Thermal Modeling of Thin-Layer Bread via Neural Networks: Experimental Validation in a Sardinian Bakery ⋆

Diego Deplano[a,1], Nicola Arridu[b], Carla Seatzu[a] and Mauro Franceschelli[a,*]

[a]Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, 09123, Italy
[b]Department of Civil, Computer Science and Aeronautical Technologies Engineering, University Roma Tre, Rome, 00146, Italy

ARTICLE INFO

Keywords:
Industry 5.0
Physics-informed learning
Data-driven identification
Thin-layer drying
Carasau
Cooling process
Agrifood.

ABSTRACT

This paper presents neural network–based methods for modeling the cooling behavior of Sardinian flatbreads, validated through experimental data collected in a Sardinian bakery. Building upon previous research on the drying–cooling dynamics of pane Carasau, which identified the Verma model as one of the most suitable thin-layer models for flatbreads, a data-driven framework is developed to tune the Verma model coefficients. Given the Industry 5.0–oriented operational setting of the bakery, the proposed framework is conceived as a human-centric decision-support tool that provides operators with scenario-dependent predictions of cooling behavior under varying initial and ambient conditions, thereby supporting informed adjustments of process settings while preserving human oversight. The goal is to learn how the model coefficients vary with boundary conditions – such as initial temperature and ambient conditions – thereby enhancing model accuracy and generalization in real industrial settings. Two identification strategies are proposed: (i) a physics-informed approach, which embeds the Verma law directly into the network training process, and (ii) a supervised approach, where the network learns the Verma model parameters from labeled cooling curves. Real thermal imaging measurements acquired in an industrial bakery are used for both training and validation. Results show that both approaches perform effectively and with similar performance, thus demonstrating the advantages of integrating physical modeling with data-driven learning.

1. Introduction

Flatbreads serve as a dietary cornerstone for nearly one-third of the global population, with the market valued at $38.8 billion in 2018 and expected to expand to $62.8 billion by 2026 (Boukid, 2022). Typically made from a simple blend of flour, water, and salt, these breads exhibit considerable diversity in shape, size, and production methods across cultures (Banooni, Hosseinalipour, Mujumdar, Taherkhani and Bahiraei, 2009). For example, the Middle Eastern Lavash, Indian Chapati, and Central American Tortillas each reflect unique culinary traditions (Salari, Tehrani and Razavi, 2015; Parimala and Sudha, 2015; Xu and Kerr, 2012). Research has historically concentrated on conventional three-dimensional breads (Hasatani, Arai, Harui, Itaya, Fushida and Hori, 1992), leaving a gap in the theoretical understanding and practical optimization of the drying and cooling processes specific to flatbreads. Notably, Salari et al. (Salari et al., 2015) is one of the few studies addressing the baking-drying kinetics of flatbreads. This shortfall in research is further compounded by the fact that many flatbreads are produced in small-scale operations with limited automation, which poses challenges in meeting the growing global demand. In these settings, even where some level of mechanization exists, much of the work still relies on outdated equipment or manual labor, which negatively impacts production efficiency and increases labor costs (Paschino, Gabella, Giubellino and Clemente, 2007).

To tackle these issues, this work studies fully data-driven models based on neural networks designed to boost production, enhance automation, and improve the reliability of less automated flatbread bakeries, with a focus on Carasau bread (Deplano, Franceschelli and Seatzu, 2023; Cadalanu, 1984; Paschino and Gambella, 2008; Cavone, Dotoli, Epicoco, Franceschelli and Seatzu, 2018; Baire, Melis, Lodi, Tuveri, Dachena, Simone, Fanti, Fumera, Pisanu and Mazzarella, 2019; Abbondio, Palomba, Tanca, Fraumene, Pagnozzi, Serra, Marongiu, Laconi and Uzzau, 2019), a traditional Sardinian flatbread known as "carta musica" (music sheets) because of its extremely thin and translucent nature (Pasqualone, 2018). The production process of Carasau begins with mixing flour, water, salt, and yeast to form the dough, which is then rolled into thin sheets, leavened, baked, cut, separated, and toasted. During baking, the rapid temperature change causes the bread to puff up due to the air expansion. After baking, the sheets undergo a crucial drying and cooling phase, where good control of the temperature

⋆This work was supported by the Italian Ministry of Enterprises and Made in Italy within the "ACCORDI PER L'INNOVAZIONE" (2021–2026), through the Project "AISAC–Tecnologie ICT e dell'Industria 4.0 per l'Analisi e l'Ingegnerizzazione di Sistemi Alimentari Complessi per la produzione di pani artigianali locali ad alto valore aggiunto" under Grant CUP: B29J23001120005- COR: 1607797.

*Corresponding author
diego.deplano@unica.it (D. Deplano);
nicola.arridu@unica.it,narridu@os.uniroma3.it (N. Arridu);
carla.seatzu@unica.it (C. Seatzu); mauro.franceschelli@unica.it (M. Franceschelli)
ORCID(s):

[1]The work of Diego Deplano was partially supported by the project e.INS Ecosystem of Innovation for Next Generation Sardinia (cod. ECS 00000038) funded by the Italian Ministry for Research and Education (MUR) under the National Recovery and Resilience Plan (NRRP) - MISSION 4 COMPONENT 2, "From research to business" INVESTMENT 1.5, "Creation and strengthening of Ecosystems of innovation" and construction of "Territorial R&D Leaders".

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/ DOI: 10.1016/j.conengprac.2026.107014

decay is necessary to prepare the product for the subsequent processing steps. In the plant examined, the sheets are first fed into a cutting machine and then into a separation machine that splits them into individual pieces. The proper operation of these machines depends on achieving sufficiently precise thermal conditions, as deviations can lead to waste through incorrect separation. Therefore, a better prediction of the temperature decay is essential for enhancing both the efficiency of automation and the quality of the final product (Ruiz-López and García-Alvarado, 2007). The framework proposed in this work enables accurate prediction of the temperature decay profile during the cooling process, helping operators understand whether the current operational conditions are appropriate for the subsequent cutting and separation phases. Indeed, this information allows the operators to adjust the oven temperature and conveyor belt speed to meet the downstream processing requirements. Although the oven and conveyor belt are automatically controlled, operators remain actively involved in supervising the process, interpreting the predicted outcomes, and taking corrective action when necessary. From this perspective, the framework reflects the human-centric vision of Industry 5.0: rather than replacing human judgment, it supports operators with useful predictive insights that strengthen their awareness of process conditions and help them make more informed decisions. In this way, the proposed framework serves as a tool to assist human expertise, contributing to more resilient, sustainable, and adaptive production systems (Hassoun, Jagtap, Trollman, Garcia-Garcia, Duong, Saxena, Bouzembrak, Treiblmaier, Para-López, Carmona-Torres et al., 2024; Melesse and Orrù, 2025).

Recent advances in Artificial Intelligence (AI) and Machine Learning have enabled diverse neural network models to optimize food production, addressing challenges in process control, quality assurance, and supply chain management. Deep learning methods have automated tasks such as ingredient selection and dynamic process simulation (e.g., cooking and dough kneading) (Huang, Zhang, Mujumdar and Li, 2025; Lamrini, Della Valle, Trelea, Perrot and Trystram, 2012; Fañanás-Anaya, López-Nicolás, Sagüés and Llorente, 2024). Early approaches using recurrent networks captured temporal and nonlinear dynamics – recently investigated in the context of monotone multi-agent systems (Deplano, Franceschelli and Giua, 2018, 2024) – but they required extensive, high-quality data and were sensitive to operational variations (Oulmelk, Srati, Afraites and Hadri, 2023; Lamrini et al., 2012; Liu, Kawaguchi, Xu and Hashimoto, 2021). To overcome these limitations, researchers have developed models that incorporate physical principles—such as physics-informed neural networks—to improve computational efficiency in inverse problems like parameter estimation and boundary condition determination (Chen, Shi, He and Fang, 2024; Qiu, Chen and Zhou, 2024). Although these advanced methods perform well in controlled settings, they still struggle with robustness in noisy, real-world environments (Mostajeran and Mokhtari, 2022; Bastianello, Deplano, Franceschelli and Johansson, 2025; Ledda, Deplano, Giua and Franceschelli, 2025). AI-driven temperature monitoring systems have made strides in food transport and quality control. By integrating multi-source data, including thermal imaging and ambient conditions, these models accurately predict temperature distributions to ensure food safety and freshness (Zou, Wu, Wang, Morales, Liu and Manzardo, 2023; Badia-Melis, Qian, Fan, Hoyos-Echevarria, Ruiz-García and Yang, 2016; Canatan, Muñoz-Carpena and Boz, 2025; Oz, Sochen, Mendlovic and Klapp, 2024; Mercier and Uysal, 2018). Hybrid approaches that combine deep learning with traditional methods, such as convolutional neural networks paired with support vector machines, further enhance prediction accuracy while maintaining computational efficiency (da Silva Cotrim, Felix, Minim, Campos and Minim, 2021).

The **main contribution** of this paper is to provide a practical methodology for predicting, in real time, the temperature decay of flatbread sheets upon exiting the oven by leveraging only their initial temperature and the ambient temperature. By modeling the temperature decay by means of the so-called *Verma model*—identified in our preliminary work as one of the most suitable thin-layer models for the temperature decay of flatbreads (Deplano et al., 2023)—this paper proposes two data-driven learning approaches to identify the parameters of the Verma model using neural networks. Once the training is completed, the neural network enables real-time computation of the Verma model parameters, which may vary with changes in the initial and ambient temperatures, and, in turn, allows real-time prediction of the temperature decay of the flatbread sheets. This significantly extends our preliminary work, in which the parameters were computed offline and not adapted in real time to changing operating conditions. The first proposed method is based on physics-informed learning, which integrates the Verma model directly into the training objective; in contrast, the second method relies on fully supervised regression based on pre-identified model parameters. The comparison demonstrates that the physics-informed approach enhances predictive consistency and resilience to noise compared to purely supervised identification. By combining thermodynamic modeling with machine learning, the proposed framework improves temperature prediction and supports timely adjustment of oven and conveyor settings to achieve the target cooling conditions for each flatbread sheet. This improves the reliability of cutting and separation, while reducing waste, energy consumption, and production losses.

Outline of the paper. The remainder of this paper is organized as follows. Section 3 presents the neural network designs and the datasets construction for the learning of cooling dynamics of flatbreads, including both physics-informed and supervised neural architectures. Section 4 reports on the comparative experimental results from real industrial data. Finally, Section 5 presents the conclusions and outlines potential directions for employing the proposed model within automated control systems for food manufacturing.

Licensed under CC BY 4.0. DOI: 10.1016/j.conengprac.2026.107014

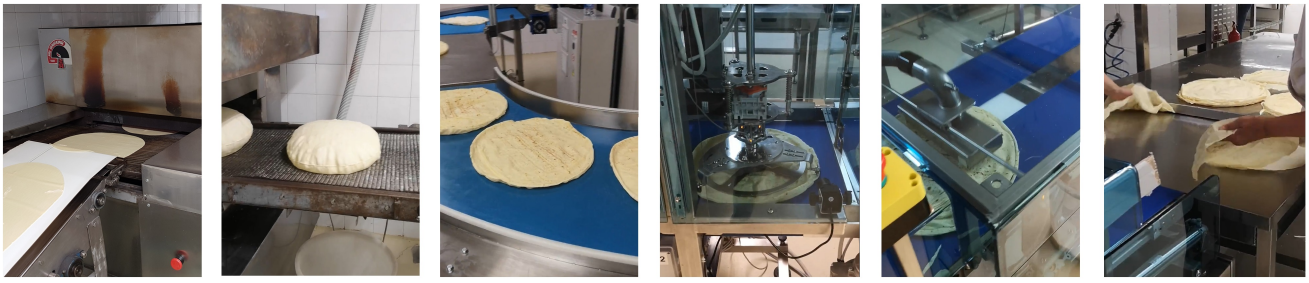


Figure 1: Pictures of Carasau production process showing the steps of baking in the oven, cooling/drying over the conveyor belt, and automated separation.

2. Notation and Nomenclature

The set of natural numbers is denoted as $\mathbb{N}$. The set of real numbers is denoted by $\mathbb{R}$ and its nonnegative subset is denoted by $\mathbb{R}_{\geq 0}$. Matrices $M \in \mathbb{R}^{n\times m}$ are represented by uppercase letters, vectors and scalars $s \in \mathbb{R}$ by lowercase letters, while sets and sequences $\mathcal{S}$ are denoted by uppercase calligraphic letters. Continuous- and discrete-times are denoted by $t \in \mathbb{R}_{\geq 0}$ and $k \in \mathbb{N}$, respectively, measured in seconds, whereas $\Delta \in \mathbb{R}_{\geq 0}$ denotes the sampling time. The index $i \in \mathbb{N}$ refers to a bread sample, $N \in \mathbb{N}$ indicates the total number of bread samples, and $n_i \in \mathbb{N}$ denotes the number of thermal measurements collected for the i-th sample. Throughout the manuscript, specific subscripts and superscripts are used to denote particular quantities:

- e: ambient quantity;
- r: rescaled quantity;
- avg: average temperature;
- max: maximum temperature;
- PHY: physics-informed learning;
- SUP: supervised learning;
- $\hat{x}$: estimation or prediction of the quantity x.

Thermal quantities in °C of the i-th bread sheet:

- $T_i(\Delta k)$: thermal 8×8 image at time Δk;
- $\tau_i^{\text{avg}}(\Delta k)$, $\tau_i^{\max}(\Delta k)$: average/maximum temperature over the hottest 3×3 submatrix at time Δk;
- $\hat{\tau}_i^{\text{avg}}(\Delta k)$, $\hat{\tau}_i^{\max}(\Delta k)$: predicted sheet temperatures;
- $\tau_{i,e}$: ambient temperature (assumed constant) during the cooling cycle of the i-th bread sheet.

Verma model $\tau_{i,r}(t)=a_i e^{-\lambda_{i,0}t}+(1-a_i)e^{-\lambda_{i,1}t}$ quantities:

- $\tau_{i,r}(t) \in [0,1]$: dimensionless temperature profile;
- $a_i \in [0,1]$: weighting coefficient of two decay modes;
- $\lambda_{i,0} \in [0,\infty)$: fast decay-rate parameter;
- $\lambda_{i,1} \in [0,\infty)$: slow decay-rate parameter;
- $\hat{a}_i$, $\hat{\lambda}_{i,0}$, $\hat{\lambda}_{i,1}$: estimated parameters.

The subscript i is omitted to denote a generic bread sheet rather than a specific sheet. Neural network quantities:

- $L \in \mathbb{N}$: number of layers;
- $n_\ell \in \mathbb{N}$: number of neurons in layer ℓ;
- $h_\ell \in \mathbb{R}^{n_\ell}$: hidden state of layer ℓ;
- $W_\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$: weight matrix of layer ℓ;
- $b_\ell \in \mathbb{R}^{n_\ell}$: bias vector of layer ℓ.
- $\alpha_\ell \in [0,1]$: parameter of the Leaky ReLU function;
- $\mathcal{D}_{\text{PHY}}^{\text{avg}}$, $\mathcal{D}_{\text{PHY}}^{\max}$, $\mathcal{D}_{\text{SUP}}^{\text{avg}}$, $\mathcal{D}_{\text{SUP}}^{\max}$: datasets.

3. Data-driven models for the cooling dynamics of Carasau bread

The production process of Carasau bread can be divided into several steps, which are pictured in Fig. 1: mixing the ingredients to obtain a dough; rolling out the dough into round sheets of about 1 mm thickness; baking the sheets in an oven and then leaving them to cool down and dry; separating the cooled and dried sheets into two separate sheets; re-baking the separated sheets. This work focuses on the automation of the *cooling* process and the subsequent separation process. In our plant, the sheets are left to cool down over a conveyor belt that brings the sheets directly into the separation machine that cuts and splits them apart. Note that, while traditionally the separation process is done manually, in our plant, it has been designed a custom separation machine that automates this process. To ensure it functions correctly, the separation machine relies on specific sheet temperature levels.

Some of us have previously identified a knowledge-driven theoretical model of the cooling dynamics, that is the so-called "*Verma model*" given next (Deplano et al., 2023):

$$\tau_r(t) = ae^{-\lambda_0 t} + (1-a)e^{-\lambda_1 t}, \qquad t \in \mathbb{R},$$

where

- $\tau_r(t) \in [0,1]$ denotes a dimensionless parameter used to rescale the temperature in the considered range, with $\tau_r(0) = 1$ and $\lim_{t\to\infty} \tau_r(t) = 0$;
- $\lambda_0, \lambda_1 \geq 0$ are the two dominant modes of the system, representing the decay rate at the beginning (fast) and at the end (slow) of the process;
- $a \in [0,1]$ is a coefficient that weights the contribution deriving from the two modes.

Given a set of parameters $\{a, \lambda_0, \lambda_1\}$ defining the Verma model, it is possible to predict the average temperature decay of the bread sheet by knowing the initial temperature $\tau(0) \geq 0$ and the ambient temperature $\tau_e \geq 0$ assumed constant $\tau_e(t) \approx \tau_e(0) =: \tau_e$ since the cooling process is much faster than ambient temperature changes) by means of the following:

$$\hat{\tau}(t) = \tau_r(t)(\tau(0) - \tau_e) + \tau_e. \tag{1}$$

Licensed under CC BY 4.0. DOI: 10.1016/j.conengprac.2026.107014

The main drawback of using the Verma model to predict the cooling behavior of flatbread lies in the difficulty of properly tuning its parameters, which vary with the initial temperature $\tau(0)$ and the ambient temperature τ_e. In this work, two alternative methods are proposed for learning the Verma model parameters using artificial neural networks (ANNs). In Fig. 2 a graphical representation of the necessary steps to construct the dataset and perform the training process with both methods is shown.

A feedforward ANNs is considered with $L \in \mathbb{N}$ layers of $n_\ell \in \mathbb{N}$ neurons for $\ell \in \{1, \cdots, L\}$ such that

$$\begin{bmatrix} \hat{a} \\ \hat{\lambda}_0 \\ \hat{\lambda}_1 \end{bmatrix} = (f_L \circ \cdots \circ f_1)\left(\begin{bmatrix} \tau(0) \\ \tau_e \end{bmatrix}\right) \tag{2}$$

where $\hat{a}$, $\hat{\lambda}_0$, $\hat{\lambda}_1$ denote the estimated parameters and $\circ$ denotes the composition operator. Each layer f_ℓ is designed as affine transformation with possibly a ReLU activation function, which is conveniently expressed using the $\mathrm{LReLU}_\alpha(x) = \max(\alpha x, x)$ operator with $\alpha_\ell \in \{0, 1\}$ as

$$f_\ell(h_{\ell-1}) = h_\ell = \mathrm{LReLU}_{\alpha_\ell}(W_\ell h_{\ell-1} + b_\ell). \tag{3}$$

Here, $W_\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ denotes the weight matrix, $b_\ell \in \mathbb{R}^{n_\ell}$ the bias vector, and the input layer is given by $h_0 = [\tau(0), \tau_e]^\top$ with $n_0 = 2$ and $n_L = 3$. The activation parameters are set as $\alpha_1 = \cdots = \alpha_{L-1} = 0$ and $\alpha_L = 1$. Let $W = \{W_\ell\}_{\ell=1}^L$ and $b = \{b_\ell\}_{\ell=1}^L$ denote the collection of all weights and biases to be learned during training.

A comparison of two different methodologies to learn the parameters $\hat{a}$, $\hat{\lambda}_0$, $\hat{\lambda}_1$ is now presented, and then an explanation of how the data have been collected and how the datasets for these two methodologies have been constructed is given. In so doing, the time is discretized by $t = \Delta k$ where $\Delta \in \mathbb{R}_{>0}$ is the sampling time and $k \in \mathbb{N}$.

1) Physics-informed learning. This strategy aims at estimating the Verma model parameters by training a feedforward ANN which minimizes the distance between the real curve $\tau_i(\Delta k)$ – sampled at discrete instants of time $k \in \mathbb{N}$ with steps $\Delta \in \mathbb{R}_{>0}$ – of the i-th sample and the curve $\hat{\tau}_i(\Delta k)$ obtained by evaluating the Verma model with the learned parameters $\hat{a}_i$, $\hat{\lambda}_{i,0}$, $\hat{\lambda}_{i,1}$, namely

$$\hat{\tau}_i(\Delta k) = (\hat{a}_i e^{-\hat{\lambda}_{i,0}\Delta k} + (1 - \hat{a}_i)e^{-\hat{\lambda}_{i,1}\Delta k})(\tau_i(0) - \tau_{i,e}) + \tau_{i,e},$$

with $\tau_{i,e}$ denoting the initial ambient temperature for the i-th sample. The loss to be minimized during the training is thus

$$\ell(W, b) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i}\sum_{k=0}^{n_i}(\hat{\tau}_i(\Delta k) - \tau_i(\Delta k))^2,$$

where $n_i \in \mathbb{N}$ is the total number of measurements for the i-th sample and $N \in \mathbb{N}$ is the total number of samples.

2) Supervised learning. This strategy aims at estimating the Verma model parameters by training a feedforward ANN which minimizes the distance between the parameters a_i, $\lambda_{i,0}$, $\lambda_{i,1}$ obtained by curve fitting of the i-th sample data and the learned parameters $\hat{a}_i$, $\hat{\lambda}_{i,0}$, $\hat{\lambda}_{i,1}$ for that sample. The loss to be minimized during the training is given by

$$\ell(W, b) = \frac{1}{3N}\sum_{i=1}^{N}\left\|\begin{bmatrix} \hat{a}_i - a_i \\ \hat{\lambda}_{i,0} - \lambda_{i,0} \\ \hat{\lambda}_{i,1} - \lambda_{i,1} \end{bmatrix}\right\|^2,$$

where $N \in \mathbb{N}$ is the total number of samples.

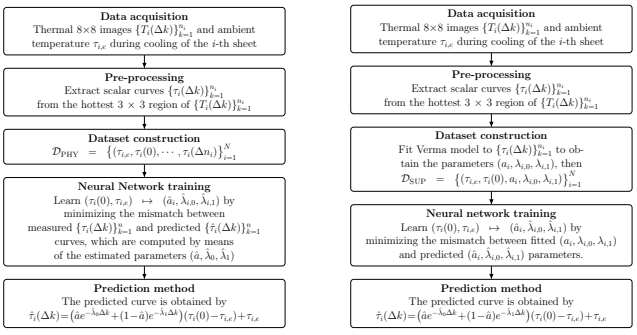


Figure 2: Graphical representation of the Verma-model parameter learning methodologies: physics-informed vs supervised.

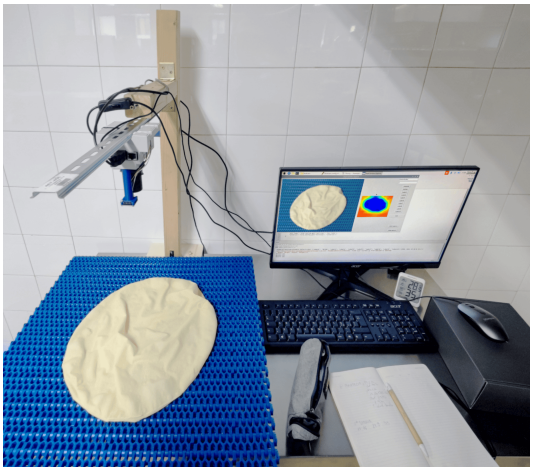


Figure 3: Setup used to capture thermal photos of the bread.

For the dataset construction, a series of measurements is carried out (see Fig. 3 for the setup) directly on a Carasau production plant in Vecchio Forno bakery in Fonni, Sardinia, Italy, using thermal imaging on $N \in \mathbb{N}$ semi-finished flatbread sheets just after the baking. The configuration of the thermal camera is visible in Fig. 4. The thermal camera is an *Adafruit AMG8833 8x8 Thermal Camera*, which has an uncertainty in the measure of 2.5°C. The camera is coupled with a *Raspberry Pi* device. An example of a Carasau bread sheet thermal measurement is given in Fig. 5. For each sheet of bread, several thermal images have been taken to monitor the decay of temperature.

Considering that a thermal image is composed by 8x8 pixels, i.e., a 8×8 matrix, let $T_i(\Delta k) \in \mathbb{R}^{8\times 8}$ be the generic matrix associated to the i-th sample, where $k \in \mathbb{N}$ refers to

Licensed under CC BY 4.0. DOI: 10.1016/j.conengprac.2026.107014

the time instant. For each bread sheet $n_i \in \mathbb{N}$ thermal photos have been collected, with a sampling time of about 1 second, thus forming the succession $\{T_i(0), T_i(\Delta), \dots, T_i(n_i\Delta)\}$. For each sequence, the 3×3 submatrix of $T_i(\Delta k)$ is identified with the highest average temperature, and then two different scalar curves from this submatrix are computed:

- The *average temperature* curve $\{\tau_i^{\text{avg}}(\Delta k)\}_{k=1}^{n_i}$;
- The *maximum temperature* curve $\{\tau_i^{\max}(\Delta k)\}_{k=1}^{n_i}$.

Both choices allow to neglect the background and the borders that are relatively colder, which may distort the previsione of the temperature decay.

For the physics-informed learning methodology, the datasets are constructed by associating the ambient temperatures to the real curves, namely,

$$\mathcal{D}_{\text{PHY}}^{\text{avg}} = \{\left(\tau_{i,e}, \tau_i^{\text{avg}}(0), \cdots, \tau_i^{\text{avg}}(n_i\Delta)\right)\}_{i=1}^{N},$$
$$\mathcal{D}_{\text{PHY}}^{\max} = \{\left(\tau_{i,e}, \tau_i^{\max}(0), \cdots, \tau_i^{\max}(n_i\Delta)\right)\}_{i=1}^{N}.$$

For the supervised learning methodology, the datasets are constructed by associating the ambient temperatures with the initial temperature and the parameters fitted to the Verma model for the curves in $\mathcal{D}_{\text{PHY}}^{\text{avg}}$ and $\mathcal{D}_{\text{PHY}}^{\max}$, respectively, namely,

$$\mathcal{D}_{\text{SUP}}^{\text{avg}} = \{\left(\tau_{i,e}, \tau_i^{\text{avg}}(0), a_i, \lambda_{i,0}, \lambda_{i,1}\right)\}_{i=1}^{N},$$
$$\mathcal{D}_{\text{SUP}}^{\max} = \{\left(\tau_{i,e}, \tau_i^{\max}(0), a_i, \lambda_{i,0}, \lambda_{i,1}\right)\}_{i=1}^{N}.$$

4. Experimental results and discussion

This section compares the two identification strategies described in the previous section for the Verma parameters – namely, a *physics-informed* approach and a *supervised* approach – under identical architectures, data splits, and evaluation protocols, interpreting their accuracy, variability, and implications for the Carasau production line.

Each dataset contains $N \sim \mathcal{O}(10^2)$ temperature-decay sequences acquired manually on semi-finished sheets immediately after baking using an 8×8 thermal camera (AMG8833, ± 2.5°C), with sampling interval ~ 1.25 s and variable sequence lengths $n_i \sim 10^{2.5}$. The datasets consist of measurements collected at various times of day (from early morning, through midday, to late evening) yielding ambient temperatures $\tau_{i,e}$ ranging over $[21, 24]$°C, initial average temperatures $\tau_i^{\text{avg}}(0)$ ranging over $[55, 82]$°C, and initial maximum temperatures $\tau_i^{\max}(0)$ ranging over $[58, 83]$°C. As mentioned above, two scalar signals $\tau_i^{\text{avg}}(\Delta k)$, $\tau_i^{\max}(\Delta k)$ are computed as the time evolution of the average/maximum over the 3×3 hottest submatrix at $k = 0$, respectively, emphasizing the core thermal region while mitigating boundary and background artifacts. An 80/20 split is adopted for the training and validation sets; the supervised sets are derived from the same bread samples, thus ensuring a like-for-like comparison.

Both approaches use the same feedforward ANN with four fully connected layers with $50 \times 50 \times 16 \times 3$ neurons, ReLU activations in all internal layers, and the following

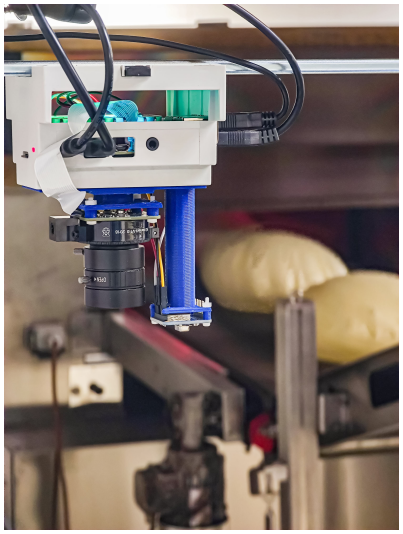


Figure 4: Custom device used to collect temperature profiles of Carasau bread. The Adafruit AMG8833 8×8 thermal camera is mounted on a blue support bracket.

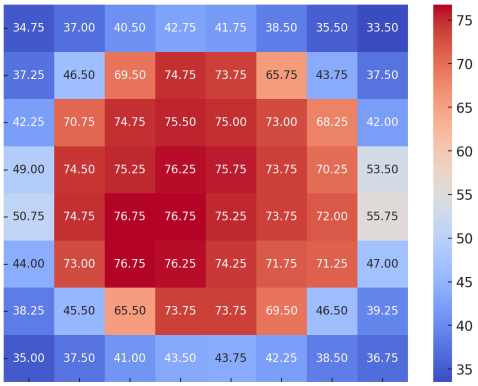


Figure 5: Surface temperature heatmap (°C) of a Carasau bread sheet at the start of the cooling phase.

activation functions in the output layer: a softplus function for the neurons associated with λ_0, λ_1 ensuring nonnegativity of the parameters, and a logistic function for the neuron associated with the convex combination parameter a ensuring it belongs to the interval $[0, 1]$. The ANNs are implemented in PyTorch (Paszke, 2019), and trained for 400 epochs ($\sim$ 0.03 s per epoch). For the sake of clarity, it is remarked that the physics-informed network minimizes trajectory mismatch between the real curve temperatures $\tau_i^{\text{avg}}(\Delta k)$, $\tau_i^{\max}(\Delta k)$ and the Verma reconstructions generated by predicted parameters $(\hat{a}, \hat{\lambda}_0, \hat{\lambda}_1)$, whereas the supervised network minimizes parameter error relative to per-curve Verma fits $(a_i, \lambda_{i,0}, \lambda_{i,1})$, contrasting end-to-end trajectory supervision with label-supervised regression.

For the sake of clarity, let us denote with $f_{\text{PHY}}^{\text{avg}}$, $f_{\text{PHY}}^{\max}$ the networks trained with the physics-informed approach and $f_{\text{SUP}}^{\text{avg}}$, $f_{\text{SUP}}^{\max}$ the networks trained with the supervised approach, and pair them into a number of configurations

Licensed under CC BY 4.0. DOI: 10.1016/j.conengprac.2026.107014

equal to the total number of experiments, that is 100, to quantify robustness across random initializations and stochastic optimization.

Performance is measured via the normalized root mean square error (NRMSE) and the mean absolute error (MAE) on the shared validation set:

$$\text{NRMSE} = \frac{1}{NT_{\max}} \sum_{i=1}^{N} \sqrt{\frac{1}{n_i} \sum_{k=0}^{n_i} \left(\hat{\tau}_i(\Delta k) - \tau_i(\Delta k)\right)^2},$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i} \sum_{k=0}^{n_i} \left|\hat{\tau}_i(\Delta k) - \tau_i(\Delta k)\right|$$

where $T_{\max}$ is the maximum temperature across all sequences, normalizing for varying trajectory ranges and lengths induced by $\tau(0)$ and τ_e.

Figure 6 illustrates representative predictions for both maximum and average temperature curves on samples. The visual agreement between both approaches and the measured data confirms the quantitative trends observed in Tables 2–1, namely that physics-informed modeling provides slightly better predictions of peak temperature decay, while both methods perform similarly for averaged curves. Tables 2–1 additionally report a baseline purely data-driven model for comparison which does not exploit the Verma structure. For this model, an auto-encoder architecture (Masti and Bemporad, 2021; Ledda et al., 2025) is used to predict the curve directly and denote it by $f_{\text{AE}}^{\max}$ and $f_{\text{AE}}^{\text{avg}}$

Table 1 summarizes the results for the maximum temperature curves across 100 configurations. In this regime, the physics-informed model clearly outperforms both the supervised approach and the purely data-driven approach, achieving a lower mean absolute error of $\sim 1°\text{C}$ against $\sim 1.5°\text{C}$ and $\sim 2°\text{C}$, respectively. These improvements indicate that embedding physical constraints effectively regularizes the model, guiding it toward more stable and physically consistent solutions, thus showing that the physical prior improves both best-case performance and stability when modeling peak temperatures. For the average temperature curves (Table 2), the physics-informed model still performs better than the supervised model but it exhibits similar performance to the baseline purely data-driven model. This suggests that the physical structure of the Verma model continues to provide a meaningful inductive bias even when the signal is smoother.

Overall, these findings highlight that the physics-informed approach is the more reliable choice. While the purely data-driven baseline falls short in accuracy for average curves, it does so as a black box, offering no insight into the underlying process and no means for operators to act on the results. By contrast, the physics-informed approach simultaneously delivers accurate temperature predictions and physically meaningful parameter estimates, making it better suited for deployment in an industrial setting where transparency, reliability, and operator interpretability are of primary concern.

Table 1
Comparison over 100 configurations for maximum curves temperature: best, worst, average, and standard deviation of NRMSE on the validation set and total/last MAE.

	NRMSE				MAE	
Model	Best	Worst	Avg	Dev	Tot	Last
$f_{\text{PHY}}^{\max}$	0.0109	0.0233	0.0145	0.0021	0.9998	0.8236
$f_{\text{SUP}}^{\max}$	0.0155	0.0361	0.0241	0.0044	1.8009	1.8885
$f_{\text{AE}}^{\max}$	0.0152	0.0419	0.0224	0.0048	1.5847	1.3843

Table 2
Comparison over 100 configurations for average curves temperature: best, worst, average, and standard deviation of NRMSE on the validation set and total/last MAE.

	NRMSE				MAE	
Model	Best	Worst	Avg	Dev	Tot	Last
$f_{\text{PHY}}^{\text{avg}}$	0.0150	0.0284	0.0201	0.0030	1.4870	1.1065
$f_{\text{SUP}}^{\text{avg}}$	0.0193	0.0401	0.0268	0.0044	2.0588	1.7059
$f_{\text{AE}}^{\text{avg}}$	0.0144	0.0326	0.0223	0.0034	1.6200	1.4227

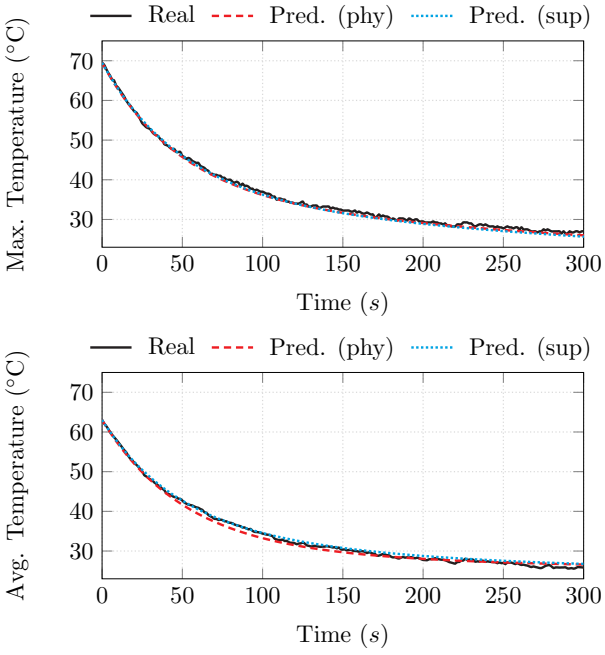


Figure 6: Representative predicted curve temperatures.

5. Conclusions

This study presented a practical methodology for the real-time prediction of temperature decay in flatbread sheets, effectively bridging the gap between theoretical thin-layer modeling and the variability inherent in industrial production. By leveraging the Verma model within a neural network framework, it has eben demonstrated that the model parameters can be dynamically identified based on varying initial and ambient thermal conditions. Two different approaches were tested: physics-informed learning and supervised regression. The results indicated that the physics-informed

Licensed under CC BY 4.0. DOI: 10.1016/j.conengprac.2026.107014

strategy provides reliable predictions, effectively enabling the robust real-time computation of decay parameters. This capability extends prior work (Deplano et al., 2023), which was limited to static, offline parameter tuning, by allowing the model to adapt to changing operating conditions. Future developments will involve coupling the identified models with automated control strategies for regulating cooling conditions and conveyor dynamics in real time, completing the transition from static modeling to predictive automation in artisanal bread production.

References

Abbondio, M., Palomba, A., Tanca, A., Fraumene, C., Pagnozzi, D., Serra, M., Marongiu, F., Laconi, E., Uzzau, S., 2019. Fecal metaproteomic analysis reveals unique changes of the gut microbiome functions after consumption of sourdough carasau bread. Frontiers in microbiology 10, 1733.

Badia-Melis, R., Qian, J., Fan, B., Hoyos-Echevarria, P., Ruiz-García, L., Yang, X., 2016. Artificial neural networks and thermal image for temperature prediction in apples. Food and Bioprocess Technology 9, 1089–1099.

Baire, M., Melis, A., Lodi, M.B., Tuveri, P., Dachena, C., Simone, M., Fanti, A., Fumera, G., Pisanu, T., Mazzarella, G., 2019. A wireless sensors network for monitoring the carasau bread manufacturing process. Electronics 8.

Banooni, S., Hosseinalipour, S., Mujumdar, A., Taherkhani, P., Bahiraei, M., 2009. Baking of flat bread in an impingement oven: modeling and optimization. Drying Technology 27, 103–112.

Bastianello, N., Deplano, D., Franceschelli, M., Johansson, K.H., 2025. Robust online learning over networks. IEEE Transactions on Automatic Control 70, 933–946.

Boukid, F., 2022. Flatbread-a canvas for innovation: A review. Applied Food Research , 100071.

Cadalanu, G., 1984. Il pane carasau. preparazione e cottura. La Ricerca Folklorica , 81–88.

Canatan, M., Muñoz-Carpena, R., Boz, Z., 2025. Continuous surface temperature monitoring of refrigerated fresh produce through visible and thermal infrared sensor fusion. Postharvest Biology and Technology 222, 113354.

Cavone, G., Dotoli, M., Epicoco, N., Franceschelli, M., Seatzu, C., 2018. Hybrid Petri nets to re-design low-automated production processes: the case study of a sardinian bakery. IFAC-PapersOnLine 51, 265–270.

Chen, J., Shi, J., He, A., Fang, H., 2024. Data-driven solutions and parameter estimations of a family of higher-order kdv equations based on physics informed neural networks. Scientific Reports 14, 1–27.

Deplano, D., Franceschelli, M., Giua, A., 2018. Lyapunov-free analysis for consensus of nonlinear discrete- time multi-agent systems, in: 2018 IEEE Conference on Decision and Control (CDC), pp. 2525–2530.

Deplano, D., Franceschelli, M., Giua, A., 2024. Stability of nonexpansive monotone systems and application to recurrent neural networks. IEEE Control Systems Letters 8, 1793–1798.

Deplano, D., Franceschelli, M., Seatzu, C., 2023. Experimental comparison of models of the drying-cooling process of flatbreads for optimized automated production: The case study of carasau bread, in: 2023 9th International Conference on Control, Decision and Information Technologies (CoDIT), IEEE. pp. 2014–2019.

Fañanás-Anaya, J., López-Nicolás, G., Sagüés, C., Llorente, S., 2024. Food cooking process modeling with neural networks. IEEE Access .

Hasatani, M., Arai, N., Harui, H., Itaya, Y., Fushida, N., Hori, C., 1992. Effect of drying on heat transfer of bread during baking in oven. Drying Technology 10, 623–639.

Hassoun, A., Jagtap, S., Trollman, H., Garcia-Garcia, G., Duong, L.N., Saxena, P., Bouzembrak, Y., Treiblmaier, H., Para-López, C., Carmona-Torres, C., et al., 2024. From food industry 4.0 to food industry 5.0: Identifying technological enablers and potential future applications in the food sector. Comprehensive reviews in food science and food safety 23, e370040.

Huang, J., Zhang, M., Mujumdar, A.S., Li, C., 2025. Ai-based processing of future prepared foods: Progress and prospects. Food Research International , 115675.

Lamrini, B., Della Valle, G., Trelea, I.C., Perrot, N., Trystram, G., 2012. A new method for dynamic modelling of bread dough kneading based on artificial neural network. Food Control 26, 512–524.

Ledda, M., Deplano, D., Giua, A., Franceschelli, M., 2025. Certification of autoencoder-based models for dynamical systems, in: 2025 IEEE 64th Conference on Decision and Control (CDC), pp. 401–406.

Liu, Y., Kawaguchi, T., Xu, S., Hashimoto, S., 2021. Recurrent neural network-based temperature control system weight pruning based on nonlinear reconstruction error. Processes 10, 44.

Masti, D., Bemporad, A., 2021. Learning nonlinear state–space models using autoencoders. Automatica 129, 109666.

Melesse, T.Y., Orrù, P.F., 2025. The digital revolution in the bakery sector: Innovations, challenges, and opportunities from industry 4.0. Foods 14, 526.

Mercier, S., Uysal, I., 2018. Neural network models for predicting perishable food temperatures along the supply chain. Biosystems engineering 171, 91–100.

Mostajeran, F., Mokhtari, R., 2022. Deepbhcp: Deep neural network algorithm for solving backward heat conduction problems. Computer Physics Communications 272, 108236.

Oulmelk, A., Srati, M., Afraites, L., Hadri, A., 2023. An artificial neural network approach to identify the parameter in a nonlinear subdiffusion model. Communications in Nonlinear Science and Numerical Simulation 125, 107413.

Oz, N., Sochen, N., Mendlovic, D., Klapp, I., 2024. Estimating temperatures with low-cost infrared cameras using physically-constrained deep neural networks. Optics Express 32, 30565–30582.

Parimala, K., Sudha, M., 2015. Wheat-based traditional flat breads of india. Critical reviews in food science and nutrition 55, 67–81.

Paschino, F., Gabella, F., Giubellino, F., Clemente, F., 2007. The level of automation of "carasau" bread production plants. Journal of Agricultural Engineering 38, 61–64.

Paschino, F., Gambella, F., 2008. Comparison between traditional and industrial plants used for production of carasau bread and evaluation of final products. Bollettino della Comunitá Scientifica in Australasia .

Pasqualone, A., 2018. Traditional flat breads spread from the fertile crescent: Production process and history of baking systems. Journal of Ethnic Foods 5, 10–19.

Paszke, A., 2019. Pytorch: An imperative style, high-performance deep learning library. arXiv preprint arXiv:1912.01703 .

Qiu, W., Chen, H., Zhou, H., 2024. Estimating the boundary conditions for 3d transient heat conduction by bidirectional long short-term memory network and attention mechanism. International Journal of Heat and Mass Transfer 233, 126042.

Ruiz-López, I., García-Alvarado, M., 2007. Analytical solution for food-drying kinetics considering shrinkage and variable diffusivity. Journal of food engineering 79, 208–216.

Salari, A., Tehrani, M.M., Razavi, S.M., 2015. Baking-drying kinetics of crisp bread: The influence of bran content and baking temperature. Iranian Food Science and Technology Research Journal 11, 225.

da Silva Cotrim, W., Felix, L.B., Minim, V.P.R., Campos, R.C., Minim, L.A., 2021. Development of a hybrid system based on convolutional neural networks and support vector machines for recognition and tracking color changes in food during thermal processing. Chemical Engineering Science 240, 116679.

Xu, S., Kerr, W.L., 2012. Modeling moisture loss during vacuum belt drying of low-fat tortilla chips. Drying Technology 30, 1422–1431.

Zou, Y., Wu, J., Wang, X., Morales, K., Liu, G., Manzardo, A., 2023. An improved artificial neural network using multi-source data to estimate food temperature during multi-temperature delivery. Journal of Food Engineering 351, 111518.

Licensed under CC BY 4.0. DOI: 10.1016/j.conengprac.2026.107014