Assessment of robustness of machine learning-assisted modelling approach to describe growth kinetics of microorganisms using Monte Carlo simulation

FATIH TARLAK – ÖZGÜN YÜCEL

Summary

Understanding the growth behaviour of microorganisms is crucial for various fields such as microbiology, food safety and biotechnology. Traditional modelling approaches face challenges in accurately capturing the dynamic and complex nature of microbial growth especially when high variation is seen. In contrast, machine learning techniques offer a promising avenue for creating more accurate and adaptable models. This study aimed to develop a new modelling method, machine learning-assisted modelling approach, and compare the robustness of machine learning-assisted and traditional modelling approaches in describing microbial growth behaviour, employing Monte Carlo simulation. The research involved subjecting both machine learning-assisted and traditional modelling approaches to 10, 50 and 500 trials. The results showed that the machine learning approach led to more robust results than the traditional modelling approach providing higher adjusted coefficient of determination (R^2_{adj}) value than 0.919 and lower root mean square error (*RMSE*) value than 0.319. These findings suggest that the machine learning-assisted modelling approach, particularly with Gaussian process regression, has the potential to serve as a highly reliable prediction method for describing the growth behaviour of microorganisms in frames of predictive food microbiology. The study provides insights into practical application of machine learning in enhancing our understanding and predictive capabilities of microbial growth dynamics.

Keywords

machine learning regression; model robustness; growth parameters; Monte Carlo simulation

Predictive microbiology represents a specialized scientific discipline employing mathematical models and computational tools to predict how microorganisms grow and survive in various environments, in particular in food [1]. This approach allows researchers, food producers and regulatory agencies to pro-actively evaluate potential hazards associated with microbial presence and activity, facilitating informed decision-making on critical aspects such as food safety, quality maintenance and shelf-life determination.

By advancing our understanding of microbial behaviour, predictive microbiology supports the development of food safety protocols and enhances quality assurance practices [2]. Utilizing mathematical models and computational tools, this field provides the food industry with data-driven decision-making capabilities, ensuring consumer safety and satisfaction. The implementation of predictive microbiology has significantly supported the industry's ability to predict and manage shelf-life of food products, leading to improvement of food safety and quality standards. Researchers gained useful insights into the complex dynamics of microbial growth in food products through predictive models [3]. These models are instrumental in assessing risks associated with microbial contamination and informing strategies for food preservation, storage and distribution [4]. Furthermore, application of predictive microbiology has enabled formulation of effective methodologies to extend the shelf-life of food products, thereby reducing food waste while maintaining consumer health and satisfaction [5].

Fatih Tarlak, Department of Bioengineering, Faculty of Engineering, Gebze Technical University, Cumhuriyet Street 2254, 41400 Gebze, Kocaeli, Turkey.
Özgün Yücel, Department of Chemical Engineering, Faculty of Engineering, Gebze Technical University, Cumhuriyet Street 2254, 41400 Gebze, Kocaeli, Turkey.

Correspondence author:
Fatih Tarlak, e-mail: ftarlak@gtu.edu.tr

A two-step modelling approach, often referred to as a two-stage or dual-stage modelling approach, involves using two separate modelling techniques or stages to analyse a complex problem or a dataset [6–8]. Each stage serves a specific purpose and builds upon the results or insights from the previous stage. The one-step modelling approach, also known as a single-stage modelling approach, involves using a single comprehensive model to analyse a dataset [9, 10]. Instead of breaking down the analysis into multiple stages, as done in the two-step approach, all relevant variables and relationships are incorporated into a single model. This approach has its own set of advantages, depending on the nature of the problem and the goals of the analysis. In this study, one-step modelling approach was followed to compare the robustness of machine learning-assisted and traditional modelling approaches in describing microbial growth behaviour [11].

In the field of predictive food microbiology, machine learning techniques have gained significant interest and attention [12]. By employing data-driven algorithms, these approaches construct models capable of predicting microbial growth, spoilage and safety in food items. Machine learning methods hold the promise of modelling complicated connections among diverse factors that influence microbial behaviour, resulting in enhanced precision in predictions and sustained food safety measures [13].

Yildirim-Yalcin et al. [14] aimed to create a predictive tool using machine learning-based regression models to predict the growth of total mesophilic bacteria in spinach. They compared these models with conventional ones like the modified Gompertz, Baranyi and Huang models, evaluating them based on statistical measures. The results revealed that machine learning models achieved higher predictive accuracy, with a minimum coefficient of determination (R^2) of 0.960 and a maximum root mean square error (*RMSE*) of 0.154. This accuracy suggested their potential as credible alternatives to traditional methods in predicting the growth of total mesophilic bacteria.

Yücel and Tarlak [15] aimed to create machine learning-based regression methods: decision tree regression (DTR), generalized additive model regression (GAMR) and random forest regression (RFR) to predict bacterial populations in beef. They used a dataset from the ComBase database (Tasmania Institute of Agriculture, Tasmania, Australia), containing 2 654 data points for *Listeria monocytogenes*, *Escherichia coli*, and *Pseudomonas* spp. in beef. Factors like temperature, salt content, water activity and acidity, crucial in assessing microbial growth or survival, were key predictors. The models' parameters were finely tuned using nested cross-validation. The evaluation, based on R^2 and *RMSE*, revealed strong predictive capabilities for all methods, with R^2 values ranging from 0.931 to 0.949 and *RMSE* values ranging from 0.597 to 0.692.

Monte Carlo simulation is a computational technique used to understand the impact of uncertainty and variability in various systems or models. It involves running numerous simulations by assigning random values to uncertain variables within a model to analyse the range of possible outcomes. By performing these iterations, Monte Carlo simulations offer insights into the probability distribution of different results, helping in decision-making and risk assessment across fields like finance, engineering or physics [16, 17].

The objective and novelty of this study was to demonstrate that the two-stage method, combining machine learning with the Baranyi model, offers enhanced robustness in predicting microbial growth behaviour. The effectiveness of this approach was validated using Monte Carlo simulations aiming to conduct a comprehensive comparison of robustness of machine learning-assisted and traditional modelling approaches. Both modelling methods were subjected to varying conditions of variability and uncertainty through 10, 50 and 500 normal random simulations. The performance differences were evaluated and highlighted, ultimately determining which approach provides more reliable and accurate predictions of microbial behaviour.

MATERIALS AND METHODS

Microbiological analysis

Growth data of *Pseudomonas* spp. were gathered from a previously published study on the growth of the microorganism on button mushrooms at various constant temperatures (4, 12, 20 and 28 °C) [18]. The experimental process involved obtaining white button mushrooms from a specific source, ensuring the mushrooms were undamaged and immediately transporting them to a laboratory at 4 °C, representing industrial storage practices. The mushrooms were placed in trays without additional packaging material and subjected to controlled temperature and humidity conditions. The chambers maintained precise temperature and humidity levels while recording data every 15 min using a data logger. For non-constant temperature conditions, dynamic temperature changes were introduced. Microbio-

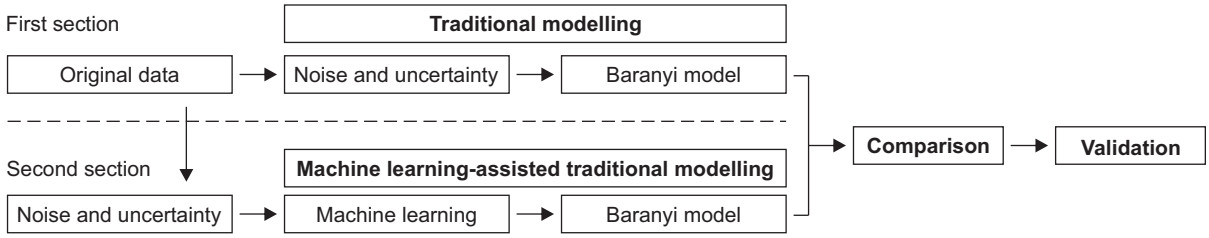

Fig 1. Flow chart of the steps followed in the current study.

logical data were collected to simulate storage, delivery and retail conditions. Sampling involved weighing and homogenizing mushrooms, followed by dilutions and enumeration of *Pseudomonas* spp. using specific culture media and incubation conditions. The study analysed a total of 297 mushroom trays across three independent batches, evaluating *Pseudomonas* spp. counts at various time points up to 240 h for each storage temperature. The results were reported as average logarithm of colony forming units per gram with standard error for each sampling point.

Modelling

The study's modelling part was divided into two sections (Fig. 1). The first section employed a traditional modelling approach, while the second section used a machine learning-assisted modelling approach. The robustness of both approaches was evaluated using a Monte Carlo simulation, where small perturbations were randomly added to the measurements to mimic measurement errors and microbial variability. For the traditional method, Baranyi model was applied to the new dataset generated by Monte Carlo simulation. The proposed method involved first applying machine learning and then using the resulting predictions as input to Baranyi model. In other words, the new dataset was fitted using the machine learning approach and these results were subsequently fed into the Baranyi model. The original data were gathered from a previously published study on the growth of *Pseudomonas* spp. on button mushrooms at various constant temperatures (4, 12, 20 and 28 °C) [18]. Monte Carlo simulation was then performed for 10, 50 and 500 simulated data with standard deviation of ± 0.25 log CFU·g^{-1} in Matlab 8.3.0.532 (R2014a) software (MathWorks, Natick, Massachusetts, USA) (Fig. 2).

Static model

In comparing machine learning and traditional modelling techniques, there are notable differences in their approaches to predicting

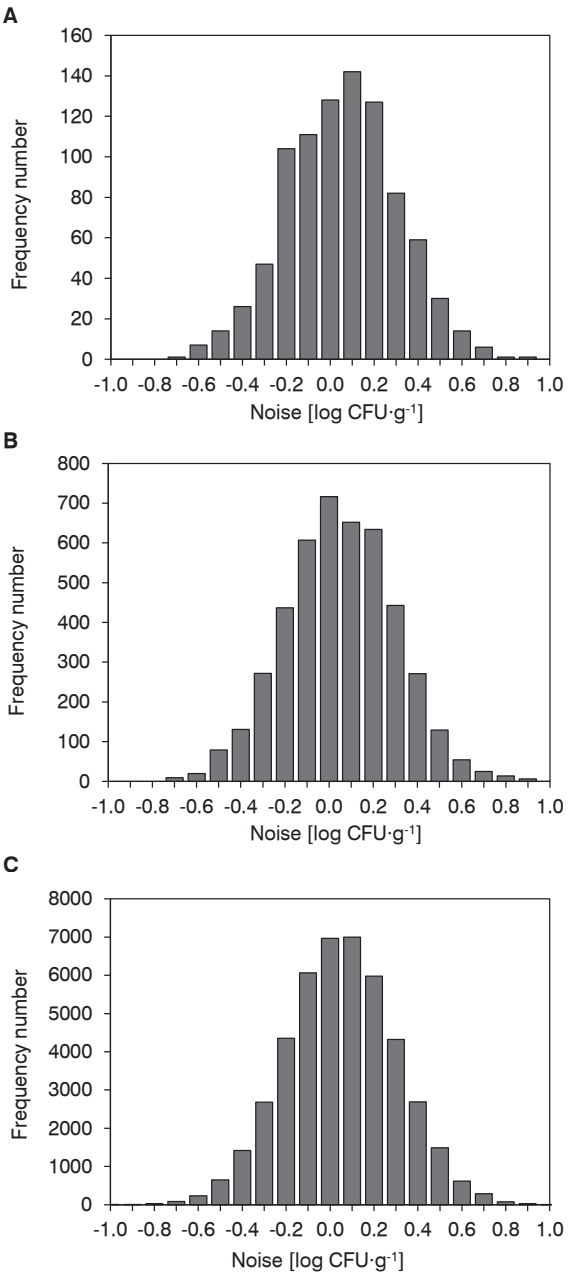

Fig. 2. Histograms of Monte Carlo simulations for uncertainty.

A – 10 times simulations, B – 50 times simulations, C – 500 times simulations.

the behaviour of microorganisms and estimating the shelf life of food products [19]. Traditional methods rely on established mathematical models and computational techniques, while machine learning utilizes algorithms to identify patterns and generate predictions based on data [20]. The strength of machine learning lies in its ability to handle complex and non-linear relationships effectively, allowing for the analysis of extensive and diverse datasets. However, it often requires substantial amounts of data for training and may lack the interpretability inherent in traditional models. Ttraditional modelling techniques offer more straightforward interpretation, often grounded in well-understood biological and chemical principles. They may be better suited for scenarios with limited data or when interpretability is important. There are numerous machine learning regression methods, but Gaussian process regression (GPR) method is the most suitable regression method to describe the behaviour of microorganisms. Therefore, GPR method was employed as a machine learning regression method.

GPR is characterized by its flexibility, complete probabilistic nature and non-parametric Bayesian approach. The method involves generating infinite-dimensional normal distributions, ultimately resulting in a multivariate Gaussian distribution. GPR builds objective functions by evaluating the distances between estimated output probability density functions derived from a provided dataset. A notable feature of GPR is its ability to maintain a high level of certainty in areas where no samples exist, even if they are significantly distant from the training data.

Baranyi model is the most commonly used primary model. It is represented by Eq. 1 and Eq. 2 [21]:

$$y(t) = y_0 + \mu_{max}F(t) - \ln\left(1 + \frac{e^{\mu_{max}F(t)} - 1}{e^{(y_{max}-y_0)}}\right) \quad (1)$$

$$F(t) = t + \frac{1}{\nu}\ln\left(e^{-\nu t} + e^{-\mu_{max}\lambda} - e^{(-\nu t-\mu_{max}\lambda)}\right) \quad (2)$$

where t is time (in hours), $y(t)$ is count of microorganisms at time t, y_0 is initial count of microorganisms, y_{max} – maximum count of microorganisms (counts are expressed as natural logarithm of colony forming units per gram), μ_{max} is maximum specific growth rate of microorganisms (expressed as natural logarithm of colony forming units per hour), λ is lag phase duration (in hours), ν is the rate of increase of the limiting substrate, assumed to be equal to μ_{max} [21].

Secondary models are used to describe the impact of various environmental conditions on the parameters of main models. These include temperature, pH, water activity (a_w), oxygen availability and concentration of additives. Secondary models are utilized after the growth data had been fitted to a primary model. The Ratkowsky model is widely used to explain the link between temperature and maximum specific growth rate [22] (Eq. 3):

$$\sqrt{\mu_{max}} = b_1(T - T_0) \quad (3)$$

where T is storage temperature (in degrees Celsius), T_0 is the theoretical lowest temperature at which microbial growth is observable (in degrees Celsius), μ_{max} is the maximum specific bacterial growth rate (expressed as unit per hour), b_1 is the regression coefficient.

Additionally, lag phase duration (λ) is defined as a function of μ_{max} with respect to temperature using Eq. 4 [23]:

$$\lambda = \frac{b_2}{\mu_{max}(T)} \quad (4)$$

where b_2 is regression coefficient, $\mu_{max}(T)$ is the function of temperature, which leads λ to be defined as a function of storage temperature.

Determination of each parameter involved utilization of the `NonLinearModel` command in Matlab 8.3.0.532 (R2014a) software. This command employs the Levenberg Marquardt algorithm. In the non-linear regression process, selecting appropriate initial values is a crucial step to accurately estimate the parameters. For the parameters y_0 and y_{max}, the minimum and maximum counts of bacterial populations across the entire temperature range were chosen as starting values, respectively. In the case of parameters b_1, b_2 and T_0, random selection of starting points might result in the estimation of parameters converging to local optima. To address this, the starting points for these parameters were deliberately chosen using the `ga` command from the Global Optimization Toolbox in Matlab. Following a successful iteration process in the non-linear regression procedure, global optimum values for the parameters were ultimately obtained.

Dynamic model

The prediction of the microbial growth under non-isothermal conditions was carried out using Eq. 5 and Eq. 6 as described by Baranyi and Roberts [21]:

$$\frac{dy}{dt} = \frac{1}{1 + e^{-Q(t)}}\mu_{max}\big(T(t)\big)\big[1 - e^{(y(t)-y_{max})}\big] \quad (5)$$

$$\frac{dQ}{dt} = \mu_{max}\big(T(t)\big) \quad (6)$$

The initial conditions to solve Eq. 5 and Eq. 6 are given in Eq. 7 and Eq. 8.

$$y(0) = y_0 \tag{7}$$

$$Q(0) = \ln q_0 \tag{8}$$

where y_0 is the initial counts of the bacterial population (expressed as natural logarithm of colony forming units per gram) and $Q(t)$ is a variable that indicates the physiological state of the bacterial population (expressed as natural logarithm of another variable $q(t)$).

μ_{max} and λ values were obtained for each isothermal condition using the Baranyi model and secondary models.

For each isothermal condition, a dimensionless variable (h_0) value was then calculated from Eq. 9:

$$h_0 = \mu_{max} \times \lambda \tag{9}$$

The average h_0 was then used to obtain the initial q_0 in the differential form of the Baranyi model using Eq. 10 [21]. This initial q_0 value is crucial for accurately modelling of the lag phase and of dynamics of subsequent growth under various environmental conditions:

$$q_0 = \frac{1}{e^{h_0} - 1} \tag{10}$$

Because μ_{max} is a function of both temperature and time, μ_{max} values estimated by the appropriate secondary model were put into the differential form of the Baranyi model. The temperature data recorded with a data logger were used to solve Eq. 5 and Eq. 6. Results were obtained using `ode45` command, which implements the fourth-order Runge-Kutta method in the Matlab software [24].

Comparison of goodness of fit of the models

RMSE and adjusted coefficient of determination (R^2_{adj}) was used to compare the models' estimate performance using Eq. 11 and Eq. 12 correspondingly:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{\left(x_{obs} - x_{fit}\right)^2}{n - s}} \tag{11}$$

$$R^2_{adj} = 1 - \left(\frac{n-1}{n-s}\right)\left(\frac{SSE}{SST}\right) \tag{12}$$

where x_{obs} represents the experimental counts of bacterial population, x_{fit} represents the fitted value, n is the number of experiments, s is the number of model parameters, SSE is the sum of squares of errors and SST is the total sum of squares.

Statistical analysis

The Wilcoxon signed-rank tests [25, 26] were employed to assess the statistical significance of differences between machine learning-assisted and traditional modelling approaches. The `signtest` command in the statistical tool of Matlab 8.3.0.532 (R2014a) software was used. Significance in statistical differences between the two modelling approaches was determined when $p \leq 0.05$.

Validation of the models used

Model validation was assessed by considering the bias (B_f) and accuracy (A_f) factors using Eq. 13 and Eq. 14, respectively [27]:

$$B_f = 10^{\frac{\sum_{i=1}^{n} \log(x_{fit}/x_{obs})}{n}} \tag{13}$$

$$A_f = 10^{\frac{\sum_{i=1}^{n} |\log(x_{fit}/x_{obs})|}{n}} \tag{14}$$

B_f is a measure of average variation between the observed and predicted values. A_f measures the average difference between the observed and predicted values by disregarding whether the difference is positive or negative. A value of 1 for B_f and A_f indicates that there is a perfect agreement between all the observed and predicted values. Mean deviation (*MD*) and mean absolute deviation (*MAD*) between the observed and predicted counts of microbial populations were also calculated to assess the prediction performance of the dynamic model as suggested by Le Marc et al. [28].

RESULTS AND DISCUSSION

The *Pseudomonas* spp. counts data derived from previously published curves by Tarlak et al. [18] related to button mushrooms stored at various temperatures (4 °C, 12 °C, 20 °C, and 28 °C) were used. These data were utilized for both a one-step modelling approach based on the Baranyi model and a machine learning approach employing Gaussian process regression. The initial bacterial counts of *Pseudomonas* spp. averaged 7.05 ± 0.14 log CFU·g^{-1} for all temperatures. The storage duration decreased as the temperature increased, ranging from 240 h to 84 h (equivalent to 10 days to 3.5 days). *Pseudomonas* spp. counts at the end of storage varied between 8.64 ± 0.13 log CFU·g^{-1} and 10.76 ± 0.05 log CFU·g^{-1}, depending on the storage temperature (Fig. 3). This showed an increased growth potential of *Pseudomonas* spp. on button mushrooms with rising storage temperature.

The capability of the conventional modelling approach was assessed through computation of

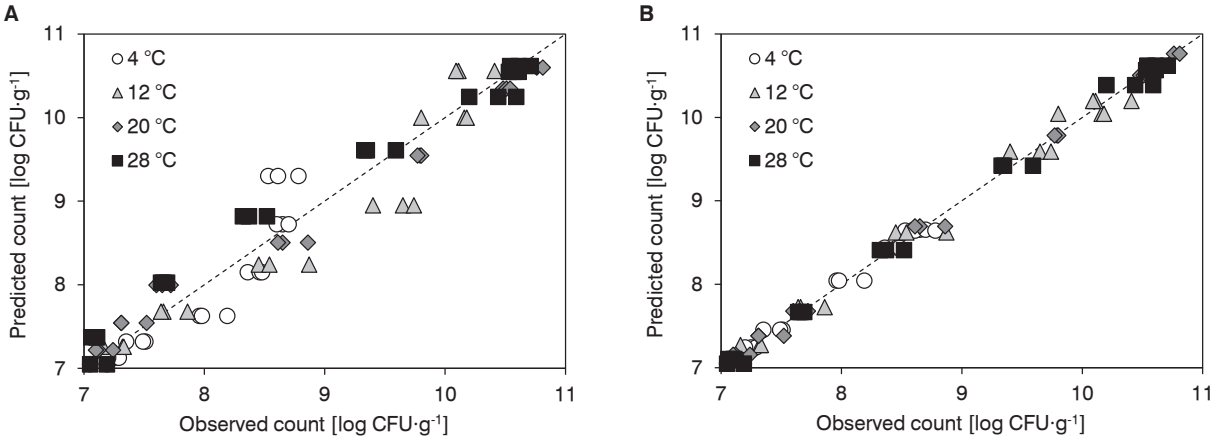

Fig. 3. Observed and predicted growth points.

A – original data for Baranyi model, B – original data for machine learning-assisted with Baranyi model.

RMSE and R^2_{adj} values. For the Baranyi model, the obtained *RMSE* and R^2_{adj} values were 0.294 and 0.950, respectively. In contrast, GPR yielded *RMSE* and R^2_{adj} values of 0.151 and 0.991, respectively. This indicated that the fitting capability of GPR surpassed that of the traditionally used Baranyi model in predicting mushroom spoilage.

The crucial parameters for characterizing the growth behaviour of microorganisms in food are the maximum specific growth rate (μ_{max}) and lag phase duration (λ). While these parameters cannot be directly determined, total counts of *Pseudomonas* spp. can be predicted using the developed model based on machine learning regression. This limitation becomes apparent when compared to traditional modelling methods in the field of predictive microbiology [29]. To address this limitation, the machine learning regression method was employed to enhance and guide the development of a more robust modelling approach.

Tab. 1 presents the values of μ_{max} and λ for *Pseudomonas* spp. on button mushrooms obtained from both the Baranyi model and the machine learning-assisted Baranyi model at each storage temperature. The μ_{max} value increased from 0.029 h^{-1} to 0.160 h^{-1} as the storage temperature rose from 4 °C to 28 °C, while λ exhibited an opposite trend, decreasing from 58.1 h to 10.5 h over the same temperature range. For the Baranyi model, the *RMSE* value was 0.294, and the R^2_{adj} value was 0.950. In contrast, the machine learning-assisted Baranyi model showed an improved performance with an *RMSE* value of 0.273 and an R^2_{adj} value of 0.956. Additionally, the Wilcoxon signed-rank test result was found to be 2.77×10^{-17}, indicating a significant difference in the prediction capability between the machine learning-assisted Baranyi model and the traditional Baranyi model. These findings collectively suggested that the used GPR before the traditional Baranyi modelling significantly enhanced the predictive

Tab. 1. Kinetic parameters and comparison of fitting capability of Baranyi model and machine learning-assisted modelling with Baranyi model.

Modelling approach	Temperature [°C]	λ [h]	μ_{max} [h^{-1}]	R^2_{adj}	*RMSE*
B-1	4	58.1 ± 9.8	0.029 ± 0.002	0.950	0.294
	12	27.6 ± 4.6	0.061 ± 0.003		
	20	16.0 ± 2.7	0.104 ± 0.005		
	28	10.5 ± 1.8	0.160 ± 0.008		
MB-1	4	57.9 ± 10.9	0.026 ± 0.002	0.956	0.273
	12	27.2 ± 5.1	0.055 ± 0.003		
	20	15.8 ± 3.0	0.095 ± 0.004		
	28	10.3 ± 1.9	0.146 ± 0.007		

B-1 – original data for Baranyi model, MB-1 – original data for machine learning-assisted Baranyi model, λ – lag phase duration, μ_{max} – maximum specific growth rate, R^2_{adj} – adjusted coefficient of determination, *RMSE* – root mean square error.

capability of the Baranyi model in describing the growth behaviour of *Pseudomonas* spp. on button mushrooms.

The Monte Carlo simulation method was employed to generate 10, 50 and 500 datasets. The values of μ_{max} and λ for *Pseudomonas* spp. on button mushrooms, derived from both the Baranyi model and the machine learning-assisted Baranyi model, are presented in Fig. 4 and Fig. 5. For the Baranyi model, the *RMSE* values ranged from 0.381 to 0.386, and R^2_{adj} values ranged from 0.919 to 0.916. In contrast, the machine learning-assisted Baranyi model provided *RMSE* values from 0.297 to 0.307 and R^2_{adj} values from 0.947 to 0.943. Statistical evaluation of the fitting capability of the two approaches revealed that the machine learning-assisted Baranyi model exhibited superior fitting capability compared to the traditionally used Baranyi model approach (Tab. 2). These results suggested that the machine learning-assisted Baranyi model approach can be reliably utilized for estimating *Pseudomonas* spp. counts on button mushrooms.

The crucial step of evaluating the dynamic model's performance is essential for validating the reliability of the estimated growth parameters of *Pseudomonas* spp. on button mushrooms. To achieve this, the growth data of *Pseudomonas* spp. on button mushrooms exposed to non-isothermal storage conditions (24 h at 4 °C and 12 h at 10 °C) were compared with the predicted growth data utilizing the differential form of the Baranyi model (Fig. 6).

Validation criteria of modelling approaches are given in Tab. 3. The machine learning-assisted Baranyi model exhibited B_f and A_f values of 1.000 and 1.019, respectively. Both B_f and A_f, close to one, indicated that the dynamic model utilized in this study possessed a high capability to predict the counts of *Pseudomonas* spp. on button mushrooms stored at varying temperatures over time. The *MD* and *MAD* values for *Pseudomonas* spp.

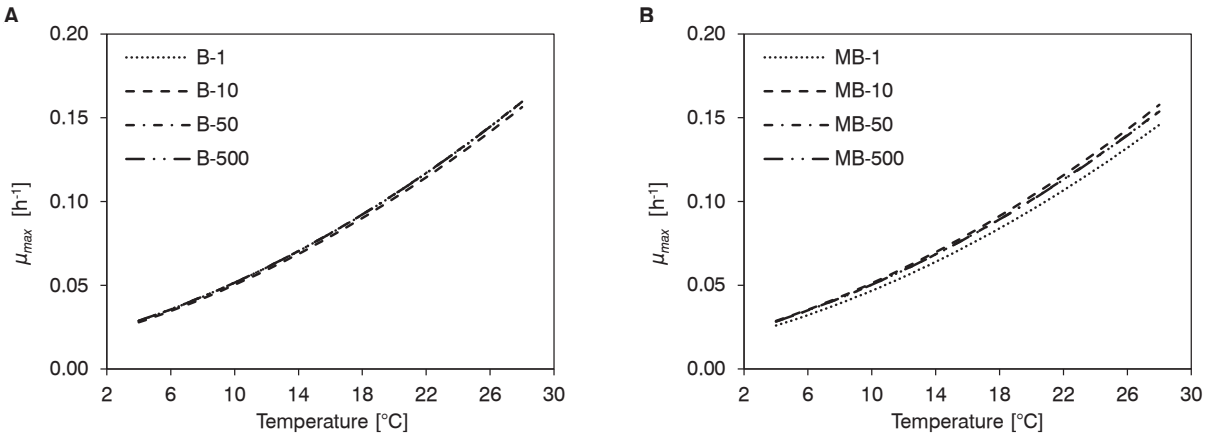

Fig. 4. Relationship between temperature and maximum specific growth rate.

A – simulated data for the Baranyi model, B – simulated data for machine learning-assisted Baranyi model.

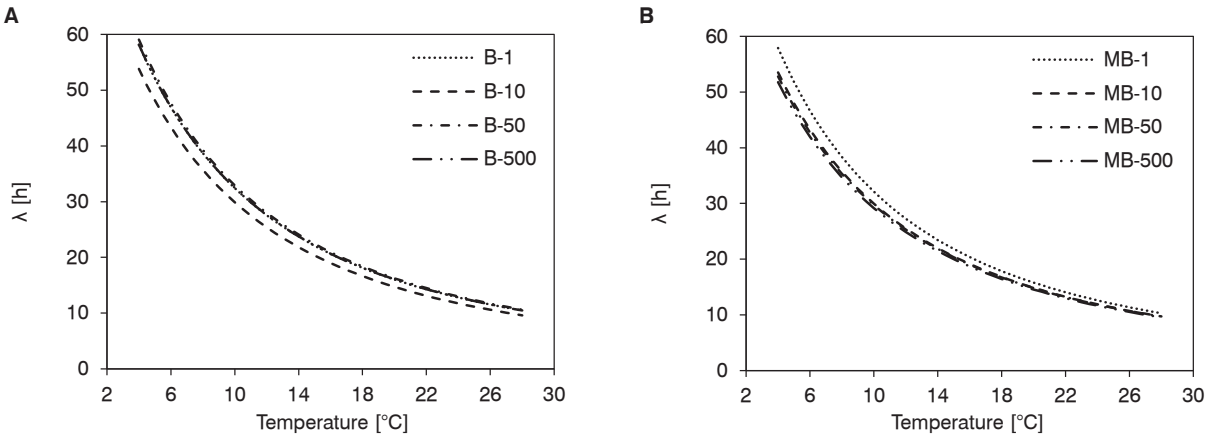

Fig. 5. Relationship between temperature and lag phase duration.

A – simulated data for the Baranyi model, B – simulated data for machine learning-assisted Baranyi model.

Tab. 2. Kinetic parameters and fitting capability of simulated data.

Modelling approach	Temperature [°C]	μ_{max} [h^{-1}]	*b*	R^2_{adj}	*RMSE*
Baranyi model					
B-10	4	0.029 ± 0.002	58.1 ± 9.8	0.950	0.294
	12	0.061 ± 0.003	27.6 ± 4.6		
	20	0.104 ± 0.005	16.0 ± 27		
	28	0.160 ± 0.008	10.5 ± 1.8		
B-50	4	0.026 ± 0.002	42.0 ± 12.9	0.956	0.273
	12	0.055 ± 0.003	18.0 ± 5.4		
	20	0.095 ± 0.004	9.9 ± 2.9		
	28	0.146 ± 0.007	6.3 ± 1.9		
B-500	4	0.029 ± 0.002	58.1 ± 9.8	0.950	0.294
	12	0.061 ± 0.003	27.6 ± 4.6		
	20	0.104 ± 0.005	16.0 ± 27		
	28	0.160 ± 0.008	10.5 ± 1.8		
Machine learning-assisted Baranyi model					
MB-10	4	0.029 ± 0.002	58.1 ± 9.8	0.950	0.294
	12	0.061 ± 0.003	27.6 ± 4.6		
	20	0.104 ± 0.005	16.0 ± 27		
	28	0.160 ± 0.008	10.5 ± 1.8		
MB-50	4	0.026 ± 0.002	42.0 ± 12.9	0.956	0.273
	12	0.055 ± 0.003	18.0 ± 5.4		
	20	0.095 ± 0.004	9.9 ± 2.9		
	28	0.146 ± 0.007	6.3 ± 1.9		
MB-500	4	0.026 ± 0.002	42.0 ± 12.9	0.956	0.273
	12	0.055 ± 0.003	18.0 ± 5.4		
	20	0.095 ± 0.004	9.9 ± 2.9		
	28	0.146 ± 0.007	6.3 ± 1.9		

B – Monte Carlo simulations for 10, 50 and 500 times with Baranyi model, MB – Monte Carlo simulations for 10, 50 and 500 times with machine learning-assisted Baranyi model.
μ_{max} – maximum specific growth rate, *b* – regression coefficient, R^2_{adj} – adjusted coefficient of determination, *RMSE* – root mean square error.

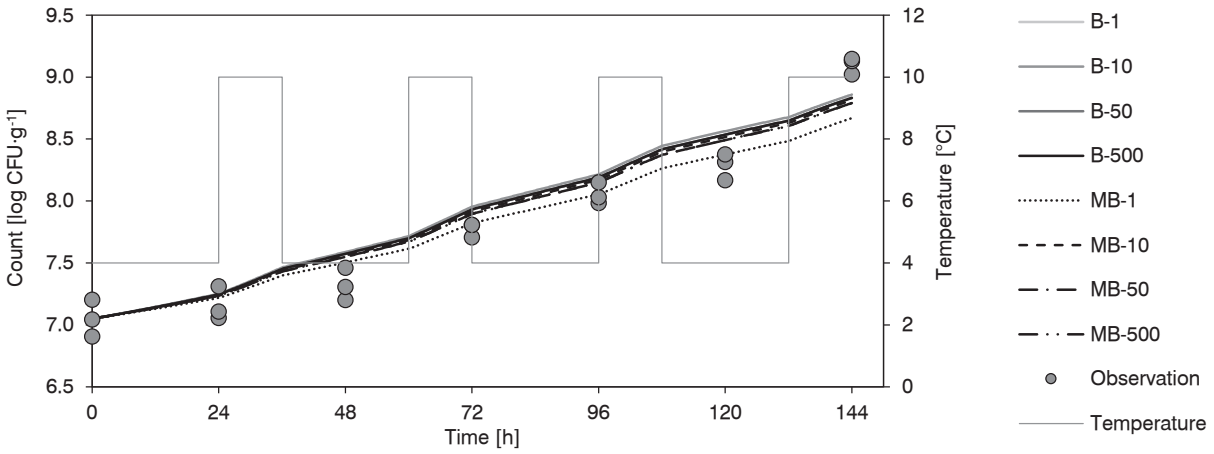

Fig. 6. Growth behaviour of *Pseudomonas* spp. on button mushrooms under dynamic temperature conditions.

B – Monte Carlo simulations for 1, 10, 50 and 500 times with Baranyi model, MB – Monte Carlo simulations for 1, 10, 50 and 500 times with machine learning-assisted Baranyi model.

Tab. 3. Validation criteria of modelling approaches.

Modelling approach	B_f	A_f	*MD*	*MAD*
Baranyi model				
B-1	1.012	1.023	0.086	0.184
B-10	1.014	1.025	0.105	0.194
B-50	1.012	1.024	0.088	0.185
B-500	1.012	1.024	0.088	0.185
Machine learning-assisted Baranyi model				
MB-1	1.000	1.019	–0.007	0.148
MB-10	1.010	1.023	0.074	0.177
MB-50	1.009	1.022	0.060	0.170
MB-500	1.008	1.022	0.059	0.169

B – Monte Carlo simulations for 1, 10, 50 and 500 times with Baranyi model, MB – Monte Carlo simulations for 1, 10, 50 and 500 times with machine learning-assisted Baranyi model.
B_f – bias factor, A_f – accuracy factor, *MD* – mean deviation, *MAD* – mean absolute deviation.

populations were –0.007 log CFU·g^{-1} and 0.148 log CFU·g^{-1}, respectively. The *MD* value of –0.007 log CFU·g^{-1} indicated that, on average, the dynamic model slightly overestimated by 0.007 log CFU·g^{-1} the observed values, while the *MAD* value of 0.148 log CFU·g^{-1} suggested that, on average, the predicted values differed by 0.148 log CFU·g^{-1} (either higher or lower) from the observed ones. In comparison, Manthou et al. [30] reported an *MD* value of –0.10 log CFU·g^{-1} and a *MAD* value of 0.22 log CFU·g^{-1} for *Pseudomonas* spp. populations on oyster mushrooms. These results collectively indicated that the models employed in the present study exhibited a superior predictive capability compared to the dynamic model previously developed for microbial contamination of oyster mushrooms. Consequently, the machine learning-assisted modelling approach utilized in this study could serve as an alternative to traditional modelling approaches for determining the number of *Pseudomonas* spp. on mushrooms.

CONCLUSION

This study demonstrated the significant potential of machine learning-assisted modelling approaches, particularly utilizing GPR, in advancing our comprehension of microbial growth dynamics. The comparison between traditional modelling methods and machine learning techniques, conducted through Monte Carlo simulations, revealed that the machine learning approach significantly increased the robustness of the models. The higher R^2_{adj} value and *RMSE* value that were obtained with the machine learning-assisted approach underscored its superiority in accurately capturing the dynamic and complex nature of microbial growth, especially in the presence of high variation. The validation metrics, including bias and accuracy factors, further support the reliability of the machine learning-assisted dynamic model. With a minimal *MD* and *MAD*, this approach emerges as a highly dependable prediction method for describing microbial growth behaviour in the realm of predictive food microbiology. This study contributed valuable insights to various fields such as microbiology, food safety and biotechnology, paving the way for more effective and adaptable models in understanding and predicting microbial growth.

REFERENCES

1. Pérez-Rodríguez, F. – Valero, A.: Predictive microbiology in foods. New York : Springer, 2013. ISBN: 9781461455202. DOI: 10.1007/978-1-4614-5520-2.
2. Tarlak, F.: The use of predictive microbiology for the prediction of the shelf life of food products. Foods, *12*, 2023, article 4461. DOI: 10.3390/foods12244461.
3. Rohilla, P. – Kumari, A. – Birania, S. – Monika.: Commercially available databases in food microbiology. In: Chhikara, N. – Panghal, A. – Chaudhary, G. (Eds): Microbes in the food industry. Beverly : Wiley-Scrinever, 2023, pp. 441–467. ISBN: 9781119775584. DOI: 10.1002/9781119776406.ch11.
4. du Plessis, F. – Goedhals-Gerber, L. L. – van Eeden, J.: Forecasting fruit export damages and enhancing food safety through risk management. Sustainability, *15*, 2023, article 15216. DOI: 10.3390/su152115216.
5. Ali, S. – Rezende, V. T. – Ullah, S. – de Paiva, E. L. – Tonin, F. G. – Abdullah – Corassin, C. H. – Oliveira, C. A. F: Food processing and challenges in the food production and quality: The foodomics approach. Food Bioscience, *56*, 2023, article 103217. DOI: 10.1016/j.fbio.2023.103217.
6. Ačai, P. – Valík, Ľ. – Medveďová, A.: One-and two-step kinetic data analysis applied for single and co-culture growth of *Staphylococcus aureus, Escherichia coli*, and lactic acid bacteria in milk. Applied Sciences, *11*, 2021, article 8673. DOI: 10.3390/app11188673.
7. Tarlak, F. – Yucel, O.: Development of a prediction software for the growth kinetics of *Pseudomonas* spp. in culture media using various primary models. Applied Food Biotechnology, *10*, 2023, pp. 1–8. DOI: 10.22037/afb.v10i1.39780.
8. Tarlak, F.: Development of a novel growth model based on the central limit theorem for the determination of beef spoilage. Applied Food Biotechnology, *8*, 2021, pp. 143–150. DOI: 10.22037/afb.v8i2.33549.
9. Tarlak, F. – Yucel, O. – Khosravi-Darani, K.:

Development of a new modelling approach and performance evaluation of meta-heuristic optimization algorithms for the prediction of kinetic growth parameters for *Pseudomonas* spp. in fish. Journal of Pure and Applied Microbiology, *16*, 2022, pp. 1263–1273. DOI: 10.22207/JPAM.16.2.55.

10. Tarlak, F. – Khosravi-Darani, K.: Development and validation of growth models using one-step modelling approach for determination of chicken meat shelf-life under isothermal and non-isothermal storage conditions. Journal of Food and Nutrition Research, *60*, 2021, pp. 76–86. ISSN: 1336-8672 (print), 1338-4260 (online) <https://www.vup.sk/download.php?bulID=2101>
11. Tarlak, F. – Costa, J. C. C. P.: Comparison of modelling approaches for the prediction of kinetic growth parameters of *Pseudomonas* spp. in oyster mushroom (*Pleurotus ostreatus*). Food Science and Technology International, *29*, 2023, pp. 631–640. DOI: 10.1177/10820132221105476.
12. Hiura, S. – Koseki, S. – Koyama, K.: Prediction of population behavior of *Listeria monocytogenes* in food using machine learning and a microbial growth and survival database. Scientific Reports, *11*, 2021, article 10613, DOI: 10.1038/s41598-021-90164-z.
13. Koyama, K. – Kubo, K. – Hiura, S. – Koseki, S.: Is skipping the definition of primary and secondary models possible? Prediction of *Escherichia coli* O157 growth by machine learning. Journal of Microbiological Methods, *192*, 2022, article 106366. DOI: 10.1016/j.mimet.2021.106366.
14. Yildirim-Yalcin, M. – Yucel, O. – Tarlak, F.: Development of prediction software to describe total mesophilic bacteria in spinach using a machine learning-based regression approach. Food Science and Technology International, published online 18 April 2023. DOI: 10.1177/10820132231170286.
15. Yücel, Ö. – Tarlak, F.: An intelligent based prediction of microbial behaviour in beef. Food Control, *148*, 2023, article 109665. DOI: 10.1016/j.foodcont.2023.109665.
16. van Boekel, M. A. J. S.: Kinetic modeling of reactions in foods. Boca Raton : CRC press, 2008. ISBN: 9781574446142.
17. Öksüz, H. B. – Buzrul, S.: Monte Carlo analysis for microbial growth curves: Monte Carlo simulations. Journal of Microbiology, Biotechnology and Food Sciences, *10*, 2020, pp. 418–423. DOI: 10.15414/jmbfs.2020.10.3.418-423.
18. Tarlak, F. – Ozdemir, M. – Melikoglu, M.: Predictive modelling for the growth kinetics of *Pseudomonas* spp. on button mushroom (*Agaricus bisporus*) under isothermal and non-isothermal conditions. Food Research International, *130*, 2020, article 108912. DOI: 10.1016/j.foodres.2019.108912.
19. Pandey, V. K. – Srivastava, S. – Dash, K. K. – Singh, R. – Mukarram, S. A. – Kovács, B. – Harsányi, E.: Machine learning algorithms and fundamentals as emerging safety tools in preservation of fruits and vegetables: a review. Processes, *11*, 2023, article 1720. DOI: 10.3390/pr11061720.
20. Khaleel, M. – Ahmed, A. A. – Alsharif, A.: Artificial intelligence in engineering. Brilliance : Research of Artificial Intelligence, *3*, 2023, pp. 32–42. DOI: 10.47709/brilliance.v3i1.2170.
21. Baranyi, J. – Roberts, T. A.: A dynamic approach to predicting bacterial growth in food. International Journal of Food Microbiology, *23*, 1994, pp. 277–294. DOI: 10.1016/0168-1605(94)90157-0.
22. Ratkowsky, D. A. – Olley, J. – McMeekin, T. A. – Ball, A.: Relationship between temperature and growth rate of bacterial cultures. Journal of Bacteriology, *149*, 1982, pp. 1–5. DOI: 10.1128/jb.149.1.1-5.1982.
23. Robinson, T. P. – Ocio, M. J. – Kaloti, A. – Mackey, B. M.: The effect of the growth environment on the lag phase of *Listeria monocytogenes*. International Journal of Food Microbiology, *44*, 1998, pp. 83–92. DOI: 10.1016/S0168-1605(98)00120-2.
24. Velugoti, P. R. – Bohra, L. K. – Juneja, V. K. – Huang, L. – Wesseling, A. L. – Subbiah, J. – Thippareddi, H.: Dynamic model for predicting growth of *Salmonella* spp. in ground sterile pork. Food Microbiology, *28*, 2011, pp. 796–803. DOI: 10.1016/j.fm.2010.05.007.
25. Wilcoxon, F.: Individual comparisons by ranking methods. In: Kotz, S. – Johnson, N. L. (Eds.): Breakthroughs in statistics : Methodology and distribution. New York : Springer, 1992, pp. 196–202. ISBN: 978-0-387-94039-7. DOI: 10.1007/978-1-4612-4380-9_16.
26. Vegdahl, A. C. – Baldwin, W. C. – Schaffner, D. W.: Growth models for *Salmonella*, *E. coli* O157:H7 and *L. monocytogenes* give different predictions for pathogen growth in cut leafy greens transportation, but are consistent in identifying higher risk conditions. Food Microbiology, *115*, 2023, article 104338. DOI: 10.1016/j.fm.2023.104338.
27. Ross, T.: Indices for performance evaluation of predictive models in food microbiology. Journal of Applied Bacteriology, *81*, 1996, pp. 501–508. DOI: 10.1111/j.1365-2672.1996.tb03539.x.
28. Le Marc, Y. – Plowman, J. – Aldus, C. F. – Munoz-Cuevas, M. – Baranyi, J. – Peck, M. W.: Modelling the growth of *Clostridium perfringens* during the cooling of bulk meat. International Journal of Food Microbiology, *128*, 2008, pp. 41–50. DOI: 10.1016/j.ijfoodmicro.2008.07.015.
29. Tarlak, F. – Yücel, Ö.: Prediction of *Pseudomonas* spp. population in food products and culture media using machine learning-based regression methods. Life, *13*, 2023, article 1430. DOI: 10.3390/life13071430.
30. Manthou, E. – Tarlak, F. – Lianou, A. – Ozdemir, M. – Zervakis, G. I. – Panagou, E. Z. Nychas, G. J. E.: Prediction of indigenous *Pseudomonas* spp. growth on oyster mushrooms (*Pleurotus ostreatus*) as a function of storage temperature. LWT – Food Science and Technology, *111*, 2019, pp. 506–512. DOI: 10.1016/j.lwt.2019.05.062.

Received 9 May 2024; 1st revised 18 June 2024; accepted 9 July 2024; published online 16 August 2024.