A primer on predictive techniques for food and bioresources transformation processes

Jason Sicard, Sophie Barbe, Rachel Boutrou, Laurent Bouvier, Guillaume Delaplace, Gwenaëlle Lashermes, Laetitia Theron, Olivier Vitrac, Alberto Tonda

Received: 8 December 2022 | Revised: 10 February 2023 | Accepted: 4 March 2023

DOI: 10.1111/jfpe.14325

COMMENTARY

Journal of Food Process Engineering WILEY

A primer on predictive techniques for food and bioresources transformation processes

Jason Sicard[1] | Sophie Barbe[2] | Rachel Boutrou[3] | Laurent Bouvier[4] | Guillaume Delaplace[4] | Gwenaëlle Lashermes[5] | Laëtitia Théron[1] | Olivier Vitrac[6] | Alberto Tonda[7]

[1]QuaPA, INRAE, France

[2]TBI, CNRS, Toulouse, France

[3]STLO, INRAE, Institut Agro, Rennes, France

[4]UMET, Université de Lille, CNRS, Centrale Lille, INRAE, Villeneuve-D'Ascq, France

[5]FARE, INRAE, UMR A 614, Université de Reims Champagne-Ardenne, Reims, France

[6]SayFood, INRAE, AgroParisTech, Université Paris Saclay, Massy, France

[7]MIA-Paris, AgroParisTech, INRAE, Université Paris Saclay, Paris, France

Correspondence
Jason Sicard, QuaPA, INRAE, St-Genès-Champanelle, 63122, France.
Email: jason.sicard@inrae.fr

Alberto Tonda, MIA-Paris, AgroParisTech, INRAE, Université Paris Saclay, 113 rue Nationale, 75013, Paris, France.
Email: alberto.tonda@inrae.fr

Abstract

To meet current societal demand for more sustainable transformation processes and bioresources, these processes must be optimized and new ones developed. The evolution of various systems (raw material, food, or process attributes) can be predicted to optimize the uses of biomass for better quality, safety, economic benefit, and sustainability. Predictive modeling can guide the necessary changes and influence industrials, governmental policies and consumers decision-making. However, achieving good predictive capability requires reflection on the models and model validation, which can be difficult. This review aims to help scientists begin to predict by presenting the techniques currently used in predictive science for food and related bioproducts. First, a guideline helps readers initiate a prediction process along with final tips and a warning about the risks involved, with a particular focus on the crucial validation step. Three broad categories of techniques are then presented: empirical, mechanistic, and artificial intelligence (or "data-driven"). For each category, the advantages and limitations of current techniques for prediction are explained in light of their current domains of applications, illustrated with literature studies and a detailed example. Thus this article provides engineering researchers information about predictive modeling which is a recent relevant development in optimization of both food and nonfood bioresources processes.

Practical applications

Predictive modeling is a recent development of much relevance in the optimization of both food and nonfood bioresources processes. The goal of this article is to guide those in research or industry who would like to start predicting. Therefore, the article is intended as a primer on prediction concepts and predictive techniques for food and non-food bioresources processing. Three categories of techniques commonly used in these fields are illustrated by various examples of current applications and a more detailed example helps to understand the implementation process. An increased ability of the global scientific body to predict the outcome of various decisions, often linked or sequential, will open new avenues for designing food products with circularity in mind: maintaining value and not creating waste in the process.

KEYWORDS
bioresource, guideline, modeling, prediction, process development

1 | INTRODUCTION

The manufacturing of food and feed and byproducts-based economic activities that exploit biotechnology and biomass to produce energy, goods or services have progressively become severe political and social concern (Bakan et al., 2022). Research in this field has grown significantly recently, with national and international governing bodies pushing for more sustainable natural resource use and less dependence on nonrenewable resources (European Commission & Directorate-General for Research and Innovation, 2017). Not surprisingly, more and more lines of research are now focused on modeling and predicting the outcomes of food and bioproducts-related processes and the qualities and properties of the end-product output.

The COVID-19 pandemic has confirmed the ability of science to prepare and guide public decision-making through predictive models. Similar outcomes are expected for the eco-efficient transformation of food and the use of bioresources. Predictive models are increasingly used in open-loop simulations to guide governments (Acs et al., 2019) and evaluate alternative policies (Schlüter et al., 2022). The bioeconomy means using renewable biological resources from land and sea, like crops, forests, fish, animals, and micro-organisms to produce food, materials and energy (European Commission, 2022). All those could benefit from scripting scenarios with efficiency, quality, safety, economy, and sustainability indicators. At the scale of the processing of food and biomass, the ability to predict the outcome of a succession of linked decisions would be beneficial at all stages of transforming agricultural biomass into food and biotechnological products, and recycling biomass (Zeng & Li, 2021). It is generally accepted that predictions can support objective decisions in many areas: minimizing cost and resources, securing investments, and scheduling supplies and deliveries (Bauwens et al., 2020). More integrated predictions supporting food safety, quality, and sustainability are usually missing or more challenging to develop. They typically require more interdisciplinary and integrative approaches than those used in engineering and more detailed ones than those supporting the public decision (Vitrac et al., 2022). They are now within reach and should significantly redefine how natural resources are used and how food and other biotechnological products are processed and distributed. In a nutshell, the technologist already has at his disposal a range of numeric predictive techniques allowing him to control the efficiency and flexibility of the transformation activities, evaluate the impacts of his choices, and find the best technological route. Innovation can be triggered virtually, and previous failures can be analyzed forensically with the help of the same predictive techniques.

The primary interest of prediction stems from its ability to remove subjectivity from decision-making while preventing adverse outcomes. In other words, the main objective is not to achieve the best choice but to direct choices away from the worst options. The methodology can be rolled out to all technology readiness levels without requiring the product to be physically present. The consequences of actions (such as modifications of recipes, processes, the distribution chain, by-product delivery) can be tested without sampling, detection limit, time, or even the existence of tested conditions or techniques (Gillet et al., 2009). Multiscale modeling opens new perspectives by virtualizing process (Touffet et al., 2021; Vitrac & Touffet, 2019) and product design (Vitrac et al., 2022). Two features offer significant breakthroughs: almost continuous descriptions from process down to molecular scales and the possibility of considering contradictory goals such as minimizing packaging waste and maximizing product shelf-life (Zhu et al., 2019). Beyond the immediate benefits of cost and development time reduction, it is thought that these techniques could ultimately lead to solutions and optimums that could not be found experimentally.

Scientific predictions have come a long way over the past few decades. Earlier, estimates were based on the fundamental laws of nature and relationships between a few measurable entities incorporated into simple models governed by a few parameters. They now mobilize complex representations involving large observational data sets and increasing knowledge, all integrated into refined process representations to provide sophisticated numerical models. In the sense of this article, predictive ability relies on three prerequisites:

- Prediction requires a scientifically sound model—even a simple one—supported by sufficient pieces of scientific evidence.
- Prediction is intentional (a priori): a predictive model is designed for a specific class of problems while providing answers even for values missing at its creation.
- Prediction can be evaluated (a posteriori), so its domain of validity can be ascertained.

Although the art of predicting is well accepted in the manufacturing industry, it remains less developed in the food industry and more generally in areas associated with the primary transformation of bioresources or the management of wastes (Erdogdu et al., 2022). Many reasons have been invoked, including the coupling of physical, chemical, and biological processes intimately coupled at different scales, the soft-matter behavior of the biological matter, the lack of mathematical and physical modeling in related curricula, the considerable variability of raw materials and the dramatic evolution of all properties during the transformation of biological matter. To accelerate innovation and encourage the ecological transition, the authors think that predicting for the bioeconomy should be a unique interdisciplinary topic. Predictions, if properly managed, could be less costly than experiments and obtained without delaying the innovation process. As engaging in a prediction process can be intimidating, we set out guidance on implementing fast and reliable predictions. We then move forward with the

main strategies for getting started, illustrated with case studies chosen to represent different classes of common problems encountered in food and bioresource processes. We conclude with future research directions on prediction.

2 | GUIDANCE ON USING PREDICTION IN YOUR RESEARCH

To help the reader get started on initiating a prediction process, we have conceived a purpose-designed guideline. The decision tree aims to define the model type to initially try to implement according to the knowledge and data available on the topic. As a first approach, we suggest assessing the fundamental knowledge (literature, expertise, data, etc.) you already have, and then drawing inspiration from Figure 1 to determine the model that, a priori, would be most suitable for you to use. This decision tree is a basic guideline to guide the user in obtaining a prediction from the available knowledge or data by choosing the right method from mechanistic, empirical, or data-driven families. Note that it does not indicate a priority or hierarchical relationship among methods. Most problems can ultimately be successfully solved with each of the presented approaches, but one may be more suitable than others. The main difficulty of the prediction does not reside in the implementation of a technique but in the necessary validation. It is therefore the incompatibility between the formulation of what is known, be it knowledge or data, and what is readily interpretable within model structure that tends to make methods unsuitable. Machine learning is a field devoted to leveraging data to create information, so it is especially useful for problems with large amounts of data. Mechanistic modeling is established in phenomenological knowledge. Empirical modeling generally uses data, information, and knowledge, but tend to be lacking in each. We have divided this last large family according to how the knowledge gaps are filled. This is a key to understanding prediction limits.

Consider where your proficiencies lie in addition to this decision tree. A learned suboptimal tool may be a better choice than the recommended one you know nothing about. The intention of this guide is to make things easier, not to force you into one way.

2.1 | Modeling process

The predictive process is always initiated by observing the phenomena to be predicted. A mathematical model is built from these phenomena to capture conceptually essential features while eliminating irrelevant detail.

The first step is to identify the relevant key variables to predict, which may require scientific knowledge depending on your focal area. Interaction with experts can be helpful to start out with good foundations. Critical properties of the variables must be evaluated. Are they qualitative or quantitative? How many levels are there, or what is the acceptable uncertainty? Are you looking for a

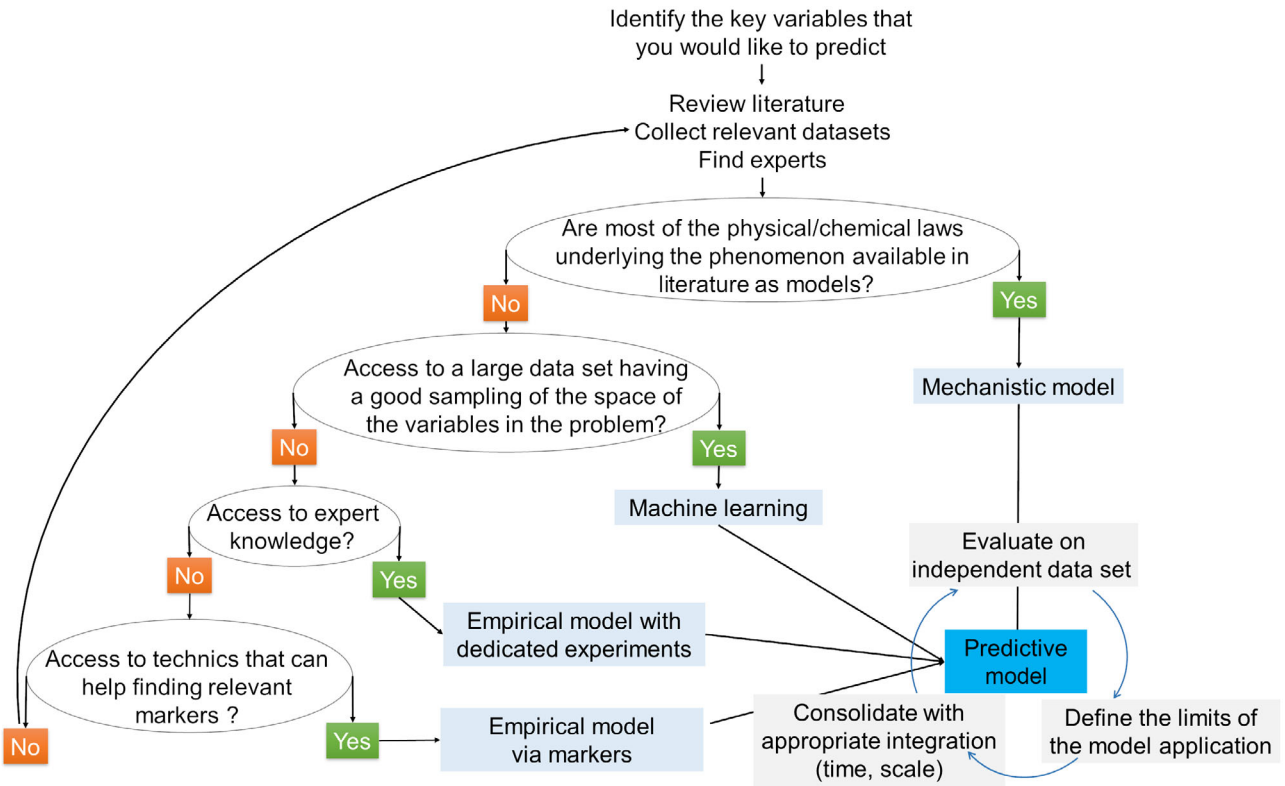

FIGURE 1 Predictive technique decision tree. What is the right approach given the data and knowledge available?

deterministic result or a probabilistic result? Do the outputs serve as feedback inputs? In the latter case, the control function parameters (e.g., proportional-integral-derivative controllers) need to be included in the modeling.

Remember the concept of using a systemic approach: try to break your problem down into simpler components. Modeling and validating unit-operation models is more reliable and resilient to changes in the process.

Thought needs to go into defining the limits of the model application. You will then need to define the model's limits, depending on what is unknown about the problem. This will determine the potential scope of the prediction, which should match your research question. The main limits concern the scope of the model, the practicability of the outputs, and the accessibility of input parameters. Limits can be strict when the model uses parts or theories that themselves have a predetermined domain of validity (mechanistic case-example: the prediction of the velocity fields in a duct for a laminar flow usually requires a Reynolds number lower than 2000) or that reflect what has been validated through the data sets, in which case validation with new data sets may extend the limits of prediction.

A model represents knowledge; this initial collecting step will therefore allow you to know what you can represent and ultimately model.

2.2 | Upgrading to prediction: Validation

When designing a model, especially a mechanistic one, it is usually best to start with a very simplified representation before adding complexity. You add complexity gradually, layer-by-layer, and only until the model output has a precision appropriate to the question asked. Going any further is futile. Redirect your efforts toward model validation and evaluation. Simply having a model is not enough to run a prediction—its reliability must be assessed first.

The results obtained and their potential applicability and limits will vary depending on the technique used. Nevertheless, validation is needed to gain confidence that the model can predict within its capabilities. It generally involves three steps of testing and improvements in an iterative loop process that continues as long as the model fails to meet its objective quality threshold:

- The qualification of the model is evaluated on a set of data that is independent from the set used to create it, in a range of situations in coherence with the limits,
- The limits of the model application are re-evaluated,
- The model is consolidated through more realistic assumptions and new developments in the methods.

The way to get out of this spiral is to meet the quality threshold, so it is essential to define this threshold appropriately. To illustrate, Guisan and Zimmermann (2000) report trade-offs and evaluation measures for habitat distribution.

When the output is qualitative and binary, simple metrics are commonly used:

- Accuracy: the ratio of correct predictions to total cases,
- Sensitivity: the ratio of correct positive predictions to total positive cases,
- Specificity: the ratio of correct negative predictions to total negative cases.

Different threshold values for these metrics result from different aversion levels to the types of risk. Likewise for a quantitative profile, setting a maximum local error reflects a different relationship to heterogeneity than a maximum error on the mean value. The worth of these relationships is to be assessed alongside the prediction objective.

3 | DESCRIPTIVE BRIEF ON PREDICTIVE MODELS AND THEIR APPLICATIONS

Prediction always requires a model able to provide a relevant output. However, from a binary check on the presence or absence of a given marker to a full-scale physical description of a process, the scientific basis of models covers several levels of complexity. Below, we provide case studies and success stories for each approach and discuss limitations and potential future evolutions.

Among the many classification possibilities, we chose to group the techniques into three broad categories, which have distinctive relation to data and knowledge.

The first category, empirical prediction, regroups the models primarily rooted in biology, life sciences, environmental science, and medicine. Their primary mathematical pillar and core assessment tool are probability and statistics. They require a moderate amount of knowledge and data and can include inputs from industry expertise or traditional craftsmanship more straightforwardly than other techniques.

The second category, mechanistic prediction, regroups the models primarily rooted in physics, chemistry, biochemistry, and biology. With underlying mechanisms and laws as a core, these techniques are closely related to mathematics, but mathematics is the language, not the primary driver. These models require extensive knowledge about the inner workings of the problem.

The third category, artificial intelligence or "data-driven" prediction, regroups the foremost models in cognitive science, computer science, and mathematical logic. These models require large amounts of data and computational resources.

The models of these three categories are usually developed and validated independently and from different mathematical roots and concepts. Their relation to causation is one of the most distinctive features. Mechanistic prediction is powered by causation, whereas correlations are considered good enough for empirical and artificial intelligence predictions. Causality is much more complex to establish than correlation. Indeed, it is common to confound undetermined

factors with biases. Causality due to many interactions is also challenging to evidence and reproduce. In this perspective, mixed approaches are likely to become the norm for the bioeconomy and the transformation of bioresources. Some other aspects, such as those related to cognitive processes (hedonic perception…) or physiological (appetite, satiety, glycemia…), may remain out of reach of mechanistic approaches. The outputs of one approach will be used as inputs of another at a letter stage of the prediction.

3.1 | Empirical prediction

3.1.1 | Description and advantages

Empiricism produces experimental observations and analyses them mathematically. Empirical approaches lie at the crossroads of machine learning approaches, which need a large amount of data, and mechanistic descriptions, making explicit the cause-and-effect relationships between the components of bio-physical-chemical processes. They mobilize a moderate amount of data and support it by an incomplete integration of phenomenological knowledge. This incomplete integration is the consequence of either a lack of knowledge about the phenomena, or a choice of simplification for practicality purposes.

Where there are knowledge gaps, empirical prediction seeks to pass directly from measured elements considered relevant by humans (introducing the "human" filter, which exists in all methods except unsupervised machine learning) to the object of interest, thus avoiding the complexity necessary for the description of mechanisms. This shortcutting makes the empirical method especially suitable when the prediction outputs must be application-ready, for instance, for economic or environmental considerations. In this type of study, statistical analysis is a pillar for supporting the validation of correlations or refuting seemingly obvious predictors, for example, appetite rating to predict energy intake (Holt et al., 2017), as the transition to causality comes from expertise (and is not always essential).

The model is often based on the definition of "markers" of the biological or physical criteria in bioresource process predictions. Once defined, these biomarkers will serve as indicators of a specific condition. The prediction is focused more on the input data than on the calculation per se, the model, or the way to analyze the data. Therefore, empirical models tend to be comparatively more straightforward, but finding an efficient marker is the challenge.

3.1.2 | Techniques used and application areas

Large scope analysis techniques

The recent rapid development of biological, physical, and chemical analytics tools, which provide indirect information on a large number of practical questions, allows us to obtain the first type of input data for empirical prediction. Some of the data obtained through these tools also lend themselves well to data-driven methods. However, they are still hampered by the cost of obtaining results, which de facto limits the amount of data and forces analysts to use expert knowledge and, hence, empirical prediction.

All of these analytics tools give access to a large domain of potential predictors. Most of them are irrelevant to the problem at hand, but some may be. The analysis revolves around finding these correlations.

The following subsections list examples of using some of the advanced analytics tools to find empirical predictors. This is aimed at researchers who may have access to these tools to get an outline of their possibilities.

Omics. The omics branches of biology, spurred by the development of advanced methods and tools, are prime users of these predictions. They statistically search for candidate genes and biomarkers associated with various kinds of target traits. Genomics can be used to predict medical problems such as hyperketonemia (Pralle & White, 2020), reproductive traits (Long, 2020), or evolutionary potential (Bay et al., 2017; Mueller-Schaerer et al., 2020). Proteomics can for example provide information on animal stress levels (Mouzo et al., 2020). In this context, analysts find correlations using between-population differences in gene expression intensity.

Omics-based predictions extensively use bio-ontologies such as the collaborative Gene Ontology knowledgebase, a directed acyclic graph that describes gene/gene product molecular functions and cellular locations and the biological process in which these functions participate. This knowledge of the biological processes adds a layer of expertise, improving confidence that any correlated predictors identified might be somehow implicated in the target processes.

Infrared spectroscopy. Mid- and near-infrared (NIR) spectroscopy has recently been successfully used to predict complex food phenotypes and an extensive set of quality traits (Bresolin & Dorea, 2020; Pralle & White, 2020; Vranic et al., 2020). NIR uses both empirical techniques and machine learning, although the proportion of machine learning increases with increasingly massive output.

Predictions are usually based on partial least squares regression (Bresolin & Dorea, 2020) and multiple linear regression models (Pralle & White, 2020) to obtain statistical indicators.

NIR has good potential for online analysis because it affordably generates real-time, nondestructive information. However, caution should be exercised in making the transition from laboratory to online prediction, as online measurement tools lack accuracy, which may substantially impact the reliability of the prediction.

Hyperspectral imaging, nuclear magnetic resonance or mass spectrometry. Other techniques such as hyperspectral imaging (Cheng et al., 2017), nuclear magnetic resonance spectroscopy or mass spectrometry (Fayeulle et al., 2019), also use statistical indicators to predict. These techniques share a similar problem of under-precision when attempting to scale up to online prediction, as affordable equipment is less discriminant (e.g., low-field magnetic resonance vs. laboratory-grade high-field magnetic resonance), and this potential limitation should be considered early in the prediction process.

Dedicated experiments approach

An alternative solution to develop an empirical prediction is to collect data through dedicated experiments. Those are based on the researcher's reasoning, knowledge and capacity to judge the a priori essential parameters and the extent of their interactions (linear, quadratic, etc.). Expert knowledge can therefore play a decisive role in designing this data collection.

Whenever experiments are performed with predictive intent and independently of the model envisioned, the crucial factor will always be the design of experiments, which should establish statistical concepts of validity, reliability, and replicability.

The following subsections list examples of some fields where this approach have generally been successful.

Predictive microbiology. The poster child of empirical prediction is primary models used for predictive microbiology, which have been very successful in describing microbial evolution as a function of time (Stavropoulou & Bezirtzoglou, 2019) and which have become essential to food risk management and assessment (Lopatkin & Collins, 2020; Valdramidis et al., 2013). These models are grounded in the basic assumption that populations of microorganisms show reproducible responses to environmental factors. Primary models are often coupled with secondary models describing the evolution of influencing factors like pH, temperature, water activity (Li et al., 2008). These secondary models often follow a mechanistic approach, as described in Section 3.2.

Resources management. The field of agricultural management is very much a user of empirical predictions for water management (Adisa et al., 2020; Pereira et al., 2020), energy management (Bersani et al., 2020; Garcia-Maraver et al., 2017; Shine et al., 2020), and biological control (Haan et al., 2020; Mills & Heimpel, 2018; Mueller-Schaerer et al., 2020), or to predict direct impacts on productivity. These productivity-focused models have been employed in a wide range of studies, from predictions looking directly at overall yield (Bekenev, 2019; Horie, 2019; Li et al., 2018; Menzel, 2021) to predictions investigating a specific cause of product loss (Ealy & Seekford, 2019; Zhou et al., 2020), typically by the impact of animal diet (Hanigan et al., 2018; Lyu et al., 2020; Święch, 2017; Trottier & Tedeschi, 2019).

Regression models are the prevalent tools (Menzel, 2021; Shine et al., 2020), but various fields have developed computed-based tools that improve efficiency and ease the display of statistical output, such as geographic information systems for maps (Adisa et al., 2020) or the "Simulation model for rice–weather relations" for rice (Horie, 2019).

Sensors-based prediction. Sensors and nondestructive measurement methods (Ealy & Seekford, 2019; Li et al., 2018) are extensively used to connect predictors to predicted outcomes, in some cases in real-time to enable in situ monitoring of the impact of an intervention. Remote sensing imagery for maps (Adisa et al., 2020; Pereira et al., 2020), Internet of Things for smart greenhouses (Bersani et al., 2020), or frost monitoring (Zhou et al., 2020), colorimetry and visible imaging (Li et al., 2018) are other examples of sensors in use.

How those can be linked together, a case study

Several quality criteria concerning animal stress level, technological quality of meat, and their relationship have been developed for prediction purposes (Mouzo et al., 2020; Théron et al., 2020). Mouzo et al. (2020) feature a proteomics study that results in protein intensities with a threshold distinguishing "control" and "stress." Théron et al. (2020) combined spectral protein fingerprints acquired by mass spectrometry and infrared to build a statistical model to distinguish the "control" and "defect" classes. To illustrate that boundaries between empirical and data-driven categories can be fuzzy, they did incorporate some supervised machine learning to find their input marker. Both animal stress levels and muscle defects predictions rely on practical noninvasive sampling: for instance, blood obtained in vivo (as opposed to organs and muscles). Identifying a precise and accessible indicator can promote short-term innovation.

3.1.3 | Limitations and relationship to mechanistic prediction

Prediction through empirical modeling does not necessarily need causation. However, this approach has limited genericity, as any intervention on the system is likely to remove the observed correlation. Moreover, as the basis for the correlation is unknown, there is no simple way to identify whether an intervention is safe. Consequently, it is difficult to translate the obtained correlations from one problem to another (Garcia-Maraver et al., 2017). This lack of genericity means a new set of experiments is likely needed whenever a factor gets modified, thus effectively starting over from near-zero and accumulating costs over time.

In situations where experiments are not practicable, a mechanistic understanding is often necessary to transpose the problem into another problem frame (e.g., time and size scaling).

Progress in expert knowledge and understanding of the phenomenon leads to hybrid empirical–mechanistic prediction (Hanigan et al., 2018; Lyu et al., 2020; Święch, 2017; Trottier & Tedeschi, 2019).

When experimental measurements accumulate, are repeated and validated, we tend to extract laws by inductive reasoning, thus giving birth to empirical laws. Like mathematical conjectures, these laws are suspected to be true due to the supporting evidence and absence of disproof, yet lack proof. The multitude of empirical laws has enabled the development of mechanistic prediction, which minimizes the need for experimental validation without ever succeeding entirely and confidently eliminating it. The following section discusses this type of prediction's current state of the art.

3.2 | Mechanistic predictions

The main steps towards mechanistic prediction on bioresource and food products covered in this manuscript are sketched out in Figure 2. The process from understanding the mechanisms to integrating the

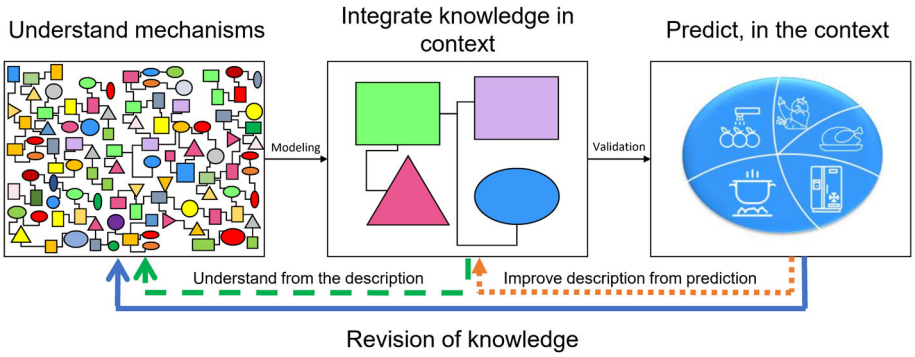

FIGURE 2 Process of knowledge revision. The prediction is progressively achieved (solid black arrows) by considering the context (orange dotted arrow) and mechanisms (solid blue arrow), the latter being possibly influenced by the former (green dashed arrow).

knowledge is progressive. It involves firstly reducing the model to only the most relevant mechanical building blocks and then evaluating the performance and limits before increasing complexity. Prediction successes and failures help forge a more accurate description of the context or understanding of the mechanisms. Furthermore, an updated description of the context may also inform on the physicochemical mechanisms.

3.2.1 | Description and advantages

Conservation laws, including mass/species, energy/heat or momentum/flow, constrain mechanistic predictions to only feasible or probable predictions. As mechanistic modeling is systemic, mechanistic predictions can apply to any system, regardless of its size, from the product scale to the entire planet. The time-frame can cover fast molecular events up to a shelf-life or the bioaccumulation of chemicals in the human body or the environment (Ho et al., 2013). Computational costs can rise quickly but this can be mitigated by associating and chaining simulations. A prediction or guess at one scale can be used hierarchically at the next scale up. Under similar conditions of independence, some mechanistic predictions can be cumulated for several steps or calculated stochastically, such as conversion rates in chemical reactions, microorganism destruction rates, or contamination rates (Vitrac & Hayert, 2005).

With the increase in computing power, mechanistic prediction moved from the experts' hands to practitioners' hands. Modern numerical simulation software are able to leave the user free to focus on the description of the problem and the choice of the physical, chemical or biological parameters. The postprocessing to transform solution fields (two-/three-dimensional [2D/3D] concentration/velocity/temperature profiles) into desired scalars (amounts, fluxes, and rates), kinetics, surface plots and movies may involve some mathematical treatment (interpolation and integration) and coding. Due to its multiple capabilities, mechanistic predictions offer the most extensive spectrum of predictions, including the capacity to reconstruct an entire real scene such as cooking or thawing in a microwave oven (Datta, 2016) that requires a combination of electromagnetism, heat, and mass transfer. The process, the product, the transformations and reactions inside the product, and the characteristics of the final product (composition, structure) can be studied separately or together.

A methodology is required to build predictive models of complex systems that include many components and phases that may evolve with time. One strategy is to combine models from different scales, such as microscopic and macroscopic models (Vitrac & Touffet, 2019), to achieve the predictions. This approach decomposes into small unit mechanisms, some well-known and others unknown. An iterative process may be necessary to achieve this goal. A recent application of this methodology is described in Section 3.2.2.

3.2.2 | Techniques used and application areas

Mechanistic predictions for solid foods and their characteristics (structure/texture, composition, nutritional value, functional value, sensory perception) during processing (drying, cooking, frying, freezing, gelation) are comparatively less developed than mechanical industries models and require more experimental validation. The chief difficulty is that the properties of solid raw food materials are highly variable and can change significantly during processing. Heat and mass transfer properties significantly change with composition (e.g., during drying) while mechanical properties evolve with structure (Datta, 2016). Shrinkage during drying/cooking, chilling injury and cell damage during freezing are challenging to capture. Image-based modeling is a powerful approach to describe solids at pore level or smaller. Examples cover realistic simulation of artificial bone cultures (Alam et al., 2016), multiphasic oil flow in French fries (Touffet et al., 2020; Vauvre et al., 2015), gas transport in fruits (Ho et al., 2010) and mechanical properties of 3D-printed cookies (Piovesan et al., 2020). Combining mechanistic predictions at different scales offers a pathway to engineering new foods (Roos et al., 2016) or develop eco-designed food packaging (Zhu et al., 2019).

Reaction–diffusion systems are widely used to predict the biochemistry of foods during storage, processing, or consumption. The completeness of the chemical reaction system varies between models, but most include at least one chemical reaction and the diffusion of an involved species. As the name "reaction–diffusion system" implies, it is already a collection of predictions, with the reaction and diffusion parts being independently tested as much as possible. As practicable, the reaction system should be defined independently of space, in zero dimension, and the applicable diffusion laws should ideally be defined without any reaction involved. The two tested models are only

coupled afterwards. Dissociating the phenomena can prove a complex task. So practitioners may find it necessary to admit hypotheses on one of the components, such as mass transport that follow Fickian diffusion or chemical reactions that follow Arrhenius law, to reduce the field of possibilities to a level suitable to the available data and accessible testing situations. In such instances, we recommend choosing hypotheses that hold valid in most—if not all—cases (anomalous diffusion, enzymatic reactions). Many examples of reaction–diffusion systems illustrate the biochemistry prediction of foods. During food storage, humidity (Zhao et al., 2019) and oxygen are major factors in the loss of product qualities, whether directly, such as produce respiration (fruits and vegetables) (Badillo & Segura-Ponce, 2020) or oxidation leading to the color shift of myoglobin in red meat, or through microorganisms (Kern et al., 2021). The very concept of transformation equates to change, and for food, biochemical reactions can play a large part in it, for example, oxidation of polyunsaturated fats at high temperature (Touffet et al., 2021) or Maillard reactions (Kocadagli & Gokmen, 2016). Likewise, digestion can be viewed as a series of transformations in which acidification is combined with the enzymatic action of pepsin in the stomach (Sicard et al., 2018).

Mechanistic predictions have been applied to a wide variety of food engineering situations (three examples in Table 1). The first example, reported by Teixeira et al. (1969), was a pioneer in food engineering. Bacterial destruction in cans was inferred numerically from the solution of the heat conduction equation in a cylinder. The reliability of the predictions led to extensions in other aspects of food safety and quality, as shown in examples 2 and 3 of Table 1.

For example, plate heat exchangers used in milk thermal treatments are prone to rapid fouling, and clean-in-place cycles produce large amounts of waste. After decades of research, fouling prediction is still a challenge for the milk industry to manage optimal production. It has been shown that knowledge of the denaturation kinetics helps predict fouling mass and should be more frequently determined (Bansal & Chen, 2006; Blanpain-Avet et al., 2016; Donato & Guyomarch, 2009; Loveday, 2016), especially for building a robust predictive deposition model. The recent model proposed by Sharma and Macchietto (2021) and improved by Alhuthali et al. (2022) offers a good compromise between computational time and prediction' quality. Figure 3 presents the scale-progression strategy applied to fouling prediction in plate exchangers. This dynamic 2D model introduced sophisticated condition-based logic in each phase and in fine enables detailed simulation of fouling and cleaning kinetics through a combination of equation-based models at different scales. Dimensionless process relations are used to solve the computationally time-consuming flow and heat transfer in the bulk. Equations were solved using differential methods for heat transfer and deposit growth in the plates and deposit layers. These models were obtained by going down dimension scales to understand critical details of the mechanisms involved. Fouling or cleaning models were refined by computational fluid dynamics simulations and molecular representations of protein aggregation. These models showed that the heating/cleaning cycle can be optimized for maximum productivity to find trade-offs between fouling and cleaning. The model was validated using various experimental data (temperature and pressure-drop, global and local mass distributions per channel) for a large variety of dairy derivatives and operating conditions. Computation time takes only a few minutes. The model has proven effective for optimizing the operation of complete plate heat exchangers, including typical heating and cleaning policies. Predicting deposit distribution on the plates with a high degree of confidence is also helpful for monitoring, diagnosis, and control purposes, and is now directly accessible to the industry.

TABLE 1 Examples of prediction in foods.

Prediction problem	Example	Risk	Benefits
Exact quantitative value $P = R$ or $P \approx R$	1. The core temperature of a product during sterilization to estimate the minimum sterilization value. (Teixeira et al., 1969)	Consumer hazard	The predictions replace difficult temperature measurements during retort sterilization and/or spore viability measurement and counting. The method paved the way to new sterilization processes: high-temperature short-time processing preserves nutrients and vitamins better while providing better food safety.
Conservative estimate with a safety margin (SF) P > R or P-R > SF Alternatively, probability p(R-P) < SF% (e.g., 5%)	2. Mass transfer of a packaging substance into food. (Vitrac & Hayert, 2005).	Consumer hazard Product recall	The predictions replace time-consuming testing (10 days) and enable extrapolation to realistic conditions (time and temperature). Only available method to evaluate recycled plastic materials in contact with food. Safety considerations and sustainability are integrated numerically for rapid prototyping of food packaging.
Causality B is an effect of A A is a cause of B The cause may be contributory, necessary, and sufficient.	3. Cause of oil uptake during deep frying and product removal from fryer. (Touffet et al., 2020)	Unhealthy food	Input for new designs (e.g., modified deep-fryers), process control strategies or formulations to reduce the cause (e.g., low-fat products).

Notes: The Prediction (P) can have various relationships to Reality (R). This table illustrates some of these possible relationships. For each case, an example is provided with the risk associated with the error in prediction and the potential benefits of correct prediction.

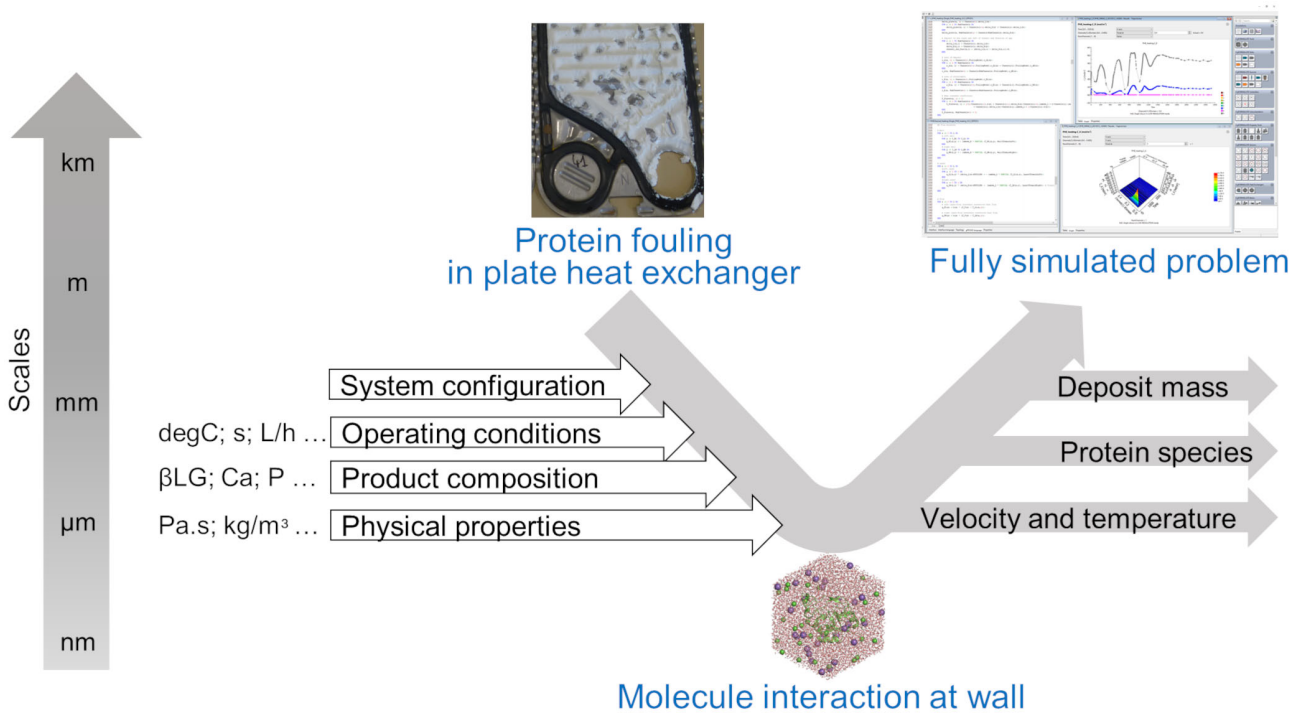

FIGURE 3 Scale-progression strategy. Also called "V strategy" in mechanistic simulations, applied to protein fouling in plate heat exchangers. Working down from the process conditions, the model describes finer and finer mechanisms, potentially down to molecular scale, until it informs on the problem. Understanding the finest-scale phenomena is then scaled back up toward process scale to evaluate the consequences and predict desired outcomes. Relevant outputs can be obtained at each step of the upscaling strategy.

3.2.3 | Limitations

Mechanistic prediction tools are developed in settings that are sometimes far removed from the transformation of bioproducts, so they require substantial adaptation efforts. Many of the developments have been obtained from various collaborative research contracts with either academics or the industry. This construction limits the dissemination and sharing of the latest data and models, creating a bottleneck for generic developments.

Note that generic commercial or open-source software packages largely fail to consider unit operations specific to food, biochemical or microbiological reactions for bio-substrates/food, macroscopic properties (e.g., food) and microscopic geometries (e.g., fibers) are poorly described. Physical, physicochemical, and thermodynamic properties remain under-reported. The natural variability of the raw materials and the evolution of the processed material also add a layer of complexity.

Among mechanical, heat, and mass transfer problems, fracture, mixing, wetting, phase separation and multiphasic flow problems with dynamic and free surfaces are not readily manageable using the numerical strategies cited above, as they involve complex descriptions at more than one length or time scale. Several extensions using particle or Lagrangian descriptions have been proposed to describe mass transfers at the food-structure scale (Vitrac & Hayert, 2020) or the coupling between oxidation reactions and anisothermal oil flow (Touffet et al., 2021). It is still an active field of research, and the available software packages are still under-intuitive and require some level of coding.

Mechanistic prediction relies on theory and physical laws to describe physical, mechanical, chemical, biochemical, and biological phenomena. So they generally struggle to incorporate craft know-how, and critical phenomena associated with food consumption, such as food perception and consumer choices, are still out of reach.

3.3 | Data-driven predictive models

3.3.1 | Description and advantages

As computational capabilities increase and more data become available, automated techniques for prediction are slowly gaining popularity. These approaches are suited for applications where the underlying fundamental mechanisms are unclear or multiple physical and chemical phenomena co-occur and lead to nontrivial interactions. The unifying idea behind these approaches is to transform the prediction problem into an optimization problem and solve it by resorting to established optimization techniques.

The classical paradigm for obtaining data-driven predictive models in this domain is illustrated in Figure 4. Meaningful features can be extracted, designed, or inferred from raw data. The resulting data set containing these features will then be split into a training and validation set to help the practitioner tune the hyperparameters of

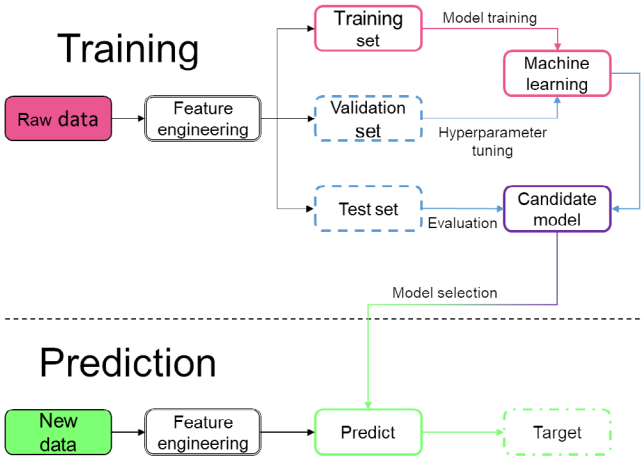

FIGURE 4 From training to prediction in machine learning. Paradigm for obtaining predictive models through machine learning and achieving predictive capability. Data-driven approaches generally require a large initial data set (pink-filled box). Data transformation or feature extraction (black-bordered box) is helpful to get more descriptive inputs that are potentially related more directly to the target. The steps to obtain models are indicated (pink-bordered box) in the "training" part. Evaluations on a validation set are often used to tune the hyperparameters of the method, while experiments on a test set are usually the final step to validate the model (dashed-blue boxes). Prediction happens when the selected model is applied to new, unseen data (green-filled box) to achieve the target (green dash-dotted box).

the machine learning method. Once the parameters are adjusted, a final check will be performed on a separate test set, and a candidate-tuned model is selected to perform predictions on unseen data.

Machine learning describes a set of computer algorithms that automatically create models and improve their performance if given more training data to learn from (Samuel, 1959). Machine learning algorithms can be applied to various tasks. The traditional taxonomy divides them into supervised, unsupervised, semi-supervised, and reinforcement learning, depending on the availability of training labels and how the training samples are generated. In the present article, we focus primarily on supervised machine learning and, particularly, on "classification" and "regression" tasks.

Machine learning techniques range from simple paradigms like linear (Pearson, 1896) and logistic regression (Cox, 1958; Gu et al., 2007) or decision trees (Breiman et al., 1984) to more complex techniques such as random forests (Breiman, 2001; Jimenez-Carvelo et al., 2019), support vector machines (Bahamonde et al., 2007; Platt, 1999) and symbolic regression (Schmidt & Lipson, 2009), and on to artificial neural networks (Hinton, 1990; Yang & Chen, 2022). New developments known as *deep learning* (Goodfellow et al., 2016) have made it possible to create larger, more effective models, although they usually require vastly more data to be properly trained.

The past decade has seen remarkable advances in deep learning fueled by big data and powerful computational resources (in particular graphics processing units), which have revolutionized many fields in science. Deep learning models are artificial neural networks with several connected layers that can learn multiple levels of data abstractions. The most common frameworks learn how to transform inputs into features amenable to predicting the corresponding outputs in a supervised way. The mapping from inputs to outputs is learned from a sizeable input–output data set by tuning each layer's parameters via a back-propagation algorithm. Once trained, the network can be used to make predictions on new and unseen data.

In deep learning, the computers learn to build useful features automatically, directly from the data. However, the automatic feature extraction carried out by deep learning methods is costly, both in computational resources and in the amount of training data. Popular examples of deep learning architectures are feed-forward neural networks (Mwaura & Kenduiywo, 2021; Rumelhart et al., 1986), convolutional neural networks (Kamilaris & Prenafeta-Boldu, 2018; Le Cun et al., 1990), recurrent neural networks (Mandic & Chambers, 2001; Okada et al., 2016), and generative adversarial networks (Goodfellow et al., 2014; Rawal et al., 2021), but there are many more.

3.3.2 | Techniques used and application areas

Predictive machine learning techniques have been widely adopted across various domains, with several success stories for bioresource transformation processes.

Techniques ranging from artificial neural networks to fuzzy logic have been successfully applied to the forecasting of fruit sensory properties (Salehi, 2020) and ripeness (B. Li et al., 2018) and fruit micronutrient content (Huang et al., 2021). Likewise, the prediction of organic solid waste treatment outcomes has been tackled using artificial neural networks, support vector machines, and ensembles of decision trees (Guo et al., 2021). Artificial neural networks have also been applied to predict higher heating values in fuels obtained from renewable biomasses (Vardiambasis et al., 2020). In other cases, machine learning models can replace costly experiments by predicting features of hypothetical products: Szczypiński et al. (2021), for example, reviews machine learning approaches to screening synthesizable organic materials.

Machine learning techniques can also provide complex metrics, such as tenderness, firmness, and springiness, from noninvasive hyperspectral imaging analysis of protein-rich foods (Cheng et al., 2017). Recent developments in deep learning also paved the way for generative models able to produce images or text. In an elegant application to food processes, Chen et al. (2019) predicted the appearance of laser-browned dough for cookies.

The detailed example concerns the deep learning methods have also brought significant advances in protein structure modeling and engineering (Gao et al., 2020; Ovchinnikov & Huang, 2021; Pearce & Zhang, 2021; Torrisi et al., 2020), which is crucial for developing biotransformation processes. Deciphering the structure–function relationships of proteins and designing macromolecules such as enzymes with novel optimized capabilities is essential to developing innovative and efficient bioprocesses for biotechnologies. The recent application of deep neural networks to protein structure prediction has

revolutionized the field. The latest version of AlphaFold (AlphaFold2; Jumper et al., 2021), developed by the Google DeepMind team has achieved unprecedented modeling accuracy in the challenging 14th Critical Assessment of protein Structure Prediction (CASP14), demonstrating its ability to predict protein structure with atomic-scale accuracy even in cases where no similar structure is known. This deep learning approach is competitive with experimental approaches and outperforms all other computational methods. It incorporates physical and biological knowledge about protein structure and exploits multi-sequence alignments. In parallel, for the protein design problem—that can be conceptually regarded as the inverse of protein structure prediction—advances in the field of deep learning have led to a promising set of approaches. Exciting results have been achieved for the redesign or de novo design of proteins by exploiting sequence or structure protein data (Ovchinnikov & Huang, 2021; Wu et al., 2021). Deep learning-based approaches can be expected to dramatically accelerate and improve success rates in protein engineering.

3.3.3 | Limitations

From a theoretical point of view, the main drawback of machine learning is arguably *overfitting*, that is, the tendency to create models that fit the training data too closely, thus involuntarily including a portion of the residual variation/noise and ultimately leading to models that deliver poor predictions on unseen data. A precise measure of the overfitting of a trained model can only be gained once new data becomes available, but it is still possible to obtain an estimate by splitting the available data into parts and training the algorithms on some parts while testing them on others that were unseen during the training process. Popular techniques for this aim include k-fold and leave-one-out cross-validation.

In most practical cases, another critical issue is getting available data in sufficient quantity and quality. Large data sets are not enough: the data should also represent all possible variations of a process of interest. A machine learning model will extrapolate poorly to unseen conditions if the available data were not representative enough. Statistical techniques such as bootstrapping (Efron, 1992) and conformal predictors (Balasubramanian et al., 2014) can indirectly estimate the reliability of the training data by computing confidence intervals around predictions. Tight confidence intervals, indicating reliable predictions, often come from relevant and representative training data.

Finally, even when the predictions provided by machine learning are satisfactory, the black-box nature of the models may prove a further drawback. Most machine learning models, especially deep learning ones, are so intricate and contain so many parameters that understanding why a model delivers a specific result which is often beyond human capabilities. However, knowing the reason behind a given prediction can be as fundamental as having a correct prediction for many real-world applications. Explainable artificial intelligence algorithms "open" the black boxes and attempt to provide at least some explanations for decisions made by machine learning models (see Adadi & Berrada, 2018 for a recent review of the field).

4 | FUTURE TRENDS

The accessibility of prediction methods, primarily mechanistic and machine learning prediction, continues to grow as more and more software solutions become available, opening up opportunities for everyone to run a prediction—including novices. Hence, prediction progressively shapes both food and bioproduct technologies and public policies. This newfound appreciation goes along with legitimate concerns regarding validity and means of control. To this end, a standardization of practices would be greatly beneficial.

Furthermore, new and current users should keep questioning the suitability of prediction even as it encompasses many applications. The areas where prediction is relevant will evolve as technologies evolve: new domains of application appear while others become obsolete. This is the case for microstructures and nanostructures previously "seen" through predictions. In contrast, today, it may be faster to observe them with an advanced microscope than to run predictions without preexisting models. Mechanistic models are liable to generate realistic outputs, but it is not easy to obtain these outputs. It is essential to understand when prediction is worthwhile.

Likewise, as needs evolve, it will sometimes be more reasonable to start with a new simple model rather than complexify a previous one.

There are good practices to avoid poor model design and misuse of models to influence decision-making. Verifying predictive capability (consistency, stability, accuracy, validity, and suitability for decision-making problems) reduces—but does not eliminate—the risk of bad decisions. Misuse of models and risk-taking bias in human decision-making must also be minimized by proper training of end-users. In the context of bioresources and food products, generalizing or extrapolating the outputs of simple models beyond the initial explored domain can lead to unreliable predictions. Conversely, coupling models from various fields of study and at different spatial and temporal scales may build resilience to changing contexts (e.g., for new raw materials, new process routes, or environmental conditions) and allows for systemic predictions. Fernandez-Mena et al. (2020) proposed innovative tools to explore the co-benefits and trade-offs of improving flow circularity in agro-food systems at different scales. Circularity and environmental performance were studied by implementing a network approach combining agent-based model showing how different sustainability solutions across a farming region could result in very different outcomes.

The so-called take-make-dispose paradigm in the agri-food sector has several limitations, including environmental degradation, resource depletion, health and safety concerns, and socioeconomic impacts (Esposito et al., 2020; Falcone et al., 2022). Innovation toward circularity involves collaboration between stakeholders (Bloise, 2019), policy changes, technological advancements, changes in consumer behavior, and business models shift (De Bernardi et al., 2022). Hybridizing the prediction capabilities is the key to reaching complex sustainable goals: resource forecasting, waste reduction, supply chains and preventive maintenance optimization.

The concept of digital twins, virtual replica of the real process operation (Verboven et al., 2020), offers a robust approach for

merging real-time data feeds and making linked decisions based on multiple criteria, in order to organize food manufacturing (Datta et al., 2022; Defraeye et al., 2021; Mangers et al., 2023). However, the fusion of measurement, prediction, and decision-support tools is still in its infancy. Further case-specific implementations must be reported to establish priorities and standards (Preut et al., 2021).

Institutions and industry working together to achieve the United Nations goal of ensuring sustainable consumption and production patterns require multidisciplinary contributions. To achieve this, lifecycle analysis can be combined with predictions from technoeconomic, process, and system dynamics models (Yao, 2017). Due to the availability of data and resources, the current contribution of analytics to achieving development goals varies by region, but the use of the presented methods is a common factor (Orhan & Guajardo, 2022). Emerging food trends, such as fortified and functional foods, additive manufacturing technologies, cultured meat, precision fermentation, and personalized food, will rely on massive amounts of information from and to the consumer that is not currently available (Hassoun et al., 2022).

Whatever the prediction process followed, every step may lead to uncertainties, and the decision-making should consider these uncertainties. It is good practice to directly evaluate these uncertainties in the modeling phase (Acar et al., 2021; Peng & Zhao, 2009). However this can prove a difficult task, especially when elementary models from different scientific branches are linked together (e.g., microbiology with heat and mass transfer in food processes) and the propagation pathways become complex (Kirchner et al., 2021). Safety factors and worse-case minimization are indirect ways to deal with uncertainty.

Ultimately, no measure can guarantee prediction accuracy. You may make no mistakes yet still fail. The effort must be proportionate to the cost of failure, and there must always be provision for risk management, so that you can try again.

AUTHOR CONTRIBUTIONS

Jason Sicard: conceptualization; visualization; writing – original draft; writing – review & editing. **Sophie Barbe**: conceptualization; visualization; writing – original draft; writing – review & editing. **Rachel Boutrou**: conceptualization; supervision; writing – original draft; writing – review & editing. **Laurent Bouvier**: conceptualization; visualization; writing – original draft; writing – review & editing. **Guillaume Delaplace**: conceptualization; visualization; writing – original draft; writing – review & editing. **Gwenaëlle Lashermes**: conceptualization; visualization; writing – original draft; writing – review & editing. **Laëtitia Théron**: conceptualization; visualization; writing – original draft. **Olivier Vitrac**: conceptualization; visualization; writing – original draft; writing – review & editing. **Alberto Tonda**: conceptualization; visualization; writing – original draft; writing – review & editing.

ACKNOWLEDGMENTS

The authors thank Véronique Santé-Lhoutellier and Elie Desmond-Le Quemener for their active participation in preliminary discussions leading to the writing of this review.

FUNDING INFORMATION

No funding was received for this manuscript.

CONFLICT OF 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.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no data sets were generated or analyzed during the current study.

ORCID

Jason Sicard https://orcid.org/0000-0002-7197-8212
Rachel Boutrou https://orcid.org/0000-0002-8373-1831
Laurent Bouvier https://orcid.org/0000-0003-1873-4179

REFERENCES

Acar, E., Bayrak, G., Jung, Y., Lee, I., Ramu, P., & Ravichandran, S. S. (2021). Modeling, analysis, and optimization under uncertainties: A review. *Structural and Multidisciplinary Optimization*, *64*, 2909–2945. https://doi.org/10.1007/s00158-021-03026-7

Acs, S., Ostlaender, N., Listorti, G., Hradec, J., Hardy, M., Smits, P., & Hordijk, L. (2019). *Modelling for EU policy support: Impact assessments*. Publications Office of the European Union. https://doi.org/10.2760/748720

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, *6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Adisa, O. M., Masinde, M., Botai, J. O., & Botai, C. M. (2020). Bibliometric analysis of methods and tools for drought monitoring and prediction in Africa. *Sustainability*, *12*(16), 6516. https://doi.org/10.3390/su12166516

Alam, T. A., Pham, Q. L., Sikavitsas, V. I., Papavassiliou, D. V., Shambaugh, R. L., & Voronov, R. S. (2016). Image-based modeling: A novel tool for realistic simulations of artificial bone cultures. *Technology*, *4*(4), 229–233. https://doi.org/10.1142/S233954781620003X

Alhuthali, S., Delaplace, G., Macchietto, S., & Bouvier, L. (2022). Whey protein fouling prediction in plate heat exchanger by combining dynamic modelling, dimensional analysis, and symbolic regression. *Food and Bioproducts Processing*, *134*, 163–180. https://doi.org/10.1016/j.fbp.2022.05.009

Badillo, G. M., & Segura-Ponce, L. A. (2020). Classic and reaction-diffusion models used in modified atmosphere packaging (MAP) of fruit and vegetables. *Food Engineering Reviews*, *12*(2), 209–228. https://doi.org/10.1007/s12393-020-09214-3

Bahamonde, A., Díez, J., Quevedo, J. R., Luaces, O., & del Coz, J. J. (2007). How to learn consumer preferences from the analysis of sensory data by means of support vector machines (SVM). *Trends in Food Science & Technology*, *18*(1), 20–28. https://doi.org/10.1016/j.tifs.2006.07.014

Bakan, B., Bernet, N., Bouchez, T., Boutrou, R., Choubert, J.-M., Dabert, P., Duquennoi, C., Ferraro, V., García-Bernet, D., Gillot, S., Mery, J., Rémond, C., Steyer, J.-P., Trably, E., & Tremier, A. (2022). Circular economy applied to organic residues and wastewater: Research challenges. *Waste and Biomass Valorization*, *13*(2), 1267–1276. https://doi.org/10.1007/s12649-021-01549-0

Balasubramanian, V. N., Ho, S.-S., & Vovk, V. (2014). *Conformal prediction for reliable machine learning. Theory, adaptations and applications (Elsevier)*. Elsevier Inc.

Bansal, B., & Chen, X. D. (2006). A critical review of milk fouling in heat exchangers. *Comprehensive Reviews in Food Science and Food Safety*, *5*(2), 27–33. https://doi.org/10.1111/j.1541-4337.2006.tb00080.x

Bauwens, T., Hekkert, M., & Kirchherr, J. (2020). Circular futures: What will they look like? *Ecological Economics*, *175*, 106703. https://doi.org/10.1016/j.ecolecon.2020.106703

Bay, R. A., Rose, N., Barrett, R., Bernatchez, L., Ghalambor, C. K., Lasky, J. R., Brem, R. B., Palumbi, S. R., & Ralph, P. (2017). Predicting responses to contemporary environmental change using evolutionary response architectures. *The American Naturalist*, *189*(5), 463–473. https://doi.org/10.1086/691233

Bekenev, V. A. (2019). Productive longevity of animals, methods of its prediction and extension. *Sel'skokhozyaistvennaya Biologiya*, *54*(4), 655–666. https://doi.org/10.15389/agrobiology.2019.4.655eng

Bersani, C., Ouammi, A., Sacile, R., & Zero, E. (2020). Model predictive control of smart greenhouses as the path towards near zero energy consumption. *Energies*, *13*(14), 3647. https://doi.org/10.3390/en13143647

Blanpain-Avet, P., André, C., Khaldi, M., Bouvier, L., Petit, J., Six, T., Jeantet, R., Croguennec, T., & Delaplace, G. (2016). Predicting the distribution of whey protein fouling in a plate heat exchanger using the kinetic parameters of the thermal denaturation reaction of β-lactoglobulin and the bulk temperature profiles. *Journal of Dairy Science*, *99*(12), 9611–9630. https://doi.org/10.3168/jds.2016-10957

Bloise, C. (2019). Collaboration in a circular economy: Learning from the farmers to reduce food waste. *Journal of Enterprise Information Management*, *33*(4), 769–789. https://doi.org/10.1108/JEIM-02-2019-0062

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC Press.

Bresolin, T., & Dorea, J. R. R. (2020). Infrared spectrometry as a high-throughput phenotyping technology to predict complex traits in livestock systems. *Frontiers in Genetics*, *11*, 923. https://doi.org/10.3389/fgene.2020.00923

Chen, P. Y., Blutinger, J. D., Meijers, Y., Zheng, C., Grinspun, E., & Lipson, H. (2019). Visual modeling of laser-induced dough browning. *Journal of Food Engineering*, *243*, 9–21. https://doi.org/10.1016/j.jfoodeng.2018.08.022

Cheng, J.-H., Nicolai, B., & Sun, D.-W. (2017). Hyperspectral imaging with multivariate analysis for technological parameters prediction and classification of muscle foods: A review. *Meat Science*, *123*, 182–191. https://doi.org/10.1016/j.meatsci.2016.09.017

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society Series B (Methodological)*, *20*, 215–232.

Datta, A. K. (2016). Toward computer-aided food engineering: Mechanistic frameworks for evolution of product, quality and safety during processing. *Journal of Food Engineering*, *176*, 9–27. https://doi.org/10.1016/j.jfoodeng.2015.10.010

Datta, A. K., Nicolaï, B., Vitrac, O., Verboven, P., Erdogdu, F., Marra, F., Sarghini, F., & Koh, C. (2022). Computer-aided food engineering. *Nature Food*, *3*(11), 894–904. https://doi.org/10.1038/s43016-022-00617-5

De Bernardi, P., Bertello, A., & Forliano, C. (2022). Circularity of food systems: A review and research agenda. *British Food Journal*, *125*, 1094–1129. https://doi.org/10.1108/BFJ-05-2021-0576

Defraeye, T., Shrivastava, C., Berry, T., Verboven, P., Onwude, D., Schudel, S., Bühlmann, A., Cronje, P., & Rossi, R. M. (2021). Digital twins are coming: Will we need them in supply chains of fresh horticultural produce? *Trends in Food Science & Technology*, *109*, 245–258. https://doi.org/10.1016/j.tifs.2021.01.025

Donato, L., & Guyomarch, F. (2009). Formation and properties of the whey protein/κ-casein complexes in heated skim milk—A review. *Dairy Science & Technology*, *89*(1), 3–29. https://doi.org/10.1051/dst:2008033

Ealy, A. D., & Seekford, Z. K. (2019). Symposium review: Predicting pregnancy loss in dairy cattle. *Journal of Dairy Science*, *102*(12), 11798–11804. https://doi.org/10.3168/jds.2019-17176

Efron, B. (1992). Bootstrap methods: Another look at the jackknife. In S. Kotz & N. L. Johnson (Eds.), *Breakthroughs in statistics: Methodology and distribution* (pp. 569–593). Springer. https://doi.org/10.1007/978-1-4612-4380-9_41

Erdogdu, F., Datta, A. K., Vitrac, O., Marra, F., Verboven, P., Sarghini, F., & Nicolai, B. (2022). Chapter 16: Mathematical modeling—Computer-aided food engineering. In P. Juliano, R. Buckow, M. H. Nguyen, K. Knoerzer, & J. Sellahewa (Eds.), *Food engineering innovations across the food supply chain* (pp. 277–290). Elsevier Inc Inc. https://doi.org/10.1016/B978-0-12-821292-9.00007-8

Esposito, B., Sessa, M. R., Sica, D., & Malandrino, O. (2020). Towards circular economy in the agri-food sector. *A Systematic Literature Review. Sustainability*, *12*(18), 7401. https://doi.org/10.3390/su12187401

European Commission. (2022). *Bioeconomy*. Retrieved January 27, 2023, from. https://research-and-innovation.ec.europa.eu/research-area/environment/bioeconomy_en

European Commission & Directorate-General for Research and Innovation. (2017). *Review of the 2012 European bioeconomy strategy*. Retrieved from. http://dx.publications.europa.eu/10.2777/8814

Falcone, G., Stillitano, T., Iofrida, N., Spada, E., Bernardi, B., Gulisano, G., & De Luca, A. I. (2022). Life cycle and circularity metrics to measure the sustainability of closed-loop Agri-food pathways. *Frontiers in Sustainable Food Systems*, *6*, 1014228. https://doi.org/10.3389/fsufs.2022.1014228

Fayeulle, N., Meudec, E., Boulet, J. C., Vallverdu-Queralt, A., Hue, C., Boulanger, R., Cheynier, V., & Sommerer, N. (2019). Fast discrimination of chocolate quality based on average-mass-spectra fingerprints of cocoa polyphenols. *Journal of Agricultural and Food Chemistry*, *67*(9), 2723–2731. https://doi.org/10.1021/acs.jafc.8b06456

Fernandez-Mena, H., MacDonald, G. K., Pellerin, S., & Nesme, T. (2020). Co-benefits and trade-offs from agro-food system redesign for circularity: A case study with the FAN agent-based model. *Frontiers in Sustainable Food Systems*, *4*, 41. https://doi.org/10.3389/fsufs.2020.00041

Gao, W., Mahajan, S. P., Sulam, J., & Gray, J. J. (2020). Deep learning in protein structural modeling and design. *Patterns*, *1*(9), 100142. https://doi.org/10.1016/j.patter.2020.100142

Garcia-Maraver, A., Mata-Sanchez, J., Carpio, M., & Perez-Jimenez, J. A. (2017). Critical review of predictive coefficients for biomass ash deposition tendency. *Journal of the Energy Institute*, *90*(2), 214–228. https://doi.org/10.1016/j.joei.2016.02.002

Gillet, G., Vitrac, O., Tissier, D., Saillard, P., & Desobry, S. (2009). Development of decision tools to assess migration from plastic materials in contact with food. *Food Additives & Contaminants: Part A*, *26*(12), 1556–1573. https://doi.org/10.1080/19440040903271355

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative adversarial nets Advances in Neural Information Processing Systems* (Vol. 27). Curran Associates, Inc.

Gu, C.-H., Li, H., Levons, J., Lentz, K., Gandhi, R. B., Raghavan, K., & Smith, R. L. (2007). Predicting effect of food on extent of drug absorption based on physicochemical properties. *Pharmaceutical Research*, *24*(6), 1118–1130. https://doi.org/10.1007/s11095-007-9236-1

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Guo, H., Wu, S., Tian, Y., Zhang, J., & Liu, H. (2021). Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: A review. *Bioresource Technology*, *319*, 124114. https://doi.org/10.1016/j.biortech.2020.124114

Haan, N. L., Zhang, Y., & Landis, D. A. (2020). Predicting landscape configuration effects on agricultural pest suppression. *Trends in Ecology & Evolution*, *35*(2), 175–186. https://doi.org/10.1016/j.tree.2019.10.003

Hanigan, M. D., White, R. R., Apelo, S. I. A., Aguilar, M., Estes, K. A., & Myers, A. (2018). Predicting post-absorptive protein and amino acid metabolism. *Revista Brasileira de Zootecnia*, *47*, e20160417. https://doi.org/10.1590/rbz4720160417

Hassoun, A., Bekhit, A. E.-D., Jambrak, A. R., Regenstein, J. M., Chemat, F., Morton, J. D., Gudjónsdóttir, M., Carpena, M., Prieto, M. A., Varela, P., Arshad, R. N., Aadil, R. M., Bhat, Z., & Ueland, Ø. (2022). The fourth industrial revolution in the food industry—Part II: Emerging food trends. *Critical Reviews in Food Science and Nutrition*, 1–31. https://doi.org/10.1080/10408398.2022.2106472. Epub ahead of print.

Hinton, G. E. (1990). Connectionist learning procedures. In Y. Kodratoff & S. R. Michalski (Eds.), *Machine learning–an artificial intelligence approach* (Vol. 3, pp. 555–610). Morgan Kaufmann. https://doi.org/10.1016/B978-0-08-051055-2.50029-8

Ho, Q. T., Carmeliet, J., Datta, A. K., Defraeye, T., Delele, M. A., Herremans, E., Opara, L., Ramon, H., Tijskens, E., van der Sman, R., Van Liedekerke, P., Verboven, P., & Nicolaï, B. M. (2013). Multiscale modeling in food engineering. *Journal of Food Engineering*, *114*(3), 279–291. https://doi.org/10.1016/j.jfoodeng.2012.08.019

Ho, Q. T., Verboven, P., Verlinden, B. E., & Nicolaï, B. M. (2010). A model for gas transport in pear fruit at multiple scales. *Journal of Experimental Botany*, *61*(8), 2071–2081. https://doi.org/10.1093/jxb/erq026

Holt, G. M., Owen, L. J., Till, S., Cheng, Y., Grant, V. A., Harden, C. J., & Corfe, B. M. (2017). Systematic literature review shows that appetite rating does not predict energy intake. *Critical Reviews in Food Science and Nutrition*, *57*(16), 3577–3582. https://doi.org/10.1080/10408398.2016.1246414

Horie, T. (2019). Global warming and rice production in Asia: Modeling, impact prediction and adaptation. *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences*, *95*(6), 211–245. https://doi.org/10.2183/pjab.95.016

Huang, X., Wang, H., Luo, W., Xue, S., Hayat, F., & Gao, Z. (2021). Prediction of loquat soluble solids and titratable acid content using fruit mineral elements by artificial neural network and multiple linear regression. *Scientia Horticulturae*, *278*, 109873. https://doi.org/10.1016/j.scienta.2020.109873

Jimenez-Carvelo, A. M., Gonzalez-Casado, A., Bagur-Gonzalez, M. G., & Cuadros-Rodriguez, L. (2019). Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity—A review. *Food Research International*, *122*, 25–39. https://doi.org/10.1016/j.foodres.2019.03.063

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kamilaris, A., & Prenafeta-Boldu, F. X. (2018). A review of the use of convolutional neural networks in agriculture. *Journal of Agricultural Science*, *156*(3), 312–322. https://doi.org/10.1017/S0021859618000436

Kern, C., Stefan, T., Sacharow, J., Kuegler, P., & Hinrichs, J. (2021). Predictive modeling of the early stages of semi-solid food ripening: Spatio-temporal dynamics in semi-solid casein matrices. *International Journal of Food Microbiology*, *349*, 109230. https://doi.org/10.1016/j.ijfoodmicro.2021.109230

Kirchner, M., Mitter, H., Schneider, U. A., Sommer, M., Falkner, K., & Schmid, E. (2021). Uncertainty concepts for integrated modeling—Review and application for identifying uncertainties and uncertainty propagation pathways. *Environmental Modelling & Software*, *135*, 104905. https://doi.org/10.1016/j.envsoft.2020.104905

Kocadagli, T., & Gokmen, V. (2016). Multiresponse kinetic modelling of Maillard reaction and caramelisation in a heated glucose/wheat flour system. *Food Chemistry*, *211*, 892–902. https://doi.org/10.1016/j.foodchem.2016.05.150

Le Cun, Y., Boser, B., Denker, J. S., Howard, R. E., Habbard, W., Jackel, L. D., & Henderson, D. (1990). Handwritten digit recognition with a back-propagation network. In D. S. Touretzky (Ed.), *Advances in neural information processing systems 2* (pp. 396–404). Morgan Kaufmann Publishers Inc.

Li, B., Lecourt, J., & Bishop, G. (2018). Advances in non-destructive early assessment of fruit ripeness towards defining optimal time of harvest and yield prediction-a review. *Plants (Basel, Switzerland)*, *7*(1), 3. https://doi.org/10.3390/plants7010003

Li, H., Xie, G., & Edmondson, A. S. (2008). Review of secondary mathematical models of predictive microbiology. *Journal of Food Products Marketing*, *14*(2), 57–74. https://doi.org/10.1080/10454440801918283

Long, J. A. (2020). The "omits" revolution: Use of genomic, transcriptomic, proteomic and metabolomic tools to predict male reproductive traits that impact fertility in livestock and poultry. *Animal Reproduction Science*, *220*, 106354. https://doi.org/10.1016/j.anireprosci.2020.106354

Lopatkin, A. J., & Collins, J. J. (2020). Predictive biology: Modelling, understanding and harnessing microbial complexity. *Nature Reviews Microbiology*, *18*(9), 507–520. https://doi.org/10.1038/s41579-020-0372-5

Loveday, S. M. (2016). β-Lactoglobulin heat denaturation: A critical assessment of kinetic modelling. *International Dairy Journal*, *52*, 92–100. https://doi.org/10.1016/j.idairyj.2015.08.001

Lyu, F., Thomas, M., Hendriks, W. H., & van der Poel, A. F. B. (2020). Size reduction in feed technology and methods for determining, expressing and predicting particle size: A review. *Animal Feed Science and Technology*, *261*, 114347. https://doi.org/10.1016/j.anifeedsci.2019.114347

Mandic, D. P., & Chambers, J. A. (2001). *Recurrent neural networks for prediction: Learning algorithms, architectures and stability*. John Wiley & Sons, Ltd.

Mangers, J., Amne Elahi, M., & Plapper, P. (2023). Digital twin of end-of-life process-chains for a circular economy adapted product design – A case study on PET bottles. *Journal of Cleaner Production*, *382*, 135287. https://doi.org/10.1016/j.jclepro.2022.135287

Menzel, C. M. (2021). A review of productivity in strawberries: Marketable yield has a linear, but inconsistent relationship with total yield, and cannot be predicted from total yield. *The Journal of Horticultural Science and Biotechnology*, *96*(2), 135–144. https://doi.org/10.1080/14620316.2020.1808086

Mills, N. J., & Heimpel, G. E. (2018). Could increased understanding of foraging behavior help to predict the success of biological control? *Current Opinion in Insect Science*, *27*, 26–31. https://doi.org/10.1016/j.cois.2018.02.013

Mouzo, D., Rodríguez-Vázquez, R., Lorenzo, J. M., Franco, D., Zapata, C., & López-Pedrouso, M. (2020). Proteomic application in predicting food quality relating to animal welfare. A review. *Trends in Food Science & Technology*, *99*, 520–530. https://doi.org/10.1016/j.tifs.2020.03.029

Mueller-Schaerer, H., Bouchemousse, S., Litto, M., McEvoy, P. B., Roderick, G. K., & Sun, Y. (2020). How to better predict long-term benefits and risks in weed biocontrol: An evolutionary perspective. *Current Opinion in Insect Science*, *38*, 84–91.

Mwaura, J. I., & Kenduiywo, B. K. (2021). County level maize yield estimation using artificial neural network. *Modeling Earth Systems and Environment*, *7*(3), 1417–1424. https://doi.org/10.1007/s40808-020-00943-2

Okada, S., Nakamoto, H., Kobayashi, F., & Kojima, F. (2016). A study on classification of food texture with recurrent neural network. In N. Kubota, K. Kiguchi, H. Liu, & T. Obo (Eds.), *Intelligent robotics and applications, Icira 2016, Pt I* (Vol. 9834, pp. 247–256). Springer International Publishing Ag. https://doi.org/10.1007/978-3-319-43506-0_21

Orhan, C. C., & Guajardo, M. (2022). Analytics in developing countries: Methods, applications, and the impact on the UN sustainable development goals. *International Transactions in Operational Research*, *29*(4), 2041–2081. https://doi.org/10.1111/itor.13018

Ovchinnikov, S., & Huang, P.-S. (2021). Structure-based protein design with deep learning. *Current Opinion in Chemical Biology*, *65*, 136–144. https://doi.org/10.1016/j.cbpa.2021.08.004

Pearce, R., & Zhang, Y. (2021). Deep learning techniques have significantly impacted protein structure prediction and protein design. *Current Opinion in Structural Biology*, *68*, 194–207. https://doi.org/10.1016/j.sbi.2021.01.007

Pearson, K. (1896). Mathematical contributions to the theory of evolution. III. Regression, heredity, and Panmixia. *Philosophical Transactions of the Royal Society of London Series A, Containing Papers of a Mathematical or Physical Character*, *187*, 253–318.

Peng, S., & Zhao, X. (2009). A review of uncertainty methods employed in water quality modeling. In *2009 international conference on environmental science and information application technology* (Vol. ii, pp. 32–35). IEEE Computer Society. https://doi.org/10.1109/ESIAT.2009.227

Pereira, L. S., Paredes, P., Melton, F., Johnson, L., Wang, T., Lopez-Urrea, R., Cancela, J. J., & Allen, R. G. (2020). Prediction of crop coefficients from fraction of ground cover and height. Background and validation using ground and remote sensing data. *Agricultural Water Management*, *241*, 106197. https://doi.org/10.1016/j.agwat.2020.106197

Piovesan, A., Vancauwenberghe, V., Aregawi, W., Delele, M. A., Bongaers, E., de Schipper, M., van Bommel, K., Noort, M., Verboven, P., & Nicolai, B. (2020). Designing mechanical properties of 3D printed cookies through computer aided engineering. *Food*, *9*(12), 1804. https://doi.org/10.3390/foods9121804

Platt, J. C. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In A. J. Smola, P. Bartlett, B. Schölkopf, & D. Schuurmans (Eds.), *Advances in large margin classifiers* (pp. 61–74). MIT Press.

Pralle, R. S., & White, H. M. (2020). Symposium review: Big data, big predictions: Utilizing milk Fourier-transform infrared and genomics to improve hyperketonemia management. *Journal of Dairy Science*, *103*(4), 3867–3873. https://doi.org/10.3168/jds.2019-17379

Preut, A., Kopka, J.-P., & Clausen, U. (2021). Digital twins for the circular economy. *Sustainability*, *13*(18), 10467. https://doi.org/10.3390/su131810467

Rawal, A., Rawat, D. B., & Sadler, B. (2021). Recent advances in adversarial machine learning: Status, challenges and perspectives. In T. Pham & L. Solomon (Eds.), *Artificial intelligence and machine learning for multi-domain operations applications iii* (Vol. 11746, 117462Q). SPIE-International Society for Optical Engineering. https://doi.org/10.1117/12.2583970

Roos, Y., Fryer, P. J., Knorr, D., Schuchmann, H. P., Schroën, C. G. P. H., Schutyser, M., Trystram, G., & Windhab, E. J. (2016). Food engineering at multiple scales: Case studies, challenges and the future—A European perspective. *Food Engineering Reviews*, *8*(2), 91–115. https://doi.org/10.1007/s12393-015-9125-z

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Salehi, F. (2020). Recent advances in the modeling and predicting quality parameters of fruits and vegetables during postharvest storage: A review. *International Journal of Fruit Science*, *20*(3), 506–520. https://doi.org/10.1080/15538362.2019.1653810

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://doi.org/10.1147/rd.33.0210

Schlüter, U., Meyer, J., Ahrens, A., Borghi, F., Clerc, F., Delmaar, C., Di Guardo, A., Dudzina, T., Fantke, P., Fransman, W., Hahn, S., Heussen, H., Jung, C., Koivisto, J., Koppisch, D., Paini, A., Savic, N., Spinazzè, A., Jeddi, M. Z., & von Goetz, N. (2022). Exposure modelling in Europe: How to pave the road for the future as part of the European exposure science strategy 2020–2030. *Journal of Exposure Science & Environmental Epidemiology*, *32*(4), 499–512. https://doi.org/10.1038/s41370-022-00455-4

Schmidt, M., & Lipson, H. (2009). Distilling free-form natural laws from experimental data. *Science*, *324*(5923), 81–85. https://doi.org/10.1126/science.1165893

Sharma, A., & Macchietto, S. (2021). Fouling and cleaning of plate heat exchangers: Dairy application. *Food and Bioproducts Processing*, *126*, 32–41. https://doi.org/10.1016/j.fbp.2020.12.005

Shine, P., Upton, J., Sefeedpari, P., & Murphy, M. D. (2020). Energy consumption on dairy farms: A review of monitoring, prediction modelling, and analyses. *Energies*, *13*(5), 1288. https://doi.org/10.3390/en13051288

Sicard, J., Mirade, P.-S., Portanguen, S., Clerjon, S., & Kondjoyan, A. (2018). Simulation of the gastric digestion of proteins of meat bolus using a reaction-diffusion model. *Food & Function*, *9*(12), 6456–6470. https://doi.org/10.1039/c8fo01120f

Stavropoulou, E., & Bezirtzoglou, E. (2019). Predictive modeling of microbial behavior in food. *Food*, *8*(12), 654. https://doi.org/10.3390/foods8120654

Święch, E. (2017). Alternative prediction methods of protein and energy evaluation of pig feeds. *Journal of Animal Science and Biotechnology*, *8*, 39. https://doi.org/10.1186/s40104-017-0171-7

Szczypiński, F. T., Bennett, S., & Jelfs, K. E. (2021). Can we predict materials that can be synthesised? *Chemical Science*, *12*(3), 830–840. https://doi.org/10.1039/D0SC04321D

Teixeira, A. A., Dixon, J. R., Zahradnik, J. W., & Zinmeister, G. E. (1969). Computer optimization of nutrient retention in the thermal processing of conduction-heated foods. *Food Technology*, *23*(16), 138–142.

Théron, L., Sayd, T., Chambon, C., Vautier, A., Ferreira, C., Aubry, L., Ferraro, V., & Santé-Lhoutellier, V. (2020). Toward the prediction of PSE-like muscle defect in hams: Using chemometrics for the spectral fingerprinting of plasma. *Food Control*, *109*, 106929. https://doi.org/10.1016/j.foodcont.2019.106929

Torrisi, M., Pollastri, G., & Le, Q. (2020). Deep learning methods in protein structure prediction. *Computational and Structural Biotechnology Journal*, *18*, 1301–1310. https://doi.org/10.1016/j.csbj.2019.12.011

Touffet, M., Allouche, M. H., Ariane, M., & Vitrac, O. (2021). Coupling between oxidation kinetics and anisothermal oil flow during deep-fat frying. *Physics of Fluids*, *33*(8), 085105. https://doi.org/10.1063/5.0055873

Touffet, M., Trystram, G., & Vitrac, O. (2020). Revisiting the mechanisms of oil uptake during deep-frying. *Food and Bioproducts Processing*, *123*, 14–30. https://doi.org/10.1016/j.fbp.2020.06.007

Trottier, N. L., & Tedeschi, L. O. (2019). Dietary nitrogen utilisation and prediction of amino acid requirements in equids. *Animal Production Science*, *59*(11), 2057–2068. https://doi.org/10.1071/AN19304

Valdramidis, V. P., Cummins, E., & Frías Celayeta, J. M. (2013). Predictive modelling of foodquality and safety. *Food Control*, *29*(2), 289. https://doi.org/10.1016/j.foodcont.2012.05.049

Vardiambasis, I. O., Kapetanakis, T. N., Nikolopoulos, C. D., Trang, T. K., Tsubota, T., Keyikoglu, R., Khataee, A., & Kalderis, D. (2020). Hydrochars as emerging biofuels: Recent advances and application of artificial neural networks for the prediction of heating values. *Energies*, *13*(17), 4572. https://doi.org/10.3390/en13174572

Vauvre, J.-M., Patsioura, A., Vitrac, O., & Kesteloot, R. (2015). Multiscale modeling of oil uptake in fried products. *AICHE Journal*, *61*(7), 2329–2353. https://doi.org/10.1002/aic.14801

Verboven, P., Defraeye, T., Datta, A. K., & Nicolai, B. (2020). Digital twins of food process operations: The next step for food process models? *Current Opinion in Food Science*, *35*, 79–87. https://doi.org/10.1016/j.cofs.2020.03.002

Vitrac, O., & Hayert, M. (2005). Risk assessment of migration from packaging materials into foodstuffs. *AICHE Journal*, *51*(4), 1080–1095. https://doi.org/10.1002/aic.10462

Vitrac, O., & Hayert, M. (2020). Modeling in food across the scales: Towards a universal mass transfer simulator of small molecules in

food. *SN Applied Sciences*, *2*(9), 1509. https://doi.org/10.1007/s42452-020-03272-2

Vitrac, O., Nguyen, P.-M., & Hayert, M. (2022). In silico prediction of food properties: A multiscale perspective. *Frontiers in Chemical Engineering*, *3*, 786879. https://doi.org/10.3389/fceng.2021.786879

Vitrac, O., & Touffet, M. (2019). Food process modeling. In I. P. Ferranti, E. M. Berry, & J. R. Anderson (Eds.), *Encyclopedia of food security and sustainability* (pp. 434–454). Elsevier. https://doi.org/10.1016/B978-0-08-100596-5.22277-4

Vranic, M., Bosnjak, K., Rukavina, I., Glavanovic, S., Pukec, N. P., Babic, A., & Vranic, I. (2020). Prediction of forage chemical composition by NIR spectroscopy. *Journal of Central European Agriculture*, *21*(3), 554–568. https://doi.org/10.5513/JCEA01/21.3.2839

Wu, Z., Johnston, K. E., Arnold, F. H., & Yang, K. K. (2021). Protein sequence design with deep generative models. *Current Opinion in Chemical Biology*, *65*, 18–27. https://doi.org/10.1016/j.cbpa.2021.04.004

Yang, R., & Chen, J. (2022). Recent application of artificial neural network in microwave drying of foods: A mini-review. *Journal of the Science of Food and Agriculture*, *102*(14), 6202–6210. https://doi.org/10.1002/jsfa.12008

Yao, Y. (2017). Models for sustainability. *Bioresources*, *12*(1), 1–3.

Zeng, X., & Li, J. (2021). Emerging anthropogenic circularity science: Principles, practices, and challenges. *iScience*, *24*(3), 102237. https://doi.org/10.1016/j.isci.2021.102237

Zhao, X., Li, W., Zhang, H., Li, X., & Fan, W. (2019). Reaction-diffusion approach to modeling water diffusion in glutinous rice flour particles during dynamic vapor adsorption. *Journal of Food Science and Technology*, *56*(10), 4605–4615. https://doi.org/10.1007/s13197-019-03925-0

Zhou, I., Lipman, J., Abolhasan, M., Shariati, N., & Lamb, D. W. (2020). Frost monitoring cyber-physical system: A survey on prediction and active protection methods. *IEEE Internet of Things Journal*, *7*(7), 6514–6527. https://doi.org/10.1109/JIOT.2020.2972936

Zhu, Y., Guillemat, B., & Vitrac, O. (2019). Rational design of packaging: Toward safer and ecodesigned food packaging systems. *Frontiers in Chemistry*, *7*, 349. https://doi.org/10.3389/fchem.2019.00349

How to cite this article: Sicard, J., Barbe, S., Boutrou, R., Bouvier, L., Delaplace, G., Lashermes, G., Théron, L., Vitrac, O., & Tonda, A. (2023). A primer on predictive techniques for food and bioresources transformation processes. *Journal of Food Process Engineering*, e14325. https://doi.org/10.1111/jfpe.14325