A rapid classification of wheat flour protein content using artificial neural network model based on bioelectrical properties

Sucipto Sucipto¹, Maffudhotul Anna², Muhammad Arwani², Yusuf Hendrawan⁴
¹,²Department of Agroindustrial Technology, Agricultural Technology Faculty, Universitas Brawijaya, Indonesia
³Halal and Qualified Industry Development (HAL-Q ID), Universitas Brawijaya, Indonesia
⁴Department of Agricultural Engineering, Agricultural Technology Faculty, Universitas Brawijaya, Indonesia
*Corresponding author, e-mail: ciptotip@ub.ac.id

Abstract

A conventional technique of protein analysis is laborious and costly. One rapid method used to estimate protein content is near infrared spectroscopy (NIRS), but the cost is relatively expensive. Therefore, it is necessary to find a cheaper alternative measurement such as measuring the bioelectrical properties. This preliminary study is a new rapid method for classified modeling of wheat flour protein content based on the bioelectrical properties. A backpropagation artificial neural network (ANN) was developed to classify the protein content of wheat flour. ANN input were bioelectrical properties, namely capacitance, and resistance and output was a type of the flour, namely hard, medium and soft flour. The result showed that the ANN model could classify the various type of flour. The best ANN model produces a mean square error (MSE) and regression correlation (R) of 0.0399 and 0.9774 respectively. This ANN model could classify the protein content of wheat flour based on the bioelectrical properties and have the potential to be used as a basic instrument to estimate the protein content.

Keywords: artificial neural network, capacitance, protein, resistance, wheat flour

Copyright © 2019 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Protein is often used as the basis for the wheat flour classification. Wheat flour consists of starch, water, protein is 70 to 75%, 12 to 14%, 8 to 16% respectively, also fat and ash are 2% and 1% respectively [1]. Gluten is the main wheat flour protein consisted of gliadin (20 to 25%) and glutenin (35 to 40%). Based on the protein content, wheat flour is divided into three groups, namely hard flour (minimum protein of 12%), medium flour (minimum protein of 10 to11%), and soft flour (minimum protein of 8 to 9%). The selection of wheat flour for foodstuffs is adjusted to its protein content. Wheat flour with a high protein contains much gluten that will determine the texture and elasticity of food products. Gliadin and glutenin give quite a large effect on the stability and dough development time [2].

On the production of wheat flour, the rapid measurement of protein content and other chemical components is very important. The conventional technique is difficult to implement because it is laborious. Some protein testing methods have been developed, such as the methods of Kjehdal, Bradford, Dumas (Nitrogen Combustion), Biuret, Lowry [3, 4], dye binding, and formol titration. Those tests take a long time, relatively expensive, need complex sample preparation, and also need chemical and laboratory expertise. One new method for predicting the quality of wheat flour is near NIRS [5], but this method is still relatively expensive.

Recently, a rapid measurement which is relatively cheap is urgently required. One alternative is dielectric spectroscopy based on bioelectrical properties of a material. Bioelectric is the electrical properties of cells or living tissue affected by metabolism compounds and ion exchange inside. Bioelectrical properties have been used to determine the quality of foodstuffs, for example, sugarcane [6-8], palm civet coffee [9], Garut citrus fruits [10], chickpea flour [11], legume flour [12], beef tallow, lard and palm oil [13, 14]. The advantages of a dielectric method with parallel plates are cheap, simple, fast and practical. This method is used to determine the quality and purity of materials quickly and practically.
The core development of the instrument is a model that was used to convert the main input became an output as fast as possible. A backpropagation ANN is one of the mathematical models that can learn input and output through iterations without any previous knowledge in the shortest time. The structure is relatively simple with connections among neurons [15]. Each neuron is processing elements that perform the converting function of input to output. ANN model has been widely used to predict the quality of ingredients and food products, including, to determine the level of eggs freshness [16], and predict wheat quality parameters [5]. The aim of this study was to develop a model that can classify the protein content of wheat flour based on the bioelectrical properties.

2. Research Method

2.1. Sample Preparation

Six different brands of wheat flour divided into three class and given symbols of A, B (Hard Flour); C, D (Medium Flour); and E, F (Soft Flour). The hard flour class consisted wheat flour which had a minimum protein content of 12%, medium flour with a protein content of 10 to 11%, and soft flour with a protein content of 8 to 9% [17]. The samples were obtained from some produced in East Java Indonesia.

2.2. Bioelectrical Properties Measurement

A Parallel plate capacitance sensor was made from a copper plate with the dimensions of 0.5x1x2cm. The samples were put in the parallel plate and connected with LCR meter (BK Precision 879B) to measure the bioelectrical properties and saved those data to the computer. Three types of bioelectrical properties were measured, namely inductance, capacitance and resistance at a frequency of 100, 120, 1000 and 10000 Hz. The bioelectrical properties were measured by two factors, the first factor is flour brand with six types of flour and symbolized by A, B, C, D, E and F. The second factor is the temperature of measurement with three levels (25, 30, 35°C) that is repeated three times. Bioelectrical properties of each wheat flour sample were measured seven times per frequency in one temperature. Therefore, the total datasets for six type of wheat flour are 1512 datasets.

2.3. Datasets Preparation

Prior to modeling in ANN, there is preparation the datasets to obtain an appropriate input in three steps namely; selection, preprocessing, and transformation data. From the measurement data of three bioelectrical properties there is a huge number of data, therefore selecting data is one of the most important steps in the preparation of datasets. Preprocessing data such as cleaning and smoothing also used to remove the instances data. A preprocessing algorithm was used to scaling the datasets. Then, the selected data was transformed using the normalization algorithm and become the input of ANN modeling.

2.4. Design of ANN Topology

ANN topology was developed using Matlab R2012a with bioelectrical properties input (L, C or R) and output (hard flour (1 0 0), medium flour (0 1 0), soft flour (0 0 1)). A Backpropagation ANN was commonly used in agricultural-based industry [18]. The ANN topology design was performed in 2 stages: training and validation. The model training stage was conducted by trial and error from the activation function, learning function, the number of hidden layers and node per hidden layer as well as the number of epochs. Trial and error of ANN topology modification can be seen in Table 1. The modification aimed to produce the best ANN topology with the lowest validation mean square error (MSE) and the highest validation correlation coefficient (R). MSE and R formula can be seen in (1) and (2) [19].

\[
R = \sqrt{1 - \frac{\sum_{i=1}^{N} \left( \frac{Q_{\text{exp}} - Q_{\text{cal}}}{Q_{\text{exp}}} \right)^2}{N}}
\]

(1)

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{Q_{\text{exp}} - Q_{\text{cal}}}{Q_{\text{exp}}} \right)^2
\]

(2)
where:
\[Q_{\text{exp}}\] = Measurement value
\[Q_{\text{cal}}\] = Prediction value
\[N\] = Measurement number of dataset

### Table 1. Modification of ANN

<table>
<thead>
<tr>
<th>Activation Function</th>
<th>Hidden layer</th>
<th>Output layer</th>
<th>Training Function</th>
<th>Node in hidden layer</th>
<th>Epoch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tansig</td>
<td>Purelin</td>
<td>Trainglm</td>
<td>10</td>
<td>1</td>
<td>1000</td>
</tr>
<tr>
<td>Tansig</td>
<td>Traingdl</td>
<td>15</td>
<td>2</td>
<td>5000</td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Purelin</td>
<td>Traingd</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tansig</td>
<td>Trainqgb</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Trainscg</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Trainbg</td>
<td>20</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Traindx</td>
<td>20</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Trainb</td>
<td>20</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Trainbr</td>
<td>20</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Trainoss</td>
<td>20</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Traincgb</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Traincf</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Traincpp</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logsig</td>
<td>Trainrp</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Results and Analysis

3.1. Preprocessing Data

As shown in Figure 1, the bioelectrical properties at a frequency of 10 kHz are biased. On the other hand, the result from preprocessing data was shown only two types of bioelectrical properties (capacitance and resistance) produce a higher coefficient correlation than all three types of bioelectrical properties. Therefore, in further discussion, both capacitance and resistance in the frequencies of 100, 120, and 1000 Hz were used to develop an ANN model. ANN model was developed only using bioelectrical properties that were obtained from the measurement at 30°C because it is approached the production of wheat flour at room temperature about 30°C.

![Figure 1. Resistance of wheat flour in various frequencies within six brands of wheat flour; A and B for Hard flour, C and D for medium flour, E and F for soft flour](image)

3.2. Effects of Frequency and Temperature on Capacitance

Effects of temperature and frequency on bioelectrical properties include capacitance and resistance at frequencies of 100 Hz, 120 Hz, 1000 Hz and temperatures of 25°C, 30°C, 35°C respectively. The mean of 7 bioelectric data measurement of each frequency on each wheat flours was depicted in Figure 2.
Figure 2. Capacitance of wheat flour in (a) 25°C, (b) 30°C, (c) 35°C in the six types of wheat flour; A and B for hard flour, C and D for medium flour, E and F for soft flour.

Capacitance is material ability to store electric charges. Bioelectric properties show the interaction of foodstuff and electric field [20]. Figure 2 show that the higher the frequency, the lower the capacitance value is. Those results are in accordance with Nelson and Trabelsi [21] stating that the higher the frequency, the lower the dielectric constant and capacitance. Higher frequency caused the rapid fluctuation of negative and positive charge in the parallel plate. When the frequency is increased, the electric current moves alternatingly in rapid motion and this motion affects the amount of capacitor charge. Before the capacitor is fully charged, the electric current is turned around so that the amount of the charge is decreased causing the capacitance value decreased [10]. Furthermore, the bioelectrical properties of wheat flour are influenced by temperature [21]. The dielectric constant decreases with increasing frequency, but increases with increasing of temperature and moisture content [11, 12]. The higher the temperature, the higher the dielectric constant value is [21].

3.3. Effects of Frequency and Temperature on Resistance

Resistance is the material ability to inhibit the flow of electrical charges. Figure 3 shown that the resistance with frequency decreased. In the higher frequency, the electrical charge that fills the capacitor becomes less which reduced the resistance of the electrical charges. Increasing external frequency will increase the speed of change in mobility of electrical charge in the material [10]. Kraft [22] reported that resistance of material was influenced by the length of dimensions, cross-sectional area, type of material and temperature. Therefore, wheat flour belongs to the insulators types cause the resistive properties in various frequency.

On the foodstuffs and agricultural products, especially in wheat flour, the water content is one of the main indicator of quality product, with a maximum of 14% [21]. Water is a very poor conductor and it is a special ion solvent that causes electrical charge flow between capacitor plates. There is no dielectric material which has perfect insulator consequently there is always leakage electrical charge between capacitor plates. Because that, the wheat flour contains water that is a very poor conductor, then its resistance increased with the rise of temperature. At high temperatures, there is high ions mobility resulting in electron collision in the material, consequently inhibits the external electrical charge to penetrate into the material.
3.4. ANN Modelling

The selection of ANN topology was carried out using trial and error. After preprocessing, the dataset was reduced to 756 bioelectric data and then become the input of ANN model. There is a total of 71.43% (540 data) for training and 28.57% (216 data) for validation. Before ANN modeling, the dataset was transformed using min-max normalization.

The transformed dataset was used to develop an ANN model. The first step was the selection of activation function which is shown that tansig algorithm used in the hidden layer and output layer produces the smallest MSE by 0.1935 and 0.088473 of the correlation coefficient (R). According to Özkran, Kun [23], the activation functions with a nonlinear type such as tansig and logsig are considered as more capable to resolve more complex problems so that the function can be used in the hidden layer. As shown in Table 2.

<table>
<thead>
<tr>
<th>Training Function</th>
<th>MSE Validation</th>
<th>R Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trainlm</td>
<td>1.2840</td>
<td>0.15811</td>
</tr>
<tr>
<td>Traingdm</td>
<td>0.1935</td>
<td>0.88473</td>
</tr>
<tr>
<td>Traingd</td>
<td>0.1939</td>
<td>0.88447</td>
</tr>
<tr>
<td>Traingdx</td>
<td>0.1676</td>
<td>0.90106</td>
</tr>
<tr>
<td>Traincgb</td>
<td>0.1169</td>
<td>0.9332</td>
</tr>
<tr>
<td>Trainscg</td>
<td>0.0939</td>
<td>0.94599</td>
</tr>
<tr>
<td>Trainbfg</td>
<td>0.3058</td>
<td>0.79448</td>
</tr>
<tr>
<td>Trainb</td>
<td>0.2860</td>
<td>0.82434</td>
</tr>
<tr>
<td>Trainbr</td>
<td>0.0582</td>
<td>0.96683</td>
</tr>
</tbody>
</table>

Trainbr learning function produces the lowest MSE of validation by 0.0582 and highest R of validation by 0.96683 compared to other learning function. Trainbr function (Bayesian regularization Backpropagation) is a network training function that updates the values of the
A rapid classification of wheat flour protein content using...

(Sucipto Sucipto)

weight and bias in accordance with Lavenberg-Marquardt optimization. This function minimizes squared and weight error combinations, then determine the correct combination to produce the best network.

Table 3 shown that ANN topology which has 20 nodes in the hidden layer 1 and 50 nodes in the hidden layer 2 provides the smallest MSE of validation by 0.0399 and highest R of validation by 0.97743. The best results are obtained using two hidden layers, but the iterations time is too long. According to Zhang, Eddy Patuwo [24], with more layer and nodes per layer, the network can solve more complex problems. However, if the number of nodes less than the complexity of problems then resulted in an underfitting condition which is referred to an ANN that can neither model the training data nor generalizes to new data, otherwise ANN model the training data too well called which are called overfitting.

![Image](image.png)

**Figure 4.** Performance (a) and Regression plot (b) in Training stage

The structure of ANN model can be seen in Figure 5 with the combination of 2–20–50-3 nodes. Two nodes inputs are capacitance and resistance, 20 nodes in the first hidden layer and 50 nodes in the second hidden layer, and also three nodes in the output layer namely hard flour, medium flour, and soft flour. The complexity data for flour classification is more than the model of effective environmental control system [26]. The final model was developed to obtain weights in each node. Weight is key element of ANN model which indicates the relative strength of the
data input or many connections that move data from one layer to another layer. Besides obtaining weight, optimal bias is also generated in the network. Each node receives input signals from other nodes in front of it through the activation function which will generate an output signal. ANN is able to learn from the examples were never known before [24].

![Figure 5. Selected ANN Topology](image)

3.5. Classification of wheat flour

Bioelectric properties of materials, especially capacitance and resistance are greatly influenced by chemical content. Therefore, the ANN model for classification of wheat flour based on the protein content is important because it is related to processing factors such as water absorption and gluten strength. That application is beneficial for bakery and another industry to produce quality products with high quality raw material. Classification of wheat flour protein based on bioelectric properties can be developed further and used to determine the wheat flour. It is related to wheat flour labeling which is rapid and affordable costs. The result of this study becomes the preliminary study for the development of a wheat flour rapid classification instrument which is reliable and affordable costs.

4. Conclusion

According to this study carried out the bioelectrical properties (capacitance and resistance) using the ANN model could classify the protein content of wheat flour. The best ANN topology to classify the wheat flour become hard, medium and soft flour of protein content is 2-20-50-3. ANN model using tansig activation function and trainbr learning function produces MSE training of 0.0097 and validation of 0.0399, and also the $R^2$ of training and validation are 0.9947 and 0.9774 respectively. This model might be useful for developing an instrument for rapid protein analysis of wheat flour.

References


