Assessment and Prediction of Free Fatty Acids Changes in Maize Based on Effective Accumulated Temperature in Large Granaries

Kai Chen¹, Wenfu Wu²,³, Yankun Lan⁴, Zhe Liu⁴, Feng Han⁴, and Yan Xu⁴

¹School of Biological and Agricultural Engineering, Jilin University, Changchun, China; ²School of Grain Science and Technology, Jilin Business and Technology College, Changchun, China; ³Warehousing Department, China Grain Reserves Group LTD. Company Jilin Branch, Changchun, China

ABSTRACT
In large granaries, maize storage period could last 3–5 years in China. Free fatty acids (FFA) content is commonly used as a sensitivity indicator of grain quality changes during storage. Samples of the stored grain are taken manually and tested in the laboratory monthly for regular quality monitoring. Although substantial labor and time are consumed, the testing way lacks of real-time. Temperature is a main factor influencing the quality of stored grain. In this study, it was analyzed that the effective accumulated temperature (EAT) had a strong correlation coefficient with FFA. Classification and regression trees (CART) and model trees (ML) in machine learning (ML) methods were used to estimate FFA by EAT based on data collected from 11 large granaries in northeastern China. While the minimum number of samples for segmentation was set as 20, the two models both had the optimal performance. The two models were evaluated by mean absolute error (MAE), root mean square error (RMSE) and the coefficient of determination (R²). The MAE, RMSE and R² of CART and MT are 1.296, 1.761, 0.759; 1.247, 1.821, and 0.741, respectively. As CART had the lower RMSE and the larger R², the model performance of CART was better. The model provided a method to estimate the changes of FFA based on the exist temperature monitoring system in large granaries as a way of stored maize quality real-time monitoring.

Introduction
Maize (Zea mays L.) as one of the most important crops in the world is widely grown for food, forage, energy, and industrial materials.¹,² In 2019, the output of maize had reached 260.78 million tons in China, while the yield of paddy and wheat were 296.93 and 237.28 million tons, respectively.³ Maize has been definitely the largest output crop for years in China.⁴ After harvesting, storage is an inevitable step for grain before eating or processing. During grain storage periods, quality and quantity loss may occur due to insufficient control manners which may result in the growth and reproduction of insects, mite, mold, and fungi.⁵,⁶ Regular quality monitoring is indispensable to insure grain storage security.

FFA is commonly regarded as a sensitivity indicator of grain quality changes during storage.⁷ Partial hydrolysis of the glycerides may occur in grain due to unfavorable storage conditions, then FFA increased.⁸ The increasing FFA means deterioration happening in quality which could contribute to deterioration of taste and smell, loss of seed vigor and reduction in processing property.⁹,¹⁰ The American Association of Cereal Chemists (AACC) has provided four methods for FFA determination including general method, rapid method for small grain, rapid method for corn and colorimetric method.¹¹ And all of them are operating through a titrimetric procedure which are precise but time-consuming.¹² Consequently, more attention should be paid to a real-time and effective FFA monitoring method.
consuming. Besides the traditional gas chromatograph method, other determining methods also have been explored. The Near infrared spectroscopy and the gas chromatography-mass spectrometer detector are used to analysis FFA in soybean and berry seed, respectively. Hui Jiang et al. used a homemade olfactory visualization sensor to quantitatively determine FFA of stored rice. If the grain samples needed to test were substantial, those methods were time-consuming and laborious. Therefore, a rapid and effective measurement of FFA has a significant meaning for stored grain in large scale.

Many grain storage researches were conducted in small-holders or laboratory scale. The conditions of grain bulk stored in the large-scale granary are quite different with the tests carried in small scale including constitutes (impurity, damaged grain), temperature gradient, management and storage period. Temperature gradients in stored grain generated from seasonal changes in ambient air temperature and solar radiation. Temperature and moisture content are two key factors that determine grain quality during storage period. Modern commercial grain storage enterprises have built a multifunctional system to insure grain storage security at least including temperature monitoring system, ventilation, cooler machine, inner circulation fumigation. Measuring grain temperature is the main method used by grain industries to monitor grain storage conditions because it is easy and inexpensive to achieve by using thermocouples. The relative humid (RH) sensor is not widely used in grain bulk by grain storage industries. A mass of grain temperature data has be obtained during the grain storage. However, the data is commonly used to monitor the storage condition or real-time control. Li, X.J. et al. had built a smart cooling-aeration system based on the temperature data, which could save energy consumption of cooling the bulk paddy stored in concrete silos. To data, less than 1% data generated in Internet of things is used and mostly for alarms or real-time control, but more can be used for optimization and prediction. It has been currently missing about how to use the data to predict the grain quality changes to insure grain storage security.

To seek the relation between the temperature data and the grain quality changes, machine learning (ML) methods could be utilized due to its effectiveness in dealing with big data. In recent years, ML methods have been applied in grain storage such as grain moisture determination, dry matter loss prediction, classification of grain storage inventory modes, prediction of grain damage in smallholders and insect populations prediction etc. ML could be an important tool used for improving grain storage security.

Classification and regression trees (CART) algorithm and model trees (MT) algorithm are both belongs to tree methods, which are widely used in data mining due to its ability of dealing with multiple variable and easy to improve. Compared with other ML methods, CART and its varieties are more acceptable to the public thanks to its explicable of the results. In this paper, models of CART and MT are proposed to predict the FFA value in stored maize based on temperature data collected from 11 large granaries in the northeast China.

Material and methods

Study area

The study maize granaries are located in Jilin Province and Heilongjiang Province, which setting in the northeast China. In 2019, the total maize output of China was 260.78 million tons, including 30.45 million tons of Jilin Province and 39.40 million tons of Heilongjiang Province according to the National Bureau of Statistics of the P.R.C., i.e., the production of maize in the two provinces accounted for 26.8% of the total. The experimental data were collected from 11 maize storage granaries including 3 in Heilongjiang Province and 8 in Jilin Province.
Temperature monitoring system of the granary

The granary is designed and built according to the national standard of P.R.C, GB 50320, 2014 Code for Design of Grain Storehouses. The length of the granary is generally 36–60 m and the width is 36–40 m. The height of grain bulk should not below 6 m and the height between the grain bulk surface and the horizontal members of the roof should not less than 1.8 m.

The temperature monitoring system deployment in the granary is regulated by the nation standard of LS/T 1203, 2002 Measurement and Control System for Condition of Stored-grain. The thermometric cables arranged in the granary are as follows: the distance between rows and columns in the horizontal direction shall not be greater than 5 m, the distance between rows and columns in the vertical direction shall not be greater than 3 m and the distance shall be away from the grain surface, granary bottom, and granary wall about 0.3–0.5 m.

Figure 1 depicts the basic working principle of temperature monitoring system. The thermometric cables with sensors are deployed in the storehouse and all the temperature data is collected at the integrated control box, then the data are transmitted through the wireless device to the work station. The data shown and stored in the work station is used as the basic information for stored grain management.

Maize quality assessment

The maize samples were taken from the granaries ranging from March 2017 to October 2020. The procedure was according to the national standard of GB 5491, 1985 Inspection of Grain and Oilseeds Methods for Sampling and Sample Reduction. The sampling quantity depends on the height and the...
area of the grain bulk. The height above 6 m is divided into 4 or 5 floors. Each floor is divided into small areas and each area should not be more than 50 m². Sampling should be taken in the four corners and the center points in each area.

The MC of maize samples was analyzed by the oven-drying method ASABE, 2017. According to the national standard of GB T 29890, 2013 Technical criterion for grain and oil-seeds storage, two provinces studied in this paper belong to the low temperature and high humidity grain storage area. The allowed highest MC of stored maize in this area is 14%, and the difference between the maximum moisture value and the minimum moisture value should not exceed 1%. Before storing in the granary, the quality of corn must be inspected one by one transport truck, and the corn must be stored by classification in strict accordance with the inspection quality standards.

The FFA value of maize samples was measured according to the national standard of GB T 20570, 2015 Guidelines for Evaluation of Maize Storage Character. At room temperature, the fatty acids in maize were extracted with anhydrous ethanol, and then titrated with potassium hydroxide standard solution, after which the FFA value was calculated. The standard stipulated that each sample should be determined twice by the same inspector. The average value was used as a result, and the difference between two values should not be greater than 2 mg KOH/100 g.

**Efficient accumulated temperature**

Temperature is a key factor for grain storage security. It has a great influence on the emergence or growth of mite and fungi in a grain bulk, and the metabolism and quality changes of the grain. Here, an efficient accumulated temperature (EAT), a temperature calculation form, is introduced as follows:

\[ A_{EAT} = \begin{cases} 
\sum_{d=0}^{n} (T_d - T_0), & T_d > T_0 \\
0, & T_d \leq T_0
\end{cases} \]  

(1)

where \( A_{EAT} \) is the EAT of the grain (°C*d), \( d \) is the storage time unit as day, \( n \) is the total days (d), \( T_0 \) is the threshold temperature (°C), \( T_d \) is the average temperature of the grain bulk in one day. Every living thing responds to changing temperature. In order to predict the matter on a quantitative basis with temperature, EAT has been used in seed germination, insect generation, crop growth, etc. The rising of FFA in grain is related to the temperature changes, thus it is reasonable to consider to build a relation between the FFA and EAT. When temperature is below 0°C, the metabolism of microorganisms stay at a low level. Therefore, \( T_0 \), the threshold temperature, is set as 0°C.

**CART and MT**

CART: The gist of tree algorithms is the extraction of meaningful subgroups characterized by homogeneous covariate values and common outcome. The traditional CART algorithm uses Shannon entropy to divide a set into binary. In this study, total variance was used to substitute for Shannon entropy. The CART is described as follows:

\[ D = \{ (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n) \} \]  

(2)

where \( D \) is a standard data set for regression problems, \( x \in \mathbb{R}^d \) is \( d \) dimensional eigenvector, \( y \in \mathbb{R} \) is a continuous random variable. If each attribute is viewed as a coordinate axis in the coordinate space, all \( D \) attributes will constitute a \( d \)-dimensional feature space and each \( x \) the eigenvector of \( d \)-dimension corresponds to a data point in the space. The objection of the CART is to divide the feature space into several subspace. And each subspace has a fixed output value, i.e., all input values in the same subspace would have a same output value.

Firstly, choosing a random attribute and traversing all possible values, then finding the optimal partition point of the attribute, which is named as \( v^* \), according to the following formula (3).
\[ v^* = \arg \min_v \left[ \min_{c_1} \sum_{x_i \in R_1(a,v)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(a,v)} (y_i - c_2)^2 \right] \]

where \( v^* \) is the optimal partition point of the attribute, \( R_1(a,v) = \{x|x \in D^a \leq v\} \), \( R_2(a,v) = \{x|x \in D^a > v\} \), \( y_i \) is the output value of data sample \( x_i \), \( c_1 \) and \( c_2 \) are the average output value of \( y_i \) corresponding to the data set of \( R_1(a,v) \) and \( R_2(a,v) \), respectively. \( c_1 \) and \( c_2 \) are calculated as follows:

\[ c_1 = \text{ave}(y_i|x \in R_1(a,v)) = \frac{1}{|R_1(a,v)|} \sum_{x_i \in R_1(a,v)} y_i \]

\[ c_2 = \text{ave}(y_i|x \in R_2(a,v)) = \frac{1}{|R_2(a,v)|} \sum_{x_i \in R_2(a,v)} y_i \]

Secondly, traversing all attributes to find the optimal partition attribute, and then dividing the feature space into two subspace according to the optimal partition point of the attribute, repeating the above steps for each subspace until the stop condition is met. In this way, a cart regression tree is generated. Assuming that the feature space is finally divided into subspace as R1, R2, ..., RM. The model formula of cart regression tree can be expressed as,

\[ f(x) = \sum_{m=1}^{M} c_m I(x \in R_m) \]

Similarly, \( c_m \) is the mean value of the output value, \( y_i \) corresponding to the sample \( x_i \) in the Set \( R_m \). The intuitive understanding of this formula is that for a given sample \( x_i \), first judge which subspace it belongs to, and then take the output value corresponding to the subspace it belongs to as the predicted value \( y_i \) of the sample.

MT: The main difference between CART and MT is the form of leaf node. The leaf node of the MT is set as piecewise linear function compared with the constant of the CART. In order to get the piecewise linear functions of the MT, some changes need to be made with the CART, e.g., allowable descent error and minimum number of samples for segmentation needed to be set. The method to get piecewise linear function of each leaf node is as follows:

\[ Y = X^T w \]

where \( X \) is the matrix of the input data, \( w \) is the regression coefficient vector. For certain input data, regression coefficient vector is calculated by finding the value which can minimize the error between the calculation of output and the real output value. The error is described as:

\[ \sum_{i=1}^{m} (y_i - x_i^T w)^2 \]

where \( y_i \) is the real output value, \( x_i \) is the input data, \( w \) is the regression coefficient vector. And matrix in the function of (7) can also be described as:

\[ (y - Xw)^T (y - Xw) \]

Calculating the derivation, then getting:

\[ X^T (Y - Xw) \]
while $X^T(Y - Xw) = 0$, we will get the result of the regression coefficient vector:

$$w = (X^TX)^{-1}X^Ty$$

(11)

where $w$ is the regression coefficient vector, $X$ is the matrix of input data, $y$ is the output value corresponding to the input data.

Performance Metrics: The performance of the model is commonly evaluated by the measures of deviation between the calculation and actual values. The measures used in this study are the mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination ($R^2$). They are shown as following:

$$MAE = \frac{1}{N} \sum_{n=1}^{N} |y_p(t) - y(t)|$$

(12)

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (y_p(t) - y(t))^2}$$

(13)

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_p(t) - y(t))^2}{\sum_{i=1}^{N} (y(t) - \bar{y})^2}$$

(14)

where $N$ is the total number of FFA value needed to be estimated, $y_p(t)$, $y(t)$ and $\bar{y}$ are the calculated value, true value, and the mean of all true values of the $i^{th}$ sample, respectively.

**Results**

In this section, the correlations between EAT and FFA were analyzed by Pearson simple correlation coefficient using BONC DSS Statistics(Version 1.0, BONC.2018). The EAT of stored maize in commercial storehouses was used to build the models of CART and MT. The MAE, MAPE, and RMSE were applied as the criteria to evaluate the performance. The building and evaluation of the models were carried out based on the Pycharm(Community Edition 2020.3, JetBrains Corp, 2020) combined with Anaconda (Individual Edition 5.3.1, Anaconda, Inc.2016).

**Correlation analysis**

Figure 2 shows the changes of FFA and EAT of maize stored in a granary which cited in Changchun City, Jilin Province. The data of the storage period started from Apr., 2018 to Sep., 2019. From the 2th to 7th month, the EAT has a significant rise from 160 to 2338.6 (°C·d). And the FFA value rose from 35.4 to 39.5 (mgKOH/100 g) during the period from 4th to 6th month. The similar FFA value of storage maize also had been obtained by other researchers. Gras, P.W. et al studied the difference between maize stored in unseal bags and seal bags. After 9 months storage, the storage maize FFA value cited in the bottom, middle, top of unsealed and sealed bags were 46.6, 56.4, 43.3; 18.6, 42.8, and 25.6 mg KOH/100 g, respectively. However, from 8th to 13th month, the EAT variance was not obvious, meanwhile, the FFA value had little changes. It means the changes of EAT, i.e, the temperature, have a significant influence on the FFA during storage.

The correlations between FFA value and EAT of stored maize were analyzed by Pearson simple correlation coefficient. It is shown in Table 1 that the EAT and FFA have a significant correlation at 0.01 level (2-tailed) up to 0.963. Therefore, it is reasonable to build a regression function between the EAT and FFA value.
Results of CART and MT

Temperature data of the stored maize are collected through the temperature monitoring system, which is shown in Figure 1, and data missing and lacking can happen accidentally. The Pauta criterion was applied to eliminate abnormal data and the linear interpolation method was used to complete the missing data, then using the data to get the EAT value. Eventually, 298 sets of data were obtained. To avoid the sample representativeness bias caused by systematic differences between the training set and the test set, 80% of the data was randomly selected to use as the training set and the rest 20% was used as the testing set. The FFA value used in the models was that the real value subtracted the initial value in order to eliminate the difference of the initial value. In fact, the models estimated the changes of FFA value of stored maize.

The allowable descent error and minimum number of samples for segmentation are the two major parameters determining the leave creation of the tree models. The allowable descent error was set as 1, it had been tested in the two models which changes had little influence. The minimum number of samples for segmentation setting need to be tested using the correlation coefficient, i.e., Pearson product-moment correlation coefficient. The relation between the correlation coefficient and MNSS value is shown in Figure 3. As the MNSS value rising, the correlation coefficient of the two models rises at first, and reaches the max value while MNSS is 20, then descends. Thus, the MNSS was setting as 20 and the two models could have the optimal performance.

How the MNSS value influences the performance of the two models is shown in Figures 4 and 5. MNSS is the minimum number of samples for segmentation, i.e, only the amount of samples reaches the MNSS value that a leaf node might be created in the tree models. This is a way of prune the tree

Table 1. The correlations between FFA value and EAT of stored maize.

| EAT (°C*d) | FFA Value (mgKOH/100 g) |
|-----------|-------------------------|
| EAT (°C*d) | 1                       |
| FFA Value (mgKOH/100 g) | 0.963**               |

**Correlation is significant at the 0.01 level (2-tailed).
Figure 3. Change in the correlation coefficient of two models with MNSS value.

Figure 4. Illustration of how changes of MNSS influence the model creating of CART. While MNSS was set as 10, 20, 30 and 40, the models were created by CART, respectively. Each blue point means the FFA value according to EAT, and the black line is the model created by CART.
models in order to minimize the influence of abnormal or irrelevant values. It can be seen that the higher MNSS value, the less functions would be created in the model, and the model would be simpler. According to Figure 3, the correlation coefficient of the two models had the optimal performance when MNSS was 20, below and above this value the correlation coefficient would descend. It can be interpreted as following: if the MNSS value is small, abnormal values might be involved in creating the model; if the MNSS is large, some helpful value would be neglected, then the performance of the model will be influenced by the MNSS value.

The Figures 6 and 7 show the outcomes of the two models, respectively, while MNSS is set as 20 in both. Firstly, traversing all possible values, then finding the optimal partition point, which is calculated according to equation (3). In the CART model, the first separation point value of EAT was 5195.487, then all values above the separation point value were stored in the leaf and the rest values were stored in the right leaf. The separation method was repeated in the two leaves, until the separation condition is not met as MNSS<20 or the allowable descent error above setting value. For example, in the second layer in Figure 6, the leaf was EAT = 7557.585, it could not be separated again, then the FFA value corresponding to the EAT values which above 7557.785 was 14.189 which stored in the left leaf node, and the FFA value corresponding to the EAT values which below 7557.785 but above 5195.487 was 9.908 which stored in the right leaf node. Finally, the CART model separation procedure was finished showing in Figure 6. The main difference in outcomes between CART model and MT was the leaf node. The leaf node of CART model was a number but the MT was a function.

**Figure 5.** Illustration of how changes of MNSS influence the model creating of MT. While MNSS was set as 10, 20, 30 and 40, the models were created by MT, respectively. Each blue point means the FFA value according to EAT, the black line is the model created by MT.
the FFA value according to EAT, the black line is the model created by MT.

Figure 6. The outcomes of model created by the CART illustrating in tree form, MNSS = 20.

Figure 7. The outcomes of model created by the MT illustrating in tree form, MNSS = 20.
Comparison results

Data (20%) was used to estimate the performance of the two models. The criterion including MAE, RMSE, and $R^2$ were computed, respectively. The smaller values of the MAE and RMSE, and the larger values of $R^2$, the better performance of the model. As shown in Table 2, the RMAE and $R^2$ of the CART were lower than the MT, but the MT had a smaller MSE value. The $R^2$ of the two models was 0.759 and 0.741, respectively. The two values were much closer. In general, the CART had a better performance than MT. However, the MT usually performs better than CART in creating model because the liner normally has a smaller difference in estimation than a number. This might be the data for building models is loosely referring to the coefficient correlations of the two were 0.871 and 0.861 when MNSS was setting as 20, respectively. In conclusion, the performances of the CART and MT were close, the CART was a little better.

Discussion

The maize is usually stored for 3–5 years in large granaries in China according to the nation grain security policy. The grain quality will descend over storage time. It is affected by the interactions of physical and biological factors during storage. The change of FFA content of grain is generally viewed as a key indicator of grain quality deterioration during storage, and it has a direct relation with the stored temperature. Temperature monitoring and controlling is an efficient way to grain quality management and widely used in large grain granary. During the winter, the cool air from the ambient is transported into the granary through fans. As shown in Figure 8, cool air goes through the input tube and ventilation cage to enter the grain, then the grain is cooled down. The ventilation cage opening rate usually is 25–35%. In summer, the temperature of grain in the top layer will rise quickly. The bottom of the grain bulk remains at low temperature. The axial fans transport the air from the bottom and inner of the bulk to the top to cool down the rising temperature through the ventilation cages and vertical tubes, shown in Figure 8. The temperature controlling manner is usually named as inner air circulation (IAC).

The thermal diffusivity of bulk grain is at a very low level according to ASABE(2008), showing as follows: corn, $10.22 \times 10^{-8}$ m$^2$/s; rice, $10.99 \times 10^{-8}$ m$^2$/s; and hard wheat, $11.5 \times 10^{-8}$ m$^2$/s. As the ambient temperature rises, the top and the sites close to the wall rise faster than the bottom and the inner of grain bulk, then the temperature gradient forms in the bulk. Taking a granary for example, the

| Model | MAE  | RMSE | $R^2$ |
|-------|------|------|-------|
| CART  | 1.296| 1.761| 0.759 |
| MT    | 1.247| 1.821| 0.741 |

Figure 8. Illustration of the ventilation system of the granary.
temperature gradient changes over time as shown in Figure 9. From Mar., 2017 to Aug., 2017, the temperature rose in the grain bulk as the proportions of the below 0°C and 0–15°C became smaller, and the high-temperature proportions including 15–25°C and above 25°C became larger. From Sept. to Feb. of the next year, the temperature of the grain bulk went down.

The temperature gradient provides the possibility for IAC, but it raises the uncertainty of the grain quality. In large granary, thermometric cables are widely used to monitor the temperature.\cite{40} For example, 60 thermometric cables that contain 420 sensors in total are normally placed in the granary which sizes is 42 m * 24 m storing the grain bulk which height is 6 m. However, the sensors system still cannot reflect all temperature condition of the granary due to its large scale and irregular self-heating points caused by complex biological conditions.\cite{41,42} Sampling in the granary for testing grain quality each time is little due to the sparsity of sample locations. Therefore, a more reliable grain bulk monitoring needs more sensors and sampling. However, sampling, testing, and arranging sensors consumes huge labor and substantial materials, which increasing the cost of storage management.

Mathematical models had been built between FFA and storage time in prior researches.\cite{43,44} But those models lack universality and might be unfeasible once the assumed conditions changed. In this study, ML methods, which had the ability to deal with big data, were used to develop a universal model for predicting the grain quality changes, i.e., the changes of FFA. The models had revealed the relationship between EAT and FFA based on data collected from 11 large granaries in two provinces. The correlation coefficients of the models were 0.871 and 0.861, respectively. Tree methods might be potentially useful, once the analytic goal is identifying strata with same covariate values and homogeneous multiple outcomes.\cite{35,45} In this study, the models performance had proven its ability in prediction and provided the possibility to obtain real-time grain quality based on the existing temperature monitoring system, which would play a beneficial role in reducing losses both in grain quality and economic.

**Conclusion**

The models of CART and MT were built using EAT to estimate the stored maize FFA changes, a key indicator of grain quality, based on data collected from large granaries sited in Heilongjiang and Jilin Province of China. It was analyzed that the correlation between EAT and FFA was 0.963. While MNSS was setting as 20, the two models could have the optimal performance corresponding to correlation coefficients was 0.871 and 0.861, respectively. The outcomes of the two models both were showed in the tree form. The MAE, MASE, and R\(^2\) were used to assess the performance of CART and MT, and the results of MAE, MASE, and R\(^2\) are shown as follows: 1.296, 1.761, 0.759; 1.247, 1.821, and 0.741, respectively. As CART had the lower RMSE and the larger R\(^2\), the
CART performed better. In large commercial granaries, temperatures generally ranged from $-10^\circ C$ to $30^\circ C$ during one year. Temperature changes result in a temperature gradient between grain layers, then different grain quality is formed. It is tough to obtain thorough grain quality by sampling. The model has provided a method to obtain FFA as the real-time grain quality indicator based on the exit temperature monitoring system. It would benefit from improving grain storage management to get quality information with less labor consuming. In further study, with the use of humidity sensors, CO$_2$ sensors, and other sensors in large grain granaries, more factors such as humidity, and CO$_2$ would be taken into consideration to improve the models. In this study, the data was collected from grain granaries in northeastern China. Substantial data need to be collected and analyzed before the models could be applied for more areas.

**Abbreviations**

The following abbreviations are used in this manuscript:
- FFA: Free Fatty Acids
- EAT: Effective Accumulated Temperature
- ML: Machine Learning
- MC: Moisture Content
- CART: Classification and Regression Trees
- MT: Model Trees
- MNSS: Minimum Number of Samples for Segmentation
- MAE: Mean Absolute Error
- RMSE: Root Mean Square Error
- $R^2$: Coefficient of Determination
- IAC: Inner Air Circulation

**Acknowledgments**

Thanks to all the grain reserve companies for providing the grain storage data in this research.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This research was supported by the National Science Foundation for Young Scientists of China, grant number 32102034.

**Data Availability Statement**

Data sharing is not applicable.

**References**

[1] CIMMYT. Maize - Global Alliance for Improving Food Security and the Livelihoods of the resource-poor in the Developing World: Proposal Submitted by CIMMYT and IITA to the CGIAR Consortium Board. Mem. Fac. Eng. Kyushu Univ. 2011, 14(5), 386–391.
[2] Guo, X. T.; Duan, X. G.; Wu, Y. Z.; Cheng, J. S.; Zhang, J.; Zhang, H. X.; Li, B. Genetic Engineering of Maize (Zea Mays L.) With Improved Grain Nutrients. J. Agric. Food Chem. 2018, 66(7), 1670–1677. DOI: 10.1021/acs.jafc.7b05390.
[3] Nation Bureau of Statistics of the P.R.C. https://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A0D0F&sj=2020 (accessed 10, October, 2021)
[4] Luo, Y.; Huang, D.; Miao, H. M.; Wu, L. P.; Zhu, J. F. Impact of Advanced Storage Facilities on Households’ Maize Storage Losses and Food Security in China. Environ. Dev. Sustain. 2021. DOI: 10.1007/s10668-021-01406-z.
[5] Sun, D. W.; Woods, J. L. Low Temperature Moisture Transfer Characteristics of Wheat in Thin Layers. Trans. ASAE. 1994, 37(6), 1919–1926. DOI: 10.10301/2013.28284.

[6] Barreto, A. A.; Abalone, R.; Gaston, A.; Bartosik, R. Analysis of Storage Conditions of a Wheat silo-bag for Different Weather Conditions by Computer Simulation. Biosyst. Eng. 2013, 116(4), 497–508. DOI: 10.1016/j.biosystemseng.2013.10.012.

[7] Gras, P. W.; Bason, M. L.; Esteves, L. A.; Sabio, G. C.; Annis, P. C.; Graver, J. Quality Changes in Maize Stored in Sealed Bag Stacks. J. Stored Prod. Res. 1990, 26(4), 199–206. DOI: 10.1006/jspr.1990.9022.

[8] Feussner, I.; Wasternack, C. The Lipoxigenase Pathway. Annu. Rev. Plant Biol. 2002, 53, 275–297. DOI: 10.1146/annurev.arplant.53.100301.135248.

[9] Stewart, R. R.; Bewley, J. D. Lipid Peroxidation Associated with Accelerated Aging of Soybean Axles. Plant Physiol. 1980, 65(2), 245–248. DOI: 10.1104/pp.65.2.245.

[10] Salman, H.; Copeland, L. Effect of Storage on Fat Acidity and Pasting Characteristics of Wheat Flour. Cereal Chem. 2007, 84(6), 600–606. DOI: 10.1094/CCH-84-6-0600.

[11] American Association of Cereal Chemists. Approved Methods of the AACC, 10th ed.; American Association of Cereal Chemistry: Minnesota, 2020.

[12] Yasumatsu, K.; Moritaka, S. Fatty Acid Compositions of Rice Lipid and Their Changes during Storage. J. Agric. Chem. Soc. Jpn. 2008, 28(5), 257–264.

[13] Igne, B.; Rippke, G. R.; Hurburgh, C. R. Measurement of Whole Soybean Fatty Acids by near Infrared Spectroscopy. J. Am. Oil Chem. Soc. 2008, 85(12), 1105–1113. DOI: 10.1007/s11746-008-1311-1.

[14] Mazurek, B.; Chmiel, M.; Gorecka, B. Fatty Acids Analysis Using Gas Chromatography-Mass Spectrometer Detector (GC/MSD) - Method Validation Based on Berry Seed Extract Samples. Food Anal. Methods. 2017, 10(8), 2868–2880. DOI: 10.1007/s12161-017-0834-1.

[15] Jiang, H.; Liu, T.; He, P. H., and Chen, Q. S. Quantitative Analysis of Fatty Acid Value during Rice Storage Based on Olfactory Visualization Sensor Technology. Sensors and Actuators B-Chemical. 2020, 309, 127816 doi: 10.1016/j.snb.2020.127816.

[16] Manandhar, A.; Milind, P.; Shah, A. An Overview of the Post-Harvest Grain Storage Practices of Smallholder Farmers in Developing Countries. Agriculture-Basel. 2018, 8, 4.

[17] Gbenou, E. S.; Adegbola, Y. P.; Hessavi, P. M.; Zossou, S. R. C.; Biaou, G. On-Farm Assessment of Maize Storage and Conservation Technologies in the Central and Northern Republic of Benin. Agriculture-Basel. 2021, 11, 1.

[18] Dowell, F. E.; Dowell, C. N. Reducing Grain Storage Losses in Developing Countries. Qual. Assur. Saf. Crops Food. 2017, 9(1), 93–100. DOI: 10.3920/QAS2016.0897.

[19] Gonzales, H. B.; Armstrong, P. R.; Maghirang, R. G. Simultaneous Monitoring of Stored Grain with Relative Humidity, Temperature, and Carbon Dioxide Sensors. Appl. Eng. Agric. 2009, 25(4), 595–604. DOI: 10.13031/2013.27466.

[20] Suleiman, R. A.; Rosentrater, K. A. Measured and Predicted Temperature of Maize Grain (Zea Mays L.) Under Hermetic Storage Conditions. Acad. J. 2016, 7(1), 1–10.

[21] Jian, F.; Jayas, D. S.; White, N. D. G. Heat Production of Stored Canola Seeds under Airtight and non-airtight Conditions. Trans. ASABE. 2014, 57(4), 1151–1162.

[22] Li, X. J.; Han, Z. Q.; Lin, Q.; Wu, Z. D.; Chen, L., and Zhang, Q. Smart cooling-aeration Guided by Aeration Window Model for Paddy Stored in Concrete Silos in a Depot of Guangzhou, China. Comput. Electron. Agric. 2020, 173, 105452. doi: 10.1016/j.compag.2020.105452.

[23] Manyik, J.; Chui, M.; Bisson, P.; Woetzel, J.; Dobbs, R.; Bughin, J., and Aharon, D. The Internet of Things: Mapping the Value beyond the Hype. McKinsey Global Institute. 7, 144 . 2015.

[24] Shen, Y.; Zhao, C. J.; Li, B.; Li, G. L.; Yin, Y. X.; Pang, B. S. Determination of Wheat Moisture Using Terahertz Spectroscopy Combined with the Tabu Search Algorithm. Anal. Methods. 2021, 13(36), 4120–4130. DOI: 10.1039/D1AY00812A.

[25] Zeymer, J. S.; Guzzo, F.; de Araujo, M. E. V.; Gates, R. S.; Correa, P. C.; Vidigal, M.; Neisse, A. C. Machine Learning Algorithms to Predict the Dry Matter Loss of Stored Soybean Grains (Glycine Max L.). J. Food Process Eng. 2021, 44(10). DOI: 10.1111/jfpe.13820.

[26] Cui, H. W.; Zhang, Q.; Zhang, J. S.; Wu, Z. D., and Wu, W. F. Classification of Grain Storage Inventory Modes Based on Temperature Contour Map of Grain Bulk Using Back Propagation Neural Network. Agriculture-Basel. 2021, 11(5), 451.

[27] Niyaboko, T.; Mvumi, B. M.; Stathers, T.; Mlambo, S., and Mubayiwa, M. Predicting Prostephanus truncatus (Horn) (Coleoptera: Bostrichidae) Populations and Associated Grain Damage in Smallholder Farmers’ Maize Stores: A Machine Learning Approach. J. Stored Prod. Res. 87(3), 2020, 101592.

[28] Driscoll, R.; Longstaff, B. C.; Beckett, S. Prediction of Insect Populations in Grain Storage. J. Stored Prod. Res. 2001, 37(4), 399.

[29] Jafio, B. A., and Kwakkel, J. H. A Novel Concurrent Approach for Multiclass Scenario Discovery Using Multivariate Regression Trees: Exploring Spatial Inequality Patterns in the Vietnam Mekong Delta under Uncertainty. Environ. Modell. Software. 145(11), 2021, 105177.
[30] Johns, H.; Bernhardt, J.; Churilov, L. Distance-based Classification and Regression Trees for the Analysis of Complex Predictors in Health and Medical Research. Stat. Methods Med. Res. 2021, 30(9), 2085–2104. DOI: 10.1177/09622802211032712.

[31] ASAE. Moisture Measurement - Unground Grain and Seeds, ASAE Standards ASAE S352.2 APR1988 (R2017); ASAE. St: Joseph, MI, 2017.

[32] Tyldesley, J. B.; A Method of Evaluating the Effect of Temperature on an Organism When the Response Is non-linear. Agric. Meteorol. 1978, 19(2–3), 137–153. DOI: 10.1016/0002-1571(78)90005-5.

[33] Milonas, P. G.; Savopoulou-Soultani, M.; Stavridis, D. G Day-degree Models for Predicting the Generation Time and Flight Activity of Local Populations of Lobesia Botrana (Den. & Schiff.) (Lep., Tortricidae) in Greece. J. Appl. Entomol. Zeitschrift für Angew. Entomol. 2001, 125(9–10), 515–518.

[34] Perez-de-Luque, A.; Flores, F.; Rubiales, D. Differences in Crenate Broomrape Parasitism Dynamics on Three Legume Crops Using a Thermal Time Model. Front. Plant Sci. 2016, 7. DOI: 10.3389/fpls.2016.01910.

[35] Segal, M. R.; Tree-Structured Methods for Longitudinal Data. J. Am. Stat. Assoc. 1992, 87(418), 407–418. DOI: 10.1080/01621459.1992.10475220.

[36] Liu, H.; Cocea, M. Semi-random Partitioning of Data into Training and Test Sets in Granular Computing Context. Granul. Comput. 2017, 2(4), 357–386. DOI: 10.1007/s41066-017-0049-2.

[37] Wu, Z. D.; Zhang, Q.; Yin, J.; Wang, X. M.; Zhang, Z. J.; Wu, W. F.; Li, F. J. Interactions of Multiple Biological Fields in Stored Grain Ecosystems. Sci. Rep. 2020, 10, 1. DOI: 10.1038/s41598-019-56847-4.

[38] Burges, H. D.; Burrell, N. J. Cooling Bulk Grain in the British Climate to Control Storage Insects and to Improve Keeping Quality. J. Sci. Food Agric. 1964, 15(1), 32–50. DOI: 10.1002/jsfa.2740150106.

[39] ASAE. Thermal Properties of Grain and Grain Products, ASAE standards D243.3 May 2003 (R2008). ASAE. St. Joseph, MI, 2008.

[40] Wu, W. F.; Cui, H. W.; Han, F.; Liu, Z.; Wu, X. M.; Wu, Z. D.; Zhang, Q. Digital Monitoring of Grain Conditions in large-scale Bulk Storage Facilities Based on Spatiotemporal Distributions of Grain Temperature. Biosyst. Eng. 2021, 210, 247–260. DOI: 10.1016/jbiosystemseng.2021.08.028.

[41] Ivova, L. S.; Sosedov, N. I.; Gerel, U.; Shvartsman, M. I.; Shatilova, T. I. Relationship between Aflatoxin Formation in Wheat Grain and self-heating and Changes in the Chemical Composition of Grain during the Growth of “Storage Molds.” Prikladnaia Biokhimiiia I Mikrobiologii 1976, 12(5), 741–749.

[42] Wang, Q. Y.; Han, F.; Wu, Z. D.; Lan, T. Y.; Wu, W. F. Estimation of Free Fatty Acids in Stored Paddy Rice Using Multiple-Kernel Support Vector Regression. Appl Sci-Basel. 2020, 10, 18.

[43] de Alencar, E. R.; Faroni, L. R. D.; Peternelli, L. A.; da Silva, M. T. C.; Costa, A. R. Influence of Soybean Storage Conditions on Crude Oil Quality. Rev. Bras. de Eng. Agricola E Ambient. 2010, 14(3), 303–308. DOI: 10.1590/S1415-43662010000300010.

[44] Junka, N.; Rattanamechaikul, C.; Wongs-Aree, C. Free Fatty Acid Deformation of Treated Black Glutinous Rice during Storage by Fluidization Drying. J. Food Process Eng. 2017, 40, e12427. DOI: 10.1111/jfpe.12427.

[45] Zhang, H.; Classification Trees for Multiple Binary Responses. J. Am. Stat. Assoc. 1998, 93(441), 180–193. DOI: 10.1080/01621459.1998.10474100.