VFAST Transactions on Mathematics
http://vfast.org/journals/index.php/VTM@ 2025 ISSN(e): 2309-0022, ISSN(p): 2411-6343
Volume 13, Number 2, July-December, 2025 *pp: 01-19*

Predictive Modeling of Parameter Variations in Conveyor Belt Dryers using Machine Learning for Improved Drying Performance

Awais Khan[1*], **Tauqeer Ahmad**[1], **Naveed Ahmad Khan**[2], **Muhammad Sulaiman**[1]

[1]Department of Mathematics, Abdul Wali Khan University Mardan, Pakistan; [2]School of Information Technology and Systems, University of Canberra, Canberra, ACT, Australia

Keywords: *Design Engineering, Tangential flow conveyor belt dryers, Food drying, Food preservation, Machine learning techniques, Artificial neural networks, Optimization algorithm, RK-4 Technique.*
Subject Classification: *62D05*

Journal Info:
Submitted:
March 28, 2025
Accepted:
December 16, 2025
Published:
December 31, 2025

Abstract This study takes an in-depth look at the mathematical model governing the operation of a tangential flow conveyor dryer operating in a co-current configuration. The Conveyor Belt Dryer (CBD) is represented as an Ordinary Differential Equation (ODE), and our research focuses on studying the influence of parameter variation and symmetry on the Rate of Exchange of Moisture Content (RMC). To achieve this, we employ a system studying framework based on artificial Neural Networks (NN) and the Levenberg-Marquardt Training (LMT) algorithm, enabling the examination of symmetry within surrogate solutions. The RK-four method is applied to generate target data points for supervised learning in the NN-LMT structure. Furthermore, we investigate various scenarios of the mathematical model related to the rate of change of moisture content. Detailed graphical representations, including histograms, absolute error plots, curve fitting graphs, and regression graphs, are employed to facilitate comprehensive explanations. Additionally, a comparative analysis between the numerical solutions obtained through the machine learning technique is provided, followed by graphical and statistical representations of the determined errors.

***Correspondence Author Email Address:**
awais.khan@awkum.edu.pk
DOI: 10.21015/vtm.v13i2.2105

1 Introduction

Conveyor dryers are the most popular device used in industrial food dehydration. In particular, it will remove water vapor particles from exotic fruits to keep them dry at room temperature [1–4]. The main purpose is to evaporate the water in solid food products and expose them to warm, dry air. lower their water content through the introduction of air. However, the drying process exposes the product to rela-

tively high temperatures, potentially impacting the quality of the final product [5, 6]. In addition, there is a risk of protein denaturation and vitamin deficiency. It is crucial to ensure that the wet-bulb temperature during drying remains appropriately lower than the dry-bulb temperature, as this is a critical requirement for the product. The possibility of maintaining a moisture content X_C above the threshold level persists until the tip of the dryer, where the liquid level is X_F. For the purposes of this study, it is assumed that X_F is greater than X_C, which is an essential premise.

The drying process involves a complex heat and mass exchange between the air and the product, which can be modeled using a differential equation [7, 8]. Solutions to these equations can be obtained either numerically or through closed-form approaches [9–11]. Over several decades, mathematical models have yielded significant insights, subsequently facilitating the development of various drying models for different food products, including generic foods [12], coconut [13], apples [14], bananas [15], pears [16], generic fruits [17], mango [18], sultanas [19], mandarin [20], kiwi [21], apricot [22], and papaya [23]. Perforated belt dryers, among the various types of food dryers, are particularly popular due to their compact size [24, 25]. Nevertheless, they pose challenges, such as elevated air temperatures, even in the final zone where the food is more susceptible to heat, and have thus been the focus of extensive research [26, 27]. In a preceding investigation conducted by Friso et al. [28], an examination of the symmetric tangential flow within a conveyor-belt dryer was undertaken. Figure (1) visually represents the setup employed in the study. By implementing a lower external temperature within the final drying chamber, this particular dryer configuration facilitates enhanced preservation of the product quality, minimizing the adverse effects associated with protein and polyphenol denaturation. However, it is worth noting that the continuous operational nature of this dryer necessitates a trade-off, whereby the reduction in temperature towards the dryer's concluding stages leads to an elongation of the overall drying period.Researchers

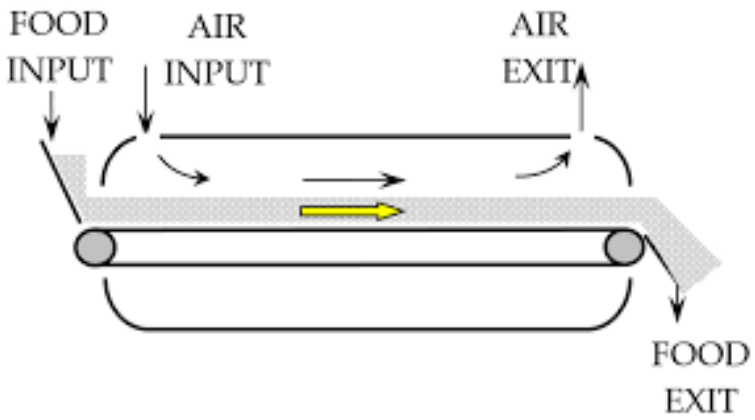

Figure 1. *Dryer for conveyor belts that uses co-current symmetric tangential flow.*

have developed a series of mathematical models [29–33], which have undergone experimental validation and have been solved using either analytical or numerical techniques. However, this particular study focuses on employ- ing machine learning techniques to solve the mathematical model (8) in order to obtain a solution for the conveyor belt dryer. This choice stems from the increasing prevalence of artificial neural networks (ANNs) in addressing a wide range of challenges across various fields such as wire coating dynamics, industries, petroleum engineering, thermal engineering, civil

engineering, and biological modeling. Inspired via the successful application of stochastic techniques in these domain names, the author's goal is to discover, check out, and behavior studies on the conveyor belt dryer version the use of ANNs. To this quit, a gentle computing approach has been advanced to optimize the mathematical model for the conveyor belt dryer, as mentioned in Khan et al. [34–37].

Several parameters play a critical position in optimizing the overall performance of the conveyor belt dryer. For example, an increase in the length of the belt leads to a higher potential for drying rate. Similarly, an increase in parameter z is expected to enhance the output [38–42]. Moreover, elevating the "Air enter temperature" is likely to result in a higher rate of dry products.These three scenarios are examined in this study, where the moisture content rate is investigated with respect to the z-axis, the rate of moisture content is analyzed in relation to the velocity of the belt (v_{belt}), and the rate of moisture content is explored in relation to the air input temperature (T_{Al}). Each of these cases is extensively investigated using a soft computing architecture [38–42]. The following points serve to emphasize key features of this study.

- The investigation encompasses three distinct parameters, where a mathematical model pertaining to the moisture content rate is examined.
- To address the aforementioned scenarios, a backpropagated Levenberg-Marquardt training (LMT) method is devised, utilizing artificial neural networks. This method is employed to train the hidden neurons and ensure the accuracy and coherence of the reference dataset of solutions.
- Extensive graphical analysis has been conducted, delving into regression, curve fitting, mean square error, and error histograms.

2 Analyzing Mathematical Problem of Drying Progress Rate Along the Belt Dryer

This section presents a mathematical derivation of the rate of change of moisture content. An axial flow conveyor dryer layout is also included in the design, along with a temperature diagram depicting air and material temperatures T_A and T_P, respectively. At the moment, the wet bulb temperature T_{WB} was assumed to be the same as T_P. Figure (2) illustrates the minuscule heat rate dq, which transfers from the atmosphere to the object via the insignificantly small area dA. The rate of heat transfer d_q from the air to the object, is extremely small in comparison to dA, the distance over which the air must travel. Consequently, d_q may be written as:

$$d_q = \alpha.dA.(T_A - T_{WB}), \tag{1}$$

where α is the coefficient of heat transfer, dA stands for the infinitesimal area, and T_A is the temperature of the atmosphere somewhere at the point of interaction with the area dA. The air's wet bulb temperature T_{WB} is considered to be the same as the commodity temperature T_{WB}. Given the possibility of an adiabatic drying process, which would occur due to the lack of thermal conduction through the walls, the fundamental shifts in temperature that occur in the air (its dry component), dT_A, may be expressed in terms of the elementary heat transfer rate, d_q. As a result, its temperature is reduced by an infinitesimal amount, or dT_A.

$$dq = -G_{Al}.c_A.dT_A. \tag{2}$$

The infinitesimal area is represented by dA, and dT_A is the infinitesimal change in air temperature whenever it makes an interface with the region dA. The mass flow rate of hot air flowing through the dryer

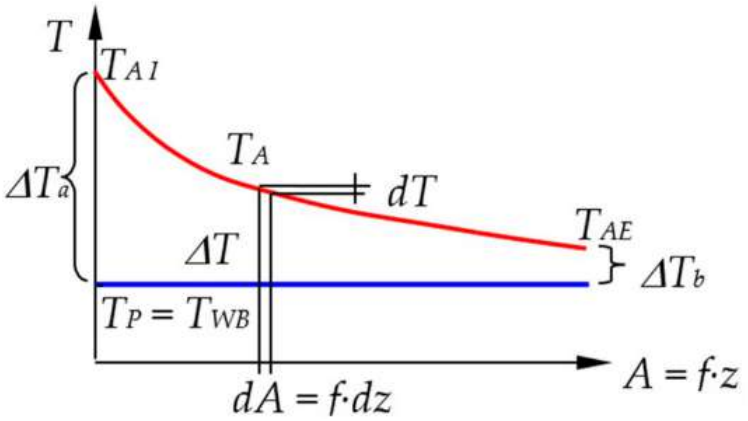

Figure 2. *Diagram showing the ultimate moisture content and air and product temperatures for the tangential-flow conveyor-belt dryer, $X_F > X_C$ is shown.*

simultaneously with the mass flow rate of dry air is denoted by the symbol G_{AI}, and the dry air specific heat is denoted by the letter c_A. In a conveyor drying system, drying takes place in the belt's direction. A new z variable has been added, and it is related to area A by a relationship found in (18) of an earlier study [**?**]: A = f.z; as a result, the figure displays dA = f.dz (3), where dz denotes an infinitesimal length (2).

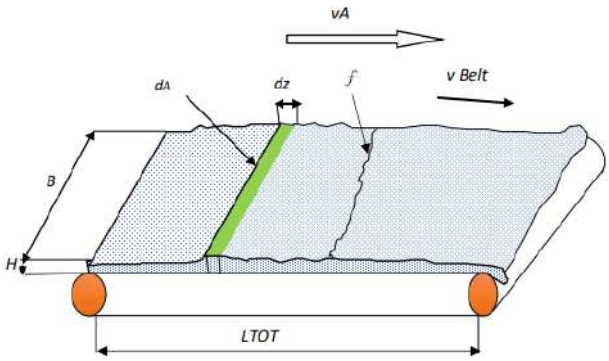

Figure 3. *Conveyor-belt dryer with tangential co-current flow: transverse dimension f; dry air velocity v_A; belt velocity v_{Belt}; total belt length L_{TOT}; product bed height (H); product bed width (B); minimal amount of material exposed to air (dA); minimal length (dz).*

In order to obtain the first-order differential equation (ODE), equate the quantities dq:

$$\frac{dT_A}{dz} = -\frac{\alpha.f}{G_{AI}.c_A}.(T_A - TWB), \tag{3}$$

figure (2) depicts the product's temperature initially T_{PI}, which is already the same as the temperature of the wet bulb T_{WB}. Quantity of heat input into the dryer's enthalpy balance to raise the product's dry mass temperature from T_{PI} to T_{WB}, where the water content is around 3%, represents a fraction of total heat energy, less than 1%, delivered by the hot air [25]. But when the mathematical model was being developed, this latter thermal energy was taken into account [28]. Consequently, the $T_{PI} = T_{WB}$ supposition was valid. It is simple to evaluate the indefinite integral of equation (3) using the variable separation method, which

yields: $-\frac{\alpha.f.z}{G_{AI}.c_A}$ = ln(T_A - T_{WB}) + C, where C is the integration constant derived from the dryer's beginning condition for z = 0, T_A = T_{AI}; where the dryer's air temperature is referred to as T_{AI}. Consequently, C = - ln(T_A - T_{WB}) and the solution is:

$$(T_A - T_{WB}) = (T_{AI} - T_{WB}).e^{-\frac{\alpha.f.z}{G_{AI}.c_A}}, \tag{4}$$

the appearance of an exponential function along the conveyor belt that has a negative exponent and contains the variable z indicates that the temperature of the wet bulb approaches the air temperature T_A asymptotically. To establish it in another way, achieving a length z = ∞ is the sole way to bring air and a product into thermo-hygrometric equilibrium, where the amount of moisture is always thought to be above the critical level. Since the conveyor-belt dryer's tangential flow procedure occurs along the variable z, instead of providing the instantaneous drying rate, Providing the drying rate of the dryer at point z of the dryer is preferred. As a result, at a position z from the dryer's start, according to the product's infinitesimal mass dm, whose drying portion will be dm_D, see Figure (3), is recognized and subjected to the airflow in an ephemeral region called dA. As a result, at position z relative to the dryer's starting point, the drying rate R [28] is changed to R= $\frac{dG_{EV}}{dm_D}$. Due to the fact that water evaporates at an infinitesimal flow rate dG_{EV} moistens the infinitesimal dry mass dm_D. The tiny area, or dA, via which the dry air transfers its heat to the product is described in earlier equality: dq = α . dA . (T_A - T_{WB}) = α . f . dz . (T_A - T_{WB}). This heat transfer rate dq causes an infinitesimal mass flow rate of water to evaporate, with the result that: $dG_{EV} = \frac{dq}{r} = \frac{f.dz.\alpha}{r}$. (T_A - T_{WB}). Where the symbol r stands for the amount of thermal energy that must be expended in order to generate one kilogram of super-heated water vapor at the temperature of the air T_A [28]. Once this relationship is plugged into a previously established drying rate specification at the generalized point z, the desired result may be attained, so:

$$R = \frac{dG_{EV}}{dm_D} = \frac{f.dz.\alpha}{r.dm_D}.(T_A - T_{WB}). \tag{5}$$

The term "infinite dry mass" refers to the subtraction of the product's entire infinitesimal mass denoted by (dm) and the infinite mass of the water that it comprises denoted by (dmW), (dmW) . (dmD): dmD = dm - dmW. At the starting spot of the dryer, it is understood what the minuscule masses of the product dm_I and water dm_{WI} are. As a result, dm_D = dm_I- dm_{WI}. The product's moisture content is known before it enters the dryer: Along with its starting volume dV_I and initial bulk density, X_I = $\frac{dm_{WI}}{dm_D}$: ρ_{BulkI} = $\frac{dm_I}{dV_I}$. Rearrange X_I and ρ_{bulkI}'s equation, then insert the previous equation to create a new equation from the prior one that provides: dm_D = ρ_{bulkI}.dV_I - dm_D.X_I. If dm_D is highlighted, we get: dm_D = $\frac{\rho_{bulkI}.dV_I}{1+X_I}$. At the beginning of the conveyor-belt dV_I is the representative symbol for the infinitesimal volume of the bulk product Figure (3), which reveals that an infinite area dA is exposed to drying air, demonstrates that an infinite bulk product with dimensions of height H_I, width B_I, and length dz has a volume dV_I = $B_I H_I$dz. So the infinitesimal dry mass dm_D is dm_D = $\frac{\rho_{bulkI}.B_I.HI.dz}{1+X_I}$. Thus, we eventually arrive at the equation:

$$R = \frac{dG_{EV}}{dm_D} = \frac{\alpha.f}{r.\rho_{BulkI.B_I.HI}}.(1 + X_I)(T_{AI} - TWB), \tag{6}$$

the spot, z units away from the dryer's intake Figure (2) where the temperature of the air is T_A corresponds to the drying rate R of equation (6). Hence, taking starting from the general definition [28]: R=$\frac{dX}{dt}$ where the moisture content on a dry basis is X, and can be written as:

$$R = -\frac{dX}{dt} = -\frac{dX}{dz}.\frac{dz}{dt} = -\frac{dX}{dz}.v_{Belt}, \tag{7}$$

taking into account equations (4), (7), and (6), the ratio among an infinitesimal length (represented by dz) and an infinitesimal time period of the product speed, v_{Belt}, can be as follows:

$$\frac{dX}{dz} = -\frac{\alpha.f}{v_{Belt}.r.\rho_{Bulkl}.B_l.H_l}.(1+X_l).(T_{Al}-T_{WB}).e^{-\frac{\alpha.f.z}{GAl.c_A}}. \tag{8}$$

The required mathematical model is presented here. From here we are establishing three cases:

- rate of moisture content change relative to the z-axis.
- rate of change in moisture content in relation to belt velocity.
- rate of moisture content change with regard to air input temperature.

Moreover the derived model (8) has been validated by [28, 29]. In the following table (3) different values are shown along the units which are very helpful for this study.

3 Comparative Analysis of Numerical Methods and Machine Learning Algorithms for Prediction

In this section, we are going to analyze two types of methods, the first is the numerical method, in which we need the mathematical model along the initial condition. Applying the mathematical framework for a particular issue and its initial state yields targeted data. The targeted data gets nearer to the real solution as the number of rounds increases. The second one is machine-learning technique, which is limited to a particular set of data. We do not need a mathematical framework of this type for the given problem. This kind consumes weights and targeted data to create a surrogate model. The number of weights in this technique depends on the number of neurons, as the number of neurons increases the number of weights increases because each neuron contains three weights. In this study, the machine learning technique is preferred because it is more efficient than the numerical method. Moreover, it is mostly applicable to real life as compared to other methods. Here we have compared the numerical and NN-LMT solution of all the three cases in table (1) and also given there difference in table (2). The error ranges in case 1 lies between 10^{-5} to 10^{-8}, for case second the error lies in 10^{-5} to 10^{-7} and for case 3 is 10^{-5} to 10^{-6}.

Table 1. *Comparison of the numerical solution and the approximate solutions provided by NNs-LMT.*

z-axis			v_{Belt}			Air input temperature		
Input	Error	Out put	Input	Target(RK-4)	Out put	Input	Target(RK-4)	Out put
0	-2.239450319	-2.239479778	0.005	-4.313935136	-4.313947153	99.5	-4.056188938	-4.056236199
0.001	-2.660664582	-2.660664656	0.006	-4.330365689	-4.330363143	100	-10.90943819	-10.90944677
0.002	-2.999642726	-2.999637387	0.007	-4.344257607	-4.344256865	102	-38.87735419	-38.87736755
0.004	-3.49197797	-3.491973967	0.008	-4.356291285	-4.356294905	105	-82.49398508	-82.49397294
0.008	-4.017344098	-4.017348131	0.01	-4.376400711	-4.37640367	109	-143.7570392	-143.7570461
0.01	-4.151087692	-4.15109073	0.013	-4.400044641	-4.400041716	111	-175.7203717	-175.7203831
0.0121	-4.241127014	-4.241131298	0.015	-4.4129407	-4.412938058	113	-208.5715746	-208.5715782
0.014	-4.293804183	-4.293810081	0.016	-4.418756829	-4.418758477	115	-242.3106479	-242.3106414
0.0147	-4.3083481	-4.308349151	0.019	-4.434243745	-4.434246561	117	-276.9375915	-276.9375787
0.015	-4.313935049	-4.313914553	0.02	-4.438866232	-4.438858284	118	-294.5840147	-294.5840183

sectionMethod for the Solution This study's methodology made utilize of artificial neural networks. When singularities do exist in equations, artificial neural networks avoid them and offer appropriate solutions. A neural network may be used to identify problem values outside of the interval, providing a general answer to the problem. The fundamental non-linear modelling capabilities, probable extrapolations, ill-conditioning, forecasts, interpolations, forecasts, usability, and robustness to noise and inadequate data

Table 2. *Difference between the numerical solution and the approximate solutions provided by NNs-LMT.*

z-axis		v_{Belt}		Air input temperature	
Input	Error	Input	Error	Input	Error
0	2.94591×10^{-05}	*0.005*	1.20×10^{-05}	*99.5*	4.72614×10^{-05}
0.001	7.35247×10^{-08}	*0.006*	2.55×10^{-06}	*100*	8.57809×10^{-06}
0.002	5.3395×10^{-06}	*0.007*	7.42×10^{-07}	*102*	1.33566×10^{-05}
0.004	4.00349×10^{-06}	*0.008*	3.62×10^{-06}	*105*	1.21429×10^{-05}
0.008	4.03306×10^{-06}	*0.01*	2.96×10^{-06}	*109*	6.88939×10^{-06}
0.01	3.03777×10^{-06}	*0.013*	2.93×10^{-06}	*111*	1.14198×10^{-05}
0.0121	4.28429×10^{-06}	*0.015*	2.64×10^{-06}	*113*	3.59512×10^{-06}
0.014	5.89815×10^{-06}	*0.016*	1.65×10^{-06}	*115*	6.53419×10^{-06}
0.0147	1.05112×10^{-06}	*0.019*	2.82×10^{-06}	*117*	1.2848×10^{-05}
0.015	2.0496×10^{-05}	*0.02*	7.95×10^{-06}	118	3.59922×10^{-06}

are used to assess the significance of the neural network technique. It is also quite robust to faults. ANN may guide the system by learning from events and applying that information to similar circumstances. Multiple tasks can be completed at once using ANN. Neural networks (NN) are among the algorithms used in deep learning. Machine learning networks that are modelled after the nervous system of humans are called neural networks. An input layer, a hidden layer, and an output layer are the parts of a neural network. The connected and interconnected hidden neurons exchange information. The algorithm (ANN) for nonlinear training is employed. It is significant due to its rapid convergence and accurate results. This process yields excellent outcomes. The Levenberg-Marquardt neural network is made up of the structural layer, the number of hidden neurons, the framework, and the arbitrary selection of output and input data for training, validation, and testing, samples. To solve the problem, one must identify each of these components. The RK-4 method was then used to solve the issue using the points produced by the "NDSolve" command in Mathematica. The step size was then set. The Matlab command window is then used to suggest the NN-LMT method and a collection of data points from the neural toolbox is used to set input and output in NN-LMT. For this study, we chose the following data set.

- Training data is 70%.
- Validation data is 15%.
- 15% data represents testing.

As demonstrated in Figure (5), the number of hidden neurons can be selected at will, and in our situation, we choose 10, which yields satisfactory results. Figure (5) demonstrates how an artificial neural network (ANN) functions and its basic operating mechanism. It is clear that the NN is composed of inputs and the weights assigned to them, then a hidden neuron, an activation function, and lastly outputs. The model we developed to solve our neural network challenge is depicted in Figure (4) utilizing the nftool command in the Matlab window. This issue has 10 neurons. You can vary the number of neurons, however, in this problem, 10 neurons produce the best results. That's why our problem has 10 neurons. Figure (4) also clearly indicates the inputs and outputs. The problem takes the form of an ordinary differential equation. The ordinary differential equations must then be solved in Mathematica using the RK-4 technique to acquire the data set. The data from the MS Excel spreadsheet is then forwarded to Matlab for usage with neural networks after being transferred to an Excel sheet. There are 10 neurons that were chosen. 70% of the values are chosen for training, 15% for testing, and 15% for validation. Operate the command numerous times. The operation is finished if the results are the most precise and best; if not, the network must be retrained and operated up to the minimal mean square error value (MSE). The stopping iteration will be determined by which iteration of the network produces the best results.

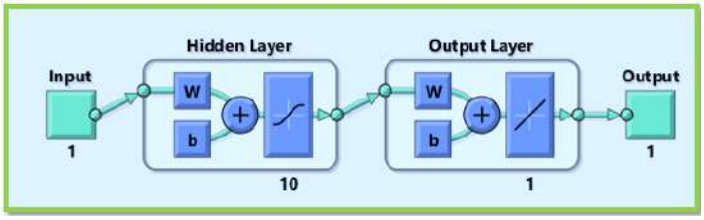

Figure 4. *A demonstration of backpropagated ANN*

4 Design Methodology

Before we get into an optimization strategy for the neuron mechanism of understanding in FFNN design, we will first run through the basic design and functionality of an FFNN.

4.1 Artificial Neural Network Using Feed Forward Neural Network

It was in the seventeenth century when mathematicians and scientists like Isaac Newton, Leibniz, Fermat, and Kepler first developed the fundamental concepts of optimization. These concepts have been reexamined. In the 1950s, utilized in digital computers with "high performance". Due to the successful outcomes of these endeavors, novel algorithms have been created and optimization theory has gained respect as a significant area of mathematics. Artificial Neural Networks (ANNs) are a kind of data processing that processes information similarly to biological nervous systems (like the brain). The primary element of this technology is the special configuration of information processing units or neurons. A network is made up of interconnected parts with I/O (input/output) capabilities and the ability to conduct local computations. The I/O potential, links to other components, and external inputs of a unit all have an impact on its output. Establishing, describing, and developing mathematical illustrations of neural networks are key topics in recent ANN investigations. Feedforward neural networks (5), which include just one or many layers of interneurons linking each neuron of the input layer to the output layer, are probably the most prominent and widely utilized kind of ANN in numerous real-world applications. "Multi-layer sensing" is simply one of the many names for it. Throughout training, weights for neural networks are modified to provide accurate input/output links. The most common foundation for adaptive learning algorithms is the width-off back technique. Composite applications of functions with mathematical expressions are multilayer feed networks. Assume the following scenario: we are given L samples, each of which has a unique set of input and output signals. It is hard to train a neural network utilizing data that is currently accessible because it is necessary for minimizing the cost function which is typically a least squares function. d Second-order optimization techniques, such as conjugate level, LM, and quasi-Newton methods, can be used to create the network. Recall that the first-order approach modifies by having the sharpest fall direction and adding momentum. They all attempt to avoid constructing the Hessian Matrix because it is a difficult implementation to do so. It is clear that the LM approach yields the highest precision, so it is worthwhile to investigate its operation and its practical applications. The LM approach has been applied perfectly in numerous instances in the recent past.

4.2 Backpropagation Procedure

The creation of backpropagation (BP) involves expanding the Widrow-Hoff learning algorithm to incorporate non-linear differentiable transfer functions and multi-layer networks. For a network to effectively categorize input carriers, correlate inputs with certain output vectors, or essentially depict a function, it must first be trained to utilize source as well as target vectors. Systems with bias, sigmoid functions, and linear output layers assess any function with a restricted amount of discontinuities. This gives a succinct synopsis of BP. Keep in mind that X is the outcome layer, 0 is the input layer, and x is the layer index. b_i^x signifies the x-th layer signal, whereas i = 0, 1,..., k[x]. Utilizing zero indices as the bias signal is simple because $b_0^x = 1$. The weights $w_{i,j}^x$, where i = 1,..., k[x], j = 0, 1,..., k[x-1], and the function of activation d^x in the x-th layer, which is composed of k[x] cells, are used to identify the i-th cell. For each $w_{i,0}^x$, the weight represents the relevant bias. Thus, the total amount of inputs is k[0], and the quantity of outcomes is k[X]. BP requires several processes. Assume the l-th trial. Then, for i = 1, 2,..., k[0], we get $b_i^0 = P_{l,i}$ source. The data array $P_{l,i}$ is used for the set of inputs, while the array $T_{l,i}$ is used for the result set for i = 1, 2,..., k[X].

- Compute the forward step of the network signal..

$$h_i^x = \sum_{j=0}^{k[x-1]} b_i^{x-1} \times w_{i,j}^x, \tag{9}$$

$$b_i^x = (h_i^x) d^x, \tag{10}$$

for x = 1,2,,.....,X and i = 1,2,....,k[X]. Keep in mind, as b_0^x=1 always. Here, we stress that the l-th sample's cost function is;

$$F_l(M) = \frac{1}{2} \sum_{i=1}^{k[X]} e_{li}^2 = \frac{1}{2} \sum_{i=1}^{k[X]} (b_i^X - J_{l,i})^2. \tag{11}$$

- Determine δ (the backward step) beginning with the resultant layer that is X-th.

$$\delta_i^X = (h_i^X)(d^X)' \times (b_i^X - J_{l,i}), \tag{12}$$

and for x = X, X-1,.....,2, We determine

$$\delta_i^{x-1} = (d^{x-1})'(h_i^{x-1}) \sum_{j=1}^{k[x]} w_{i,j}^x * \delta_i^x, \tag{13}$$

for i = 0, 1,....., k[x-1].

- Determine $F_l(M)$'s derivative in relation to M,

$$g_{l,i,j}^x = \delta_i^x \times a_j^{x-1}, \tag{14}$$

for x = 1,2,.....,X, i = 1,2,....,k[x], j = 0,1,....., k[x-1]. These derivatives correspond precisely to the gradient's terms

$$g_i(M) = \nabla_M F_l(M). \tag{15}$$

Once we've obtained the result of $g_i(M)$, we may use the widely recognized incremental learning technique to update M

$$(w_{i,j}^x)^{new} = (w_{i,j}^x)^{old} - \alpha g_{l,i,j}^x, \tag{16}$$

where the learning rate parameter is α.

Throughout the incremental learning process, the aforementioned phases are repeated from the first sample to the last sample, and each stage represents a distinct epoch. It takes numerous epochs to create a good network. The nonlinear least squares issue associated with neural network training can be resolved using a class of nonlinear least squares techniques. One of them uses batch learning instead of incremental learning, therefore the cost function is the sum of the sample cost functions.

$$d = \sum_{l=1}^{L} F_l(M) = \frac{1}{2} \sum_{l=1}^{L} \sum_{i=1}^{k[X]} e_{li}^2. \tag{17}$$

So

$$g(M) = \nabla_M d(M) = \sum_{l=1}^{L} g_l(M), \tag{18}$$

$$g_{i,j}^x = \sum_{l=1}^{L} g_{l,i,j}^x. \tag{19}$$

4.3 LM Algorithm

The iterative LM strategy can be used to find a multidimensional operation that can be written as the total squared of non-linear exact-valued functions. It is now the established approach for resolving non-linear least-squares problems in numerous sectors. LM may be contrasted with both the Gauss-Newton method and the steepest descent. When the current result differs significantly from the actual result, the strategy operates similarly to the steepest descent strategy; it will coincide gradually and slowly. A Gauss-Newton approach is one in which the current solution is significantly near to the ideal one. After that, an overview of the information-based LM approach is given. However, bear in mind that a thorough examination of the LM method is outside the purview of this work, and more in-depth analyses are advised for those who are interested. Because it works, this method is generally accepted as the best for training artificial neural networks. Similar to quasi-Newton techniques,The LM methodology was created to get close to the second-order training rates despite requiring the Hessian matrix to be calculated. When the efficiency function's shape is formed by a combination of squares, that is;

$$f(M) = \frac{e^J.e}{2}, \tag{20}$$

where $e^J = [e_{1,1}, e_{2,1}, ..., e_{k[X],1}, e_{1,2}, ..., e_{k[X],q}]$, q = k[X].L and M is made up of every single network's weights. The well-known recurring computation serves as the foundation for Newton's approach of reducing the efficiency of the function.

$$M_{i+1} = M_i - H^{-1} \nabla d(M), \tag{21}$$

whereas ∇ d(M) is the gradient of d(M) and d(M) = $\frac{e^J e}{2}$, then

$$\nabla d(M) = T^J(x).e. \tag{22}$$

Thus, the following can be used to characterize the Hessian array:

$$H(x) = T^J(x)T(x) + S(x), \tag{23}$$

the first derivatives of the network errors concerning weights and biases are contained in the Jacobian matrix T, and A vector for network errors is e. If it can be stated that S(x) is the minimal with respect to the Jacobian product, then this equation can be used to approximate the Hessian matrix.

$$H(x) \approx T^J(x)T(x), \tag{24}$$

Consequently, the Gauss-Newton algorithm is;

$$M_{i+1} = M_i - [T^J(x)T(x)]^{-1}J^J(x)e. \tag{25}$$

One disadvantage of this method is that the reduced Hessian matrix may need to be invertible. One possible solution to this issue is to use a modified Hessian matrix,

$$H(x) \approx \mu I + T^J(x)T(x), \tag{26}$$

the identity matrix is expressed by I, and μ illustrates the parameter that makes H(x) definitively positive and so invertible. The LM method takes into account the latest modification made to H(x).

$$M_{i+1} = M_i - T^J(x)e[T^J(x)T(x) + \mu kI]^{-1}. \tag{27}$$

As μ_k is the new way to represent μ, it is evident that μ is not static throughout the algorithm's execution. The technique's performance depends on the selected value of μ as it controls convergence reliability and speed by making sure the Hessian may be turned. For supervised learning, the Levenberg-Marquardt

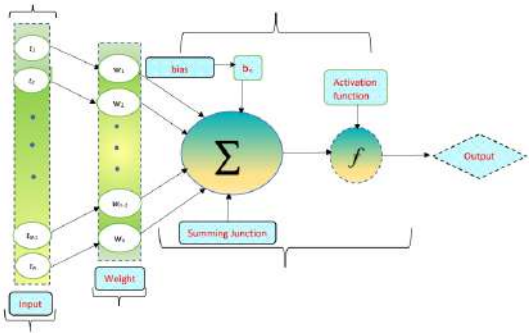

Figure 5. *Single-neuron architecture of artificial neural networks.*

technique processes data for training, validation, and testing using the reference data set from the first step. Regression analysis, mean square error, and histogram errors are performance metrics used to evaluate the convergence and accuracy of the design scheme. Figure (6) depicts the NN-LMT technique's operating mechanism as well as a step-by-step methodology.

5 Results and Discussion

We conducted experiments to evaluate the efficacy of the suggested method and investigate the impact of changes in various parameters of the moisture content rate of the conveyor belt dryer. By change in parameters, three different cases arise which are illustrated in Figure (7). Figure (8) illustrates how the training outcomes of ANN algorithms are constructed graphically. The graphical representation makes it

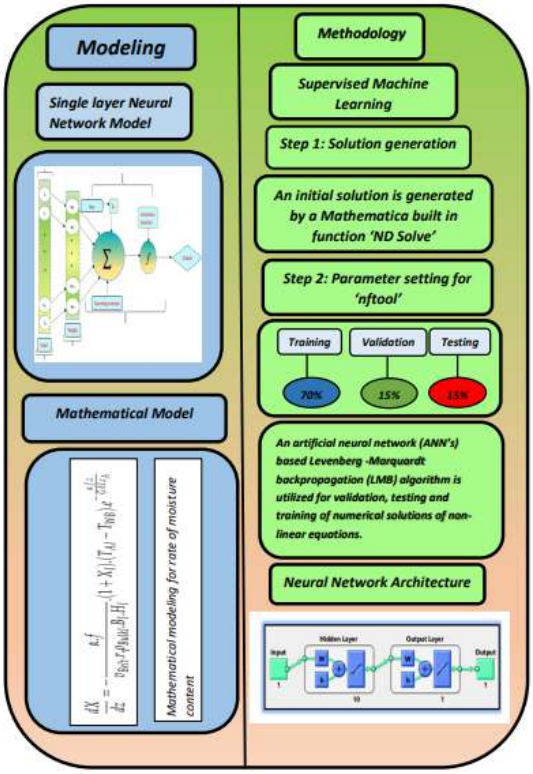

Figure 6. *Mechanism of the NN-LMT technique's design for numerically solving the mathematical model of moisture content rate.*

clear that the initial high values of MSE begins to decrease, once the total quantity of iterations increases. Further research indicates that the ideal line, represented by dotted lines in the (870, 42, and 1000) epochs, is reached by the lines created using the data acquired throughout the ANN's validation, training, and testing phases. The model's learning is stated to be optimal when the ANN reaches the smallest (MSE) outcome during the (870, 42, 1000) epochs, signifying the end of the training phase following multiple rounds of repetition. This process shows that the ANN models have finished their incredibly efficient training step. Utilizing the data values obtained from the numerical approach, the probability of validating, testing, and training a set of data are 15%, 70%, 15%, and 15%, accordingly.The process of convergence of the (MSE) function is shown in Figure (8). For each case performance values lie around 1.6777×10^{-11} at 870 epochs and time has taken 03 seconds, 5.0012×10^{-12} at 42 epochs, time consumed 01 seconds, and 1.1945×10^{-10} at 1000 epochs at 05 second. Investigating the effect of changing the rate of moisture content by changing parameters is shown in Figure (9). These variables indicate how quickly the conveyor belt removes water drops from the food.

The training states of ANN models-gradient, mu, and validation checks-are represented graphically in the Figure (10). The charts display the variation in the gradient coefficient using an epoch number that rises steadily over time. Be mindful of changes in the mu values, which indicate ANN weight fluctuations, and how close to zero the final gradient correlation coefficients are. These results show that as the number of epochs increases, the values of the least gradient coefficient continue to drop. While mu is between 10^{-08}, 10^{-11} and 10^{-07} with epochs (870, 42, 1000), the related values of training states are 1.9064×10^{-11},

Table 3. *Details of the pilot dryer's geometry and some of its operations*

Quantity	*Representation*	*Units*	*Values*
Air section	A_A	m^2	0.15
Alfalfa input moisture content (W.B)	Y_1	%	65.4±1.3
Thermal energy [28]	r	$kj.kg^{-1}$	2617
Form factor.Convective heat transfer coefficient [28]	F.α	$W.m^{-3}.K^{-1}$	5144
Width of belt	B_l	m	0.3
Alfalfa input bulk density	ρ_{Bulkl}	kgm^{-3}	197±7.5
Length of Belt	L_{TOT}	m	6
Alfalfa input moisture content (D.B)	X_l		1.892±0.110
Alfalfa bulk height	H_l	m	0.05
Transverse dimension.Convect.heat trans.coeficient	f.α=B_1.H_1.F.α	$W.m^{-1}.K^{-1}$	77.16
Air mass flow rate [28]	G_{Al}	$kg.s^{-1}$	0.354

5.1443×10^{-12} and 1.0573×10^{-10}. It has been demonstrated that after multiple test runs, the errors derived by ANN models progressively approach ideal and optimum values. These ANN training findings show that the created ANNs effectively finished their training assignments. Examining the error histogram is essential for determining how effectively ANN models perform. Moreover statistical performance of the gradient for each case is further demonstrated in Figure (10), and the values of its gradient can be found in the aforementioned region 9.9811×10^{-08}, 9.417×10^{-08}, and 0.00093748 respectively. Additionally, the mu, a control parameter for the neural network training method and a validation failure mode in all instances of the rate of moisture content model, is demonstrated in these figures and also displayed in Table (4). For each case, the value of the mu is 10^{-08}, 10^{-11} and 10^{-07} respectively.

Figure (11) illustrates graphically how the outputs deducted from the objectives are known as mistakes calculated from the proposed MLP network frameworks. The targeted data and approximate solution fit well together and exhibit the fewest errors, as shown in Figure (11). The error values of A(t) range between 10^{-05} and 10^{-07}. Similarly, the range of the error values S(t) is between 10^{-06} and 10^{-07}; for p(t) it ranges from 10^{-05} and 10^{-06}. The data set contains 70% training, 15% validation, and 15% testing. Different colors are used in each case to signify it. Blue represents trained data, which displays 160 out of 350 points in the first case, 110 out of 300 points in the second case, and 200 out of 500 points in the third case. Likely 30, 20, and 50 points are represented by the green color of verified data, whereas 30, 15, and 20 points are shown by the red color of tested data. The vertical line represents the zero error in each case.

In Figure (12), outcomes from the ANN models' training stages are shown. The x-axis of the graph displays the goal values, while the y-axis displays a visual representation of the ANN output forecasts. The data points gathered during the training phase are represented graphically as being on the blue compatibility (fit) line. The solid line represents the best-fitting linear regression line between the output values and target values, and the values of R depict their relationship. The output values and target values have a clear linear connection, as indicated by the regression analysis R = 1 used throughout this computation. These findings show that there is a link between the output values and target values and that ANN models have successfully completed the training phase with relatively low error levels. These outcomes show that the proposed ANN models' assessment phase (R = 1) was successfully finished with extremely low error rates. The acquired information conclusively shows that all (five) MLP network frameworks were designed to generate extremely precise predictions.

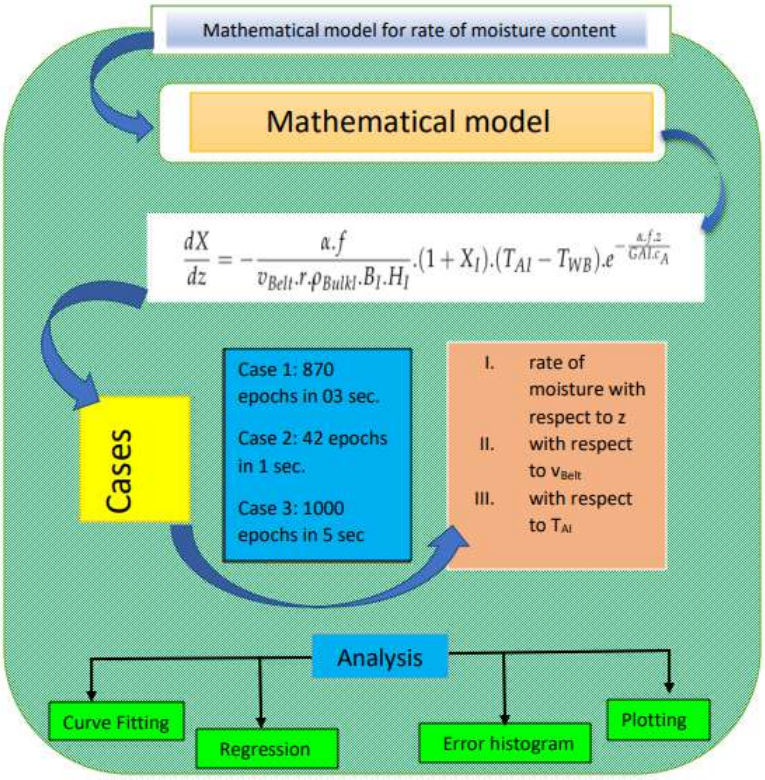

Figure 7. *Different cases for rate of moisture content*

Table 4. *Statistical analysis for the given problem*

				MSE					
Cases	Neurons	Training	Validation	Testing	Gradient	Mu	Epoches	Regression	Time
1	10	1.9064×10^{-11}	1.6777×10^{-11}	1.9757×10^{-11}	9.9811×10^{-08}	1×10^{-08}	870	1	3sec
2	10	5.1443×10^{-12}	5.0012×10^{-12}	5.7927×10^{-12}	9.4171×10^{-08}	1×10^{-11}	42	1	1sec
3	10	1.0573×10^{-10}	1.1945×10^{-10}	$1.4488e\times10^{-10}$	9.3748×10^{-04}	1×10^{-07}	1000	1	5sec

6 Conclusion

This study centered on investigating the mathematical version that describes the moisture content material fee in the conveyor belt dryer. The model offers insights into the velocity at which water droplets are extracted from the food in the course of the drying technique. Additionally, the version consists of variables including belt velocity, air entry temperature, and the z-axis, that can affect the drying rate. By using a supervised learning method with the NN-LMT-based computing method, the influence of these parameters on the drying rate becomes quantified. The findings, illustrated within the accompanying figures, reveal that an increase in belt speed results in a more glide of product through the dryer, consequently improving the dehydration rate. Moreover, the presence of high air temperature and low relative humidity allows rapid dehydration. The variable z, as formerly mentioned, is intricately connected to the area and its manipulation can provide multiplied space for food substances to go into, thereby augmenting the dehydration rate. The method's robustness, perfection, and efficacy were assessed through the use of a comprehensive graphical evaluation that made use of methods such as regressions, error histograms, imply square errors (MSE), and computational complexity evaluation. These consequences together contribute to our knowledge of optimizing the conveyor belt dryer's overall performance and display the viability of the evolved soft computing approach for addressing this problem. Future research endeavors might also further refine the version and discover extra parameters to enhance the efficiency of the drying procedure in conveyor belt dryers.

[performance for varying z] [performance for varying v_{belt}] [performance for varying T_{Al}]

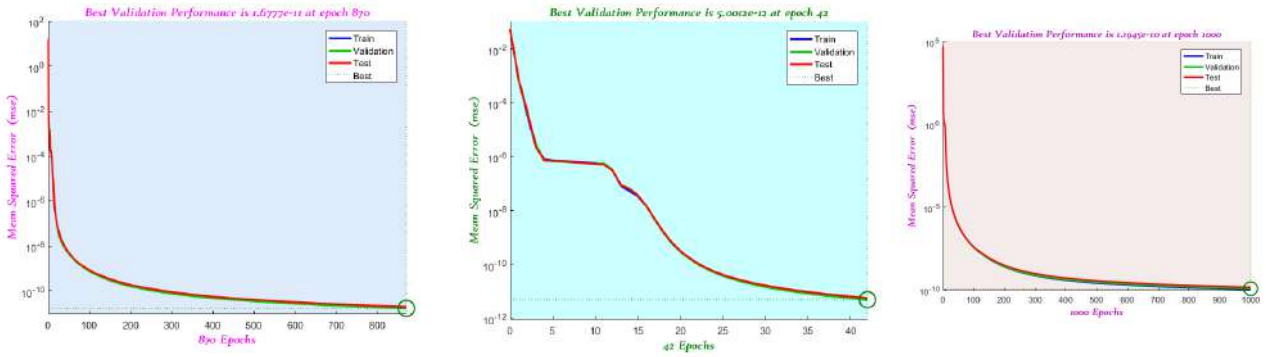

Figure 8. The NN-LMT's average square error for the moisture content rate

[Fitting plot for varying z] [Fitting plot for varying v_{belt}] [Fitting plot for varying T_{Al}]

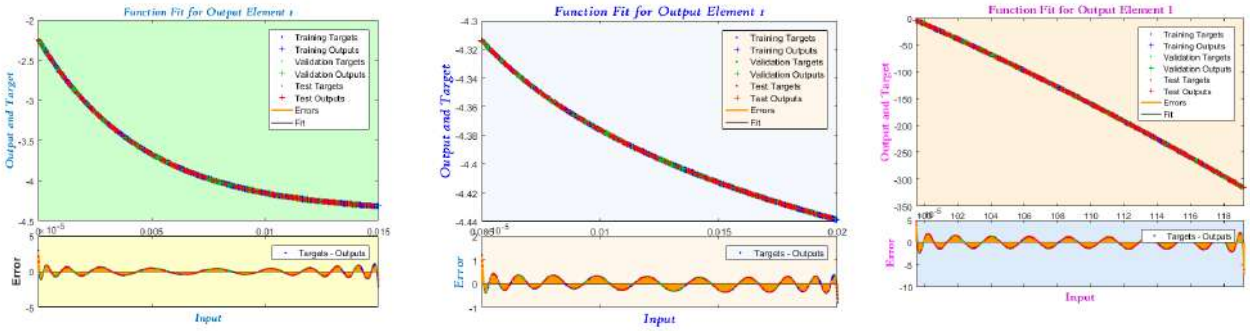

Figure 9. Analyzing the differences between the approximate solutions produced by NN-LMT and the numerical solution.

[Training plot for varying z] [Training plot for varying v_{belt}] [Training plot for varying T_{Al}]

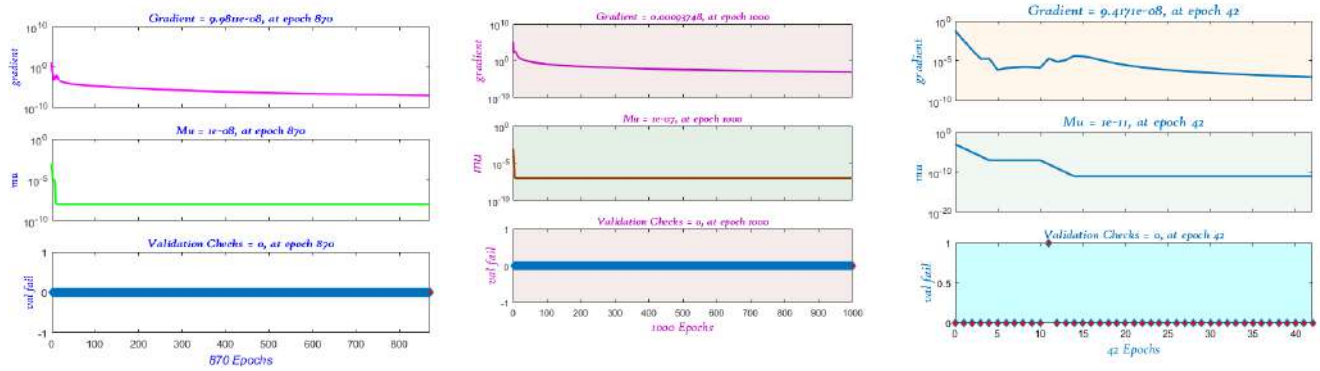

Figure 10. Measuring the NN-LMT method's performance in predicting moisture content rate by its mean square error.

[error histogram plot for varyingz] [error histogram plot for varyingv_{belt}] [error histogram plot for varying T_{AI}]

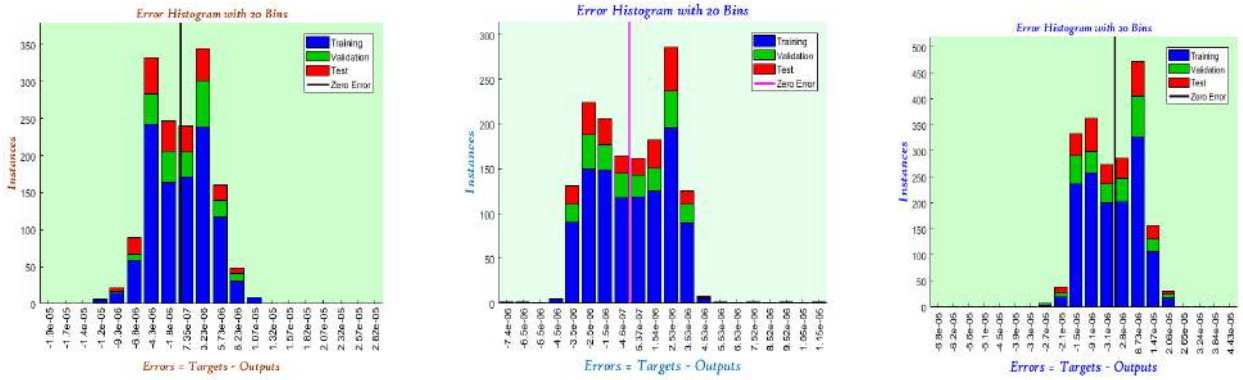

Figure 11. The desired data and the approximative solution are taken into consideration while analyzing the error histogram.

Table 5. *Abbreviation and their description*.

Abbreviation	***Description***
ODE	*Ordinary Differential Equation*
CBD	*Conveyor-Belt Dryers*
NN	*Neural Network*
MLP	*Multiple-Layer Perceptron*
RMC	*Rate of Moisture Content*
LMT	*Levenberg-Marquardt Training*
ANN	*Artificial Neural Network*

[Regression plot for varyingz] [Regression plot for varyingv_{belt}] [Regression plot for varying T_{AI}]

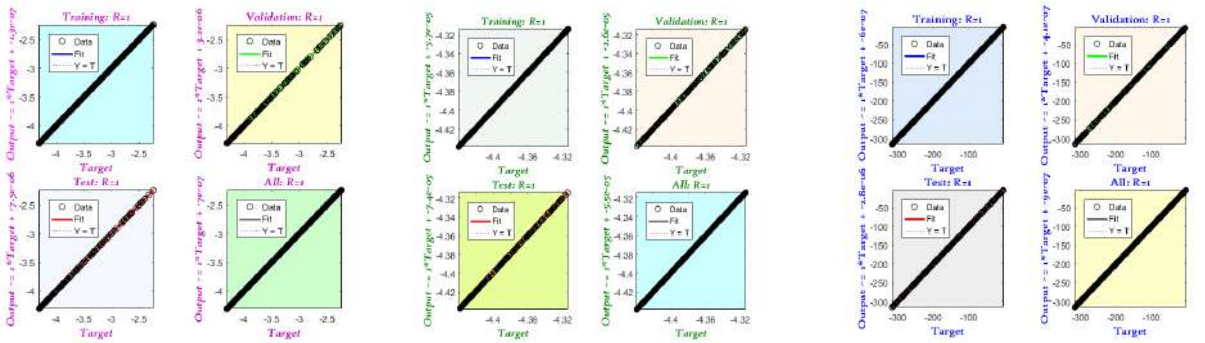

Figure 12. Regression analysis of the rate of moisture content.

7 Abbreviations and descriptions

Table (5) represents the descriptions of important abbreviations used in this study.

Author Contributions

Awais Khan:Supervison, Writing- Original draft preparation, Methodology.**Tauqeer Ahmad**: Software, Validation, Reviewingand Editing. **NaveedAhmed Khan**: Conceptualization. **Muhammad Sulaiman**: Investigation, Data curation, Visualization.

Compliance with Ethical Standards

It is declare that all authors don't have any conflict of interest. It is also declare that this article does not contain anystudies with human participants or animals performed by any of the authors. Furthermore, informed consent wasobtained from all individual participants included in the study.

References

[1] Fernandes, F. A. N., Rodrigues, S., Law, C. L., Mujumdar, A. S. (2011) 'Drying of exotic tropical fruits: a comprehensive review', *Food and Bioprocess Technology*, 4(2), pp. 163–185.

[2] Kaleta, A., Górnicki, K., Winiczenko, R., Chojnacka, A. (2013) 'Evaluation of drying models of apple (var. Ligol) dried in a fluidized bed dryer', *Energy Conversion and Management*, 67, pp. 179–185.

[3] Márquez, C. A., De Michelis, A. (2011) 'Comparison of drying kinetics for small fruits with and without particle shrinkage considerations', *Food and Bioprocess Technology*, 4(7), pp. 1212–1218.

[4] Tzempelikos, D. A., Mitrakos, D., Vouros, A. P., Bardakas, A. V., Filios, A. E., Margaris, D. P. (2015) 'Numerical modeling of heat and mass transfer during convective drying of cylindrical quince slices', *Journal of Food Engineering*, 156, pp. 10–21.

[5] Bezerra, C. V., da Silva, L. H. M., Corrêa, D. F., Rodrigues, A. M. C. (2015) 'A modeling study for moisture diffusivities and moisture transfer coefficients in drying of passion fruit peel', *International Journal of Heat and Mass Transfer*, 85, pp. 750–755.

[6] García-Alvarado, M. A., Pacheco-Aguirre, F. M., Ruiz-López, I. I. (2014) 'Analytical solution of simultaneous heat and mass transfer equations during food drying', *Journal of Food Engineering*, 142, pp. 39–45.

[7] Ertekin, C., Firat, M. Z. (2017) 'A comprehensive review of thin-layer drying models used in agricultural products', *Critical Reviews in Food Science and Nutrition*, 57(4), pp. 701–717.

[8] Fernando, W. J. N., Low, H. C., Ahmad, A. L. (2011) 'Dependence of the effective diffusion coefficient of moisture with thickness and temperature in convective drying of sliced materials. A study on slices of banana, cassava and pumpkin', *Journal of Food Engineering*, 102(4), pp. 310–316.

[9] Friso, D. (2015) 'A mathematical solution for food thermal process design', *Applied Mathematical Sciences*, 9(6), pp. 255–270.

[10] Friso, D., Baldoin, C. (2015) 'Mathematical modelling and experimental assessment of agrochemical drift using a wind tunnel', *Applied Mathematical Sciences*, 9(110), pp. 5451–5463.

[11] Friso, D. (2019) 'An approximate analytic solution to a non-linear ODE for air jet velocity decay through tree crops using piecewise linear emulations and rectangle functions', *Applied Sciences*, 9(24), p. 5440.

[12] Aversa, M., Curcio, S., Calabro, V., Iorio, G. (2007) 'An analysis of the transport phenomena occurring during food drying process', *Journal of Food Engineering*, 78(3), pp. 922–932.

[13] Ramsaroop, R., Persad, P. (2012) 'Determination of the heat transfer coefficient and thermal conductivity for coconut kernels using an inverse method with a developed hemispherical shell model', *Journal of Food Engineering*, 110(1), pp. 141–157.

[14] Askari, G. R., Emam-Djomeh, Z., Mousavi, S. M. (2013) 'Heat and mass transfer in apple cubes in a microwave-assisted fluidized bed drier', *Food and Bioproducts Processing*, 91(3), pp. 207–215.

[15] Baini, R., Langrish, T. A. G. (2007) 'Choosing an appropriate drying model for intermittent and continuous drying of bananas', *Journal of Food Engineering*, 79(1), pp. 330–343.

[16] Guinè, R. P. F. (2008) 'Pear drying: Experimental validation of a mathematical prediction model', *Food and Bioproducts Processing*, 86(4), pp. 248–253.

[17] Castro, A. M., Mayorga, E. Y., Moreno, F. L. (2018) 'Mathematical modelling of convective drying of fruits: A review', *Journal of Food Engineering*, 223, pp. 152–167.

[18] Barati, E., Esfahani, J. A. (2011) 'A new solution approach for simultaneous heat and mass transfer during convective drying of mango', *Journal of Food Engineering*, 102(4), pp. 302–309.

[19] Sabarez, H. T. (2014) 'Mathematical modeling of the coupled transport phenomena and color development: Finish drying of trellis-dried sultanas', *Drying Technology*, 32(5), pp. 578–589.

[20] Akdaş, S., Başlar, M. (2015) 'Dehydration and degradation kinetics of bioactive compounds for mandarin slices under vacuum and oven drying conditions', *Journal of Food Processing and Preservation*, 39(6), pp. 1098–1107.

[21] Kaya, A., Aydın, O., Dincer, I. (2008) 'Experimental and numerical investigation of heat and mass transfer during drying of Hayward kiwi fruits (Actinidia Deliciosa Planch)', *Journal of Food Engineering*, 88(3), pp. 323–330.

[22] Bon, J., Rosselló, C., Femenia, A., Eim, V., Simal, S. (2007) 'Mathematical modeling of drying kinetics for apricots: influence of the external resistance to mass transfer', *Drying Technology*, 25(11), pp. 1829–1835.

[23] Lemus-Mondaca, R. A., Zambra, C. E., Vega-Gálvez, A., Moraga, N. O. (2013) 'Coupled 3D heat and mass transfer model for numerical analysis of drying process in papaya slices', *Journal of Food Engineering*, 116(1), pp. 109–117.

[24] Friso, D. (2018) *Ingegneria dell'industria Agroalimentare (Food Engineering Operations)*. Padova, Italy: CLEUP.

[25] Geankoplis, C. J. (1993) *Transport Process and Unit Operations*, 3rd edn. New York: Prentice-Hall.

[26] Salemović, D. R., Dedić, A. Đ., Ćuprić, N. Lj. (2017) 'Two-dimensional mathematical model for simulation of the drying process of thick layers of natural materials in a conveyor-belt dryer', *Thermal Science*, 21(3), pp. 1369–1378.

[27] Farias, R. P., Santiago, D. C., Holanda, P. R. H., Lima, A. G. B. (2004) 'Drying of grains in conveyor dryer and cross flow: A numerical solution using finite-volume method', *Revista Brasileira de Produtos Agroindustriais, Campina Grande*, 6(1), pp. 1–16.

[28] Friso, D. (2020) 'Conveyor-belt dryers with tangential flow for food drying: Mathematical modeling and design guidelines for final moisture content higher than the critical value', *Inventions*, 5(2), p. 22.

[29] Gul, N., Mashwani, W. K., Aamir, M., Aldahmani, S., Khan, Z. (2023) 'Optimal model selection for k-nearest neighbours ensemble via sub-bagging and sub-sampling with feature weighting', *Alexandria Engineering Journal*, 72, pp. 157–168.

[30] Ali, A., Khan, Z., Khan, D. M., Aldahmani, S. (2024) 'An optimal random projection k nearest neighbors ensemble via extended neighborhood rule for binary classification', *IEEE Access*, 12, pp. 61401–61409.

[31] Khan, Z., Ali, A., Aldahmani, S. (2024) 'Feature selection via robust weighted score for high dimensional binary class-imbalanced gene expression data', *Heliyon*, 10(19).

[32] Aldahmani, S., Zoubeidi, T. (2020) 'Graphical group ridge', *Journal of Statistical Computation and Simulation*, 90(18), pp. 3422–3432.

[33] Khan, Z., Ali, A., Khan, D. M., Aldahmani, S. (2024) 'Regularized ensemble learning for prediction and risk factors assessment of students at risk in the post-COVID era', *Scientific Reports*, 14(1), p. 16200.

[34] Friso, D. (2021) 'Conveyor-belt dryers with tangential flow for food drying: Development of drying Odes useful to design and process adjustment', *Inventions*, 6(1), p. 6.

[35] Huang, W., Jiang, T., Zhang, X., Khan, N. A., Sulaiman, M. (2021) 'Analysis of Beam-Column Designs by Varying Axial Load with Internal Forces and Bending Rigidity Using a New Soft Computing Technique', *Complexity*, 2021(1), p. 6639032.

[36] Alhakami, H., Kamal, M., Sulaiman, M., Alhakami, W., Baz, A. (2022) 'A machine learning strategy for the quantitative analysis of the global warming impact on marine ecosystems', *Symmetry*, 14(10), p. 2023.

[37] Khan, N. A., Sulaiman, M., Alshammari, F. S. (2022) 'Heat transfer analysis of an inclined longitudinal porous fin of trapezoidal, rectangular and dovetail profiles using cascade neural networks', *Structural and Multidisciplinary Optimization*, 65(9), p. 251.

[38] Khan, N. A., Sulaiman, M., Tavera Romero, C. A., Alarfaj, F. K. (2021) 'Theoretical analysis on absorption of carbon dioxide (CO_2) into solutions of phenyl glycidyl ether (PGE) using nonlinear autoregressive exogenous neural networks', *Molecules*, 26(19), p. 6041.

[39] Khan, A., Sulaiman, M., Alhakami, H., Alhindi, A. (2020) 'Analysis of oscillatory behavior of heart by using a novel neuroevolutionary approach', *IEEE Access*, 8, pp. 86674–86695.

[40] Sulaiman, M., Salhi, A., Khan, A., Muhammad, S., Khan, W. (2018) 'On the theoretical analysis of the plant propagation algorithms', *Mathematical Problems in Engineering*, 2018(1), p. 6357935.

[41] Khan, N. A., Sulaiman, M., Kumam, P., Bakar, M. A. (2021) 'Thermal analysis of conductive-convective-radiative heat exchangers with temperature dependent thermal conductivity', *IEEE Access*, 9, pp. 138876–138902.

[42] Sulaiman, M., Masihullah, M., Hussain, Z., Ahmad, S., Mashwani, W. K., Jan, M. A., Khanum, R. A. (2019) 'Implementation of improved grasshopper optimization algorithm to solve economic load dispatch problems', *Hacettepe Journal of Mathematics and Statistics*, 48(5), pp. 1570–1589.