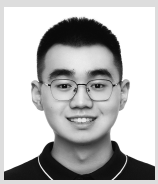
Temperature Interpolation of Grain in Storage: A Hybrid Deep Learning Approach Based on CNN and Attention Mechanism



Zhongke Qu[1], Ke Yang[2], Zhaolin Gu[3], Yue Li[4], Xuemei Jiang[4]
1 Doctoral student, School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, China
2 Master student, School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, China
3 Professor, School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, China
4 Engineer, Sinograin Chengdu Storage Research Institute Co. Ltd., China

穀物倉庫の温度管理が悪いと、せっかく収穫した穀物が無駄になる。中国では収穫後の穀物ロスで総生産量の5％が失われている。これを防ぐために深層学習のアルゴリズムを用いて温度管理を最適化する方法を探った。

Abstract

Grain storage temperature is a crucial indicator for food safety. Acquiring high-precision temperature distribution data within the grain bin holds significant importance for grain situation analysis and decision-making. However, constrained by practical constraints and cost considerations, the sensor network deployed in the granary is insufficient to fully reflect the temperature distribution since the temperature values in areas without sensors remain undetermined. To address this issue, this study has designed an algorithm specifically for interpolating the temperature field in granaries, combining convolutional neural network (CNN), attention mechanism, and Multilayer Perceptron (MLP). CNN is utilized to capture local spatial features in temperature data, the attention mechanism is employed to focus on key and sensitive temperature areas of the granary, and MLP is used for deep feature fusion. In a granary in Shaanxi, China, the performance of the proposed method was evaluated after comparison with five machine learning and deep learning models. The results indicate that the proposed method significantly outperforms other comparative models, showcasing superior performance. This study successfully addresses the challenge of acquiring temperature distribution data inside granaries, providing an effective and feasible solution to enhance grain storage safety and reduce unnecessary losses.

Keywords grain storage, grain security, post-harvest loss, deep learning, temperature interpolation

Background

By 2050, the global population is expected to reach 9.1 billion, requiring approximately 70% additional food production to sustain the entire human race[1].In recent years, against the backdrop of the Russian-Ukrainian war, the COVID-19 pandemic, and extreme weather events, the increasing demand for food production has raised concerns, with food insecurity rates rising after decades of decline[1].

Over the past few decades, most countries have focused on increasing food production as a policy to address the continuously growing food demand. However, post-harvest loss (PHL) is a critical issue that has not received the necessary attention, with less than 5% of research funding allocated to this issue in recent years. A survey of 1,890 grain merchants in 54 regions across 9 provinces in China shows that grain storage conditions have the greatest impact on post-harvest grain loss[2]. 5% of the country's total grain output is due to losses during post-harvest storage[3]. Based on the grain output over 2023, the loss during the storage phase reached 34.7 billion kilograms, equivalent to the output of 43,500 km2 of high-yield farmland, enough to feed 860 million people for a year. Therefore, reducing food losses during the storage phase is one of the most practical and effective ways to ensure China's food security.

Grain temperature is one of the important indicators for monitoring the storage status. Ensuring accurate

measurement of the temperature is foundational to maintaining grain safety[4]. Temperature can influence the prevalence of pests and microorganisms within grain storage. If even a minor heating event occurs within the grain and isn't addressed promptly, it can lead to drastic humidity and temperature fluctuations inside the storage, resulting in grain condensation, mold formation, and subsequently, substantial grain losses[5].

At present, grain storages across China have implemented arrays of temperature sensors within the silos, enabling real-time monitoring of internal grain temperatures. Due to practical constraints and setup costs, only a select number of sensors can be placed. This doesn't provide a comprehensive view of the temperature distribution within the silo, and monitoring is mainly limited to the areas equipped with sensors. Areas without sensors could be critical zones where grains undergo respiration and heat generation. Grain is a poor conductor of heat, leading to slow heat transfer within the grain pile. If certain sections experience temperature anomalies due to rapid pest infestations or mold growth, by the time adjacent temperature sensors detect these changes, it could potentially be too late. Therefore, it's essential to utilize spatial interpolation methods to address the gaps in data from sampling points. By establishing a robust interpolation system, we can identify areas with temperature anomalies, implement preventive measures, and aim to minimize grain losses while guaranteeing the quality and safety of the stored grain.

This study delves deep into the characteristics of grain temperature spatiotemporal sequence data, such as high dimensionality, non-linearity, strong spatiotemporal correlation, and dynamic changes, establishing a valid and efficient grain temperature interpolation model. In the realm of grain storage, our study pioneers an in-depth exploration and introduces a grain temperature interpolation method based on CNN, Attention, and MLP. This method excels in processing non-linear spatial data and doesn't have particular prerequisites for sample data distribution and assumptions. Through convolutional operations, CNNs are able to discern the relationships between proximate temperature points in the granary, leading to the creation of feature maps with localized spatial relevance. These feature maps can effectively identify key spatial patterns such as temperature fluctuations, local hotspots, or cold spots. Multiple such feature maps are stacked and combined to form more complex feature representations. Through several hidden layers, MLP is designed to understand the intricate functional relationships among various features. This multi-layer learning strategy allows the model to capture more complex and abstract information from the extracted features, such as the global distribution of the temperature field or long-term trends in temperature changes. In this model, the role of the attention mechanism is particularly evident. Through weight allocation, it can automatically identify which areas (or features) are more critical during interpolation[6]. For example, when a region exhibits abrupt temperature spikes or dips, the attention mechanism intensifies its concentration on that specific area, ensuring precise interpolation outcomes. Validations have shown that this method enhances the precision of spatial interpolation techniques.

Study area and data source

Hanzhong City is located at the geographical center of China's territory, belonging to the subtropical climate zone. The Qinling Mountains to the north serve as a shield, preventing cold fronts from easily entering, resulting in a mild and moist climate. In the ecology areas of stored grain in China, it belongs to the fourth category: North of China medium hot and dry ecology area of stored grain, which makes it an advantageous location for grain silos, as depicted in Fig. 1. The left image shows the seven ecology areas of stored grain in China, while the right image displays the latitude and longitude coordinates of Hanzhong City, Shaanxi Province.

As the predominant staple food crop worldwide, rice sustains more individuals than any other crop. Forecasts from the International Rice Research Institute (IRRI) project a 50% escalation in global rice demand by 2050[7]. To address this, the present study employs rice temperature data from the Hanzhong No. 25 granary,

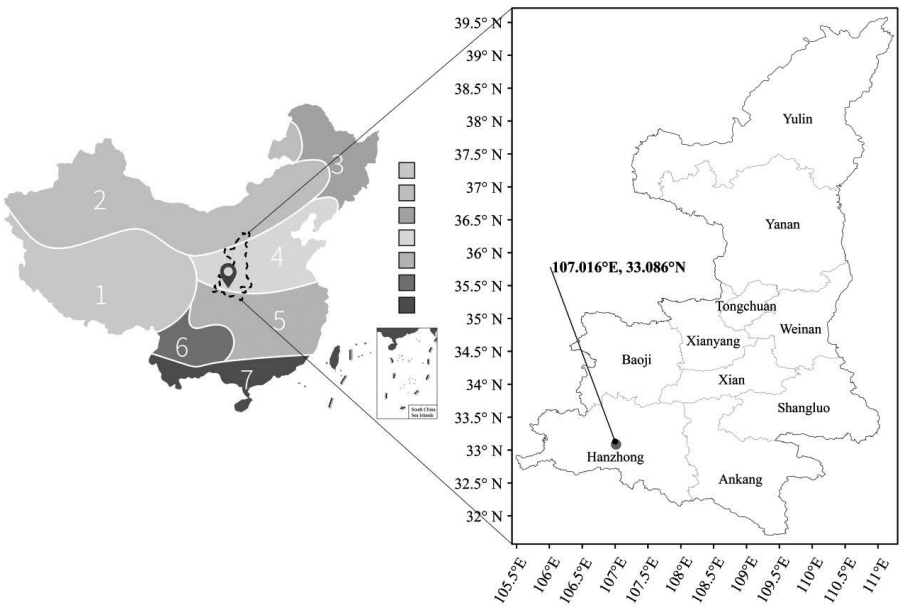


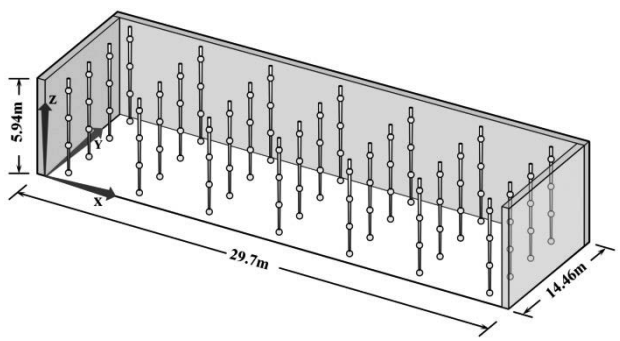
Fig. 1. Granary location selection

managed by China Grain Reserves Group Company.

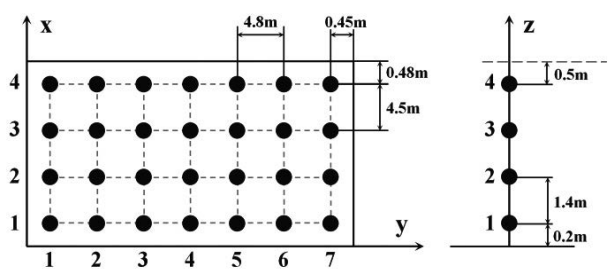
The Hanzhong silo, measuring 29.7 meters in length, 14.46 meters in width, and with a grain surface height of 4.9 meters, is a common type of tall flat granary found in China. The surroundings outside the granary and the internal grain storage environment are shown in Fig. 2. and Fig. 3. respectively. The sensor network is regularized arranged across four layers inside the granary, with 28 sensors on each layer, totaling 112 sensors for the entire bin. Four rows of temperature measurement cables are arranged from south to north with a longitudinal distance of 4.5m between them. Additionally, seven columns are placed from east to west with a lateral spacing of 4.8m. Vertically, there are 4 temperature measuring points, with the topmost point being 0.5m from the grain surface, the bottommost point 0.2m from the warehouse floor, and a distance of 1.4m between each measuring point. The sensor network layout is shown in Fig. 4. Given that grain are poor conductors of heat, the transfer of heat in and out is relatively slow[8], leading to data collection once a week. The optimal sampling time is set at 10 a.m. daily, as measurements taken at this moment minimize deviations caused by environmental temperature fluctuations, especially since, by 10 a.m. in Hanzhong, the temperature tends to stabilize, with a diminished influence from the cold air of the night or the warmth of the daytime, ensuring a more reliable temperature reading. A total of 15,120 temperature data points, corresponding to 135 representative days, were compiled following the exclusion of erroneous records and the imputation of missing values. In chronological order, the first 90% of the data will be used for model training and the remaining 10% will be used for model testing.

Fig. 2. The external environment of the granary

Fig. 3. The internal storage environment of the granary



(a) Three-dimensional view of sensors layout



(b) Two-dimensional view of sensors layout

Fig. 4. Sensor network layout in the granary

Proposed method

This section delineates the architecture of the proposed grain temperature interpolation method and its primary components. For improved feature extraction and prediction performance, we amalgamate Convolutional Neural Networks (CNN), Attention Mechanism, and multi-layer perceptron (MLP) into a unified framework, introducing a novel hybrid grain temperature interpolation method termed CAMNN. Figure 5 depicts the proposed method comprising 4 basic blocks: Data collection Block, CNN Block, Attention Block and MLP Block. In the following, each block of the proposed method is described in detail:

(1) Data collection block

Through the grain information management system, we collected temperature readings from 112 sensors for 135 representative days inside the granary. Moreover, we gathered other feature data that might influence the temperature: sensor bin temperature, bin temperature, external temperature, external humidity, and the whole bin average temperature. Due to issues like sensor damage and aging, as well as connection problems between sensors and receiving devices, there might be anomalies and zero values in the data, necessitating further data cleaning. The nulls and outliers above 45˚C or below -20˚C will be replaced with the average temperature of the layer. After data cleaning, the original temperature data and feature data were transformed from two-dimensional array data into tensor format, facilitating the subsequent use of automatic differentiation (Autograd) and optimizer (Optimizer) for model training and evaluation. The inputs for the model are x, y, z, $f_1, f_2, \ldots f_n$, where x, y, z are the coordinates of the temperature sensor, and $f_1, f_2, f_3 \ldots f_n$ are external features respectively. The naming convention for temperature measurement points is: (row index, column index, vertical layer index). For instance, (2,4,3) indicates the 2nd row, 4th column, and the 3rd temperature measurement point from the bottom of the grain. The primary goal is to create a correlated mapping between the sensor locations, external characteristics, and sensor temperature readings, represented as $T = f(x, y, z, f_1, f_2, \ldots f_n)$.

(2) CNN block

After preprocessing the collected data, it is input into the model. The input data matrix is X, $X \in \mathbb{R}^{b \times c \times s}$, where b, c and s respectively represent the number of data samples (batch size) captured in one training, the number of features (channels), and the length of sequence data (sequence length). The values of batch size and channels in this model will be discussed in detail in subsequent chapters. As the data originates from various sensors at a singular point in time, the sequence length

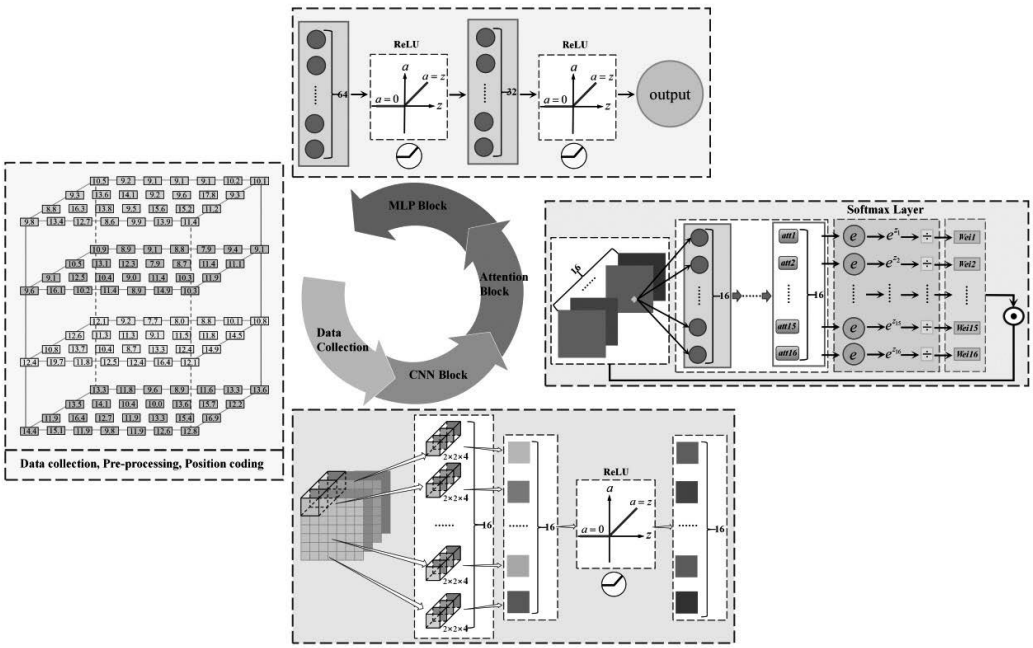


Fig. 5. Illustrative architecture of CAMNN.

is set to 1. Initially, the input data goes through the convolutional layer. This layer aims to execute convolution operations on every input channel, capturing the local dependencies among various features. By sliding the convolutional kernel across the input data, CNN can proficiently recognize and capture local features from temperature information, which is especially important for spatial data like temperature fields, as local temperature distributions are often influenced by micro-environmental factors inside the granary, such as airflow and humidity[9]. The mathematical representation of the convolution operation is:

$$O_{\text{conv}}[i,j] = \sum_{k=1}^{KS} X[i, j-k] \times W_{\text{conv}}[k] + b_{\text{conv}} \quad (1)$$

where O_{conv} denotes the output of the convolution; X is the input matrix; i indicates the batch index; j denotes the position in the sequence; W_{conv} is the weight of the convolution layer; b_{conv} is the bias of the convolution layer; KS indicates the number of Kernel Size. The size of each output feature map $oMapN$ for each convolutional layer and the number of trainable parameters for each convolutional layer satisfies the following relationship:

$$oMapN = \left(\frac{(iMapN - CWindow)}{\text{CInterval}} + 1 \right) \quad (2)$$

$$CParams = (iMap \times CWindow + 1) \times oMap \quad (3)$$

where $iMap$ represents the size of each input feature map; $CWindow$ is the size of the convolutional kernel; $CInterval$ indicates the stride of the convolutional kernel in the preceding layer, with a value of 1 in this model; $oMap$ signifies the count of output feature maps in each convolutional layer; $iMap$ denotes the number of input feature maps, with 1 indicating the bias that's shared within a single output map. Assume in the convolutional layer, the output value of the k-th neuron in output feature map n is x_{nk}^{out}, and x_{mh}^{in} represents the output value of the h-th neuron in the input feature map m, then the aforementioned meets the subsequent equation:

$$x_{nk}^{\text{out}^{\circ}} = f_{\text{cov}}(x_{1h}^{\text{in}^{\circ}} \times w_{1(h)n(k)} + x_{1(h+1)}^{\text{in}^{\circ}} \times w_{1(h+1)n(k)} + x_{1(h+2)}^{\text{in}^{\circ}} \times w_{1(h+2)n(k)} + \cdots + b_n) \quad (4)$$

where f_{cov} denotes the activation function used in the convolution operation. $w_{1(h)n(k)}$ represents the weight associated with the h-th neuron on the input feature map m and the k-th neuron on the output feature n map . To facilitate further processing and computation, superfluous dimensions are eliminated using the squeeze operation. The dimensionality of the data is changed from Batch Size × Channels × 1 to Batch Size × Channels:

$$O_{\text{squeezed}}[i,j] = O_{\text{conv}}[i,j,1] \tag{5}$$

Only the first value from the third dimension is chosen. As the convolution yields an output of Batch Size × hidden_channels × 1, redundancy is observed in the third dimension. There might be non-linear relationships in the sensor data. For instance, the response of certain sensors might depend on specific thresholds or ranges of other sensors. Owing to its non-linear characteristics, ReLU can assist the CNN model in grasping and understanding these non-linear relations. As a result, the ReLU activation function was employed to capture intricate features and patterns.

$$O_{\text{relu}}[i,j] = \text{ReLU}\big(O_{\text{squeezed}}[i,j]\big) \tag{6}$$

$$\text{ReLU}(x) = \max(0,x) \tag{7}$$

(3) Attention block

The attention mechanism serves to pick out the most pertinent parts for interpolation from the various local features extracted by the CNN. This allows the model to dynamically shift its focus depending on the intricacy of the input, rather than giving equal importance to all features. The primary objective of the attention module is to assign a weight to the data from each sensor, indicating its significance to the interpolation goal. Thus, we can weight the information of each sensor based on their importance, resulting in a comprehensive feature representation. Initially, a parameterized linear function, f, is employed to transform O_{relu}, computing the raw attention scores for every input data. In this context, f represents a fully connected layer.

$$s(q,k) = A = tanh\big(f(O_{\text{relu}})\big) = tanh(O_{\text{relu}}W_a + b_a) \tag{8}$$

where $W_a \in \mathbb{R}^{c \times d_a}$ and $b_a \in \mathbb{R}^{d_a}$. d_a is a hyperparameter that represents the dimension of the attention weights. Next, the tanh function is used to add non-linearity to Z, resulting in the final attention scores $s(q,k)$. To amplify the differences between the weights and ensure that the sum of the weights is 1, the softmax function is applied to normalize the attention scores $s(q,k)$:

$$s(q,k)' = \frac{\exp\left(s(q,k)_i\right)}{\sum_j \exp\left(s(q,k)_j\right)} \tag{9}$$

Through calculating weights, the attention mechanism identifies the input sections the model should "concentrate" on. By multiplying the normalized attention weights with the feature maps from the CNN output, a weighted feature representation is achieved:

$$O_{\text{att}} = O_{\text{relu}} \odot s(q,k)' \tag{10}$$

Where $\odot$ represents the Hadamard product (element-wise multiplication). In order to mitigate the overfitting risk associated with a substantial increase in input dimensions, the post-weighted features were consolidated into a comprehensive representation:

$$O_{\text{int}} = \sum_{i=1}^{c} O_{\text{att}\,i} \tag{11}$$

where $O_{\text{att}\,i}$ denotes the weighted output for the i^{th} feature. The calculated weights are applied to the corresponding inputs, and through weighted summation, the model can focus more on those input parts with higher weights. In this block, attention scores are computed using a linear layer, bypassing the use of encoding and decoding structures. The structures of encoding and dec+oding primarily originate from the sequence-to-sequence (Seq2Seq) model, which typically performs better in tasks that require capturing long-distance dependencies, such as machine translation or text summarization[10]. In this particular task, the emphasis lies on the data's spatial configuration and neighboring

relations (considering granaries as suboptimal heat conductors). The process of encoding and decoding could potentially add unwarranted intricacy. The linear layer we used can directly compute weights for input data, allowing the model to focus more directly on input parts related to interpolation. Moreover, with fewer parameters, it's less prone to overfitting.

(4) MLP block

The MLP serves as the model's concluding module, tasked with amalgamating data from the preceding two layers to execute the ultimate spatial interpolation. It merges the local features filtered out by the attention mechanism into a global description, and then interpolates the temperature field inside the granary[11]. In the input layer, the MLP receives weighted features from the attention layer. These features encompass temperature information from various sensor locations within the granary, as well as the significance highlighted by the attention mechanism.

$$I_{mlp} = O_{\text{int}} \tag{12}$$

Hidden layers can capture complex patterns and nonlinear relationships in the temperature field, aiding in more accurate interpolation of positions without sensor placement. Each neuron in the hidden layer calculates the sum of weighted inputs and implements a nonlinear transformation through an activation function. Denote L as the count of hidden layers. The description for the l^{th} hidden layer is:

$$H_l = \sigma\left(I_{mlp_l} W_{h_l} + b_{h_l}\right) \tag{13}$$

where I_{mlpl} is the input for the l^{th} hidden layer. For the first layer, $I_{mlp1} = I_{mlp}$, and for other layers, $I_{mlp_l} = H_{l-1}$. W_{h_l} and b_{h_l} are the weight and bias of the l^{th} hidden layer, respectively. σ is the ReLU activation function. Finally, the ultimate interpolation is executed through the output layer. W_o and b_o represent the weight and bias of the output layer, respectively.

$$O_{\text{mlp}} = H_L W_o + b_o \tag{14}$$

Initially, CNN is employed to extract local spatial features. Subsequently, these features are processed using the attention mechanism, facilitating the judicious distribution of computational resources based on sensor readings, particularly when computational capabilities are constrained. This ensures heightened focus on significant regions. The MLP integrates all information by capturing the more complex interdependencies between features. Additionally, based on limited sensor data, it can accurately interpolate the temperature at any position inside the granary, enabling monitoring of the entire temperature field of the granary.

Experimental setup

To build the deep neural network models, all experiments are conducted via Keras and Pytorch. The model training and fitting are conducted on a workstation with a configuration of NVIDIA GeForce RTX3080 and Intel(R) Xeon(R) Gold 6253CL CPU @ 3.10GHz. The Pycharm is used as the development tool. Python 3.5 is used as the program language. Pytorch is used as the neural network learning framework. In experiment, Adam serves as the optimizer to optimize the neural network parameters, with an initial learning rate set to 0.1. Mean square error (MSE) is utilized as a loss function that can be back-propagated to update the weights and biases. The batch size is 20, and the number of training epochs is 500. In the experiments, packages such as Keras, numpy, pandas, matplotlib and so on will be used. This study employs ten-fold cross-validation for model training and validation. The dataset is divided into ten equal parts, with nine parts serving as the training set and one part as the validation set. The training set is utilized for model training, parameter fitting, and determining spatial neural network weight coefficients. The validation set does not participate in the training process and is used to evaluate the model's generalization capability. Through cross-validation, results from training on varied groupings are averaged, mitigating the sensitivity related to data segmentation.

Evaluation metrics

In this research, three evaluation metrics—root-mean-square error (RMSE), mean absolute error (MAE), and R-squared (R^2)—are utilized to assess the

performance of the proposed model and to equitably compare it with alternative approaches[12]. The formulas for RMSE, MAE, and R^2 are presented below:

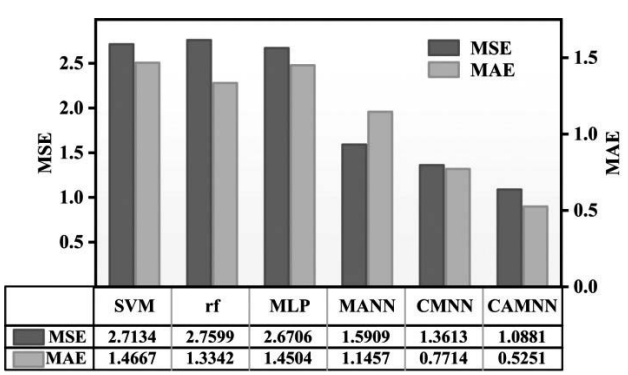
$$MSE = \frac{1}{n}\sum_{i=1}^{n}(t_i - \hat{t}_i)^2 \quad (15)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|t_i - \hat{t}_i| \quad (16)$$

In the aforementioned formulas, n denotes the number of interpolation data points; the interpolated value is $\hat{t}_i$; the actual value is t_i, and $\bar{t}_i$ is the average of actual value. RMSE and MAE are prevalent evaluation metrics for regression problems in this research. Both RMSE and MAE serve as indicators of model performance, with their values ranging from zero (indicative of optimal performance) to bigger.

Comparing between of different models

Comparison experiments were conducted to analyse the effects of various models including support vector machine (SVM), random forest (RF), multilayer perceptron (MLP), MANN (MLP-Attention) and CMAA (CNN-MLP) in the point temperature interpolation. Fig.6 illustrate the comparison of the proposed method and other models.



	SVM	rf	MLP	MANN	CMNN	CAMNN
MSE	2.7134	2.7599	2.6706	1.5909	1.3613	1.0881
MAE	1.4667	1.3342	1.4504	1.1457	0.7714	0.5251

Fig. 6. Comparision of MSE and MAE of CAMNN with other models.

From the figure, it can be seen that deep learning-based methods have a lower average error and more substantial predictive potential than shallow machine learning-based methods. Broadly speaking, all the deep learning models outperform the traditional models, and the proposed method has the best performance. This fact is that the shallow machine learning methods structure is relatively simple. Conversely, When addressing the temperature interpolation challenge in grain storage, deep learning techniques possess the ability for automatic feature learning, high-dimensional data processing and a specific structural design. Therefore, deep learning-based approaches are more adaptable in modeling complicated and nonlinear temperature interpolation.

The SVM model has the worst prediction performance in all regions, with the highest RMSE and MAE values. RF is relatively better than SVM, but the information of surrounding points cannot be used in point temperature interpolation. That is why there is no dazzling performance on the SVM.

For the MLP, the multiple neurons can capture more linear combinations and the nonlinear transformation of ReLU makes the result better. However, In the face of intricate data distributions or pronounced non-linearities, a mere MLP might fall short in capturing the details and patterns. Additionally, being a feedforward neural network, MLP struggles to effectively discern feature interactions or handle spatial data. As can be seen from the figure, the results of MLP leave much to be desired.

An ablation study has been considered to evaluate the performance. By introducing the Attention Mechanism and CNN, the overall performance of the interpolation model is significantly improved. The CAMNN model has the lowest error values compared with simple hybrid models and MLP single because the proposed model addresses the corresponding problem by using the advantages of CNN and Attention mechanism. This improvement is attributed to the effectiveness of the proposed model combination utilizing the strengths of each model. The reasons for the differences in the performance of these models can be attributed to the following points: 1) In grain temperature field modeling, three aspects are paramount: capturing spatial characteristics, emphasizing pivotal data, and

discerning non-linear relationships. The proposed model adeptly manages all three, yielding the topmost accuracy. 2) The CMNN model, while proficient in identifying spatial characteristics and non-linear relationships, doesn't give adequate attention to crucial data, leading to slightly reduced accuracy. 3) The MANN model, even with its ability for highlighting key data and understanding non-linear patterns, lacks in recognizing spatial characteristics. Given that grains remain relatively stable during storage and the temperature changes induced by their biochemical reactions are minimal compared to the impacts of surrounding points, models that don't account for spatial features tend to have significantly lower accuracy.

Overall, the MAE and MSE of CAMNN are 0.5251 and 1.0881 respectively, achieving reductions of 64.20% and 59.90% over SVM, 60.64% and 60.57% over RF, 63.80% and 59.26% over MLP, 54.17% and 31.60% over MANN, 31.93% and 20.07% over CMNN, respectively.

Visualization Analysis

Furthermore, we present the three-dimensional spatial distribution of interpolated results for each implemented method to show the results more clearly and highlight our proposed method's overall benefits over other methods. As shown in Figure 7 and Figure 8, at the surface of the grain pile and near the walls, the temperature of the grain is more dynamic due to their proximity to the external environment. Consequently, these locations exhibit larger temperature fluctuations. Conversely, the central region experiences minimal temperature variations, owing to the poor thermal conductivity of the grain. During the summer season, the temperature difference between the grain's periphery and the central part of the pile gradually increases, accurately depicting the temperature distribution pattern of grain storage, commonly referred to as the "cold core and hot surface" phenomenon.

The performance of various methods exhibits some differences, as can be analyzed from the interpolation result plots: 1) The spatial distribution of grain temperature exhibits a clear stratification phenomenon, with vertical temperature gradients greater than horizontal ones. This consideration was also taken into account when designing the sensor network, with smaller vertical spacing than horizontal spacing. From the interpolation plots generated by CAMNN, it is evident that the vertical trend is better represented. This is attributed to the attention mechanism capturing this crucial trend, resulting in interpolation outcomes that incorporate more information from points in the vertical direction. 2) Inside the grain pile, the temperature fluctuation are quite stable and exhibit the characteristic of uniformity, meaning that in the absence of external disturbances, the temperature differences among internal nodes are not significant, and all models exhibit good performance. 3) In the figure, the first row illustrates the fusion models, and the second row depicts the non-fusion models. Several interpolation methods yield satisfactory predictive accuracy, yet the fusion models are more precise in predicting the temperature at the surface of the heap and those close to the walls and floor. The proposed method, in particular, reproduces values that are closer to the true values, notably at points susceptible to external disturbances, as shown by the arrow in part d. 4) As shown in the black rectangular area in part b of the figure, CMAA outperforms MANN, which is due to the CNN architecture is inherently adept at handling data with spatial correlations. The temperature distribution on the surface displays spatial patterns similar to images, and CNN can effectively capture these local details. The black boxed dots in part c indicate that MANN is superior to CMNN. This superiority stems from the significant temperature changes at points near both the wall and the grain surface/ground. There are fewer such points in the training set that reflect this characteristic, and the Attention mechanism excels at emphasizing key features and efficiently selecting and concentrating information. 5) At the grain surface layer, as shown in the black rectangular area of part d, MLP, RF, and SVM all perform inadequately, with noticeably less accuracy than the fusion models. The reason for this is that fusion models, by integrating various network types, offer higher model complexity and flexibility, better adapting to the characteristics of the data. 6) In areas of extreme high and low temperatures, as well as

where there are large temperature fluctuations, marked by the red and black arrows in part d, the non-fusion models are not sufficiently effective. The reason for this underperformance is that the CNN is adept at capturing local spatial characteristics, a vital aspect for recognizing these points. Meanwhile, the Attention mechanism aids in the model focus on key temperature changes, enhancing its ability to recognize anomalies. The non-fusion models might lack sensitivity to abrupt changes, particularly in scenarios where there is an inadequate coverage of these extreme cases in the training data.

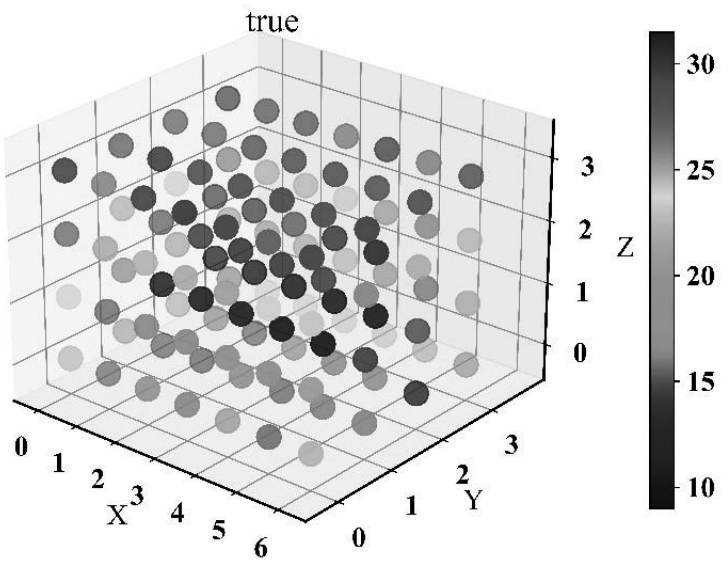


Fig. 7. Measured values of the Hanzhong granary temperature sensor network on August 25, 2022

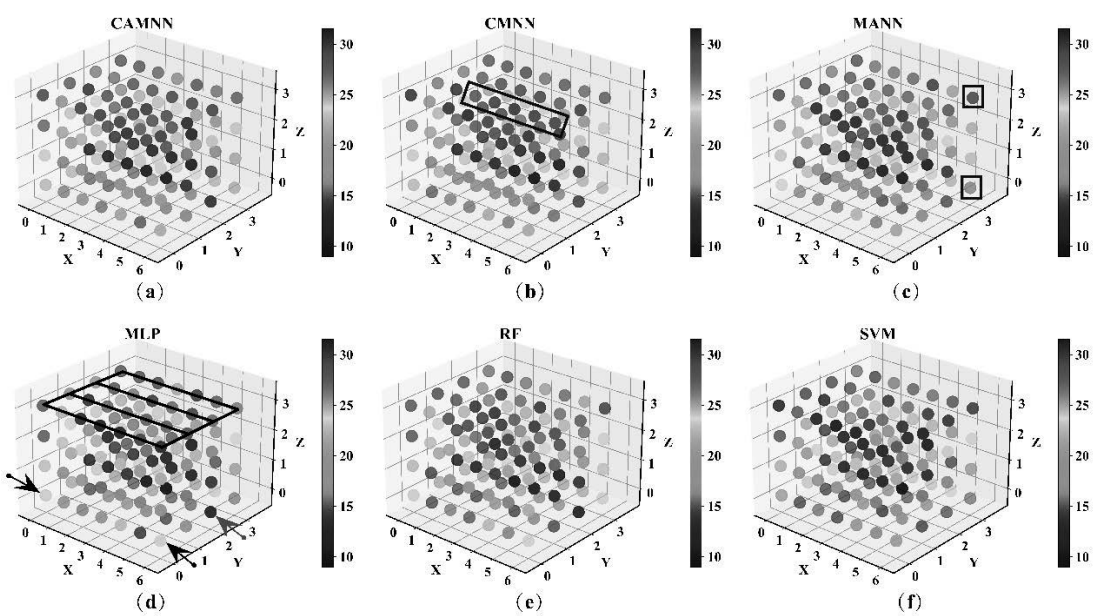


Fig. 8. Interpolation values of corresponding sensor nodes of the Hanzhong granary on August 25, 2022

Conclusion

The process of grain storage encompasses various elements, including the internal airflow within the grain heap, fluctuations in its temperature and humidity, and changes in the external storage conditions, all of which culminate in a distinctive grain storage ecosystem. The temperature fluctuations inside the grain pile are intricate, and it's impractical to represent the entire temperature distribution solely with data from a limited number of sensors. China boasts a rich variety of grains and numerous types of granaries. It's challenging to effectively measure and calculate the internal parameters of the grain pile, making the temperature distribution and interpolation particularly difficult. At present, there's a lack of comprehensive research on interpolation algorithms for grain temperature fields. Only a few scholars have applied some simple models, such as the BP algorithm based on deep learning and traditional spatial interpolation algorithms like Kriging. These models, however, fall short in accuracy and versatility, overlooking the unique attributes of grains and often failing to meet practical engineering demands. We presented a hybrid neural network based on CNN and MLP with the attention mechanism in this study. Leveraging the local feature extraction of CNN, dynamic selection of Attention, and global information integration of MLP, we achieved precise spatial interpolation of the temperature distribution within the grain pile. Such a design ensures that the model has high accuracy and strong adaptability, allowing it to automatically learn and adjust weights to cater to the data distribution and characteristics of various grain types in different regions and different size of granaries.

The interior of the grain pile is a dynamic system that evolves over time, composed of biological factors, non-biological factors, and environmental factors. The actual environment of grain storage are highly complex. This study solely implements spatial interpolation algorithms based on sampled data, neglecting the temperature variations in grain under extreme scenarios such as mold or insect damage. For future work, 1) we plan to construct a small-scale granary for experimental purposes to cultivate mold for testing, collect data, and validate the model. 2) Incorporate more features, such as the structure and materials of the granary, to further improve the accuracy of interpolation. 3) To better understand the decision-making process of the model, we consider introducing model interpretability tools, such as SHapley Additive exPlanations (SHAP) or Local Interpretable Model-agnostic Explanations (LIME). 4) Ultimately, develop a real-time monitoring and early warning system that automatically sends an alarm when the model interpolates that the grain temperature may reach a dangerous level.

References

1) Kumar D, Kalita P (2017) Reducing Postharvest Losses during Storage of Grain Crops to Strengthen Food Security in Developing Countries. Foods 6:8. https://doi.org/10.3390/foods6010008

2) Chen X, Wu L, Shan L, Zang Q (2018) Main Factors Affecting Post-Harvest Grain Loss during the Sales Process: A Survey in Nine Provinces of China. Sustainability 10:661. https://doi.org/10.3390/su10030661

3) Liu P, Ding L (2017) Study on Reducing Food Losses and Ensuring Food Security. Rural Economy and Science-Technology 28:5–7. https://doi.org/10.3969/j.issn.1007-7103.2017.05.003

4) Krishnan P, Ramakrishnan B, Reddy KR, Reddy VR (2011) Chapter three - High-Temperature Effects on Rice Growth, Yield, and Grain Quality. In: Sparks DL (ed) Advances in Agronomy. Academic Press, pp 87–206

5) Wu W, Cui H, Han F, et al (2021) Digital monitoring of grain conditions in large-scale bulk storage facilities based on spatio-temporal distributions of grain temperature. Biosystems Eng 210:247–260. https://doi.org/10.1016/j.biosystemseng.2021.08.028

6) Niu Z, Zhong G, Yu H (2021) A review on the attention mechanism of deep learning. Neurocomputing 452:48–62. https://doi.org/10.1016/j.neucom.2021.03.091

7) Rezk NG, Hemdan EE-D, Attia A-F, et al (2023) An efficient IoT based framework for detecting rice disease in smart farming system. Multimed Tools Appl 82:45259–45292. https://doi.org/10.1007/s11042-023-15470-2

8) Ziegler V, Paraginski RT, Ferreira CD (2021) Grain storage systems and effects of moisture, temperature and time on grain quality - A review. J Stored Prod Res 91:101770. https://doi.org/10.1016/j.jspr.2021.101770

9) Kattenborn T, Leitloff J, Schiefer F, Hinz S (2021) Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. Isprs J Photogramm Remote Sens 173:24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010

10) Peng X, Shu W, Pan C, et al (2022) DSCSSA: A Classification Framework for Spatiotemporal Features Extraction of Arrhythmia Based on the Seq2Seq Model With Attention Mechanism. IEEE Trans Instrum Meas 71:1–12. https://doi.org/10.1109/TIM.2022.3194906

11) Desai M, Shah M (2021) An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). Clinical eHealth 4:1–11. https://doi.org/10.1016/j.ceh.2020.11.002

12) M H, M.N S (2015) A Review on Evaluation Metrics for Data Classification Evaluations. International Journal of Data Mining & Knowledge Management Process 5:01–11. https://doi.org/10.5121/ijdkp.2015.5201