytoplankton, zooplankton, and dissolved matter, are imperative. While numerous modeling tools and techniques are available for fishpond, it is evident that sophisticated biophysical models are essential for design and planning of pond aquaculture ecosystems. However, the case-specific validation of these models for individual ponds is a highly expensive and time-consuming process. Therefore, testing and improving the reusability of the existing modeling frameworks are crucial tasks.

In light of this, our work aims:

- to analyze the reusability of a formerly developed and validated reference biophysical fishpond model; and
- to enhance the reusability of this model for a broader range of differently managed fishponds.

Testing and improvements included appropriate reduction and extension of the selected reference model to describe different pond management cases and a hypothetically modified scenario for the aquatic food web. Having improved and validated the model, it was tested for scaling-up for another fishpond.

2. Materials and methods

2.1. Short description of the reference model

The reference model used in this study was developed by Varga et al. in 2020 as a part of a decision support tool (DSS) for the ClimeFish project to assess the impact of climate change on fishpond aquaculture.

The representation of the food web followed the structure, published by Svirezhev et al. (1984), but instead of using choice functions, the relative availability-based selection of alternative feeds and preys was used. Moreover, the biomass-based fish growth was replaced for the individual fish mass calculation, according to the anabolic and catabolic processes in sense of Bolte et al. (2000).

The model was implemented for the SzegedFish Ltd. fish farm site, and was validated using 10 years of data from the farm pond register. The dynamic process model of medium complexity represented a simplified food web with predator–prey interactions, involving common carp, bighead carp, zooplankton, phytoplankton, benthos and detritus. In the pond water, the model considered dissolved oxygen, nitrogen, and phosphorus components, as well as a solid mass of feed (corn) and manure. The modelled life-related processes of fish (common carp and bighead carp) included individual catabolism and anabolism-based feed consumption and growth, respiration, faecal and ammonium excretion and mortality. For the other species, biomass growth, the catabolized food, respiration, and the decay were calculated. Life processes of phytoplankton described the photosynthesis, the nitrogen and phosphorus limited growth, and the oxygen production. The model took into consideration the managerial interventions (stocking, feeding, manuring, water level control) and a set of meteorological data.

As a limitation of the reference model, determined by the limited availability of measured data, the following parts were considered only by a series of assumptions, as follows:

- In the absence of explicit data for solar radiation and humidity, only estimated values were used;
- In the absence of appropriate zooplankton, phytoplankton, benthos and detritus measurement the initial values were estimated from published data, and this part of model was not validated, comprehensively.

It was concluded that, despite the limited amount of data available, the approximately validated, but well-structured balance model made it possible to study the effects of climate change on fishpond ecosystems and on carp growth.

2.2. Applied modeling framework

The reference model was implemented in the previous version of Programmable Process Structures (Varga et al., 2017; Varga & Csukas, 2017).

In the current study we applied the up-graded, consolidated version of Programmable Process Structures (PPS) Varga and Csukás (2022). In PPS, the unified process models of different process systems are generated from one general state and one general transition *meta*-prototypes. The *meta*-prototypes are prepared for distinguished input and output of model-specific, additive conservational measures and of signals. Also, the *meta*-prototypes give a template to define spatial and temporal scales and various parameters in the prototype programs. The generation of the actual models involves the multiplication of these *meta*-prototypes to describe the process net of the actual problem under investigation. Actually, it results in a special net structure that consist of the actual state and transition elements, describing the structure of the investigated process system. The functionalities behind this structure are ensured by the case-specifically defined, functional program prototypes that can also be derived from the two general *meta*-prototypes. These program prototypes determine the locally executable programs of the investigated model. During the execution, the actual state and transition elements are calculated by the programs of their associated state or transition prototypes. The communication between the state and transition elements is solved by uniform connections, while the generated model is executed by a generally usable kernel program. The PPS framework represents a unified multi-disciplinary methodology to combine holistic structural and local functional characteristics in modeling and simulation-based problem-solving of various process systems. PPS is implemented in the declarative, logical language of SWI-Prolog. In particular, unification and lists of functors in logical programming AI language help the effective (and reusable) representation and execution of the models. PPS has proven to be a successful approach for constructing models, simulating, designing, planning, and evaluating strategies in agro-environmental process systems (Varga, 2022; Varga et al., 2022, Varga et al., 2023). The reusable, locally executable code prototypes offer advantages, especially for describing large, multi-scale systems with standardized components. The fact that most variables are local, promotes code reusability and simplifies variable naming within local programs.

2.3. First series of pilot experiments to check and improve the reference model

For testing and improving the reference model's reusability, datasets from pilot ponds were used, and subjected to various experimental conditions.

During the 7-month growing season from 1 April to 31 October in 2021 and 2022, carp rearing experiments were carried out in closely monitored fishponds to generate data on food web dynamics. The experiments were conducted in earthen ponds with a surface area of 10,000 m^2 and a depth of two meters at the site of the Hungarian University of Agriculture and Life Sciences (Institute of Aquaculture and Environmental Safety, Research Centre for Aquaculture and Fisheries, MATE AKI HAKI, Szarvas, Hungary). These ponds, assigned the codename CS2, CS3, CS6, and CS7, were used to generate experimental data. The geo-locations of these ponds are illustrated in the Mendeley database (Sharma et al., 2024).

Two (CS6, CS7) and three (CS2, CS3, CS6) ponds were stocked with second year common carp in 2021 and 2022, respectively. The ponds were operated under various feeding and fertilization regimes throughout the production season to track the impact of different nutrient management scenarios on pond food web. Table 1 provides details on the stocked fish quantities, feeding rate, fertilizer input, and stocking and harvest timings. The raw measurement data, along with the comprehensive feed and fertilizer input for five fishponds, can be found in "RawData_2021.xlsx" and "RawData_2022.xlsx" respectively, which are located in the "Measured_Data" folder of the Mendeley database (Sharma et al., 2024). The process for gathering the data is outlined below:

- Water samples from the ponds were collected twice a week, and analyzed for ammonium (mg/dm^3), nitrate (mg/dm^3), nitrite (mg/dm^3), orthophosphate (mg/dm^3) and chlorophyll-a (mg/dm^3) using to standard analytical methods;
- In 2021, dissolved oxygen (mg/dm^3) and water temperature (°C) were both measured manually twice a day, using multi-parameter water quality meter. In 2022, sensors (Aquaread AP7000) were placed in the fishponds to measure these parameters hourly.
- Zooplankton biomass (cm^3/100 dm^3) was monitored twice a week using a 50 μm mesh plankton net. For each sample, 100 dm^3 of pond water was filtered and concentrated to 100 cm^3. All samples were preserved in formaldehyde, then settled in a centrifuge tube and biomass was measured after 24 h.
- Meteorological data, including air temperature (°C), wind speed (m/s), precipitation (mm/day), and solar radiation (W/m^2), were collected from the Agromet Solar automatic meteorological station, located approximately 1 km away from the ponds in Szarvas.
- Detailed information on the schedule and amount of feed (in kg) supplied to the ponds was recorded throughout the production period;
- Strategy (date and quantity) for adding organic manure and/or inorganic fertilizer in each pond was recorded. Additionally, laboratory measurements for the manure composition were recorded;
- We recorded the number and weight (in kg) of the fish that were stocked and harvested, as well as the increase in weight, determined through fish sampling under different pond managerial practices.

2.4. Overview of case studies involved in reusability check and improvement

Considering the limited availability of validating measurements and the other limitations of the reference model (see Section 2.1), this model has been adapted to different cases from pilot experimental ponds. Moreover, in certain situations, fishponds receive both organic manure and inorganic fertilizers to boost natural food production for fish growth, while this extra pool of nitrogen from inorganic fertilizer was not considered in the previous model. In addition, the reference model did not take into account the dry matter content and the nitrogen and phosphorus content of the dry matter when describing the breakdown of the manure. To capture also these parts, we planned a series of pilot pond experiments.

Accordingly, in this work, the existing reference model (Szeged-Fish) of a typical production pond was used to investigate the characteristics of a series of pilot-scale experiments. These experiments covered the characteristics of a wider range of production ponds. The reference model was used to generate appropriately reduced (2021CS6) and extended (2022CS6 and 2022CS2) computational models of these pilot case studies. At the same time, a step-by-step reusability check of these underlying pilot models was carried out with the knowledge of the

Table 1
Datasets used for reusability check and improvement of the model.

Source of dataset	Szeged-Fish farm	Controlled pilot experiments at MATE HAKI					Hypothetically extended case
Dataset code	Szeged-Fish	2021 CS6	2022 CS6	2022 CS2	2021 CS7	2022 CS3	Cyano
Type of model	Reference model	Reduced model[8]	Extended model[9]	Extended model[9]	Reduced model[8]	Extended model[9]	Distinguish eukaryotes and cyano-bacteria
Role of model	Reference model	To test and improve the reusability of the reference model (Section 3.3)			To validate the improved model (Section 3.4)		Hypothetical extension (Section 3.5)
Date of stocking	01.04.2011	26.05.2021	16.05.2022	16.05.2022	26.05.2021	16.05.2022	02.05.2021
Date of harvest	31.10.2011	09.09.2021	14.09.2022	21.09.2022	07.09.2021	28.09.2022	02.09.2021
Stocking density, kg/ha	376	101	200	200	201	200	200
Feed input[1] t/ha/season	2.2	no	0.6[2]	0.7[2]	no[2]	0.9[2]	0.9
Manure[3], t/ha/season	1	no	5[4]	5[4]	11[5]	9[5]	9
Inorganic fertilizer[6], kg/ha/season	no	no	no	200[7]	no	no	no
Distinguished eukaryotes + cyano-bacteria	no	no	no	no	no	no	Initial cyano-bacteria concentration: 0.5 %[10]

[1] Wheat.
[2] Daily feed portions corresponded to 0.5; 1; 2; 2 and 1 % of estimated biomass weight in May, June, July, August, and September, respectively.
[3] Cow manure.
[4] In two installments. For further information, the reader is referred to the Mendeley database (Sharma et al., 2024).
[5] In four installments. For further information, the reader is referred to the Mendeley database (Sharma et al., 2024).
[6] Ammonium nitrate.
[7] In two installments.
[8] Reduced model (highlighted by green) describes extensive fish production with low stocking rates and without external nutrient supply.
[9] Extended model describes intensive fish production with a higher nutrient supply, with manure, and with optional inorganic fertilizer.
[10] Based on expert estimation

measured data from the respective pilot units, beginning with the simpler cases and working our way up to the more complex ones. After each step, we refined and tested the parameters or details of the model as necessary. These modifications have been adjusted for further processing. Finally, the evaluated and improved model was validated using data from two additional pilot cases (2021CS7 and 2022CS3) and underwent several rounds of testing and validation calculations. As a result, a cohesive model applicable to a broader range of ponds, including natural ponds and those that are intensively fed and fertilized, was established.

The data used in the reference model, the plan for the pilot experiments and the hypothesized model extension are listed in Table 1.

The last row and column in Table 1. refer to the specifications considered to build a hypothetically extended scenario, where the eukaryotes and cyanobacteria groups are distinguished instead of an integrated consideration of phytoplankton.

2.5. Second series of validating experiments

To account for sampling and measurement errors (Section 3.2, point 9), a number of formerly conducted set of pilot pond experiments were also included. In course of the ARRAINA project (Advanced Research Initiatives for Nutrition and Aquaculture, N288925-EU FP7) three feed types (crop, fish oil- and plant oil-based) were tested, using common carp monoculture at the MATE AKI HAKI, Szarvas, Hungary. Two earthen ponds with parallel experiments (average area: 1772 m^2, depth: 1.3 m) were assigned to each diet type. Water inflow was provided from the nearby oxbow lake of river Körös. On average 5288 individuals/ha, with an average body weight of 59.8 g were stocked into the lakes in 2014. As usual with the semi-intensive technology, cow manure was used to enhance the natural production of the ponds. Each pond individually received 450 kg of manure in the season.

In view of the availability of parallel pilot experiments, we also used these data for additional validation of the improved model.

The detailed experimental data are summarized in the RawData_2014.xlsx in the Mendeley database (Sharma et al., 2024).

2.6. Evaluation of measurement and simulation errors

For data filtering, particularly concerning dissolved oxygen, total inorganic nitrogen, and phosphorus, we utilized Matlab® Data Cleaner for pre-processing. The chosen smoothing method was the moving median, using a smoothing factor of 0.25. We opted for the moving median as it effectively handled the presence of significant outlier data and fluctuating values in the raw measurements. The cleaned data can be accessed in the Mendeley database (Sharma et al., 2024) under the file name "Filtered_DO_N_P_2021_2022.xlsx".

In general, the measurement error during experimentation is defined as the sum of sampling and non-sampling error (EC CROS Knowledge repository, 2023).

In the first series of experiments (Section 2.3) in the absence of knowledge about measurement errors (i.e., in the absence of parallel experiments), we described the average deviation of simulated and calculated data using normalized root mean square error (NRMSE, %) (Chai & Draxler, 2014), as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{i=1}^{N}\left(\mathrm{x_i} - \mathrm{x_i^*}\right)^2} \tag{1}$$

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{\mathrm{x}_{max} - \mathrm{x}_{min}} * 100 \tag{2}$$

where RMSE is the root mean square error, in the unit if the original values, x_i is the i^{th} value of the observed time series, x_i* is the i^{th} value of the calculated time series (model prediction), i is the i^{th} variable, N is the number of data points, NRMSE is the normalized root mean square error, %, and x_{max} and x_{min} are the maximum and minimum values of the observed time series. The measured data, including sampled and harvested carp biomass, measurements of water quality parameters (e.g., dissolved oxygen, total inorganic nitrogen, phosphorus), and phytoplankton and zooplankton biomass, were used for NRMSE calculations. This error measure summarizes the errors made during sampling, measurement, model mapping, and simulation.

In the second validation series of pilot experiments (Section 2.5), there were 2–2 parallel measurements for 3 case studies. For these experiments, we characterized the differences in the measurements by the standard deviation (Hayat Khan, 2011), according to the expression

$$SD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{3}$$

where SD is the standard deviation, dimension is expressed in the same units as the original values, x_i is each individual data point in the set, μ is the mean, and N is the total number of data points.

It should be noted that the rigorous statistical approach was constrained by the limited number (2) of parallel experiments, as well as by the possibly non-normalized distribution of the data.

For the comparison of error regarding the measured vs. calculated data compared to the SD of measurements, RMSE values were also calculated for the six additional experiments, based on Eq.1.

The calculation of RMSE and NRMSE, as well as SD values were prepared in the MS Excel file of the simulation output, and in the SD_of_parallels_2014.xlsx, respectively, and both are available in the Mendeley database (Sharma et al., 2024).

3. Results and discussion

3.1. Simplified overview of the investigated fishpond model

A simplified structural overview of the studied model is illustrated in Fig. 1. According to the stocking of pilot ponds, we removed bighead carp from the original reference model. Also, due to its relatively limited proportion in the carp diet, as well as its unknown initial concentration and parameters, the benthos was not represented in the base model.

The simplified structure shows the contribution of different nutrients (feed, as well as zooplankton produced from phytoplankton and/or detritus, influenced by the added manure and inorganic fertilizer) to the diet of the produced carp. Given their significant role in natural pond food production, the model also takes into account the nitrogen and phosphorus content, in addition to the total mass.

The primary nutrient sources for carp are feed and zooplankton. Within the ecosystem's food web processes, solar radiation-driven synthesis produces phytoplankton biomass from CO_2, H_2O, N and P, while releasing O_2 into the water. Additionally, there is an equilibrium-driven kinetic transport of O_2 (and also CO_2) between the water and the atmosphere.

In this simplified model, 'available detritus' refers to the cumulative organic solid phase resulting from the ongoing decomposition of manure and other allochthonous sources such as uneaten feed, faecal matter from species, and decomposition from species mortality. The detritus contains N and P from these components, and part of these nutrients are continuously released into the dissolved nutrient pool in water. Furthermore, zooplankton also consumes detritus, creating a feedback loop in the food chain. Depending on the movement of fish or windy weather, some of the detritus may be sedimented, while some of settled matter may be resuspended.

The reference model did not account for the significant sedimentation and resuspension events, as the low manuring rate in this pond model limited the detritus levels within a narrow range. However, this factor has been included in the improved model (see further details in Section 3.3).

3.2. Workflow of improvements, validation and testing of the model

Considering the scope of the previously developed reference model, we concluded that it was designed for a relatively intensive fishpond management, with high stocking densities and feeding rates, and low manuring rates. Accordingly, the role of the natural food web was underrepresented. The step-by-step workflow of checking, improving, validating and testing of the reference model was as follows:

1. We analyzed the limitations of the reference model. We found that the main limitations (lack of measurements of solar radiation and humidity, as well as for plankton and detritus concentrations, etc.) were related to the functionalities of the natural food web. Therefore, they did not have a critical effect on the results (i.e., the biomass of the fish produced), but these features need special attention in the present work, in the knowledge of having more data on this segment of the model.
2. We considered the 5 available pilot pond experiments (controlled pilot experiments at MATE AKI HAKI in Table 1) and selected the

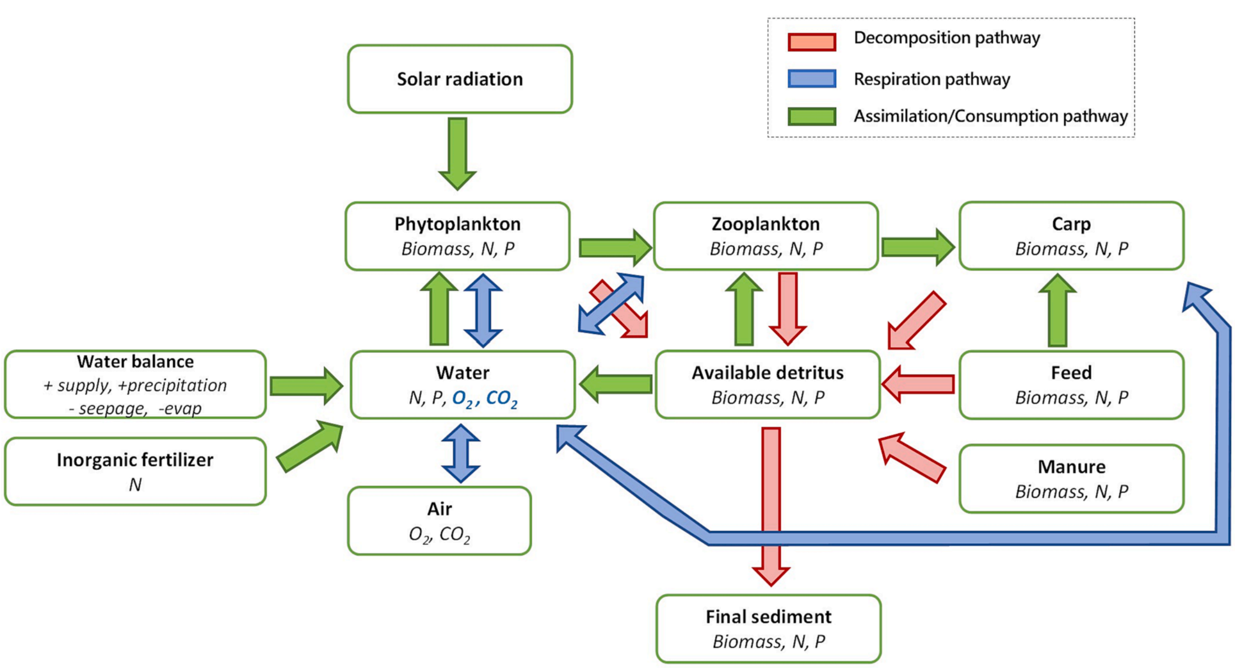

Fig. 1. Structural overview of the investigated models.

Table 2
Datasets used for additional validation.

Source of dataset	Controlled pilot experiments at MATE HAKI					
Dataset code	2014 64	2014 62	2014 61	2014 53	2014 63	2014 51
Type of model	Extended model[1]					
Role of model	To validate the improved model (Section 3.7)					
Date of stocking	04.04.2014					
Date of harvest	13.11.2014					
Stocking density, kg/ha	333	370	313	310	287	282
Feed input, t/ha/season	5.2[2]	5.0[2]	6.0[3]	5.4[3]	4.2[4]	5.2[4]
Manure[5], t/ha/season	2.5	2.6	2.5	2.4	2.8	2.5
Inorganic fertilizer, kg/ha/season	no	no	no	no	no	no
Distinguished eukaryotes + cyano-bacteria	yes	yes	yes	yes	yes	yes

[1] Extended model describes intensive fish production with a higher nutrient supply, with manure, and with optional inorganic fertilizer.
[2] Winter wheat for the first five days after stocking, afterwards feed, containing fish oil. The daily amount of the feed was calculated as 0.6–3.5% of the metabolic body weight (MBW%: $kg^{0.8}$). The feeding protocol during the weeks of the experiment was summarized as: 1–3: 0.6%, 4–5: 1%, 6–10: 1.5%, 11–13:2%, 14–16: 3%, 17–25: 3.5%, 26–32: 1.4%.
[3] Winter wheat for the first five days after stocking, afterwards feed, containing vegetable oil. The daily amount of the feed was calculated as 0.6–3.5% of the metabolic body weight (MBW%: $kg^{0.8}$). The feeding protocol during the weeks of the experiment was summarized as: 1–3: 0.6%, 4–5: 1%, 6–10: 1.5%, 11–13:2%, 14–16: 3%, 17–25: 3.5%, 26–32: 1.4%.
[4] Winter wheat for the first five days after stocking, afterward cereal. The daily amount of the feed was calculated as 0.6–3.5% of the metabolic body weight (MBW%: $kg^{0.8}$). The feeding protocol during the weeks of the experiment was summarized as: 1–3: 0.6%, 4–5: 1%, 6–10: 1.5%, 11–13:2%, 14–16: 3%, 17–25: 3.5%, 26–32: 1.4%.
[5] Cow manure.

most extensive, one average manured and one manured + fertilized experiments for the improvements to represent a wide range, while the remaining two cases were dedicated to validation. Accordingly, we have defined the following steps.

3. The parameters of the natural pond were calibrated for the case without feeding and manuring.
4. Once the modified parameters were fixed; the model was revised for the case with feeding and manuring.
5. After the changes were fixed, the model was extended for the additional use of inorganic fertilizer.
6. The model was validated for the reduced case with manuring, but without feeding.
7. The model was validated for the extended case with feeding and manuring.
8. Case CS7 was extended and parameterized to distinguish between eukaryotes and cyanobacteria in the phytoplankton group.
9. The model was tested by further validation considering sampling and measurement errors for a previously executed set of pilot pond experiments (Table 2).
10. The model was tested by scaling-up to a large production pond, with knowledge of stocking, feeding, and manuring for the approximate calculation of yield.

3.3. Stepwise reusability check and improvement of the reference model

We used the validated parameters and programs of the reference model as a starting point for the subsequent cases. According to the above strategy, we began with a basic reduced case study (2021CS6) and progressed through a stepwise approach for two additional cases (2022CS6 and 2022CS2).

The first pilot experiment (2021CS6, see Table 1), was a natural pond, representing low stocking density with no external nutrient input, typical of extensive fish production. To generate this simplified scenario, unnecessary state elements (i.e., feed, manure) and transition elements (i.e., feeding, manuring, uneaten feed, manure decomposition, etc.) were omitted. The unnecessary individual state and transition elements of PPS model can be switched off (by setting the 'Yes/No' existence of the given element to 'No', see the native simulation files of Lake_G_prot.graphml in the Mendeley database), or the user can set zero initial conditions and zero rate determining parameters in the respective user input. The comparison of measured and simulated values revealed the need for specific changes in the model. Calculation formulas and parameters for both the reference and modified models are available in the Calculation_formulas_and_parameters.xlsx of Mendeley database (Sharma et al., 2024) for detailed insights into these changes. The significant observations and actions were the following:

- Considering the actually available, site-specific solar radiation data, the program prototype, describing the phytoplankton related processes ("prot_t_phytop") was upgraded to use the actually available radiation data for the calculation;
- The initial phytoplankton, zooplankton, detritus, N and P concentrations of the model were estimated on the basis of the early-season measurements. Given the fluctuating and potentially inaccurate measurements, these initial conditions were fine-tuned through systematic simulations. It is important to highlight that these pond ecosystem models are extremely sensitive to the initial phytoplankton, zooplankton, detritus, N and P concentrations (Janse, 2005). This can be understood by comparing the proportions of these components in large volumes of water with the changes in the production and consumption of the given components during the season. Therefore, updating these initial conditions was crucial for accuracy.
- Natural pond food web models, lacking external feeding, exhibit high sensitivity to initial phytoplankton, zooplankton, and detritus levels due to a positive feedback loops. Increases in phytoplankton and detritus consumption by zooplankton, coupled with increases in detritus biomass (due to zooplankton decomposition) intensifies this feedback (Fath & Halnes, 2007). In addition, availability-driven consumption of phytoplankton and detritus by zooplankton generates extra detritus, increasing turbidity, that decreases phytoplankton synthesis and potentially reduces oxygen level, so adversely affects carp growth. Aligning initial values with calculated rates during testing was crucial to mitigate this issue. For example, the refined initial conditions for 2021 CS6 were established, resulting in initial phytoplankton concentration (C_{Phytop}) = 20 kg/ha, initial

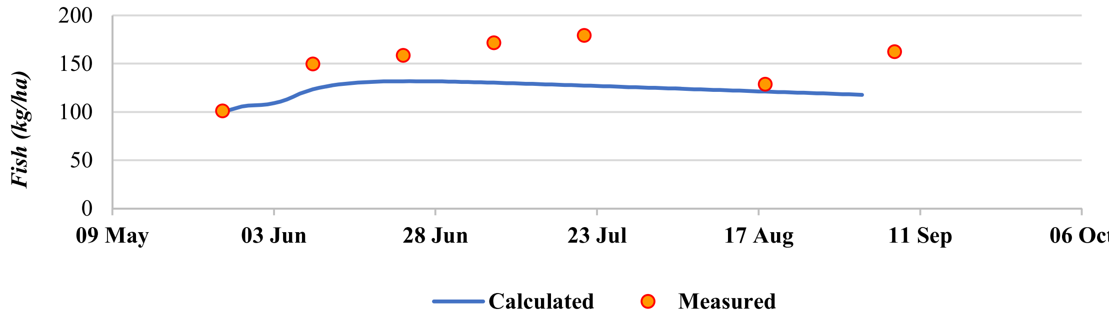

Fig. 2. Fish biomass in pilot experiment pond 2021CS6.

zooplankton concentration (C_{Zoop}) = 40 kg/ha, N concentration (C_N) = 2.06E-04 kg/m^3, and P concentration (C_P) = 4.20E-05 kg/m^3. (Note that the model's sensitivity to initial N, P and plankton concentrations was less noticeable when the external feed was dominant within the reference conditions.)

The measured and simulated data of the refined model are compared in Figs.S1.1-S1.7 of the supplementary material. Here, only two significant features are discussed in more detail. In Fig. 2 the measured and calculated fish biomass is compared. Both data sets showed that under given condition, after an initial period of weight gain, catabolic processes became dominant and fish lose weight. In addition, the measured values were higher than the simulated values because starving fish could find additional emergency food resources (e.g., zoo-benthic organisms) besides the zooplankton (Jurajda et al., 2016), which were not considered in the model.

In addition, the model did not follow the rapid increase in chlorophyll *a*-based measurements at the end of the season (see Fig. S1.5 in the supplementary file), highlighting the need for model refinement. This symptom also occurred in other pilot experiments, leading to an extension of the model with a hypothesis of the temperature-driven appearance of cyanobacteria (as presented in Section 3.5).

Fig. 3 shows the measured and simulated total inorganic nitrogen (TIN) concentrations. A decreasing trend can be observed in both the measured and simulated data, due to the lack of manure or inorganic fertilizer supply. However, this decrease did not limit the photosynthesis associated with the phytoplankton growth-.

The next pilot experiment (2022CS6, see Table 1) tested a semi-intensive management case of pond production. Fish were stocked at a rate of 200 kg/ha and were fed at an average feeding rate of 0.6 t/ha/season, together with a manuring rate of 5 t/ha/season. The manuring level in this case was significantly higher than in the reference model. Consequently, the first simulation showed a rapid increase in detritus concentration (which couldn't be validated due to the lack of available measurements). This amount of detritus had several effects on the modelled process, including:

- increased water turbidity, which inhibits photosynthesis and reducing phytoplankton biomass and oxygen production;
- increased detritus and decreased phytoplankton concentrations, resulting in zooplankton feeding on detritus with excess oxygen consumption;
- decreased oxygen levels prevented weight gain in fish and zooplankton, resulting in excess detritus;
- finally, this feedback had a significant negative impact on fish production.

The above malfunctions were identified through simulations and inspired the following improvements to the model. The prototype program "prot_t_detritus" was extended to consider permanent sedimentation of a certain fraction of detritus, together with the associated amount of N and P (see Fig. 1). Consequently, when the detritus concentration exceeded the availability limit (D_{min}), the sedimentation rate increased in proportion to the amount available. We formulated these processes in the following equations for total sediment (Eq.4, based on Svirezhev et al., 1984) and sedimented N and P loads (own formulation, Eqs.5 and 6).

$$DSed = -1{*}Sed{*}\max((D - D_{min}), 0){*}Area{*}DT \tag{4}$$

$$DSedN = \frac{DSed}{D{*}Area}{*}ND{*}Area{*}10000{*}Depth \tag{5}$$

$$DSedP = \frac{DSed}{D{*}Area}{*}PD{*}Area{*}10000{*}Depth \tag{6}$$

where:

Area is the surface area of the pond (ha), while 10,000 represents the conversion for m^2/ha;

D is the concentration of available (suspended) detritus (kg/ha);

Depth is the depth of the pond water (m);

D_{min} is the lower limit concentration of available (suspended) detritus (kg/ha);

DSed is the amount of sedimented detritus; (kg).

DSedN is the amount of sedimented nitrogen; (kg).

DSedP is the amount of sedimented phosphorus; (kg)

DT is the time step of the model (day);

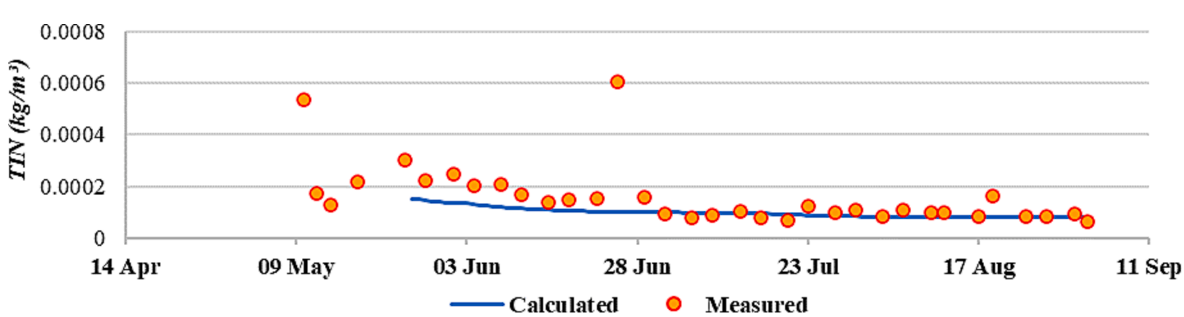

Fig. 3. The total inorganic nitrogen (TIN) concentration in the pilot experiment pond 2021CS6.

ND is the detritus-related nitrogen concentration (kg/m^3)
PD is the detritus-related phosphorus concentration (kg/m^3); and
Sed is the sedimentation rate coefficient (1/day).

Eq. (4) determines the amount of sedimented detritus, while Eqs. (5) and (6) calculate the related sedimentation of nitrogen and phosphorus.

During the cyclical stepwise improvements, the parameters Sed and D_{min} were determined to be 0.6 1/day and 132 kg/ha respectively.

As the previous reference model overlooked the details of the manure composition, in this work we applied the available actual composition and re-identified the decomposition coefficient in the prototype program of "prot_manure_decomp", formulated according to the following relationships:

$$DM = Alpha*M*Area*DT \tag{7}$$

$$DN = DM*Dry*Ncont \tag{8}$$

$$DP = DM*Dry*Pcont \tag{9}$$

where:
Alpha is the rate of decomposition, (Svirezhev et al., 1984), (1/day);
Area is the surface area of the pond (ha);
Ncont is the concentration of nitrogen in dry manure, based on lab measurement (kg/kg);
Pcont is the concentration of phosphorus in dry manure, based on lab measurement (kg/kg);
DM is the amount of the decomposed manure (kg);
DN is the amount of the decomposed nitrogen (kg);
DP is the amount of the decomposed phosphorus (kg);
DT is the time step (day);
M is the concentration of the manure (kg/ha); and
Dry is the dry matter content of the manure, based on lab measurements (kg/kg).

For actual experiments, simulations used Dry = 0.421 kg/kg, Ncont = 0.139 kg/kg, and Pcont = 0.0526 kg/kg. from laboratory measurements. After the stepwise identification, the parameter Alpha = 0.2 1/day was verified.

A complete set of calculation formulas, parameters, and corresponding program codes are available in the file "Calculation_formulas_and_parameters.xlsx" in the Mendeley database of Sharma et al., 2024, in the form of declarative code.

The measured and calculated data of the refined model are compared in Figs. S1.8-S1.14 of the supplementary material. Here, only two commented features are illustrated. Fig. 4 shows the measured and calculated fish biomass for the experimental pond 2022CS2. In view of the higher recorded value, it should be emphasized that the operators found a substantial amount of trash fish (134 kg) during harvest – along with the originally stocked carp. These fish had entered the pond through the inlet water supply. A similar situation was also observed in pilot experiments 2022CS2 and 2022CS3.

Fig. 5 shows the change in the calculated detritus concentration. It shows the increase after manure input, followed by sedimentation, converging to a limit value expressing an average amount of detritus, suspended in the pond.

In the third pilot experiment (2022CS2, see Table 1), in addition to the feed and manure, two doses of inorganic fertilizer (ammonium-nitrate) were added to the pond. Considering the rapid dissolution of ammonium-nitrate, its nitrogen content appears directly in the water, while the manure-originated nitrogen and phosphorus are considered separately and, part of them, can be trapped with the removed sediment as was described by Eqs. 4–6. The simulated and measured data of the refined model are shown in Figs. S1.15 – S1.21 of the supplementary material. The stepwise increase in the total inorganic nitrogen (TIN) concentration is shown in Fig. 6. In the model with concentrated parameters, the component dissolves rapidly (causing the sharp steps in the calculated concentrations).

Fig. 7 shows the change in phytoplankton concentration. Similar to the previous 2021CS6 experiment, the increase in phytoplankton concentration was only significant in the last period of the season. It is important to note that in this 2022CS2 case, the laboratory and sensor measurements were complementary. Zero values in the sensor measurements during the second period of the season indicate a faulty operation of the sensor (measurement limit problem), while the increased amount of phytoplankton at the end of the season was detected by the laboratory measurements. This anomaly is discussed in Section 3.5.

3.4. Validation of the improved model

After conducting the reusability tests and refining the reference model, based on previous case studies, we proceeded to validate the improved model using data from two additional pilot experiments (2021CS7 and 2022CS3, see Table 1). The structure and parameters, including all previous improvements, remained fixed for these experiments. The full set of calculation formulae and parameters are available in the "Calculation_formulas_and_parameters.xlsx" file in the Sharma et al., 2024 Mendeley database. The native model files for the simulations are also available there.

In the first validation experiment (2021CS7), we intensified the natural food chain involving phytoplankton, zooplankton, and carp by quadrupling the manuring pattern (i.e., 3 + 2 + 3 + 3 t/ha manure). The measured and calculated data are presented in Figs. S2.1 – S2.7 in the supplementary material. Fig. 8 shows the comparison of the measured and simulated biomass of carp.

Despite multiple manuring, the carp biomass showed only a modest increase of 70 kg. This observation was understood by comparing nitrogen, phosphorus, phytoplankton, and zooplankton concentration in the unmanured (2021CS6) with manured (2021CS7) pilot experiments, both without feed, which shows that:

- the increased nitrogen levels are only partially reflected in increased phytoplankton levels, because solar radiation limits phytoplankton production, while multiple manuring also increases water turbidity (Terziyski et al., 2007), reducing the available radiation;

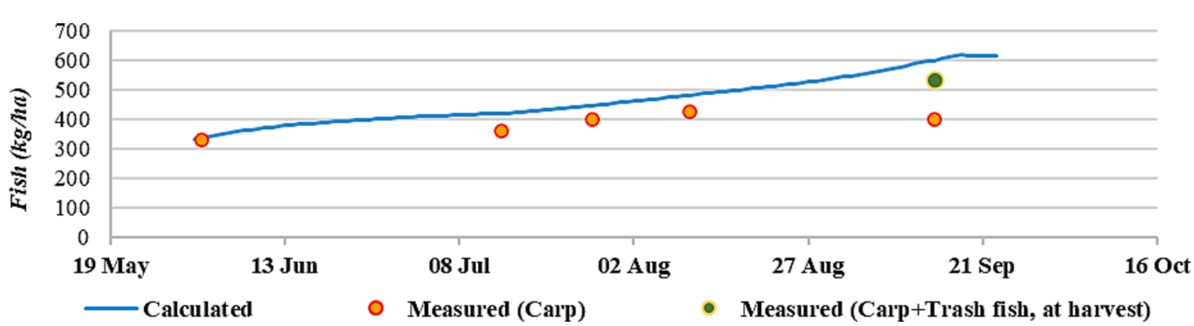

Fig. 4. Common carp biomass for the pilot experiment pond 2022CS6.

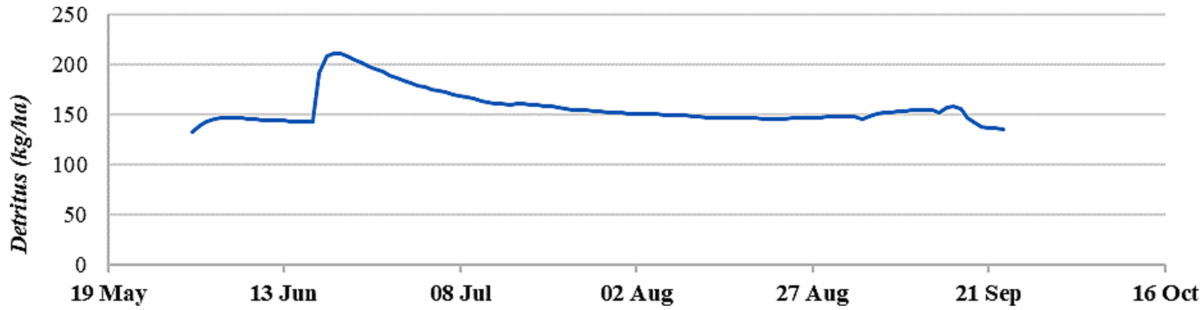

Fig. 5. Calculated detritus concentration in the pilot experiment pond 2022CS6.

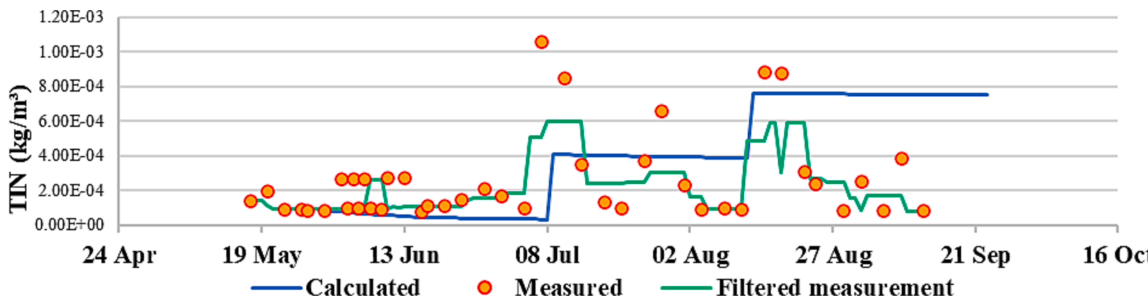

Fig. 6. Total inorganic nitrogen (TIN) concentration in case of stepwise fertilizer supply for the pilot experiment pond 2022CS2.

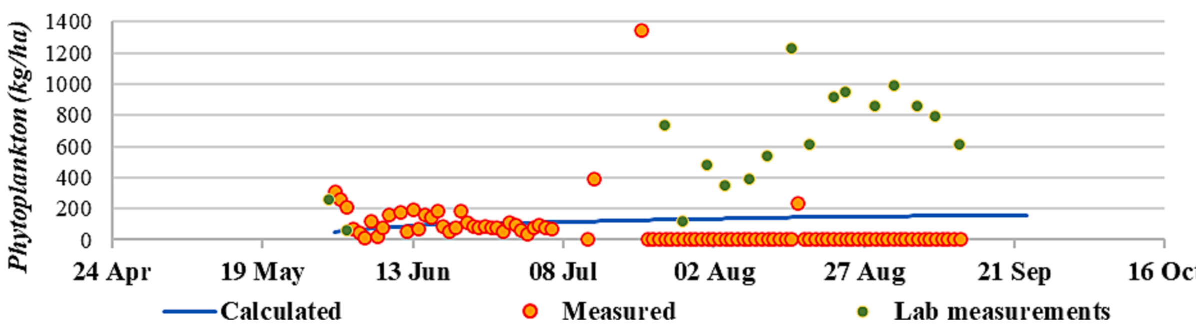

Fig. 7. Phytoplankton concentration for the pilot experiment pond 2022CS2.

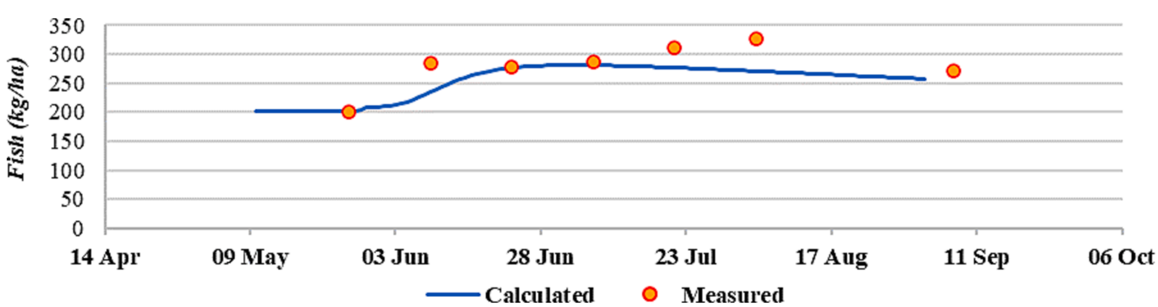

Fig. 8. Carp biomass for pilot experiment pond validation case 2021CS7.

- the increased phytoplankton did not lead to a proportional increase in zooplankton, prompting a further investigation into their relationship;
- moreover, a measured, but un-modelled increase in phytoplankton concentration was observed towards the end of the season (see Fig. 9).

These issues are discussed in more detail in Section 3.5, where an advanced model explores the hypothetical differentiation of eukaryotes and cyanobacteria, taking into account the reduced effect of cyanobacterial consumption on zooplankton.

In the second experimental pond (2022CS3) used for validation, an initial manure application of 4 t/ha was made before stocking, 1 t/ha after stocking, followed by two additional applications of 2 t/ha.

Detailed results of this validation case are presented in Figs. S2.8 – S2.13 in the supplementary material. Fig. 10 shows the data comparing measured and simulated individual carp biomass. Notably, the higher measured values are consistent with the operator's observation that a significant amount of trash fish (192 kg/ha), accompanied the originally stocked carp in the fishponds.

It is important to note that in this case (2022CS3, Fig. 11), as well as in other pilot experiments during 2022, a peak in measured dissolved

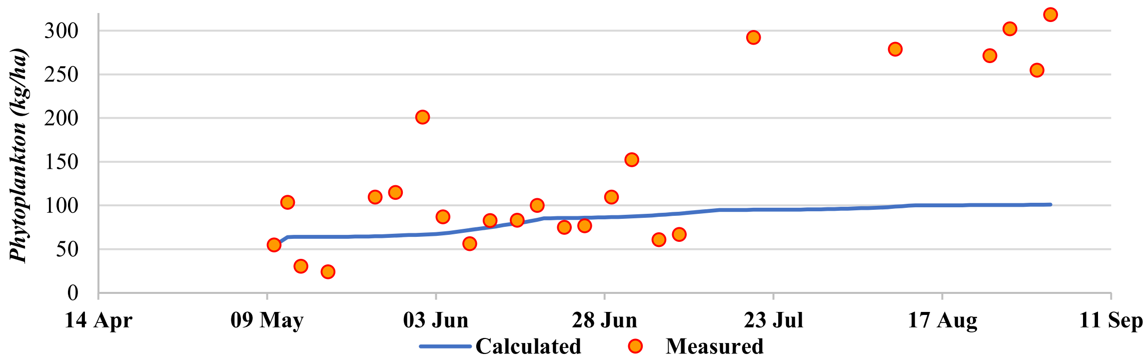

Fig. 9. Phytoplankton concentration for pilot experiment pond validation case 2021CS7.

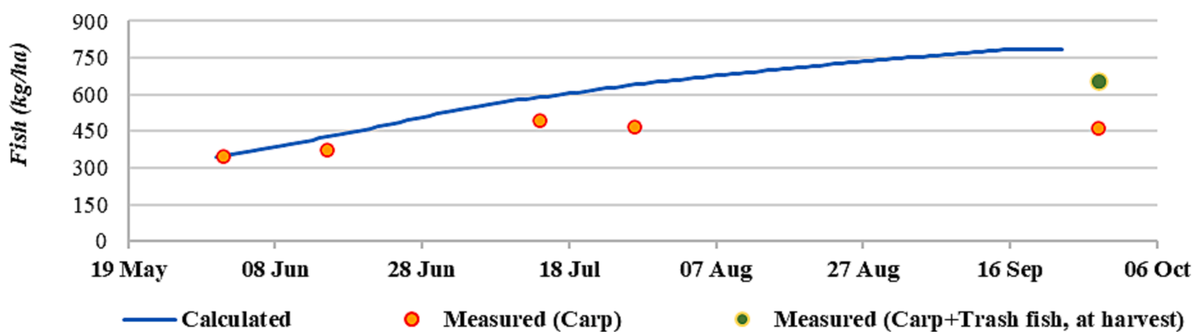

Fig. 10. Carp biomass in pilot experiment pond validation case 2022CS3.

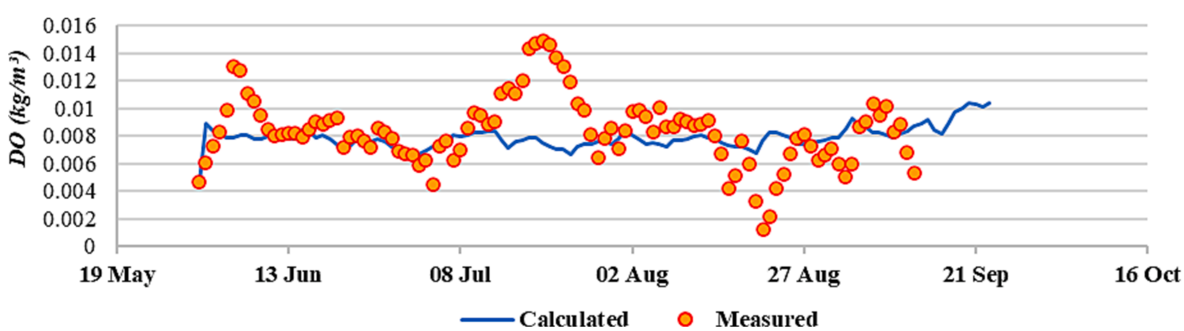

Fig. 11. Dissolved oxygen (DO) concentration in pilot experiment pond validation case 2022CS3.

oxygen (DO) was observed in the middle of the season. This phenomenon can be attributed to temporary activation of the paddlewheel aerators during warm days. This was not included in the model as no recorded data were available for inclusion.

3.5. *Extended model for distinction between eukaryotic and cyanobacterial groups of phytoplankton*

As described in sections 3.3 and 3.4, the refined and validated model still shows a notable anomaly that remains unexplained within the investigated model. In particular, the increase in phytoplankton concentration during the last period of the season defies current interpretation. To address this, possible explanations and model improvements are proposed in line with the existing literature (e.g., Potužák et al., 2007; Jeppesen et al., 2011):

- The phytoplankton group can be divided into the groups of eukaryotes and cyanobacteria;
- At the beginning of the season, eukaryotes dominate and coexist with a minimal presence of cyanobacteria;
- The growth kinetic coefficients of eukaryotes and cyanobacteria are different;
- Cyanobacteria have a higher minimum, optimum, and maximum growth temperatures (T_{min} = 22 °C, T_{opt} = 28 °C, T_{max} = 36 °C) compared to eukaryotes (T_{min} = 9 °C, T_{opt} = 24 °C, T_{max} = 34 °C) (Lürling et al., 2013);
- The presence of cyanobacteria reduces the appetite of zooplankton (Fulton and Paerl, 1987). In the availability-driven food web model, the zooplankton consumption coefficient is higher for eukaryotes than for cyanobacteria.

Considering these hypotheses, the process model underwent the following refinements (calculation formulas and parameters are available in the file “Calculation_formulas_and_parameters.xlsx” file in Sharma et al., 2024):

- The state element “s_phytop” was replaced by the state elements “s_cyano” and “s_eukar”;
- The initial concentration of cyanobacteria (s_cyano) was estimated to be 0.05 % (based on estimates from fishpond experts) of the total initial phytoplankton biomass;

- The transition element (“t_phytop”) and prototype (“prot_t_phytop”) were replaced by the modified code of the transition elements (“t_cyano” and “t eukar”) and prototypes (“prot_t_cyano” and “prot_t eukar”), respectively;
- Instead of the maximum production rate coefficient for phytoplankton (i.e., 20 1/day), through systematic identification the values for eukaryotes and cyanobacteria were adjusted to, 20 1/day and 3 1/day, respectively;
- In the prototype program “prot_t_zoop”, the calculation of the competitive consumption rate kinetics was defined as follows
- for eukaryotes (E) with a maximum rate of 1.6 1/day with an availability ratio of E/(E + C + D), where E, C, and D refer to eukaryotes, cyanobacteria and detritus, respectively,
- for cyanobacteria (C) with a maximum rate of 0.2 1/day with an availability ratio of C/(E + C + D),
- for detritus (D) with a maximum rate of 0.5 1/day with an availability ratio of D/(E + C + D).

The detailed simulation, focusing on the amount of eukaryotes and cyanobacteria produced, is presented in the following parts of the supplementary material: in Figs. S3.1 – S3.2 for case 2021CS6, Figs. S3.3 – S3.4 for case 2022CS6, Figs. S3.5 – S3.6 for case 2022CS2, Figs. S3.7 – S3.8 for case 2021CS7 and in Figs. S3.9 – S3.10 for case 2022CS3.

A notable aspect of the results obtained for the case of 2022CS3 is illustrated in Figs. 12-14. The comparison of simulated and measured phytoplankton concentrations in Fig. 12 shows that separate consideration of eukaryotes and cyanobacteria leads to improved results.

Fig. 13 shows the measured and simulated zooplankton concentrations for the extended model based on the hypothetical scenario. In Fig. 14 the calculated concentrations of eukaryotes and cyanobacteria show that the increasing presence of cyanobacteria is likely influenced by warmer temperatures along the season. Growth of cyanobacteria was initiated slowly by the first warm peak in mid-June, and was accelerated by the second one. The positive feedback stopped at the beginning of cooling down in mid-August.

Investigation of this model extension indicated its potential for improved performance, although validation through subsequent trials is still required. In addition, this case study highlights the ability of Programmable Process Structures to easily generate an extended model.

3.6. Overview of the practical results with the refined model

Detailed simulation results are summarized in sections S1-3 of the Supplementary material, while comprehensive input files, including structural descriptions, input data, parameters, and local program prototypes for PPS, are available in the Mendeley database (Sharma et al., 2024), together with the corresponding simulation output files. Table 3 summarizes the practical results of the cases studied.

As the fully extensive carp production rate remains very limited (see 2021CS6), it can be concluded that extensive use of an ecosystem-based food web requires additional manuring. However, the efficiency of the phytoplankton → zooplankton → carp chain seems to be limited by solar radiation, so the excess amount of manure does not significantly increase fish production.

In feed-based production, the contribution of natural nutrients is influenced by the additional manure or fertilizer. However, solar radiation limits their effect and excess amounts of manure or fertilizer increase fish biomass degressively, while increasing nutrient emissions.

The error analysis of the improved model, presented in Table 3, shows the comparison between simulated and measured values.

The improved model showed an overall satisfactory performance. In validation ponds 2021CS7 and 2022CS3, the model tended to over- or underestimate some variables. NRMSE values range from 1.5 % to 58.4 %.

It is to be noted that large NRMSE values contain both sampling & measurement and model errors. The correct evaluation of model error and further development of the model require the knowledge of sampling and measurement errors.

3.7. Consideration of measurement errors by a second series of validating pilot measurements

The significant discrepancies between measured and calculated characteristics are common in real-world agricultural and aquacultural processes (Pandey & Hardaker, 1995). The erroneous estimations can be partly explained by the different erroneous methods of sampling and measurement used during the experiments. Mainly, spatial and temporal sampling emerged as critical factors for such discrepancies (Bellocchi et al., 2010). Accurately capturing processes with distributed parameters within a given area relies on appropriate spatial sampling (Wang et al., 2012). Temporal sampling, on the other hand, takes into account the varying measured values at a specific spatial point over time (Gómez-Dans et al., 2022). Unfortunately, representative sampling is often hampered by the high costs and significant manpower required. While measurement errors directly from the equipment are typically small, errors from the sampling process itself can be significant, especially if inadequate equipment is used or limits are missed (Espig et al., 2020). The above discussion is visually presented in Fig. 15, emphasizing the position and significance of the calculated error measures between the measured and simulated data.

As the conclusions of modeling studies are highly based on measured values, it is crucial to assess the quality of the measurements (Tarkkonenand & Vehkalahti, 2005). Furthermore, to distinguish errors in measurement, as well as in model mapping and simulation, it is essential to understand errors that occur during sampling and measurement. To make a first, modest step toward this understanding, the model was tested by further validation with the consideration of sampling and measurement errors for a formerly executed set of pilot pond experiments (described in Section 2.5), where two parallel experiments were made for three pilot pond case studies. These experiments were designed to test three types of feed, so we estimated the respective ‘feed conversion rate’ parameters for carp’s diet.

First, we calculated the standard deviation (SD), characterizing the differences in the measurements of the 3*2 parallel experiments defined

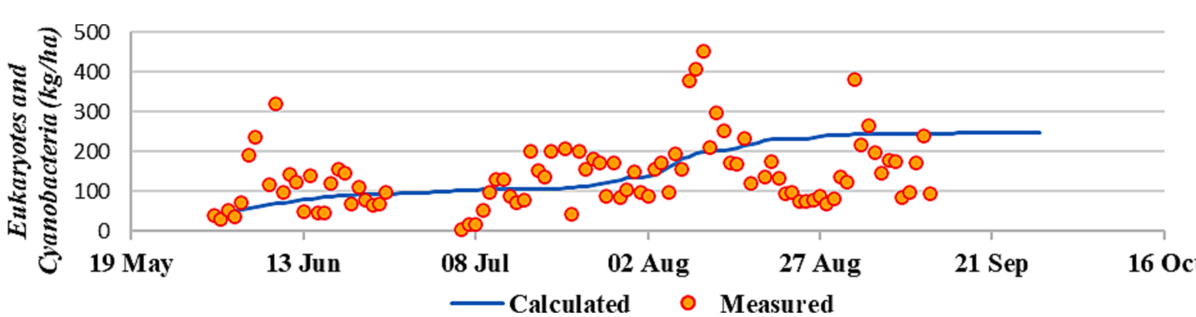

Fig. 12. The sum of eukaryotes and cyanobacterial concentrations for pilot pond experiment 2022CS3.

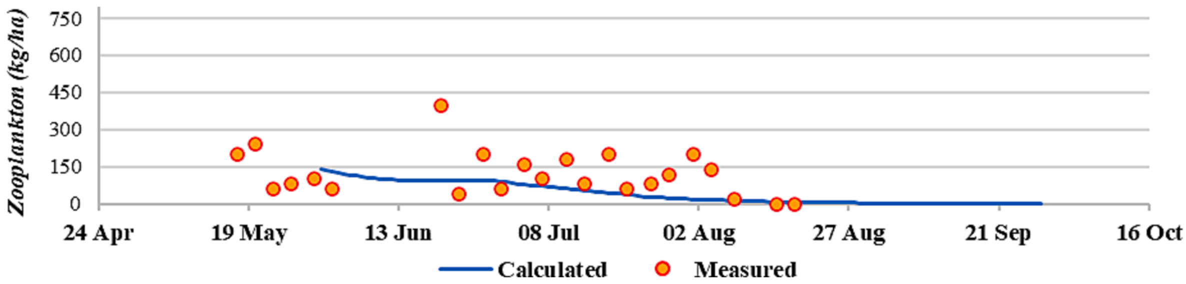

Fig. 13. Zooplankton concentrations for pilot experiment pond case 2022CS3.

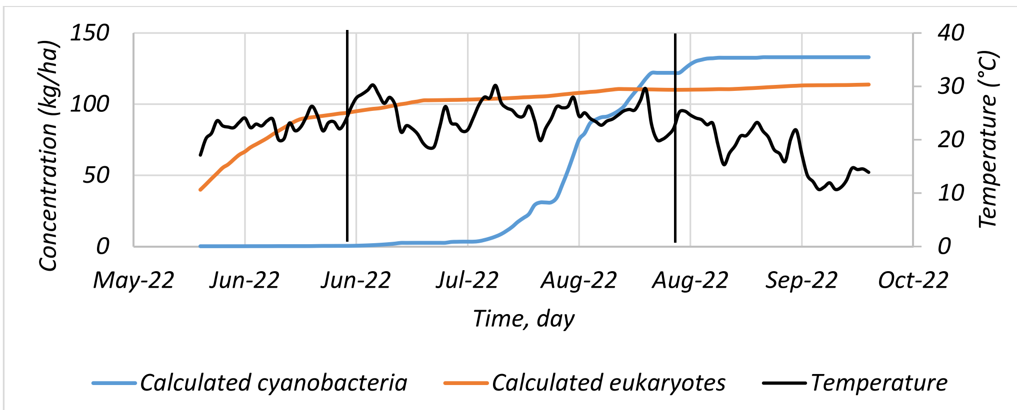

Fig. 14. Eukaryotes and cyanobacteria concentrations in pilot experiment pond 2022CS3, illustrating also the temperature.

Table 3
Summarized overview of case studies.

Characteristics	Pilot experiment				
	2021CS6	2022CS6	2022CS2	2021CS7	2022CS3
Feeding t/ha/season	no	0.6	0.7	no	0.9
Manuring, t/ha/season	no	5 (4 + 1)	5 (4 + 1)	11 (3 + 2 + 3 + 3)	9 (4 + 1 + 2 + 2)
Fertilizing, kg/ha/season	no	no	200 (100 + 100, ammonium nitrate)	no	no
Stocking density, kg/ha	101	200	200	201	200
Harvested fish biomass, measured, kg/ha	162	536	638	270	655
Harvested fish biomass, calculated, kg/ha	117	615	638	257	782
Harvested fish biomass, with hypothetic cyanobacteria consideration, calculated, kg/ha	117	569	639	271	829
NRMSE, carp biomass, %				21	34.6
NRMSE, zooplankton biomass, %				38	1.5
NRMSE, phytoplankton biomass, %				22	11.6
NRMSE, DO, %				24	19
NRMSE, TIN, %				28.6	34
NRMSE, PO4-P, %				26.8	58.4

by Eq.3 in Section 2.6. The calculated results are summarized in Table 4.

Next, we ran the simulation model to calculate the 3*2 experiments and calculated the error between the simulated and the averaged measured values by means of the error measure (RMSE) defined by Eq.1 in Section 2.6. This measure represents the deviation in the unit of the original measurement value. These values can be seen in Table 5.

SD (in the context of parallel measurements) and RMSE (in the context of the comparing of measured and simulated data) provide different evaluations of their respective sets of observations (Meyer, 2012). In our study, given the very limited and not comprehensive measurements available, we used these two, different error measures to demonstrate the possible causal relationship between these two phenomena, namely (i) the difference between the parallel measurements and (ii) the difference between the measured and simulated data. It should be emphasized that the input data for the simulation are the same for all (actually both) parallel experiments, while the reason for the differences between the parallel measurements is hidden (invisible) to the model. Nevertheless, the error (e.g., RMSE) between measured and simulated data includes the error of the parallel experiments. This overlapping interaction makes model identification and validation more difficult. At first look, increasing the number of parallel pilot experiments may seem to be a solution, but only with huge labor demand and cost to consider. However, there is a deeper problem, namely the inherent differences between the pilot ponds, studies in parallel. Unfortunately, we do not have any information on the “history” of these pilot ponds, e.g. on the volume, the chemical and biological composition of the residual water and re-suspendable sediment, on the pond-specific seepage, etc. These conditions can lead to increased fish mortality and

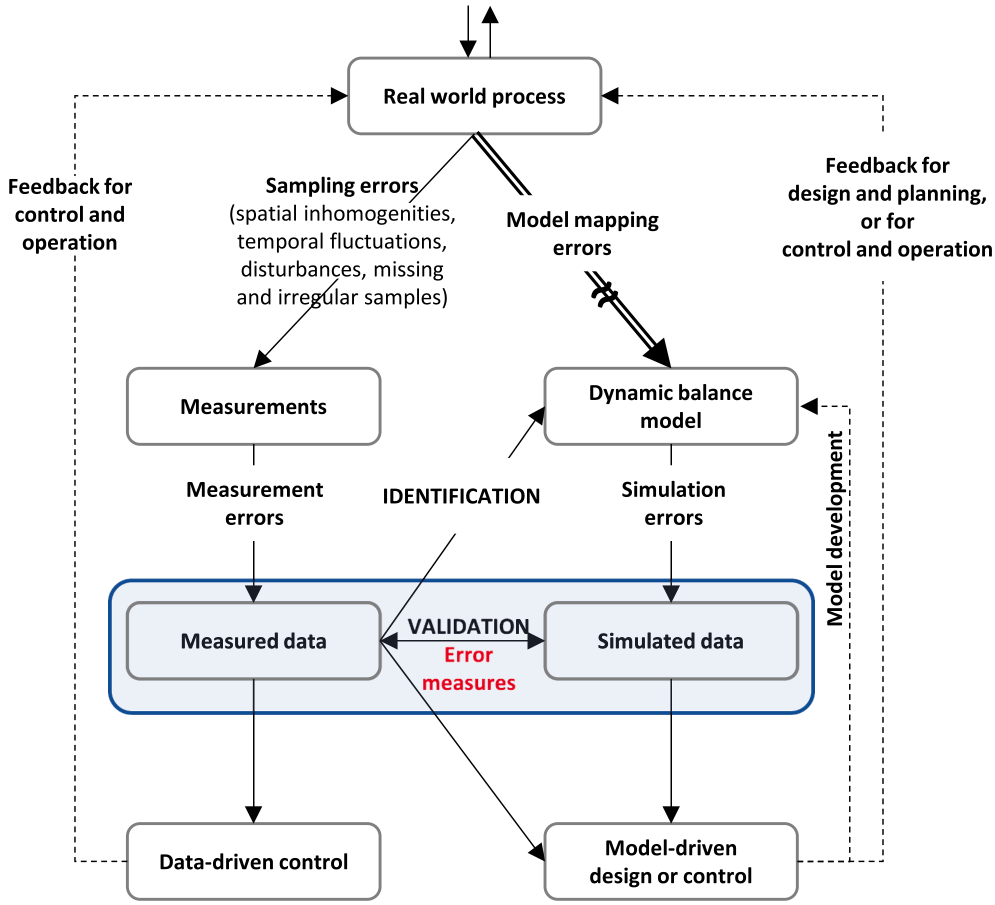

Fig. 15. Position and meaning of errors.

Table 4
Ranges of SD values regarding the parallel experiments (in the same unit as the respective parameter).

Parameter	SD					
	Experiments 64 and 62		Experiments 61 and 53		Experiments 63 and 51	
	Min	Max	Min	Max	Min	Max
Dissolved oxygen, mg/l	0.01	1.53	0.00	1.76	0.00	2.71
Nitrogen, mg/dm^3	8.3E-05	2.96E-01	1.67E-04	2.79E-01	2.75E-04	1.74E-01
Phosphorus, mg/dm^3	5.2E-09	8.75E-04	5.51E-09	1.78E-03	5.72E-09	7.25E-03
Carp, g/piece	0.17	71.88	0.01	141.00	0.22	125.25
Zooplankton, kg/ha	7.34	798.75	17.24	2377.89	45.65	108.79
Phytoplankton, kg/ha	2.72	413.61	0.00	344.18	5.44	379.14
Detritus, kg/ha	0.00	76.66	0.00	111.80	0.00	202.74

Table 5
RMSE values for the measured/calculated data for the 3*2 parallel experiments.

Parameter	ID of experiments						Range of RMSE (in the same unit as the respective observation value)	
	63	51	64	62	61	53	min	max
Dissolved oxygen, mg/l	2.17	2.13	1.73	1.94	2.32	2.60	1.73	2.60
Nitrogen, mg/dm^3	0.20	0.20	0.20	0.21	0.12	0.12	0.12	0.21
Phosphorus, mg/dm^3	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Carp, g/piece	40.01	44.81	150.20	116.38	64.03	40.37	40.01	150.20
Zooplankton, kg/ha	509.43	531.67	1 150.19	1 137.17	60.58	82.92	60.58	1150.19
Phytoplankton, kg/ha	258.20	256.07	236.19	237.97	262.60	257.57	237.53	261.89
Detritus, kg/ha	85.72	92.12	127.98	127.33	111.08	113.40	85.72	127.98

other hidden side effects on the food web and chemical composition. It is worth emphasizing again, that the dynamic simulation of the given non-linear system with positive feedback loops is very sensitive to the initial conditions, if these conditions do not fit appropriately (i.e., contradict) it can affect the calculated functionalities.

In line with the above discussion, in case of the food web elements (carp, zooplankton, phytoplankton and detritus), both model calculations and parallel measurements show a larger error. For example, in

Table 6
Simulated fish biomass in the up-scaled model.

Characteristics	Pond ID and Year				
	BIX 2013	BIII 2014	BII 2015	BVIII 2014	BVIII 2015
Area, ha	24	123	141	58	58
Feeding, t/ha/season	1.538	1.053	1.304	1.422	1.573
Manuring, t/ha/season	8.3	4.1	2.8	3.4	2.6
Fertilizing, kg/ha/season	no	no	no	no	no
Stocking density, kg/ha	370	338	333	248	247
Harvested fish biomass, measured, kg/ha	1360	1080	975	1266	1159
Harvested fish biomass, calculated, kg/ha	1336	1138	1103	1027	1067
Relative difference, $\frac{Calculated - Measured}{Measured}$*100, %	−1.76	5.37	13.13	−18.88	−7.94

case of zooplankton, model predicts with a range of 60.58 – 1150.19 kg/ha RMSE. Similarly in the measurement side, experiments are characterized by a larger range (7.34 – 2377.89 kg/ha).

The RMSE values for carp biomass showed slightly larger differences between experiments ID64 vs. ID62 and ID61 vs. ID53. To investigate this, we mined the incomplete historical data from 2014 which revealed that variations in carp biomass data (ID61 vs. ID53) were due to higher fish mortality in a particular pond. The deeper reason may be related to the unknown history of the pond.

Considering the former discussion about missing pond history, future work should include numerous consecutive years and comprehensive modeling of historical processes throughout the entire calendar year. Also the simulations would highly benefit from data from a continuous logbook recording in addition to the sensor and laboratory measurements (extended by the data for the realistic initial and boundary conditions).

3.8. Test of model-based scaling-up

When using the model for simulation-based scaling-up, it is necessary to actualize case-specific data such as:

- the number and biomass or individual mass of fish stocked;
- the composition and quantity of feed and the feeding strategy applied;
- the composition and quantity of manure (and/or inorganic fertilizer) and the dosing strategy applied; and
- possible estimates of the initial conditions such as zooplankton, phytoplankton, detritus, and water quality representing concentrations,

while all other model parameters and program prototypes can be used according to the previously validated reusable model.

To demonstrate this scaling-up test, we used the available limited part of the previously collected data from a production fish farm in Biharugra (Hungary) (46°55′32.6″N 21°33′04.4″ E) for the years 2014 – 2016. The data were extracted from specific pond sites with designated codes: BIX2013, BIII2014, BII2015, BVIII2014, and BVIII2015. Data on pond area, stocked and harvested fish biomass and feed and manure applied were used from this source. Similar meteorological conditions as in the pilot experiments were assumed. The missing data and parameters were taken from the analyzed pilot case 2022CS3.

The production-based and calculated results for the harvested carp biomass are summarized in Table 6. The input and output files of the model are collected in the "Simulations_Scaling_up" folder of the Mendeley database (Sharma et al., 2024).

In addition to the estimated approximate fish production, the upscaled simulation can provide estimates of the other characteristics and environmental impacts of the production pond (e.g., for the required water supply, nutrient emission, O_2 production, and for CO_2 sequestration, etc.).

4. Limitations and strengths of the results

In parts discussed above, the proposed workflow for the reusability check-based improvement of a formerly developed reference model combines additional knowledge, coming from two sources. One of them contains new data, from complementary experiments, carried out within and beyond the scope of the original model. The other source is the reasoning about the possible reductions and extensions of the underlying conceptual model. In an effective combination, these two types of model refinement could be carried out in parallel. However, in this work the modeling had to use the results of previous experiments, controlled by actual practical purposes, but not by reusability. This situation leads to two types of limitations (and weaknesses), resulting from the experimental and modeling efforts.

One part of the experimental limitations is caused by the lack of information about the possible measurement errors. Most of these errors is associated with sampling, which is one of the most critical issues in agricultural and aquacultural measurements. First of all, these systems cannot be characterized by concentrated parameters, and even in certain cases by limited mixing amongst some compartments, on the one side. However, the cost and manpower needed for multiple, parallel measurements have to be taken into consideration, on the other. This contradiction can be resolved only by a compromise, where computational modeling helps to make rational decisions. For example, in our case, parallel measurement for initial conditions of some critical components (such as phytoplankton, zooplankton) in various points of the site would support the model development, highly. It should be followed by frequently repeated samplings from one to two points for shorter periods during the experiments. Of course, the success of this strategy would need the interaction with modeling before and during the experimentation.

The non-sampling related experimental limitations of the actual measurements resulted.

- from missing series of measurements (e.g., for suspended detritus in the 2021 and 2022 pilot experiments);
- from lack of measurements, especially in the first part of the experiments (i.e., before manuring and stocking);
- from faulty sensor data (e.g., in 2022CS2 for phytoplankton concentration);
- from unregistered, missing information about some "hidden" phenomena such as switching on of aerator on warm days, or trash fish input.

The limitations of the applied method for improving reusability come mainly from the fact that it requires the systematic incorporation of experts' reasoning. In contrast to the easily automated sensor data-driven machine learning model improvements (utilized for the solution of well-defined local control problems), the 'a priori' model-based reusability improvement (utilized for longer-term design and planning) requires the evolution of conceptual models and the description of new functionalities that is not easily automated.

To address these limitations in the future, the continuous collaboration of field experts, sampling and measurement staff, and model developers should be realized, from experimental design through interim discussions to evaluation of results.

The strength of the study lies in its systematic model reuse strategy, achieved through stepwise reusability checks and improvements, integrating existing model knowledge with additional measurements. The adaptable framework of Programmable Process Structures enabled automatic model generation and simulation, enhancing reusability testing and improvement. This experience will guide the future development of ecosystem-oriented agricultural and aquacultural models, particularly in pilot experimentation-based and process model-assisted scale-up of production units.

5. Conclusions

The reusability of a process model involves a broader application beyond its original scope, including reduced models of specific subprocesses and extended models with new features. Model-based scaling up, which is fundamental to design and planning, is a special case of reusability. Models, validated by small-scale (e.g., pilot) experiments can approximately predict the characteristics of larger production units. Furthermore, planning uses the knowledge, gained from pilot tests to represent essential process characteristics over a larger space and longer time horizon. Both the scaling-up of pilot experiments and the design and planning of complex process systems require first principles-based dynamic simulation models.

Complex agro-environmental processes combine natural ecosystems with human activities, making comprehensive experimentation expensive and laborious. The construction of reusable models from pilot experiments is therefore essential.

This work presents a stepwise workflow for reusability testing and improvement by re-implementing and refining a mechanistic biophysical fishpond model, constructed using the framework of Programmable Process Structures.

The model architecture of the Programmable Process Structures helped to solve the reusability check-based improvement with the available limited amount of relatively erroneous data. This was supported by the unified state and transition elements and their unified model prototypes, which follow the building blocks of the real-world processes being modelled. Accordingly, all of the parameter and functionality (i.e., program code) changes require only local intervention without reasoning about holistic side-effects in the applied mathematical construct or computational tool. Moreover, the unnecessary model elements and prototype programs can be turned off, while new model elements and prototype programs can be conveniently added by local intervention. The price of this flexibility is that the improvement process based on the reusability check cannot be automated, especially in the case of the actually limited data availability. However, the lessons learned from the expert-driven execution of such a difficult task inspire further methodological developments towards more and more automated solutions.

Starting from the reference model, a series of case studies represented a wide range of scenarios from the simple, extensive use of the natural food web to the intensive cases with feeding, manuring, and fertilizing in freshwater pond aquaculture. In the first phase, the model was tested and refined on a subset of pilot-scale experiments covering a wider range of characteristics. In the second phase, the refined model was validated using measured data from additional pilot case studies. This approach resulted in an improved model for fishponds following a wider range of production practices. The systematic improvement process also revealed hidden processes, and a hypothesis-based section distinguishing between eukaryotic and cyanobacterial phytoplankton groups was incorporated into the improved model.

After analyzing the overall simulation errors, the model was tested through a further validation, taking into account sampling and measurement errors from a previously conducted set of pilot pond experiments, where two parallel experiments were performed for three pilot pond case studies.

Finally, the refined model showed satisfactory results when applied to the scale-up of large-scale production ponds with very limited input data availability.

The reusability checking procedure highlighted the critical need for well-designed and careful experiments, given the inherent uncertainty associated with sampling and measurement errors in pilot experiments. It underlines the importance of involving model experts and preliminary modeling to establish the sampling and measurement strategy. The plan should include a limited but necessary number of locally distributed and parallel samples to gain accurate knowledge of measurement errors. In addition, as food web models are very sensitive to the initial conditions, consideration of pond history and a comprehensive set of initial concentrations are extremely important.

CRediT authorship contribution statement

P. Sharma: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **G. Gyalog:** Data curation, Investigation, Methodology, Project administration, Resources, Supervision, Validation. **L. Berzi-Nagy:** Data curation, Investigation, Methodology, Validation. **F. Tóth:** Data curation, Investigation. **Z. Nagy:** Data curation, Investigation. **B. Halasi-Kovács:** Funding acquisition, Resources. **D.L. Fazekas:** Data curation, Investigation. **D. Mezőszentgyörgyi:** Funding acquisition, Resources. **B. Csukas:** Formal analysis, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **M. Varga:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Software, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research has been performed in the scope of the European Union's Horizon 2020 research project EATFISH (grant no. 956697). The research work of Monika Varga was supported by the Bolyai János Research Scholarship of the Hungarian Academy of Sciences, Hungary. Data used in this article were generated in projects funded by the Hungarian Ministry of Agriculture (project no. HaGF/4/2021 and HaGF/11/2022, ARRAINA project N288925-EU FP7).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2024.108664.

References

Arnaud, E., Laporte, M.-A., Kim, S., Aubert, C., Leonelli, S., Miro, B., Cooper, L., Jaiswal, P., Kruseman, G., Shrestha, R., Buttigieg, P.L., Mungall, C.J., Pietragalla, J., Agbona, A., Muliro, J., Detras, J., Hualla, V., Rathore, A., Das, R.R., Dieng, I., Bauchet, G., Menda, N., Pommier, C., Shaw, F., Lyon, D., Mwanzia, L., Juarez, H., Bonaiuti, E., Chiputwa, B., Obileye, O., Auzoux, S., Yeumo, E.D., Mueller, L.A.,

Silverstein, K., Lafargue, A., Antezana, E., Devare, M., King, B., 2020. The Ontologies Community of Practice: A CGIAR Initiative for Big Data in Agrifood Systems. Patterns 1 (7), 100105.
Aubin, J., Baizeau, V., Jaeger, C., Roucaute, M., Gamito, S., 2021. Modeling trophic webs in freshwater fishpond systems using Ecopath: towards better polyculture management. Aquac. Environ. Interact. 13, 311–322. https://doi.org/10.3354/AEI00406.
Bellocchi, G., Rivington, M., Donatelli, M., Matthews, K., 2010. Validation of biophysical models: issues and methodologies. A Review. Agronomy for Sustainable Development 30 (1), 109–130. https://doi.org/10.1051/AGRO/2009001.
Bolte, J., Nath, S., Ernst, D., 2000. Development of decision support tools for aquaculture: the POND experience. Aquac. Eng. 23 (1–3), 103–119. https://doi.org/10.1016/S0144-8609(00)00049-2.
Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. Geosci. Model Dev. 7 (3), 1247–1250. https://doi.org/10.5194/GMD-7-1247-2014.
Chary, K., Brigolin, D., Callier, M.D., 2022. Farm-scale models in fish aquaculture – An overview of methods and applications. Rev. Aquac. 14 (4), 2122–2157. https://doi.org/10.1111/RAQ.12695.
Christensen, V., Walters, C. J., Pauly, D., & Forest. D. (2008). Ecopath with Ecosim 6 User Guide, November. https://ecopath.org/downloads/.
Cohen, A.R., Chen, G., Berger, E.M., Warrier, S., Lan, G., Grubert, E., Dellaert, F., Chen, Y., 2022. Dynamically Controlled Environment Agriculture: Integrating Machine Learning and Mechanistic and Physiological Models for Sustainable Food Cultivation. ACS ES and T Engineering 2 (1), 3–19. https://doi.org/10.1021/acsestengg.1c00269.
Aquaculture Advisory Council. (2021). The provision of ecosystem services by European aquaculture. 10.1038/461472a.
de Kok, J.L., Engelen, G., Maes, J., 2015. Reusability of model components for environmental simulation – Case studies for integrated coastal zone management. Environ. Model. Softw. 68, 42–54. https://doi.org/10.1016/J.ENVSOFT.2015.02.001.
Donatelli, M., Cerrani, I., Fanchini, D., Fumagalli, D., & Rizzoli, A.-E. (2012). Enhancing Model Reuse via Component-Centered Modeling Frameworks: the Vision and Example Realizations. In R. Seppelt, A. A. Voinov, S. Lange, & D. Bankamp (Eds.), International Congress on Environmental Modelling and Software. Managing Resources of a Limited Planet: Pathways and Visions under Uncertainty, Sixth Biennial Meeting, (pp. 1185–1192). iEMSs 2012. https://scholarsarchive.byu.edu/iemssconference/2012/Stream-B/140.
Dong, S.-L., Dong, Y.-W., Cao, L., Verreth, J., Olsen, Y., Liu, W.-J., Fang, Q.-Z., Zhou, Y.-G., Li, L.i., Li, J.-y., Mu, Y.-T., Sorgeloos, P., 2022. Optimization of aquaculture sustainability through ecological intensification in China. Rev. Aquac. 14 (3), 1249–1259.
Durden, J.M., Luo, J.Y., Alexander, H., Flanagan, A.M., Grossmann, L., 2017. Integrating “Big Data” into Aquatic Ecology: Challenges and Opportunities. Limnol. Oceanogr. Bull. 26 (4), 101–108. https://doi.org/10.1002/LOB.10213.
Espig, M., Finlay-Smits, S.C., Meenken, E.D., Wheeler, D.M., Sharifi, M., 2020. Uncertainty in and around biophysical modelling: insights from interdisciplinary research on agricultural digitalization. Royal Society Open. Science 7 (12), 201511.
European Commission, Directorate-General for Research and Innovation, Baker, L., Cristea, I., Errington, T. (2020). Reproducibility of scientific results in the EU: scoping report, (W,Lusoli,editor) Publications Office. https://data.europa.eu/doi/10.2777/341654.
Fath, B.D., Halnes, G., 2007. Cyclic energy pathways in ecological foodwebs. Ecol. Model. 208 (1), 17–24. https://doi.org/10.1016/J.ECOLMODEL.2007.04.020.
Fulton III, R.S., Paerl, H.W., 1987. Effects of colonial morphology on zooplankton utilization of algal resources during blue-green algal (Microcystis aeruginosa) blooms 1. Limnol. Oceanogr. 32 (3), 634–644. https://doi.org/10.4319/lo.1987.32.3.0634.
Gómez-Dans, J.L., Lewis, P.E., Yin, F., Asare, K., Lamptey, P., Aidoo, K.K.Y., Maccarthy, D.S., Ma, H., Wu, Q., Addi, M., Aboagye-Ntow, S., Doe, C.E., Alhassan, R., Kankam-Boadu, I., Huang, J., Li, X., 2022. Location, biophysical and agronomic parameters for croplands in northern Ghana. Earth Syst. Sci. Data 14 (12), 5387–5410. https://doi.org/10.5194/ESSD-14-5387-2022.
Gyalog, G., Tovar, J. P. C., & Békefi, E. (2022). Freshwater Aquaculture Development in EU and Latin-America: Insight on Production Trends and Resource Endowments. Sustainability 2022, Vol. 14, Page 6443, 14(11), 6443. 10.3390/SU14116443.
Hagiwara, H., Mitsch, W.J., 1994. Ecosystem modeling of a multi-species integrated aquaculture pond in South China. Ecol. Model. 72 (1–2), 41–73. https://doi.org/10.1016/0304-3800(94)90145-7.
Hayat Khan, M., 2011. Standard Deviation. In: Lovric, M. (Ed.), International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_535.
Holzworth, D.P., Huth, N.I., de Voil, P.G., 2010. Simplifying environmental model reuse. Environ. Model. Softw. 25 (2), 269–275. https://doi.org/10.1016/j.envsoft.2008.10.018.
Horváth, L., Tamás, G., Seagrave, C. (Eds.), 2002. Carp and Pond Fish Culture. Wiley.
Janse, J.H., 2005. Model studies on the eutrophication of shallow lakes and ditches. Wageningen University Vol. 24, Issue 1.
Jeppesen, E., Kronvang, B., Olesen, J.E., Audet, J., Søndergaard, M., Hoffmann, C.C., Andersen, H.E., Lauridsen, T.L., Liboriussen, L., Larsen, S.E., Beklioglu, M., Meerhoff, M., Özen, A., Özkan, K., 2011. Climate change effects on nitrogen loading from cultivated catchments in Europe: Implications for nitrogen retention, ecological state of lakes and adaptation. Hydrobiologia 663 (1), 1–21. https://doi.org/10.1007/s10750-010-0547-6.
Jorgensen, S.E., 1995. Development of models for fishpond management. Acta Hydrobiologica 37 (Suppl. 1).
Jurajda, P., Adámek, Z., Roche, K., Mrkvová, M., Štarhová, D., Prášek, V., Zukal, J., 2016. Carp feeding activity and habitat utilisation in relation to supplementary feeding in a semi-intensive aquaculture pond. Aquac. Int. 24 (6), 1627–1640. https://doi.org/10.1007/S10499-016-0061-6/TABLES/2.
Kochba, M., Diab, S., Avnimelech, Y., 1994. Modeling of nitrogen transformation in intensively aerated fish ponds. Aquaculture 120 (1–2), 95–104. https://doi.org/10.1016/0044-8486(94)90225-9.
Liu, J., Peng, C., Dang, Q., Apps, M., Jiang, H., 2002. A component object model strategy for reusing ecosystem models. Comput. Electron. Agric. 35 (1), 17–33. https://doi.org/10.1016/S0168-1699(02)00067-4.
Lokers, R., Knapen, R., Janssen, S., van Randen, Y., Jansen, J., 2016. Analysis of Big Data technologies for use in agro-environmental science. Environ. Model. Softw. 84, 494–504. https://doi.org/10.1016/J.ENVSOFT.2016.07.017.
Lürling, M., Eshetu, F., Faassen, E.J., Kosten, S., Huszar, V.L.M., 2013. Comparison of cyanobacterial and green algal growth rates at different temperatures. Freshw. Biol. 58 (3), 552–559. https://doi.org/10.1111/J.1365-2427.2012.02866.X.
Mathisen, B. M., Haro, P., Hanssen, B., Björk, S., & Walderhaug, S. (2016). Decision Support Systems in Fisheries and Aquaculture: A systematic review. 10.48550/arxiv.1611.08374.
Meyer, T. (2012). Root Mean Square Error Compared to, and Contrasted with, Standard Deviation. Technical note. Surveying and Land Information Science, Volume 72, Number 3, September 2012, pp. 107-108(2), Retrieved on Nov 16, 2023 from https://www.ingentaconnect.com/content/aags/salis/2012/00000072/00000003/art00002?crawler=true.
Ogawa, H., Mitsch, W.J., 1979. Modeling of power plant impacts on fish populations. Environ. Manag. 3 (4), 321–330.
Palásti, P., Kiss, M., Gulyás, Á., & Kerepeczki, É. (2020). Expert knowledge and perceptions about the ecosystem services and natural values of Hungarian fishpond systems. Water (Switzerland), 12(8). 10.3390/W12082144.
Palma, R., Reznik, T., Esbrí, M., Charvat, K., Mazurek, C. (2016). An INSPIRE-Based Vocabulary for the Publication of Agricultural Linked Data. In: Tamma, V., Dragoni, M., Gonçalves, R., Ławrynowicz, A. (eds) Ontology Engineering. OWLED 2015. Lecture Notes in Computer Science, vol 9557. Springer, Cham. 10.1007/978-3-319-33245-1_13.
Pandey, S., Hardaker, J.B., 1995. The role of modelling in the quest for sustainable farming systems. Agr. Syst. 47 (4), 439–450. https://doi.org/10.1016/0308-521X(95)92109-J.
Plesser, H.E., 2018. Reproducibility vs. Replicability: A brief history of a confused terminology. Frontiers. Neuroinformatics 11, 76. https://doi.org/10.3389/FNINF.2017.00076/BIBTEX.
Popp, J., Békefi, E., Duleba, S., Oláh, J., 2019. Multifunctionality of pond fish farms in the opinion of the farm managers: the case of Hungary. Rev. Aquac. 11 (3), 830–847. https://doi.org/10.1111/RAQ.12260.
Potužák, J., Hůda, J., Pechar, L., 2007. Changes in fish production effectivity in eutrophic fishponds - Impact of zooplankton structure. Aquac. Int. 15 (3–4), 201–210. https://doi.org/10.1007/s10499-007-9085-2.
Prinsloo, J.F., Schoonbee, H.J., 1984. Observations on fish growth in polyculture during late summer and autumn in fish ponds at the Umtata Dam Fish Research Centre, Transkei. Part I: The use of pig manure with and without pelleted fish feed. Accessed at Jan 10, 2023 from Water SA 10 (1), 15–23. https://journals.co.za/doi/pdf/10.10520/AJA03784738_1272.
Repar, L., Šestak, M., & Copot, D. (2023). Towards Trusted Data Sharing and Exchange in Agro-Food Supply Chains: Design Principles for Agricultural Data Spaces. Sustainability 2023, Vol. 15, Page 13746, 15(18), 13746. 10.3390/SU151813746.
Rosen, M. A., Carolan, M. S., Luyckx, M., & Reins, L. (2022). The Future of Farming: The (Non)-Sense of Big Data Predictive Tools for Sustainable EU Agriculture. Sustainability 2022, Vol. 14, Page 12968, 14(20), 12968. 10.3390/SU142012968.
Schaap, D.M.A., Novellino, A., Fichaut, M., Manzella, G.M.R., 2022. Data management infrastructures and their practices in Europe. Ocean Science Data: Collection, Management, Networking and Services 131–193. https://doi.org/10.1016/B978-0-12-823427-3.00007-4.
Sharma, Priya; Gyalog, Gergő; Berzi-Nagy, Laszló; Tóth, Flórián; Nagy, Zoltan; Halasi-Kovács, Béla; Fazekas, Dorottya; Mezőszentgyörgyi, Dávid; Csukás, Béla; Varga, Mónika (2024), “Data for reusability check-based refinement of a biophysical fishpond model”, Mendeley Data, V2, doi: 10.17632/837f4mvpmb.2.
Stella, T., Francone, C., Yamaç, S.S., Ceotto, E., Pagani, V., Pilu, R., Confalonieri, R., 2015. Reimplementation and reuse of the Canegro model: From sugarcane to giant reed. Comput. Electron. Agric. 113, 193–202. https://doi.org/10.1016/J.COMPAG.2015.02.009.
Svirezhev, Y.M., Krysanova, V.P., Voinov, A.A., 1984. Mathematical modelling of a fish pond ecosystem. Ecol. Model. 21 (4), 315–337. https://doi.org/10.1016/0304-3800(84)90066-8.
Talari, G., Cummins, E., McNamara, C., O’Brien, J., 2022. State of the art review of Big Data and web-based Decision Support Systems (DSS) for food safety risk assessment with respect to climate change. Trends Food Sci. Technol. 126, 192–204. https://doi.org/10.1016/J.TIFS.2021.08.032.
Tarkkonenand, L., Vehkalahti, K., 2005. Measurement errors in multivariate measurement scales. J Multivar Anal 96 (2005), 172–189. https://doi.org/10.1016/j.jmva.2004.09.007.
Terziyski, D., Grozev, G., Kalchev, R., Stoeva, A., 2007. Effect of organic fertilizer on plankton primary productivity in fish ponds. Aquac. Int. 15 (3–4), 181–190. https://doi.org/10.1007/S10499-007-9086-1/FIGURES/3.

Varga, M., 2022. Developing Plant Models of Reduced Complexity by Chemical Process Engineering Way of Thinking. Hungarian Journal of Industry and Chemistry 50 (1), 57–65. https://doi.org/10.33927/HJIC-2022-09.

Varga, M., Csukas, B., 2017. Generation of extensible ecosystem models from a network structure and from locally executable programs. Ecol. Model. 364, 25–41. https://doi.org/10.1016/J.ECOLMODEL.2017.09.014.

Varga, M., Berzi-Nagy, L., Csukas, B., Gyalog, G., 2020. Long-term dynamic simulation of environmental impacts on ecosystem-based pond aquaculture. Environ. Model. Softw. 134, 104755 https://doi.org/10.1016/j.envsoft.2020.104755.

Varga, M., Csukás, B., 2022. Foundations of Programmable Process Structures for the unified modeling and simulation of agricultural and aquacultural systems. Information Processes in Agriculture, in Press,. https://doi.org/10.1016/j.inpa.2022.10.001.

Varga, M., Prokop, A., Csukas, B., 2017. Biosystem models, generated from a complex rule/reaction/influence network and from two functionality prototypes. Biosystems 152, 24–43. https://doi.org/10.1016/J.BIOSYSTEMS.2016.12.005.

Varga, M., Gyalog, G., Raso, J., Kucska, B., Csukas, B., 2022. Programmable Process Structures of Unified Elements for Model-Based Planning and Operation of Complex Agri-environmental Processes. Springer Optimization and Its Applications 184, 223–249. https://doi.org/10.1007/978-3-030-84152-2_11/FIGURES/20.

Varga, M., Csukas, B., Khanal, S., Bakshi, B.R., 2023. Lessons from the biosphere for the anthroposphere: Analysis of recycling structures of conservational measures. Resour. Conserv. Recycl. 192, 106919 https://doi.org/10.1016/J.RESCONREC.2023.106919.

Wang, J.F., Stein, A., Gao, B.B., Ge, Y., 2012. A review of spatial sampling. Spatial Statistics 2 (1), 1–14. https://doi.org/10.1016/J.SPASTA.2012.08.001.