OhmNet: Advanced neural network-based viscosity prediction of sauces for efficient Ohmic heating processing.

JAVED, Tasmiyah, AKMAL, Muhammad <http://orcid.org/0000-0002-3498-4146>, BREIKIN, Timofei, MILLMAN, Caroline <http://orcid.org/0000-0003-4935-0477> and ZHANG, Hongwei <http://orcid.org/0000-0002-7718-021X>

OhmNet: Advanced neural network-based viscosity prediction of sauces for efficient Ohmic heating processing

Tasmiyah Javed, Muhammad Akmal, Timofei Breikin, Caroline Millman, Hongwei Zhang

PII: S1466-8564(25)00183-3

DOI: https://doi.org/10.1016/j.ifset.2025.104099

Reference: INNFOO 104099

To appear in: *Innovative Food Science and Emerging Technologies*

Received date: 1 April 2025

Revised date: 30 May 2025

Accepted date: 25 June 2025

Please cite this article as: T. Javed, M. Akmal, T. Breikin, et al., OhmNet: Advanced neural network-based viscosity prediction of sauces for efficient Ohmic heating processing, *Innovative Food Science and Emerging Technologies* (2024), https://doi.org/10.1016/j.ifset.2025.104099

OhmNet: Advanced Neural Network-Based Viscosity Prediction of Sauces for Efficient Ohmic Heating Processing

Tasmiyah Javed[a]*, Muhammad Akmal[a], Timofei Breikin[a], Caroline Millman[a], and Hongwei Zhang[a]*

[a] *Advanced Food Innovation Centre, Sheffield Hallam University, Sheffield, S1 1WB, UK.*

T.Javed@shu.ac.uk, m.akmal@shu.ac.uk, t.breikin@shu.ac.uk, c.e.millman@shu.ac.uk, h.zhang@shu.ac.uk.

Abstract

Industrial food processes such as Ohmic Heating (OH) are gaining popularity due to their lower carbon emissions and improved energy efficiency. The effectiveness of OH largely depends on the electrical conductivity, physical properties, and rheological characteristics of the food product, with dynamic viscosity directly influencing the fluid flow, residence time, and heating rate in a Continuous Flow Ohmic Heating (CFOH) system. Therefore, accurate prediction of viscosity during CFOH processing is crucial for optimising heating efficiency and maintaining the desired output temperature, ultimately reducing energy consumption and operational costs. To address this challenge, this study introduces *OhmNet* - an advanced Neural Network (NN)-based predictive model designed to accurately estimate the dynamic viscosity of tikka sauce during OH, offering a robust solution for viscosity prediction in CFOH applications. The predictive model has been developed using real-time data obtained from heating experiments, where viscosity measurements were recorded using a rheometer at varying target temperatures. To achieve the optimal configuration of *OhmNet*, three different approaches were explored: separate network development for each target temperature, a transfer learning-based neural network, and a one-hot encoding-based unified neural network model. These approaches were systematically evaluated through a grid search for hyperparameter tuning to identify the most accurate and robust dynamic viscosity predictive model during Continuous Flow Ohmic Heating. The resulting *OhmNet* model demonstrates high performance and reliability, achieving a Mean Squared Error (MSE) of 0.002, a Mean Absolute Error (MAE) of 0.025, and a coefficient of determination (R^2) equal to 0.99. This optimal configuration of *OhmNet* offers a powerful tool for enhancing process efficiency and control in industrial food processing applications. In the future, the model can be seamlessly integrated with advanced process controllers for precise temperature control and power consumption optimisation, driving sustainable and energy-efficient food processing applications.

Keywords: viscosity prediction; tikka sauce; Artificial Neural Network (ANN); Ohmic Heating (OH).

1. Introduction

The physical and rheological properties of food products play a fundamental role in the design and optimisation of industrial food processes. These properties are critical not only for ensuring process efficiency but also for preserving quality, safety, and organoleptic attributes of final products (Rai et al., 2005). Among them, viscosity is an essential parameter due to its direct influence on flow behaviour, heat transfer efficiency, residence time, and energy requirements during thermal processing (Said Toker et al.). These process parameters are especially relevant in the context of emerging processing technologies such as Ohmic Heating (OH).

Ohmic Heating is an advanced thermal processing technique that leverages the electrical conductivity of food to generate internal heat through resistance. Compared to conventional heating, OH offers faster

and more uniform heating, lower energy consumption, and reduced carbon emissions, making it attractive for sustainable food production (Silva et al., 2022). A specialised application of OH, known as Continuous Flow Ohmic Heating (CFOH), is particularly suitable for processing pumpable food products such as soups, slurries, juices, and sauces. In CFOH systems, viscosity becomes a controlling factor for key operational parameters, including flow rate, voltage input, target temperature, and energy consumption (Manzoor et al.).
However, viscosity measurement during CFOH remains a challenge. Conventional rheometry is labour-intensive, offline, and often impractical in real-time settings. While empirical models such as the Arrhenius and power-law equations have historically been used to estimate viscosity, they fall short when dealing with the non-linear and dynamic behaviour of complex food fluids under real processing conditions (Karaman et al., 2014; Peleg, 2017). Likewise, traditional regression models (e.g., linear or polynomial) fail to capture the non-Newtonian behaviour of materials like sauces and lack robustness across different formulations and heating profiles (Oroian, 2013).

To address these limitations, the food science community has increasingly turned to Machine Learning (ML) techniques, particularly Artificial Neural Networks (ANNs), which excel in modelling complex, nonlinear relationships between input features and output responses. In the domain of food processing, ML models have been applied to predict viscosity across diverse applications, ranging from fruit purees to dairy gels and starch-based systems. For example, (Perera et al., 2025) proposed a grey-box soft sensor for predicting melt viscosity in polymer extrusion processes by combining physics-based models with Recurrent Neural Networks (RNNs), achieving an improvement of approximately 95% over radial basis function neural networks (RBFNN). Similarly, a notable study by (Heidari et al., 2016) used multilayer perceptron (MLP) networks to predict the viscosity of fruit purees, achieving high accuracy by tuning hyperparameters such as the number of neurons and hidden layers (Heidari et al., 2016). This study also emphasised the importance of hyperparameter tuning and model architecture optimisation to reduce prediction errors and improve model robustness.

Table 1 presents a comparative summary of recent ML-based studies on viscosity prediction in food systems. These studies span a variety of food matrices and heating technologies, including conventional, microwave, and ohmic heating, and deployed algorithms ranging from feedforward ANNs to hybrid physics-informed models. Notably, while ML has been explored for several thermal processing methods, applications in OH and especially CFOH remain sparse. Research by (Silva et al., 2020) is one of the few studies that investigated ML for OH processes, but focused more on electrical conductivity and microbial inactivation than rheological modelling.

While machine learning models, particularly neural networks, have been previously employed in engineering disciplines to estimate viscosity and other fluid properties, their integration into real-time Ohmic Heating (OH) applications in the food industry remains largely unexplored. Most existing works focus on static rheological models under ideal lab conditions or use simulated data with limited industrial relevance. Additionally, despite numerous studies demonstrating the potential of ML techniques for viscosity prediction, most of these models are either domain-specific or lack generalisability across different processing conditions. Many existing models do not account for the dynamic nature of viscosity changes during continuous flow processing, limiting their application in real-time monitoring systems. Nevertheless, the challenge is to develop a predictive model that can generalise viscosity predictions over a wide range of temperature profiles while maintaining robustness and accuracy.

These trends reveal a clear research opportunity to develop a specialised, ML-based predictive model for viscosity in the CFOH process. Such a model must account for the unique heating behaviour of ohmic systems, the non-Newtonian nature of food fluids, and the need for real-time adaptability. To bridge this gap, this research introduces OhmNet - a feedforward ANN designed to predict the dynamic viscosity of tikka sauce during CFOH. This work utilises real-time experimental data collected via a Modular Compact Rheometer (MCR 302), calibrated to mimic the flow and thermal dynamics of a

pilot-scale CFOH system. The dataset comprises three heating profiles (75°C, 90°C, and 100°C), capturing the nonlinear temperature-viscosity trend of non-Newtonian fluids.

To develop OhmNet, three distinct approaches were considered:

1. Separate Network Approach: Independent networks were trained on each dataset, optimising the model architecture for each temperature range.
2. Transfer Learning Approach: A pre-trained model was fine-tuned on subsequent datasets to transfer knowledge from one temperature range to another.
3. One-Hot Encoding Unified Model: A single network was trained using combined datasets with one-hot encoding to represent different temperature ranges.

These approaches were systematically trained and evaluated over three different datasets with various temperature ranges using hyperparameter tuning to identify the optimal configuration in terms of training function, number of hidden layers, and neuron count. Performance metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and coefficient of determination (R^2) were used to evaluate model accuracy. This robust OhmNet may be integrated with data-driven controllers to adapt the CFOH voltage inputs according to the variations in process conditions, maintaining high accuracy throughout extended operations.

Therefore, this study positions OhmNet as the first data-driven, domain-specific viscosity prediction tool for CFOH, capable of real-time integration with process controllers for adaptive voltage regulation. By enabling robust and accurate viscosity estimation, OhmNet supports the broader goal of achieving sustainable, intelligent food processing systems in the food industry.

Study (Year)	Heating Methods	ML Model	Input Features	Food sample	Performance	Observation/ Novelty
(Siejak et al., 2024)	Ambient lab	Decision Tree (DT) Random Forest (RF)	Colour indices (L*,a*,b*); pH; concentration; electrical conductivity	Pectin solution samples	DT: $R^2 \approx 0.999$; RMSE = 0.108 RF: $R^2 \approx 0.998$; RMSE = 0.294	First ML approach for pectin viscosity. DT/RF vastly outperformed single neural networks. Demonstrated rapid and accurate viscosity estimation from easy-to-measure physical properties, aiding in quality control.

Study (Year)	Heating Methods	ML Model	Input Features	Food sample	Performance	Observation/ Novelty
(Dahl et al., 2025)	High moisture extrusion	Random Forest (RF)	Formulation composition (pea/faba protein %, starch, pectin, cellulose, carrageenan levels); moisture content; processing cluster features	140 formulations of plant-protein/polymer mixes (meat analog)	Single-output RF accurately predicted linear viscoelastic parameters (e.g. G′) but struggled on non-linear. Multi-output RF (using large-deformation inputs) achieved $R^2 \approx 0.94$ on non-linear viscosity parameters	Integration of formulation and rheology: Used clustering & variable importance to reduce features. Multi-output RF captured interdependency of rheological metrics, yielding accurate viscosity and yield stress predictions for complex plant mixes. Highlights data needs for non-linear regimes.
(Yang et al., 2025)	Conventional (no active heating; properties measured post-processing)	Combined optical Monte Carlo simulation + ML (regression/ANN)	Optical scattering and absorption parameters (simulated via Monte Carlo); particle size and concentration proxies	Apple puree with varying particle size (light–particle interactions)	Reported high predictive accuracy (noted model captured viscosity and viscoelastic moduli trends).	Hybrid simulation + ML approach to infer viscosity from optical behaviour. Provided mechanistic insight: how microstructure (particle–light interactions) affects puree viscosity. Useful for non-contact viscosity estimation in fruit purees.
(Lie-Piang et al., 2023)	Conventional heating	Spline regression Random Forest (RF) Neural Network (NN)	Ingredient blend ratios (pea vs. lupin fractions); basic composition data (protein, starch, fibre content)	Multi-crop protein ingredients (pea, lupin) – various blend formulations	All models could predict key functional properties. The unified model (both crops) had a slightly higher error than crop-specific models. RF was effective, but needed more data for complex traits.	Predict *techno-functional* properties (viscosity, gelation, emulsion stability, foaming) from low-refined ingredients. Showed ML can generalise across crop types, supporting sustainable ingredient selection (trade-off: ~5–10% higher error for a universal model).

Study (Year)	Heating Methods	ML Model	Input Features	Food sample	Performance	Observation/ Novelty
(Batista et al., 2021)	Conventional (milk pasteurisation & fermentation for yogurt)	Artificial Neural Network (feed-forward ANN)	Process conditions (fermentation temperature/time); formulation (solids %, stabiliser levels)	Non-fat yogurt experiments under varied formulation & incubation conditions	ANN models accurately predicted texture metrics. For viscosity ("ANN-VIS"), the typical prediction error was low (R^2 ~ 0.95 reported)	Predicted rheology & texture in dairy: Developed separate ANN sub-models for viscosity (flow index) and TPA texture. Enabled optimisation of yogurt formulation by linking ingredients/process to final viscosity and mouthfeel.
(Rocha et al., 2020)	Ohmic Heating (pasteurization of cheese curd)	Multiple ML methods (e.g., PLS, ANN)	Voltage gradient (0, 4, 8, 12 V/cm) vs. conventional heating, process time; product properties (e.g., dispersion viscosity profile, temperature)	Minas Frescal (fresh cheese) – ohmically heated vs conventionally heated batches	Developed Models showed "good agreement" with experimental results. Able to identify how ohmic parameters influence the viscosity profile and sensory attributes.	First ML in the ohmic processing of food. Showed ohmic heating yields distinct rheological outcomes, and ML could link process settings to viscosity and sensory drivers (e.g., juiciness, colour). Demonstrated feasibility of real-time viscosity prediction under ohmic treatment for process optimisation.
(Chen et al., 2019)	Hydrothermal (high-pressure thermal pretreatment)	Feed-forward ANN	Temperature, residence time, solid concentration, particle size	Microalgae slurry (Chlorella) undergoing hydrothermal liquefaction	Reported that empirical correlations failed, but ANN predicted viscosity well, where experiments showed non-linear behaviour	Pioneered ANN for high-pressure/high-temp slurry viscosity. Enabled real-time estimation of slurry viscosity during biomass pretreatment, improving scale-up design. Showed ML can capture complex interactions (e.g., thermal cell disruption vs. viscosity) better than classical models.

Study (Year)	Heating Methods	ML Model	Input Features	Food sample	Performance	Observation/ Novelty
(Perera et al., 2025)	Extrusion (polymer melt analog for food extruders)	Hybrid DNN + physics ("grey-box" soft sensor)	Extruder settings (screw speed, barrel temps, etc.); physics model's viscosity estimate as input feature	Non-food: Polymer melt in extruder (real-time monitoring scenario)	NRMSE ≈ 0.22% (virtually zero error), ~95% error reduction vs. prior pure-ML model (RBF network), Outperformed standalone MLP and LSTM networks in real-time viscosity tracking.	Combined a first-principles extrusion model with a correcting DNN. Real-time prediction of melt viscosity achieved. Illustrates state-of-the-art approach – merging domain knowledge with ML, which could inspire next-generation food process models.

2. Materials and Methods

2.1. Data collection and Preprocessing

The viscosity (mPa.s) versus temperature (°C) data for the tikka sauce, prepared using a proprietary formulation, were collected using a Modular Compact Rheometer (MCR 302, Anton Paar). The rheometer was calibrated to match the operating parameters of a pilot-scale Continuous Flow Ohmic Heating (CFOH) system to ensure the experimental measurements were representative of real-time processing conditions. In particular, the flow rate was fixed at 1 L/min, which had been previously validated as the optimal rate for starch activation and effective cooking in the heating chamber. This calibration ensured accurate simulation of CFOH dynamics during viscosity profiling.

Each experimental run involved heating the sauce from ambient temperature to the three final target temperatures: 75 °C, 90 °C, and 100 °C. Data were collected over a period of 500 seconds for each run, with a sampling rate of 1 data point per second. This yielded approximately 4000 data points per dataset, with temperature ranges spanning 20°C to 95°C. The flow rate setting corresponded to a shear rate range of 0 – 42.45 s^{-1}, calculated using the relation in Equation 1, which ensured relevance to actual fluid dynamics under processing conditions.

$$\dot{\gamma} = {}^{8V_f}\!/_{\pi R^3} \qquad (1)$$

where V_f is the volumetric flow rate (m^3/s) and R is the pipe radius (m) (Icier & Bozkurt, 2009).

Temperature-viscosity data pairs collected during each run were used to construct three distinct datasets, representing different thermal processing scenarios. This enabled robust training, validation, and testing of the predictive model while allowing assessment of its generalisation across varying final temperature conditions. The observed data exhibited a smooth, nonlinear decline in viscosity with increasing temperature, typical of shear-thinning behaviour in non-Newtonian food materials like tikka sauce.

Similarly, the heating rate of the rheometer was set to 8.114 °C/min, determined using the heat transfer (Equation 2) to ensure consistency between electrical power input and thermal gain of the fluid during CFOH:

$$Q = mc_p\Delta\theta \qquad (2)$$

where Q is the input power (W), m is the mass of substance in the heating chamber at a particular instance (kg), c_p is specific heat (J/kg °C), and $\Delta\theta$ is the heating rate (°C/min).

Before being used for model training, all raw viscosity data underwent thorough preprocessing. This preprocessing involved identifying and removing outliers caused by transient fluctuations or instrument noise. Outliers were identified based on deviations beyond three standard deviations from the local trend. Additionally, duplicate records with identical temperature and viscosity values from repeated measurements were eliminated to prevent data redundancy and bias. The cleaned datasets were then normalised using Min-Max Scaling to map temperature and viscosity values into a standardised range of 0 to 1 using Equation 3. This normalisation approach not only mitigated the risk of convergence issues during neural network training but also ensured that each feature contributed proportionately to error minimisation (Raju et al., 2020).

$$p_{norm} = \frac{p - p_{min}}{p_{max} - p_{min}} \qquad (3)$$

where, p_{norm} is the normalised data point, p is the data point at a particular instant, p_{min} and p_{max} are the minimum and maximum values for each data point of the variables (temperature or viscosity) in the dataset.

Following normalisation, the cleaned datasets were randomly split into training, validation, and testing subsets in a 70%:15%:15% ratio. The training set was utilised to optimise the neural network weights, while the validation set was employed for hyperparameter tuning and to prevent overfitting. The testing set was kept completely independent to evaluate the final model performance. Each of the three temperature-target datasets was split separately following this ratio. In the combined modelling approaches (described in subsequent sections), the unified data from all three temperature runs were concatenated, shuffled, and split in the same 70:15:15 proportion.

Figure 1 provides a clear and systematic overview of the entire process, from data collection and preprocessing to modelling the neural networks. It also highlights how the predicted viscosity will be integrated into the voltage input controller used for the CFOH plant, demonstrating the practical application of the developed model.

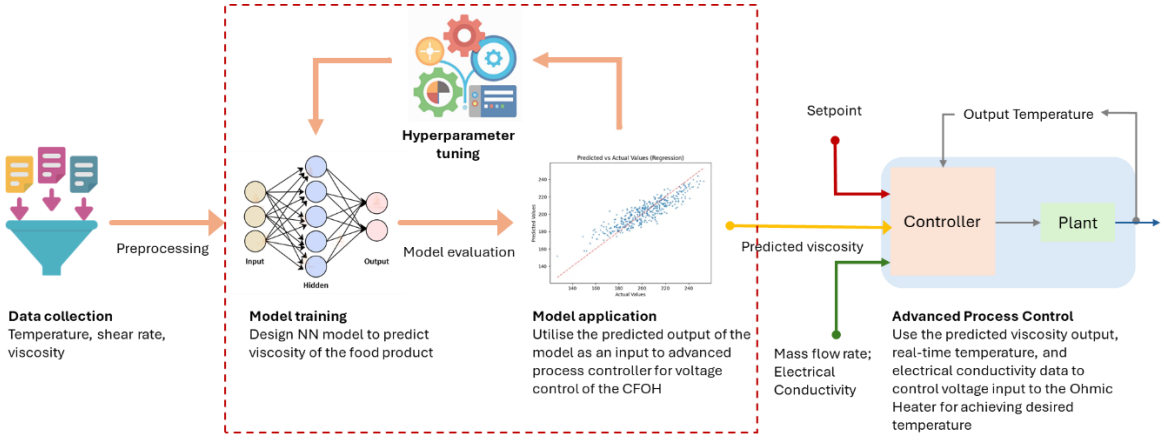


Figure 1: Schematic diagram of the viscosity prediction model and its integration with the controller.

2.2. Neural Network Architecture Design

To predict sauce viscosity from temperature, a feed-forward artificial neural network (ANN) – termed *OhmNet* was developed. Three modelling strategies were explored to leverage the available datasets:

2.2.1. Separate Single-Dataset Networks:

In this approach, three independent ANN models were trained separately, one for each targeted temperature dataset (75 °C, 90 °C, 100 °C). Each model was designed to learn the specific viscosity–temperature relationship corresponding to its respective heating profile. By treating each dataset in isolation, this approach allowed the model to capture unique characteristics of each heating profile or composition batch. Training separate models provided a performance baseline and served as a benchmark to evaluate whether a model trained on one temperature range could generalise to another. It is commonly used in applications where data segments differ structurally or contextually (Bhagya Raj & Dash, 2022; Liu et al., 2023). Additionally, each of these networks underwent comprehensive hyperparameter tuning, exploring configurations with 1, 2, and 3 hidden layers, using different training functions and varying the number of neurons in each hidden layer to optimise performance. This strategy enables insight into whether a model trained on one condition generalises poorly compared to others, justifying the need for more flexible architectures (Abinaya et al., 2024).

2.2.2. Sequential Transfer Learning:

Transfer learning has gained traction in domains with limited data, offering a means to reuse learned representations from one dataset on another related task (Pan, 2010; Tan et al., 2018). The transfer learning paradigm was employed where a single neural network was first trained on the 75 °C dataset, then fine-tuned sequentially on the 90 °C dataset, and finally on the 100 °C dataset. In each stage, the previously trained network's weights served as the initial weights for the next dataset, and only a few epochs of additional training, with a smaller learning rate, were performed to adjust the model to the new data (Abdalla et al., 2019). This progressive fine-tuning allowed the model to transfer learn data trends from one heating condition to the next, under the assumption that the general functional relationship between temperature and viscosity is similar across datasets. The transfer learning strategy aimed to improve data efficiency and generalisation by utilising the largest dataset first and adapting to incremental differences in the subsequent datasets (e.g., extended temperature range up to 100 °C). This approach leverages the shared underlying rheological structure across different heating runs to improve generalisability while reducing training time and overfitting risk (Guo, Y. et al.).

2.2.3. One-Hot Encoding Unified Model:

One-hot encoding is a widely adopted technique for representing categorical variables in neural networks (Bishop, 1995). This approach combined all data into one large dataset, and a single unified ANN was trained to handle all scenarios. To inform the network of which heating profile a data point came from, we introduced a one-hot encoded input to represent the dataset identity. Specifically, an extra binary input node was added for each of the three conditions (75, 90, 100 °C), set to 1 for the corresponding condition and 0 for the others. This way, a sample from the 75 °C-target dataset would have the [75°C] indicator = 1 (and others 0), whereas a sample from the 100 °C run would have the [100°C] indicator = 1, etc. This one-hot label acts as a contextual flag so that the network can implicitly learn any shift or scaling in the viscosity–temperature relationship between the different final temperatures. The unified model training used the entire pooled dataset with the one-hot feature, with the expectation that it could generalise across the full temperature range (ambient to 100 °C) within one network. This approach tests the model's capacity to internalise a broad temperature–viscosity mapping, offering scalability and simplicity if performance remains high (Goodfellow et al., 2016). Figure 2 illustrates the three modelling strategies for optimising the proposed *OhmNet* for predicting sauce viscosity.

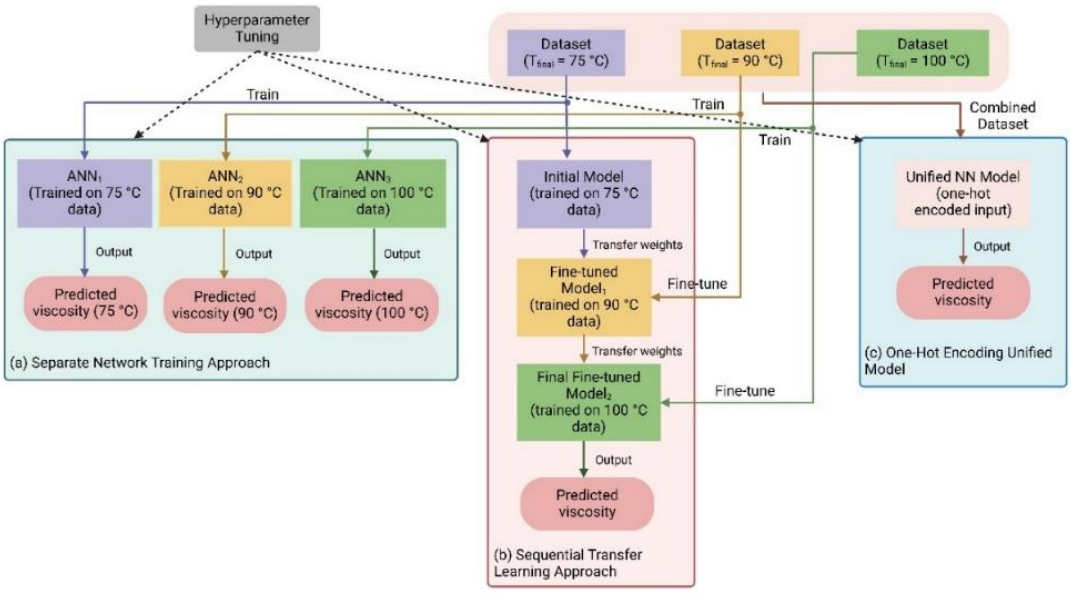


Figure 2: Block diagram illustrating the three modelling strategies for viscosity prediction

All three modelling approaches explored in this study were implemented using a consistent base architecture: a multi-layer feed-forward artificial neural network (ANN) with one input neuron (temperature) and one output neuron (predicted viscosity). The design allowed for flexibility in depth, enabling configurations with one, two, or three hidden layers. The number of neurons in each hidden layer was systematically varied using fixed values of 4, 7, and 9, selected to span from a compact network to a moderately complex one, given the single-input/single-output structure of the problem.

The architecture of *OhmNet* was intentionally kept shallow, with 1 to 3 hidden layers and relatively few neurons per layer. This decision was driven by the low-dimensional nature of the dataset, temperature being the sole independent variable, and the smooth, monotonic nature of the viscosity-temperature relationship typically observed in non-Newtonian fluids like tikka sauce. Given these conditions, simpler network architectures proved sufficient to model the nonlinear dependencies effectively while maintaining computational efficiency and avoiding overfitting (Tan et al., 2018). A grid search across architecture configurations confirmed that more complex networks did not significantly improve performance.

To ensure that the design choices were both empirically justified and theoretically sound, an established formula (Equation 4) was used to guide the initial neuron count, N_h:

$$N_h = 2(\sqrt{(Y+2)})X \qquad (4)$$

where X is the number of input neurons and Y is the number of output neurons (Huang, 2003).

For our case ($X = 1$, $Y = 1$), the equation suggests an initial estimate of 3 – 4 neurons. Based on this, the search space was extended to 7 and 9 neurons to capture more complex nonlinearities where needed. In multi-layer configurations, symmetric sizes (e.g., [7, 7]) and asymmetric layouts (e.g., [4, 7, 9]) were tested. The best-performing architecture was a three-hidden-layer configuration with 7, 9, and 4 neurons, respectively, demonstrating a strong balance between predictive power and generalisability.

The activation functions used in the hidden layers were log-sigmoid (logsig), introducing the necessary nonlinearity, while the output layer used a purelin (linear) function to support real-valued viscosity

outputs. This setup is common in regression-based neural models, particularly where the outputs are continuous and unbounded (Mirarab Razi et al., 2014).

To train the OhmNet models, three widely-used optimisation algorithms were compared: resilient backpropagation (*trainrp*), Levenberg–Marquardt (*trainlm*), and Bayesian regularisation (*trainbr*). The resilient backpropagation algorithm updates weights based only on the sign of the gradient, which can be more robust to noisy gradients. The Levenberg–Marquardt algorithm is a fast second-order optimisation method that often yields quicker convergence for moderate-sized networks, and has been reported to perform well in viscosity prediction tasks (Afrand et al., 2016). Bayesian regularisation modifies the training objective to include a weight penalty, effectively preventing overfitting by implementing an automatic regularisation. It typically yields slightly slower training but can improve generalisation on limited data (Naidu et al., 2020).

A grid search was conducted across combinations of layer depths, neuron counts, and training algorithms. Each training algorithm was run in batch mode with an initial learning rate set in the range of 0.01 to 0.001. We empirically tuned the learning rate and observed that higher values (around 0.01) sped up initial learning but risked oscillation, whereas lower values (0.001) gave more stable, albeit slower, convergence. Ultimately, a learning rate of 0.005 was adopted as a good compromise for most runs, and in cases of transfer learning fine-tuning, an even smaller rate (0.001) was used to gently adjust the pre-trained weights. Each network was trained for a maximum of 1000 epochs, with early stopping if the validation error did not improve for 20 consecutive epochs to prevent overfitting. Weight initialisation was randomised, and we performed 5 repeats for each configuration to mitigate any effects of random initialisation on the results. The algorithm outlining the general modelling strategies and the optimisation of network design parameters is presented in Table 1.

This structured and empirically driven approach confirmed that a relatively simple ANN architecture could successfully capture the underlying viscosity dynamics of tikka sauce across the CFOH process, validating the suitability of *OhmNet* for real-time predictive control.

Table 1: Optimisation and development of the OhmNet employing three modelling strategies.

	Algorithm: OhmNet Development and optimisation	
1	**Require:** Temperature-viscosity datasets (D_T)	
	Preprocess data	
2	Normalise dataset	
3	**for** $i = 1 : D_T$ **do**	% temperature range (T_r)
4	**for** $j = 1 : A_t$ **do**	% training approach (A_t)
5	**for** $k = 1 : F_t$ **do**	% training function (F_t)
6	**for** $l = 1 :$ *HL* configuration **do**	% hidden layer (HL)
7	**for** $m = 4 : N_n$ **do**	% neuron count per layer (N_n)
8	NN setup with input (temperature); output (viscosity)	
9	Initialise NN weights and biases	
10	Select T_r and F_t	
11	Split data (0.70:0.15:0.15)	
12	**for** p = 1 : P_m **do**	% NN configuration (P_m)
13	Fit NN model to training data	
14	Validate the model on validation data	
15	Store performance metrics (MSE, MAE, R^2)	
16	**end for**	
17	Calculate average accuracy metrics on validation data	
18	**end for**	
19	**end for**	
20	**end for**	
21	**end for**	
22	**end for**	
23	Test the best model on unknown data	
24	Calculate final performance metrics (MSE, MAE, R^2)	
25	Compare final performance metrics of all models	

26	Select optimal configuration
27	Arrange network models in descending order based on performance metrics
28	**return** OhmNet

2.3. Performance Evaluation and Hyperparameter Optimisation

To ensure robust and generalisable model development, we implemented a systematic hyperparameter optimisation and performance evaluation framework for *OhmNet*. The architecture search included exhaustive combinations of hidden layers (1 to 3), neuron counts per layer (4, 7, and 9), and training algorithms—resilient backpropagation (*trainrp*), Levenberg–Marquardt (*trainlm*), and Bayesian regularisation (*trainbr*). Each configuration was assessed through a grid search strategy, guided by validation performance, to determine the optimal model for each of the three prediction approaches explored in this study. This grid search strategy ensured the influence of each architectural choice and training method on the performance of the developed ANN.

Model performance was evaluated using three standard regression metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). These metrics are widely recognised for regression problems in food engineering and rheological property modelling (Icier & Bozkurt, 2009; Said Toker et al.). MSE, which penalises larger errors, was used as the primary objective during training. MAE provides an intuitive sense of average prediction error in physical units, and R^2 quantifies how well the model captured variance in viscosity (Barradas Filho et al., 2015).

During training, the objective was to minimise MSE. Validation MSE was monitored for early stopping, while final model selection among candidates was based on a combination of the lowest validation MSE and the highest validation R^2 (to ensure the model not only has low error but also captures variance well). We also report MAE for an intuitive measure of error magnitude, following common practice in viscosity predictions (Icier & Bozkurt, 2009).

Further, to mitigate the risk of overfitting, especially critical in data-driven rheological modelling, we adopted several best-practice strategies . First, each dataset was randomly partitioned into training (70%), validation (15%), and testing (15%) subsets. The training set was used for learning model weights, the validation set for early stopping and hyperparameter tuning, and the test set for independent performance assessment, completely unseen during training or tuning. Early stopping was applied with a threshold of 20 epochs, terminating training when validation MSE no longer improved, thereby preventing overtraining.

To account for the stochastic nature of neural network training, we performed Monte Carlo repeated averaging. Each network configuration was trained five times with different random seeds and dataset splits. This approach is endorsed in the literature as a robust alternative to k-fold cross-validation when dealing with time-series or continuous input data structures (Barradas Filho et al., 2015; Peleg, 2017). The final performance scores were averaged to ensure that results were not artifacts of favourable random initialisations.

Further, architectural simplicity was enforced by restricting models to shallow networks (1–3 hidden layers) with modest neuron counts, aligning with prior studies in viscosity prediction of fruit juices and food systems, which found minimal performance gains from deeper models while increasing overfitting risk (Bhagya Raj & Dash, 2022; Chen et al., 2021; Goyal, 2014). For configurations using the *trainbr* algorithm, Bayesian regularisation was embedded in the loss function to penalise excessive weight magnitudes and enhance generalisation.

Notably, k-fold cross-validation was not employed in this study due to the continuous and monotonic nature of the temperature-viscosity data. Arbitrary partitioning inherent to k-fold approaches could violate the physical continuity of the data, thus undermining both interpretability and learning performance. Instead, our combination of repeated

train/validation/test splits, early stopping, regularisation, and conservative architecture design provided a principled and effective safeguard against overfitting, while maintaining the physical relevance of the data patterns.

The described performance evaluation and validation strategy aligns with recent literature in viscosity prediction using ANNs, where similar practices have yielded highly accurate models for fruit juice, nanofluid, and biodiesel systems under varying conditions (Barradas Filho et al., 2015; Chen et al., 2021; Pourramezan et al., 2024)

3. Results & Discussion

3.1. Comparative Performance of Modelling Approaches & Configurations

The *OhmNet* model was evaluated using three different modelling strategies (Separate Networks, Transfer Learning, and One-Hot Encoding) across various network architectures and training algorithms. The three modelling approaches exhibited distinct performance trends. Overall, the one-hot encoding approach delivered the highest accuracy, followed by the transfer learning approach, with the separate-networks approach slightly lagging. This indicates that leveraging a unified network with categorical (one-hot) inputs or transfer of learned features can improve generalisation compared to training individual networks in isolation. The multi-task learning effect likely improves generalisation by knowledge sharing between related tasks (Hu et al.). Table 2 and figure 3 illustrate the performance of the best configuration from each approach in terms of mean squared error (MSE), mean absolute error (MAE), and R^2.

Table 2: Quantitative results of the best performing networks of the three ANN modelling approaches

Approach	***MSE***	***MAE***	***R^2***
Separate Networks	0.0035	0.040	0.970
Transfer Learning	0.0022	0.030	0.985
One Hot Encoding	0.0020	0.025	0.990

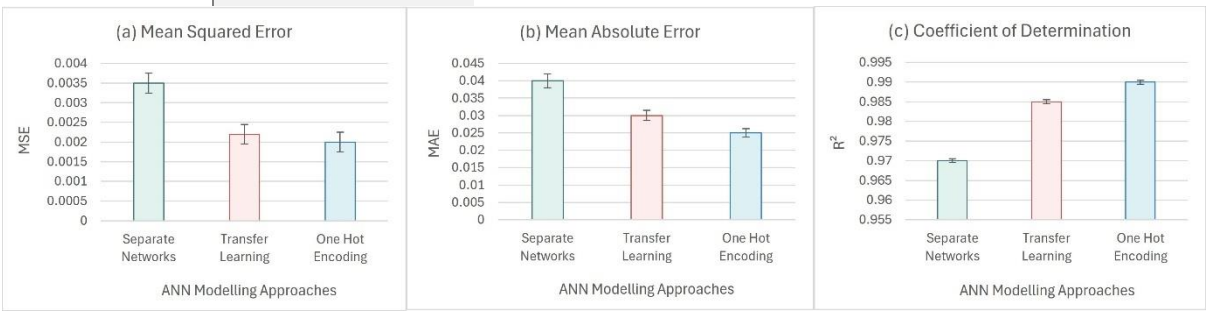


Figure 3: Performance of the best-performing network from each. The one-hot encoding model shows the lowest MSE and MAE and the highest R^2, indicating superior accuracy.

Several factors contribute to the superior performance of the one-hot encoding approach. Firstly, the unified model was trained on the entire dataset with a categorical indicator, effectively increasing the training sample size and allowing the network to learn general viscosity patterns applicable across all subsets. Common dependencies could be learned from all data, while the one-hot input neuron enabled the network to adjust predictions for each specific category. This resembles a multi-task learning

scenario in which related viscosity prediction share representation learning. The result is improved with overall prediction accuracy. This is in line with previous findings that combining data from related cases can enhance model performance. For instance, Rai *et al.* developed a single ANN model to predict the viscosity of various fruit juices (orange, peach, pear, etc.) as a function of concentration and temperature, and achieved a low prediction error using a unified network (Rai et al., 2005). Their unified model handled multiple juice types with high accuracy, analogous to the proposed *OhmNet* one-hot network handling multiple categories in one model. In this study, the one-hot network similarly benefited from the broader data coverage, yielding high R^2 and low error across all categories.

In contrast, the 'Separate Networks' approach showed more variable results. Some individually tuned networks achieved a good fit on their specific data, but overall they tended to underperform the one-hot model when comparing test-set accuracy across conditions. This is because each separate network was trained on a smaller subset of data, making it more prone to overfitting or underfitting for that specific scenario.

The 'Transfer Learning' strategy produced intermediate results. The trend learned from one dataset was used to initialise training on another. This improved the learning efficiency and final accuracy on smaller datasets, narrowing the performance gap relative to the one-hot model. Indeed, the transfer-learned models consistently outperformed the purely separate models on subsets with limited data, confirming that transfer learning is particularly effective when data for certain tasks is limited (Lee & Lee, 2023). In our results, transfer learning improved the R^2 of the weakest separate network by several points (from 0.92 to 0.95), bringing it closer to the unified model's performance. These trends show that under ohmic heating, a unified modelling technique may maximise common patterns in the viscosity behaviour of tikka sauce, therefore providing an advantage in predicting accuracy and consistency across all configurations.

Another notable trend is the effect of model complexity (hidden layer configuration) on each approach. Increasing the number of hidden layers/neurons generally improved prediction accuracy for all approaches, but the magnitude differed. Figure 4 shows the best R^2 achieved by each approach as network complexity increased from a single hidden layer ([4]) to two layers ([4,7]) and three layers ([7,9,4]). All approaches saw R^2 rise with more complex architectures, reflecting the network's enhanced capacity to capture nonlinear viscosity relationships. Notably, the unified one-hot model benefited the most from larger architectures. The R^2 increased from 0.950 (1 layer) to 0.990 (3 layers). The transfer learning approach also reached high accuracy with the deepest network (R^2 = 0.985). Whereas, the separate models improved more modestly, with R^2 levelling off around 0.96–0.97 at the largest architecture, perhaps due to limited data per network. This suggests that the one-hot and transfer learning approaches could leverage the added complexity more fully, since they effectively had more data or better initial weights to train the larger networks. In contrast, a very complex network in the separate approach risked overfitting each small subset, yielding diminishing returns.

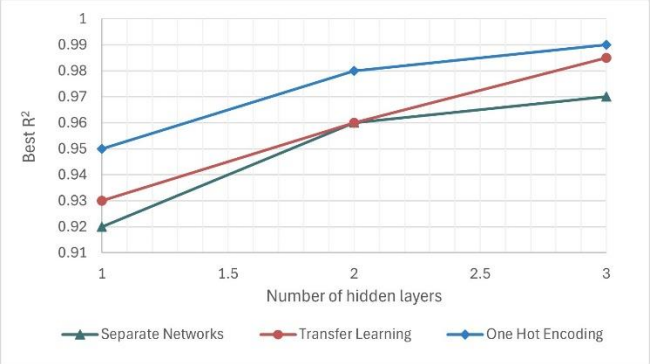


Figure 4: Trends in prediction R^2 (test data) as model complexity increases (1 hidden layer to 3 hidden layers) for each approach.

3.2. Impact of Training Algorithms on Model Performance

Across all approaches and network configurations, the choice of training algorithm had a notable impact on performance. Across all three approaches, networks trained with Levenberg–Marquardt (*trainlm*) or Bayesian regularisation (*trainbr*) consistently outperformed those trained with the simpler resilient backpropagation (*trainrp*). Models using *trainrp* converged more slowly and often yielded higher final errors (MSE and MAE), suggesting difficulty in optimising the complex mapping of temperature and viscosity with this method. In contrast, the *trainlm* algorithm often achieved the lowest training and validation errors among the three, alligning with its reputation for fast convergence and high accuracy in function approximation tasks (Pourramezan et al., 2024).

Many of the best models in our tests were obtained with *trainlm*, which frequently drove MSE to very low values. The *trainbr* algorithm, while sometimes slightly slower to converge, demonstrated comparable generalisation performance. Notably, *trainbr*-trained networks tended to have less divergence between training and testing error, indicating robust modelling. In some experiments, *trainbr* achieved prediction R^2 values as high or higher than *trainlm* for the same architecture, echoing findings in other studies where Bayesian regularisation provided the most accurate predictions (Guo, Q. et al., 2023).

Figure 5 compares the performance of the training algorithms in terms of average R^2 achieved across the developed models. The Levenberg–Marquardt algorithm yielded the highest average R^2 (0.955) and a very high peak R^2 (near 0.99) in at least one configuration. This aligns with literature reports that *trainlm* tends to find very accurate solutions for regression problems, often achieving the lowest MSE and highest correlation among training methods (Heidari et al., 2016).

In this study, *trainlm* drove the models to the lowest MSE values for nearly every network configuration, confirming its efficiency in minimising error. For example, using *trainlm*, the one-hot network with [7,9,4] layers reached an MSE of only 0.0020 (Table 1), whereas the same architecture trained with *trainrp* stalled at a higher MSE (0.0030). This gap is consistent with findings by Heidari *et al.*, who reported that Levenberg–Marquardt training achieved a lower MSE (on the order of 10^{-5}) for nanofluid viscosity prediction compared to gradient-descent-based training (which yielded MSE 10^{-3}), corresponding to a high R^2 (0.99998 for LM vs 0.9994 for a standard backprop) in their case (Heidari et al., 2016).

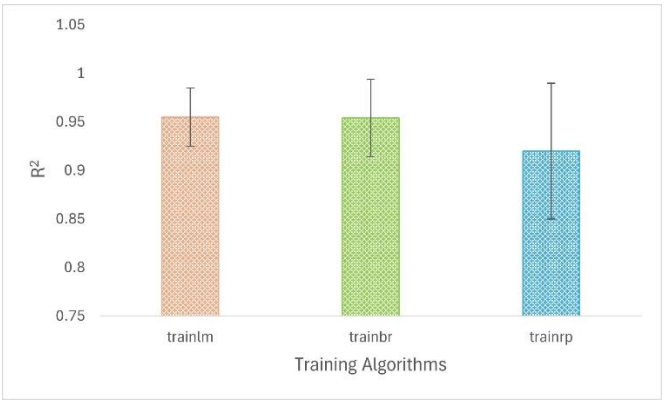


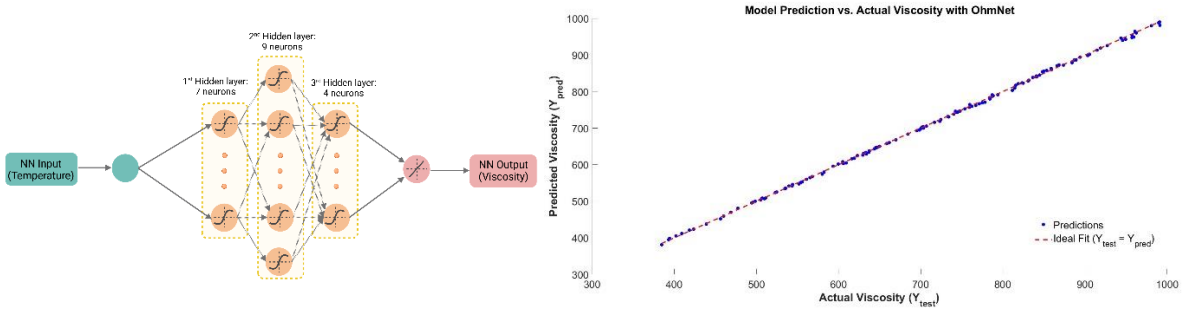
Figure 5: Average test R^2 achieved by each training algorithm (*trainlm:* Levenberg–Marquardt, *trainbr:* Bayesian regularisation, *trainrp:* resilient backpropagation).

The Bayesian regularisation training (*trainbr*) also performed strongly, with an average R^2 virtually identical to *trainlm* (Figure 5). *Trainbr* typically reached slightly higher final error than *trainlm* when models were not overfitting; however, it proved advantageous in scenarios prone to overfitting. In the separate-networks approach, the deepest architecture [7,9,4] tended to overfit when trained with *trainlm* (yielding a lower test R^2 of 0.93 despite very low training error). Applying Bayesian regularisation on that same architecture prevented overfitting by effectively penalising excessive weights, resulting in a higher test R^2 (0.97). This was the one case where *trainbr* outperformed *trainlm* in our results. Such behaviour is supported by previous studies, such as, Naidu *et al.*, reported that that *trainlm* and *trainbr* are the top-performing algorithms for regression tasks, significantly outperforming gradient-descent-based methods (Naidu et al., 2020). In our context, trainbr's built-in regularisation improved generalisation for complex models, making it the best choice for the separate approach's largest network and yielding the overall best separate-network performance (Table 2). Apart from that scenario, *trainbr* and *trainlm* were usually within a narrow margin of each other in accuracy, both outperforming *trainrp*.

For the viscosity prediction of tikka sauce, using these advanced training functions was crucial to minimise error; the choice between them came down to a trade-off between absolute error minimisation (*trainlm*) and slightly better generalisation stability (*trainbr*). In practical terms, this means that careful selection of the training algorithm can yield measurable improvements in model accuracy.

3.3. Best-Performing Network Configuration (OhmNet) Analysis

From the extensive grid of experiments, we identified a best-performing configuration (*OhmNet*). The top model was achieved using the 'One-Hot Encoding' strategy trained with the Levenberg–Marquardt (*trainlm*) algorithm, with a network architecture of three hidden layers ([7,9,4]), shown in figure 6. This particular setup attained the highest R^2 and lowest error metrics among all tested models. To quantify its performance: it reached an R^2 of 0.990 on the validation/test set, alongside a MSE of 0.0020 and a corresponding MAE of 0.025. Moreover, the *OhmNet* can predict the viscosity for all categories with only 0.2% error and capture 99% of the variance in the data. In practical terms, this means the model's viscosity predictions were very close to the experimental measurements across the range of heating conditions. Thus, the *OhmNet* not only excelled in absolute terms on its own test set, but also provided a solution applicable to all variations, which is a significant advantage for practical deployment.



(a) (b)

Figure 6: (a) Schematic architecture of the proposed OhmNet model trained with One-Hot Encoding approach for viscosity prediction. (b) Model prediction vs. actual viscosity prediction. The network comprises three hidden layers with 7, 9, and 4 neurons respectively, trained using the *trainlm* training algorithm. Log-sigmoid (*logsig*) activation functions are used in all hidden layers, while a linear (*purelin*) function is applied at the output layer.

Examining more closely the reasons this arrangement worked so effectively may credit a number of elements for its effectiveness. Firstly, the [7,9,4] architecture provides a large number of neurons and thus a high capacity to approximate complex nonlinear functions. Viscosity as a function of multiple input variables (temperature, composition, etc.) can be highly nonlinear. The deeper network can form hierarchical feature representations. First layer neurons might capture the base effects of temperature or concentration on viscosity, while deeper neurons capture interactions between variables. Simpler architectures ([4] or [4,7]) may underfit these relationships, whereas the [7,9,4] network can fit them more accurately given sufficient data. Indeed, we observed monotonic improvements in training and validation performance as we increased hidden layer complexity, until the point of reaching [7,9,4]. The superior performance of the [7,9,4] topology alligns with neural network theory that a network with more neurons/layers can model more complex functions, provided overfitting is controlled.

Secondly, the one-hot multi-task approach meant that the network was effectively trained on a larger dataset covering multiple conditions, which acts as a form of data augmentation and regularisation. The shared hidden layers had to learn features that are useful for predicting viscosity under all given conditions of formulation in the tikka sauce, leading to more robust feature learning. This explains the model's strong generalisation – it was less prone to fitting noise in any single condition because it always had to satisfy multiple scenarios simultaneously during testing.

Thirdly, using the *trainlm* algorithm allowed the model to converge to a very accurate solution. The Levenberg–Marquardt training facilitated quick adjustment of the large number of weights in the two hidden layers, efficiently finding a set of parameters that minimised MSE. We observed that this best model converged in relatively few epochs and the final training error was extremely low, indicating that *trainlm* successfully found a near-optimal fit in the weight space for this architecture. Additionally, although *trainlm* tends to overfit if unchecked, in this case, the risk was mitigated by the multi-task nature of the one-hot model and by early stopping on a validation set. This high level of performance is significant for food process modelling, as it suggests the model can accurately predict viscosity changes in tikka sauce during ohmic heating, potentially enabling precise control and optimisation of the heating process in real applications.

Given these reasons, it becomes clear why the *OhmNet* model excelled. It combined high model capacity, ample training data covering all scenarios, and a powerful training algorithm. Each of these elements was necessary to optimise all training, validation, and testing conditions. If unoptimised, a large network could overfit a small dataset, and an efficient optimiser cannot fix underfitting due to insufficient model complexity. However, combined optimisation steps produced a model that fits the viscosity data with high fidelity.

It should be noted that while the *OhmNet* model is the overall best, the transfer learning approach's top model was nearly as good, and would be a strong alternative if a unified model were not desired. The transfer learning model's success suggests that if one had to deploy separate models for each scenario (for practical or interpretability reasons), training them with an initial phase on a large dataset (or on a related task) is highly beneficial. In our case, pre-training on the full combined data and then fine-tuning for each specific category allowed the transfer models to almost reach the unified *OhmNet* model's accuracy. This approach capitalises on a similar principle – shared learning – but in a sequential manner. Literature on property prediction models also supports multi-task training or transfer learning that often yields better generalisation than isolated training. In materials and food processing domains, where data for certain products may be scarce, transfer learning is emerging as a powerful tool to improve predictive models (Lee & Lee, 2023).

3.4. Implications for Neural Network Design in Food Viscosity Prediction

The findings from this evaluation have important implications for designing neural network models in food rheology and similar domains. Integrative modelling approaches like the one-hot encoded *OhmNet* demonstrate clear advantages in scenarios where multiple product formulations or processing conditions are involved. Rather than developing and calibrating separate models for each variation of a sauce or each operating condition, a single well-designed network can handle all cases by including categorical inputs. This not only streamlines model development but, as shown by the performance of the one-hot approach, can improve predictive accuracy by pooling data, effectively making the training dataset richer and the learned model more robust.

In practical terms, a food manufacturer aiming to predict viscosity for different recipes of a sauce (mild, medium, spicy tikka sauce, etc.) under various heating profiles could deploy one unified *OhmNet* model rather than a suite of specialised models. The unified model would be easier to update and validate, and it would ensure consistency in predictions across products.

The success of the transfer learning strategy also offers a valuable guideline: when facing a new product or condition with limited data, the model can transfer knowledge from existing trained models. For instance, if a new type of sauce is introduced, instead of training a viscosity model from scratch, one could start with the pretrained model (trained on existing sauces) and fine-tune it on the new sauce's data. The results showed that this approach yields better initial accuracy than a standalone model on small data and approaches the performance of a fully trained multi-task model. This implies that neural network design for food applications should consider knowledge reuse by building a base model on a wide range of data (perhaps spanning many products) and adapting it, which can save time and data collection efforts while still providing high accuracy.

Another implication is the importance of choosing appropriate training algorithms and regularisation techniques for food property prediction networks. The contrast in performance between *trainlm*/*trainbr* and *trainrp* in our study underscores that algorithm choice can make or break a model's effectiveness. For complex, non-linear food processes like ohmic heating, second-order optimisation methods or Bayesian regularisation helped in navigating the error surface to find better minima. In practice, this means that developers of such models should leverage algorithms known to yield high performance, especially if the dataset size is moderate and the problem requires capturing subtle effects. Regularisation (*trainbr*) proved useful in controlling overfitting for larger networks – this suggests that techniques like Bayesian regularisation or early stopping, or alternative methods like dropout, could be beneficial when expanding network complexity for food viscosity models. Model complexity should be matched with training strategy, i.e., if a very large network is used, robust training/regularisation becomes crucial; otherwise, a slightly smaller network might be preferable to ensure the model generalises well.

These results contribute to the broader understanding of applying ANNs in food engineering. They demonstrate that an appropriately structured neural network can achieve high predictive accuracy in a challenging task like real-time viscosity prediction, which often involves dynamic changes and complex physicochemical interactions. This level of performance approaches that of experimental measurement accuracy, indicating that such models could be reliably used for process monitoring or control (e.g., to adjust heating in an ohmic process to reach a target temperature). The robustness of the best model means food technologists can trust its predictions across a range of conditions, making it a practical tool in the power control of the advanced process control of OH systems.

4. Conclusion

Continuous Flow Ohmic Heating (CFOH) has emerged as a highly efficient and sustainable technology for industrial food processing. It offers significant advantages over conventional heating methods, including high energy efficiency and reduced carbon footprint. These benefits make CFOH an attractive option for large-scale operations aiming to lower energy consumption and environmental impact without compromising processing throughput or product quality.

A key factor in realising the full benefits of CFOH is the precise control of food product quality. Viscosity is one of the vital parameters that govern how the product flows and how heat is distributed through it during ohmic processing. Accurate real-time viscosity prediction is therefore essential for maintaining process efficiency and stability. It enables operators to anticipate changes in flow behaviour, adjust pumping rates or electrical input parameters proactively, and ensure consistent heating.

Therefore, in this study, we developed an ANN model called *OhmNet* to predict the dynamic viscosity of tikka sauce under CFOH conditions. Three different ANN training strategies were explored for model development to identify the most effective learning approach. These strategies included training separate networks for different operating conditions, applying a transfer learning technique to leverage knowledge from one dataset to another, and using one-hot encoding to train a single unified network capable of handling multiple process scenarios. Each approach provided a unique way to capture the complex relationship between processing parameters (such as temperature) and the sauce viscosity.

Among the tested approaches, the one-hot unified modelling combined with the Levenberg–Marquardt backpropagation algorithm (*trainlm*) and a three-hidden-layer network architecture (with 7, 9, and 4 neurons in the respective layers) delivered the best performance. This optimised *OhmNet* configuration achieved an impressively low prediction error and an excellent fit to experimental data, as reflected by MSE of 0.002, MAE of 0.025, and R^2 of 0.99. These results indicate that *OhmNet* can accurately capture the rheological dynamics of the sauce during CFOH, outperforming the other training strategies tested. The high accuracy and low error metrics suggest that the model generalises well and can reliably predict viscosity changes in real time as processing conditions vary.

The strong performance of *OhmNet* has important implications for real-time process control in ohmic heating systems. Integrating this data-driven predictive model into a closed-loop process controller for CFOH setup can serve as a "soft sensor" for viscosity, providing instantaneous estimates without the need for manual sampling or off-line measurements. In practice, an *OhmNet*-based monitoring system could continuously feed viscosity predictions into the process controller, which then fine-tunes input power and flow rates. This would allow for much finer heating precision, ensuring the sauce reaches the desired temperature consistently and uniformly. Moreover, such real-time adjustments help avoid energy wastage by preventing overprocessing; the system would use only the necessary amount of electrical energy to achieve target conditions, thereby improving overall energy efficiency. In essence, deploying *OhmNet* in this manner enables smarter CFOH operations that can maintain product quality while minimising energy usage and processing time.

Author Contributions: Tasmiyah Javed: Methodology, Software, Validation, Formal analysis, Writing – original draft. **Muhammad Akmal:** Supervision, Writing - Review & Editing. **Timofei Breikin:** Supervision, Writing - Review & Editing. **Caroline Millman:** Supervision, Resources, Funding acquisition, Writing - Review & Editing. **Hongwei Zhang:** Supervision, Conceptualisation, Funding acquisition, Formal analysis, Resources, Writing - Review & Editing.

Data Availability Statement: The data are available upon request to the corresponding authors.

Acknowledgments: The authors would like to thank Premier Foods, Ohm-E Technology (UK), and the AFIC team at Sheffield Hallam University for the guidance and resources received when performing experiments.

Funding sources: This research was supported by Innovate UK under Grant number 10074002.

Conflicts of Interest: The authors declare no conflicts of interest.

References

Abdalla, A., Cen, H., Wan, L., Rashid, R., Weng, H., Zhou, W., & He, Y. (2019). Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure. *Computers and Electronics in Agriculture, 16*710.1016/j.compag.2019.105091

Abinaya, S., Panghal, A., Kumar, S., Kumari, A., Kumar, N., & Chhikara, N. (2024). Artificial Intelligence (AI) in Food Processing. *Nonthermal Food Engineering Operations,* , 1–54.

Afrand, M., Nazari Najafabadi, K., Sina, N., Safaei, M. R., Kherbeet, A. S., Wongwises, S., & Dahari, M. (2016). Prediction of dynamic viscosity of a hybrid nano-lubricant by an optimal artificial neural network. *International Communications in Heat and Mass Transfer, 76*, 209. 10.1016/j.icheatmasstransfer.2016.05.023

Barradas Filho, A. O., Barros, A. K. D., Labidi, S., Viegas, I. M. A., Marques, D. B., Romariz, A. R., de Sousa, R. M., Marques, A. L. B., & Marques, E. P. (2015). Application of artificial neural networks to predict viscosity, iodine value and induction period of biodiesel focused on the study of oxidative stability. *Fuel, 145*, 127–135.

Batista, L. F., Marques, C. S., dos Santos Pires, A. C., Minim, L. A., Soares, N. d. F. F., & Vidigal, M. C. T. R. (2021). Artificial neural networks modeling of non-fat yogurt texture properties: effect of process conditions and food composition. *Food and Bioproducts Processing, 126*, 164–174.

Bhagya Raj, G., & Dash, K. K. (2022). Comprehensive study on applications of artificial neural network in food process modeling. *Critical Reviews in Food Science and Nutrition, 62*(10), 2756–2783.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford university press.

Chen, H., Fu, Q., Liao, Q., Xiao, C., Huang, Y., Xia, A., Zhu, X., & Kang, Z. (2019). Rheokinetics of microalgae slurry during hydrothermal pretreatment processes. *Bioresource Technology, 289*, 121650.

Chen, H., Fu, Q., Liao, Q., Zhu, X., & Shah, A. (2021). Applying artificial neural network to predict the viscosity of microalgae slurry in hydrothermal hydrolysis process. *Energy and AI, 4*, 100053.

Dahl, J. F., Schlangen, M., van der Goot, A. J., & Corredig, M. (2025). Predicting rheological parameters of food biopolymer mixtures using machine learning. *Food Hydrocolloids, 160*, 110786.

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

Goyal, S. (2014). Artificial Neural Networks in fruits: A comprehensive review. *International Journal of Image, Graphics and Signal Processing, 6*(5), 53.

Guo, Q., He, Z., & Wang, Z. (2023). Prediction of monthly average and extreme atmospheric temperatures in Zhengzhou based on artificial neural network and deep learning models. *Frontiers in Forests and Global Change, 6*10.3389/ffgc.2023.1249300

Guo, Y., Shi, H., Kumar, A., Grauman, K., Rosing, T., & Feris, R.SpotTune: Transfer Learning through Adaptive Fine-tuning.

Heidari, E., Sobati, M. A., & Movahedirad, S. (2016). Accurate prediction of nanofluid viscosity using a multilayer perceptron artificial neural network (MLP-ANN). *Chemometrics and Intelligent Laboratory Systems, 155*, 73. 10.1016/j.chemolab.2016.03.031

Hu, Z., Zhao, Z., Yi, X., Yao, T., Hong, L., Sun, Y., & Chi, E. H.Improving Multi-Task Generalization via Regularizing Spurious Correlation.

Huang, G. (2003). Learning capability and storage capacity of two-hidden-layer feedforward networks. *IEEE Transactions on Neural Networks, 14*(2), 274–281.

Icier, F., & Bozkurt, H. (2009). Ohmic Heating of Liquid Whole Egg: Rheological Behaviour and Fluid Dynamics. *Food and Bioprocess Technology, 4*(7), 1253. 10.1007/s11947-009-0229-4

Karaman, S., Yilmaz, M. T., Kayacier, A., Dogan, M., & Yetim, H. (2014). Steady shear rheological characteristics of model system meat emulsions: Power law and exponential type models to describe effect of corn oil concentration. *Journal of Food Science and Technology*, 10.1007/s13197-014-1434-3

Lee, M., & Lee, I. (2023). Transfer Learning with Deep Neural Network toward the Prediction of Wake Flow Characteristics of Containerships. *Journal of Marine Science and Engineering, 11*(10)10.3390/jmse11101898

Lie-Piang, A., Hageman, J., Vreenegoor, I., van der Kolk, K., de Leeuw, S., van der Padt, A., & Boom, R. (2023). Quantifying techno-functional properties of ingredients from multiple crops using machine learning. *Current Research in Food Science, 7*, 100601.

Liu, Z., Wang, S., Zhang, Y., Feng, Y., Liu, J., & Zhu, H. (2023). Artificial intelligence in food safety: A decade review and bibliometric analysis. *Foods, 12*(6), 1242.

Manzoor, A., Jan, B., Shams, R., Ul, Q., & Rizvi, E. H.*Ohmic heating technology for food processing: a review of recent developments*

Mirarab Razi, M., Kelessidis, V. C., Maglione, R., Ghiass, M., & Ghayyem, M. A. (2014). Experimental Study and Numerical Modeling of Rheological and Flow Behavior of Xanthan Gum Solutions Using Artificial Neural Network. *Journal of Dispersion Science and Technology, 35*(12), 1793. 10.1080/01932691.2013.809505

Naidu, K., Ali, M. S., Abu Bakar, A. H., Tan, C. K., Arof, H., & Mokhlis, H. (2020). Optimized artificial neural network to improve the accuracy of estimated fault impedances and distances for underground distribution system. *Plos One, 15*(1)10.1371/journal.pone.0227494

Oroian, M. (2013). Measurement, prediction and correlation of density, viscosity, surface tension and ultrasonic velocity of different honey types at different temperatures. *Journal of Food Engineering, 119*(1), 167. 10.1016/j.jfoodeng.2013.05.029

Pan, S. (2010). Q.: A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering, 22*(10), 1345–1359.

Peleg, M. (2017). Temperature–viscosity models reassessed. *Critical Reviews in Food Science and Nutrition, 58*(15), 2663. 10.1080/10408398.2017.1325836

Perera, Y. S., Li, J., & Abeykoon, C. (2025). Machine learning enhanced grey box soft sensor for melt viscosity prediction in polymer extrusion processes. *Scientific Reports, 15*(1)10.1038/s41598-025-85619-6

Pourramezan, M., Rohani, A., & Abbaspour-Fard, M. H. (2024). Machine Learning-Based Predictions of Metal and Non-Metal Elements in Engine Oil Using Electrical Properties. *Lubricants, 12*(12)10.3390/lubricants12120411

Rai, P., Majumdar, G. C., Dasgupta, S., & De, S. (2005). Prediction of the viscosity of clarified fruit juice using artificial neural network: a combined effect of concentration and temperature. *Journal of Food Engineering, 68*(4), 527. 10.1016/j.jfoodeng.2004.07.003

Raju, V. G., Lakshmi, K. P., Jain, V. M., Kalidindi, A., & Padma, V. (2020). Study the influence of normalization/transformation process on the accuracy of supervised classification. Paper presented at the *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT),* 729–735.

Rocha, R. S., Calvalcanti, R. N., Silva, R., Guimarães, J. T., Balthazar, C. F., Pimentel, T. C., Esmerino, E. A., Freitas, M. Q., Granato, D., & Costa, R. G. (2020). Consumer acceptance and sensory drivers of liking of Minas Frescal Minas cheese manufactured using milk subjected to ohmic heating: Performance of machine learning methods. *LWT, 126,* 109342.

Said Toker, Ö, Tahsin Yilmaz, M., Karaman, S., & Kayacier, A.Adaptive Neuro-fuzzy Inference System and Artificial Neural Network Estimation of Apparent Viscosity of Ice-cream Mixes Stabilized with Different Concentrations of Xanthan Gum.10.3933/ApplRheol-22-63918

Siejak, P., Przybył, K., Masewicz, Ł, Walkowiak, K., Rezler, R., & Baranowska, H. M. (2024). The Prediction of Pectin Viscosity Using Machine Learning Based on Physical Characteristics—Case Study: Aglupectin HS-MR. *Sustainability, 16*(14), 5877.

Silva, R., Rocha, R. S., Guimarães, J. T., Balthazar, C. F., Ramos, G. L. P., Scudino, H., Pimentel, T. C., Azevedo, E. M., Silva, M. C., & Cavalcanti, R. N. (2020). Ohmic heating technology in dulce de leche: physical and thermal profile, microstructure, and modeling of crystal size growth. *Food and Bioproducts Processing, 124*, 278–286.

Silva, R., Rocha, R. S., Ramos, G. L. P., Xavier-Santos, D., Pimentel, T. C., Lorenzo, J. M., Campelo, P. H., Silva, M. C., Esmerino, E. A., & Freitas, M. Q. (2022). What are the challenges for ohmic heating in the food industry? Insights of a bibliometric analysis. *Food Research International, 157,* 111272.

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. Paper presented at the *Artificial Neural Networks and Machine Learning*–

ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27, 270–279.

Yang, Y., Chen, X., Wang, X., Pan, L., & Lan, W. (2025). Rheological modelling of apple puree based on machine learning combined Monte Carlo simulation: Insight into the fundamental light-particle structure interaction processes. *Food Chemistry*, *464*, 141611.

Declaration of Interest Statement

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.

Highlights

- The proposed One-Hot Encoded ANN model (*OhmNet*) predicts sauce viscosity accurately against changing temperature during Continuous Flow Ohmic Heating (CFOH).
- Transfer learning improves ANN performance for viscosity modelling.
- Three neural network strategies were evaluated for CFOH process control.
- ANN model enables energy-efficient, real-time Ohmic Heating control.