Temperature Forecasting for Stored Grain: A Deep Spatiotemporal Attention Approach

Shanshan Duan, *Student Member, IEEE*, Weidong Yang, Xuyu Wang, *Member, IEEE*, Shiwen Mao, *Fellow, IEEE*, and Yuan Zhang

***Abstract*—The development of Internet-of-Things (IoT) technology promotes the advances of grain condition detection and analysis systems. Temperature monitoring is a main element to maintain grain quality, and effective control of grain temperature is crucial to safe storage of grain. In this article, an encoder–decoder model with attention mechanism is proposed to accurately forecast the temperature of stored grain. Considering that the points on the gradient direction of the temperature surface have a great influence on the temperature of the target point, the Sobel operator is used to extract the local characteristics of the target point. In addition, considering the correlation structure in the sensory data, the attention mechanism is used to extract the global features of the target point. The extracted spatial features are fed into long short-term memory (LSTM) networks to obtain the long-term state information of spatial factors. LSTM unit and convolutional neural network are used to encode the spatial features of the target points. Taking meteorological factors as the external input of the decoder, temporal attention mechanism and LSTM unit are used to complete the decoding process and realize the prediction of grain temperature in the future. The results with real grain storage data show that the proposed model outperforms several schemes, including Kalman-modified the least absolute shrinkage and selection operator (Kalman-modified LASSO), temporal graph convolutional network (T-GCN), LSTM, CNN-LSTM, and convolutional LSTM (Conv-LSTM), with considerable gains.**

Manuscript received January 19, 2021; revised April 17, 2021; accepted May 3, 2021. Date of publication May 7, 2021; date of current version November 19, 2021. This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFD0401001; in part by the Open fund of Key Laboratory of Grain Information Processing and Control under Grant KFJJ-2017-201; in part by the Program for Science and Technology Innovation Talents in Universities of Henan Province under Grant 19HASTIT027; in part by the Innovative Funds Plan of Henan University of Technology Plan under Grant 2020ZKCJ02; in part by the Natural Science Foundation of Henan Province under Grant 182300410036; in part by NSF under Grant ECCS-1923163; and in part by the Wireless Engineering Research and Education Center at Auburn University, Auburn, AL, USA. This article was presented in part at the IEEE 2020 International Conference on Parallel and Distributed Systems (ICPADS 2020), Hong Kong, China. *(Corresponding authors: Weidong Yang; Yuan Zhang.)*

Shanshan Duan and Weidong Yang are with the Henan Key Laboratory of Grain Photoelectric Detection and Control and the College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450066, China (e-mail: 18903838857@163.com; yangweidong@haut.edu.cn).

Xuyu Wang is with the Department of Computer Science, California State University at Sacramento, Sacramento, CA 95819 USA (e-mail: xuyu.wang@csus.edu).

Shiwen Mao is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: smao@ieee.org).

Yuan Zhang is with the Key Laboratory of Grain Information Processing and Control and the College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450066, China (e-mail: zhangyuan@haut.edu.cn).

Digital Object Identifier 10.1109/JIOT.2021.3078332

***Index Terms*—Attention mechanism, deep learning, grain temperature, meteorological metrics, spatiotemporal modeling.**

I. Introduction

UP TO 2050, the world cultivated land will increase from nearly 0% to nearly 1% every year and the annual growth rate of world medial grain production may be as high as 1.43% to meet the world's food needs [1]. The predicted world grain inventory is 876 million tons by the end of 2021. It was also estimated that China's wheat output would be 134 million tons in 2020. According to the organization for economic co-operation development (OECD) and the food and agriculture organization (FAO) of the United Nations (OECD-FAO) agricultural outlook 2020–2029, the COVID-19 pandemic may further undermine food security. In 2020/2021, wheat inventories will increase further, most of which will come from China [2], [3]. The storage period (three to five years) of wheat in China is longer than that in developed countries. However, estimates suggest that poor grain storage management in developing countries results in 20%–50% postharvest losses of grains grown [4], which is a huge waste. Consequently, ensuring safe food storage is of critical importance.

Usually mold and pest are the two major causes of loss of stored food, while the main cause of mold and pest growth in grain pile is high temperature and high humidity of grain [5]. To ensure high quality of grain storage, it is necessary to monitor and predict the temperature and moisture levels properly to prevent the growth of mold and insects. The Internet of Things (IoT) has been more and more widely used in precision agriculture in recent years [6]–[8]. With the development of the IoT, temperature sensors, humidity sensors, and gas sensors are widely deployed in grain storage, which allow closely monitoring the state of stored grain to a certain extent, which plays an important role for safe food storage. It is necessary to control the key storage factors that affect the grain quality. For instance, a low-cost and nondestructive detection system based on Wi-Fi is used to estimate the humidity and mildew conditions of stored wheat with machine learning algorithms [9]–[11]. A network of smart sensors is used to cover and monitor grain storage area, which achieves effective coverage, high detection accuracy, and message transmission ability with its chain-based structure that is not affected by various obstacles in the grain storage facility [12]. The finite element method (FEM) and other numerical simulation methods have

been applied to predict all kinds of thermal physics and mass transfer in the process of grain storage. However, it is usually difficult and complex to determine the parameters of such methods [13], [14]. It would be interesting to develop more effective storage parameter analysis and prediction methods with more advanced techniques from time-series analysis and machine learning, in particular, deep learning.

There are many factors affecting the safety of grain storage. Since temperature sensor networks are widely deployed in grain storage systems and their wired transmissions are usually reliable, grain temperature data are usually available and complete for many storage systems. In addition, the heat gradient in stored grain pile leads to natural convection flow and moisture migration in bulk grain. Therefore, the knowledge of temperature distribution is the basis of the research. Most of the prior works attempt to ensure safe storage of grain by understanding the trend of temperature change inside the grain pile in advance. Some statistical models have been proposed to forecast the grain temperature, such as the least square method, time-series regression, and so on. In addition, machine learning models are utilized to achieve enhanced temperature predication performance. Considering that temperature, humidity, and other storage parameters are affected by the relevant meteorological conditions, machine learning algorithms [e.g., support vector machine (SVM) and adaptive boosting (AdaBoost) [9], [15], [16]] are utilized to forecast the average temperature of the upper layer of stored wheat where historical grain temperature data and the corresponding meteorological data are utilized [15], [16]. Nevertheless, those models ignore the temporal and spatial correlation of temperature in the grain pile. Therefore, there is still considerable room for improvement in machine learning-based grain temperature prediction.

In this article, a temporal and spatial approach based on the attention mechanism is proposed to forecast the grain temperature in a granary [17]. The proposed approach is motivated by several observations. First, we perform a preliminary analysis of the grain temperature data collected from a specific granary and the weather data of the corresponding area. Considering the grain temperature data measured by each sensor as a target time series, we examine the influence of relevant weather factors on grain temperature. We find that there is a considerable correlation between grain temperature and the external meteorological factors. Second, there is considerable spatial correlation among the temperature data collected from neighboring sensors deployed at different positions in the granary. Moreover, each temperature time series itself also exhibits considerable autocorrelation and self-similarity over time. Such correlations motivate us to explore the temporal and spatial interdependencies in grain temperature data for better prediction. Third, the temporal attention mechanism has been shown effective to adaptively select the relevant state of the encoder for improving temperature forecasting performance.

The proposed spatial and temporal approach is based on the attention mechanism and consists of four main components: 1) spatial feature extraction, where the Sobel operator is used to extract local features and the spatial attention mechanism is used to extract global features of the target sensor locations; 2) the encoder: an artificial recurrent neural network (RNN) deep learning model, i.e., long short-term memory (LSTM), is incorporated to obtain the long-term state information of spatial features, and the Inceptions modules are used to extract all the spatial information to complete the coding process; 3) the decoder: the temporal attention mechanism is used to adaptively capture the relevant information and combine it with the corresponding weather factors to achieve the final temperature prediction; and 4) the selection of model parameters and the evaluation of prediction results by using real grain storage data sets. It can be seen from the experimental results that the prediction accuracy of the proposed approach is superior to the several other schemes used in our comparison study, including Kalman-modified least absolute shrinkage and selection operator (Kalman-modified LASSO), a temporal graph convolutional network (T-GCN), LSTM, convolutional neural network (CNN)-LSTM, and convolutional LSTM (Conv-LSTM).

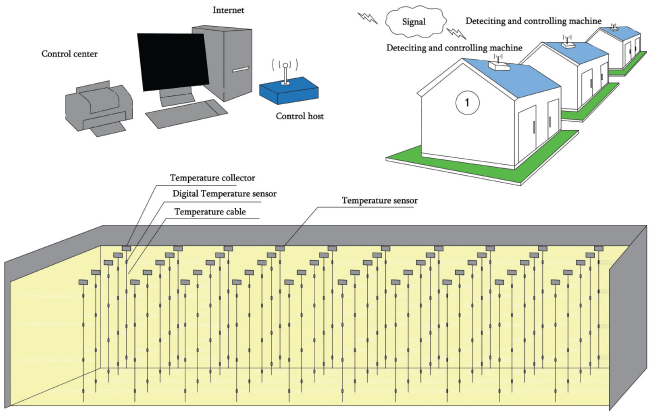

Fig. 1. Grain condition measurement sensor network and the collection and control system. Note that this is an illustrative plot. The sensor network used in this article consists of 200 sensors deployed in ten rows, five columns, and four layers.

In the remainder of this article, Section II provides the problem description and preliminary analysis of grain temperature data. Section III introduces the proposed temporal and spatial prediction approach. Section IV presents our performance evaluation of the proposed model using real granary data. Section V reviews the related work. Finally, Section VI summarizes this article.

II. Problem Description and Preliminary Analysis

A. Data Set and Problem Statement

A grain measurement and control system based on the IoT is designed and deployed to collect food condition data in a granary, as shown in Fig. 1. The IoT system for grain monitoring and control consists of computer technology, network technology, and electronic technology to detect grain, data transmission, data storage, and analysis, and to improve the grain situation of the system through the control of grain storage equipment. The system is composed of a host computer, a transmission interface, a slave computer, sensors for condition of stored grain, thermometric cable and other hardware, and grain monitoring software. The host computer is a computer

that has been installed with the software of grain situation measurement and control system. It controls the actions of all kinds of equipment by sending commands and receives the detection signals. It has the functions of signal correction, data display, storage, statistical analysis, management of network connection, etc. The transmission interface is a device that receives and transmits signals between the host computer and the slave computer. The slave computer receives the commands of the host computer, returns the grain information and the status information of the grain condition control equipment collected by the sensors to the host computer, and executes the control actions. A thermometric cable is a special cable for detecting grain temperature, which is usually composed of temperature sensor, wire, tensile steel wire, and sheath. The temperature sensor is a DS18B20 digital temperature sensor produced by a Dallas, TX company. The absolute value of temperature error is not more than 0.5°. In the IoT system, the control host is connected to the Internet and transmits the grain information to the central control center. The control center can remotely operate the grain condition control equipment in the granary following the data analysis (e.g., start the ventilation system or the air conditioning system).

In this article, 200 temperature sensors are arranged in a 3-D array in a large wheat warehouse. In the tall square granary, the temperature sensor layout principle is that the distance between the horizontal and horizontal temperature measuring cables should be no more than 5 m, the distance between the vertical cables should be no more than 2 m, and the distance from the cables to the grain surface, granary bottom, and granary wall should be within 0.3–0.5 m. More details on data collection and preprocessing of the IoT system can be found in [17], but omitted here for brevity.

According to their locations, the sensors can be organized into three planes: 1) the *x*-*y* plane; 2) the *x*-*z* plane; and 3) the *y*-*z* plane. The historical data collected from the 200 temperature sensors in the past T days are denoted as $\mathbf{X} = (\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^{200}) \in \mathbb{R}^{200\times T}$, where $\mathbf{x}^i \in \mathbb{R}^{1\times T}$ is the sensory data from the sensor i. The objective is to combine the sensory temperature data with the corresponding historical weather data to predict the temperatures at each target location in the future τ days (i.e., τ-step ahead forecast), which are denoted as $\hat{\mathbf{y}}^i = (\hat{y}^i_{T+1}, \hat{y}^i_{T+2}, \ldots, \hat{y}^i_{T+\tau})^\top \in \mathbb{R}^\tau$, with a temporal and spatial distribution approach.

The complex process of mass and heat transfer in the ecosystem of stored grain challenges the accurate prediction of grain temperature. Several key problems must be solved to accurately predict the temperature of stored grain by using the spatiotemporal model and historical data. First, the temperature of stored grain piles is affected by external meteorological factors. For example, when the grain in the granary is dry and free of insects, the activities of the organisms in the grain pile will be extremely weak, and the heat generated has no significant impact on changing the grain temperature. The normal grain temperature changes with the external temperature, which affects the air temperature in the grain bin, and the air temperature change in the grain bin, in turn, affects the grain temperature. The driving force of heat transfer is the temperature gradient developed throughout the grain mass due to the effects of ambient temperature, solar radiation, and other meteorological factors. The temperature gradient induces natural convection currents of varying speeds that also transfer heat. Second, grain is a poor conductor of heat, and the natural convection of air in grain pile is usually very weak. Therefore, although the change of grain temperature is affected by external factors, it has its special characteristics. The grain temperature is affected by the temperature from the surface to the center and from the outside to the inside, developing to the depth gradually and slowly.

Therefore, the change of grain temperature lags far behind that of the air temperature and bin temperature. The daily lowest and highest temperature values usually occur 1–2 h later than the lowest and highest temperature values of the warehouse, respectively. A similar trend is observed for the annual variation of grain temperature, which changes with the air temperature, but with a big lag. In the rising season of air temperature, the grain temperature gradually rises, but is usually lower than the air temperature. In the falling season of air temperature, the grain temperature also gradually drops, but is usually higher than the air temperature. In general, the occurrence of the lowest and the highest values in a year is usually delayed for one to two months compared with the lowest and the highest values of the air temperature, respectively. In a certain area, the air temperature, bin temperature, and surface grain temperature all show a periodic law of daily and annual patterns, but sometimes the periodicity is not very regular.

B. Correlation Structure Analysis of Grain Temperature Data

In this article, the temperature time series collected by sensors deployed at different locations in the stored grain pile are first analyzed from the perspective of temporal and spatial correlations. Our main findings are presented in the remainder of this section.

1) Autocorrelation and Self-Similarity in the Time Domain: The autocorrelation coefficient of the temperature samples collected by a sensor can be defined as [18]

$$\rho = \frac{\sum_{t=1}^{l}(x_t - \bar{x})\left(x_{t+\xi} - \bar{x}\right)}{\sqrt{\sum_{i=1}^{l}(x_t - \bar{x})^2}\sqrt{\sum_{i=1}^{l}\left(x_{t+\xi} - \bar{x}\right)^2}} \tag{1}$$

where x_t and $x_{t+\xi}$ are the observations of grain temperature at time t and $t+\xi$, respectively; $\bar{x}$ is the mean value of temperature; and l is the time length (or, number of samples). The autocorrelation analysis results of temperature samples are shown in Table I. The natural ordinal number in the first column of the table is the value of the lag period ξ from 1 to 20. The column *AC* is the computed autocorrelation value, the *PAC* column is the estimated partial autocorrelation coefficient value, the *Q-stat* column is the value of q-statistic, and the *Prob* column is the probability that the value of the q-statistic is greater than the Q-value calculated by the samples. The original hypothesis of the Q statistic is that the sequence is nonautocorrelation. If the p value is greater than 1% significance level, the original hypothesis is accepted, that is, the sequence is not autocorrelated. If the *Prob* value is less than a given significance level (1%), the original hypothesis is rejected, i.e., there is autocorrelation in the time series. The

TABLE I
AUTOCORRELATION IN THE TIME DOMAIN

ξ	AC	PAC	Q-Stat	Prob
1	0.998	0.998	1261.1	0.000
2	0.994	−0.387	2512.8	0.000
3	0.989	−0.034	3753.1	0.000
4	0.984	0.048	4981.5	0.000
5	0.979	0.043	6198.7	0.000
6	0.974	0.021	7405.3	0.000
7	0.970	0.017	8601.9	0.000
8	0.965	0.023	9789.2	0.000
9	0.961	−0.020	10967.	0.000
10	0.957	−0.030	12136.	0.000
11	0.953	0.000	13296.	0.000
12	0.948	−0.044	14445.	0.000
13	0.944	−0.046	15584.	0.000
14	0.938	−0.013	16711.	0.000
15	0.933	−0.017	17827.	0.000
16	0.928	−0.025	18930.	0.000
17	0.922	−0.029	20020.	0.000
18	0.916	−0.036	21098.	0.000
19	0.910	−0.030	22162.	0.000
20	0.903	−0.043	23212.	0.000

Prob values in Table I are all zero, which indicate that the time series of grain temperature samples have self-correlation. The *AC* column shows that the autocorrelation values are quite high (i.e., all above 0.9).

We also find that the temperature time series of stored grain exhibit the characteristics of partial and overall similarity under different time scales, namely, *self-similarity*. In this article, we use the Hurst exponent H to measure the self-similarity of stochastic processes. When $0 < \mathrm{H} < 0.5$, the time series has long-term correlation, but the overall trend in the future is contrary to that of the past, i.e., the anti-persistence. When $\mathrm{H} = 0.5$, the time series is uncorrelated. When $0.5 < \mathrm{H} < 1$, the time series has the characteristics of long-term correlation, that is, the process is self-similar. Rescaled range (R/S) analysis is usually used to obtain the value of H [19]–[21]. The process of calculating the Hurst parameter by rescaled range analysis (R/S) is as follows.

Step 1: The temperature time series $\{x_1, x_2, \ldots, x_n\}$ is divided into g nonoverlapping subsequences with length r, denoted as $\{x_{i1}, x_{i2}, \ldots, x_{ir}\}$, $i = 1, 2, \ldots, g$.

Step 2: The mean value of each subsequence is calculated as

$$\bar{x}_i = \frac{1}{r}\sum_{j=1}^{r} x_{ij}, \quad i = 1, 2, \ldots, g. \tag{2}$$

Step 3: The deviation of each subsequence is calculated as

$$y_{ij} = x_{ij} - \bar{x}_i, \quad i = 1, 2, \ldots, g, j = 1, 2, \ldots, r. \tag{3}$$

Step 4: The accumulated deviation of each subsequence is calculated as

$$z_{ij} = \sum_{k=1}^{j} y_{ik}, \quad i = 1, 2, \ldots, g, j = 1, 2, \ldots, r. \tag{4}$$

Step 5: The range each subsequence is calculated as

$$R_i = \max\left(z_{ij}\right) - \min\left(z_{ij}\right), \quad i = 1, 2, \ldots, g, \quad j = 1, 2, \ldots, r. \tag{5}$$

Step 6: The standard deviation of each subsequence is calculated as

$$S_i = \sqrt{\frac{1}{r-1}\sum_{j=1}^{r}\left(x_{ij} - \bar{x}_i\right)^2}, \quad i = 1, 2, \ldots, g. \tag{6}$$

Step 7: Calculate the value of the rescaled range as

$$RS_i = \frac{R_i}{S_i}, \quad i = 1, 2, \ldots, g. \tag{7}$$

Step 8: The R/S values of the calculated subsequences are averaged as

$$(R/S)_r = \frac{1}{g}\sum_{i=1}^{g} RS_i. \tag{8}$$

Step 9: Increase the value of r and repeat the previous steps to obtain the rescaled range $(R/S)_r$ of logarithmic temperature series over time span of different length r. According to the definition of the Hurst index H, it describes the proportional relationship between $(R/S)_r$ and r^H, namely

$$(R/S)_n = C \times n^H. \tag{9}$$

Step 10: Taking logarithm on both sides of (9), $\log(r)$ is used for a linear regression of $\log((R/S)_r)$. The intercept of the regression equation is the constant C in (9), and the slope is the Hurst exponent H.

As shown in Fig. 1, the sensor nodes are deployed in four vertical layers from the top to the bottom of the granary. We calculate the Hurst parameter for the average temperature of each layer. The average temperature time series of each layer is denoted as $X_{a1}^{\text{layer}}, X_{a2}^{\text{layer}}, \ldots, X_{a1264}^{\text{layer}}$, for layer $= 1, 2, 3, 4$. The time series of average temperature in each layer is divided into nonoverlapping sections and the Hurst parameter is calculated, respectively. The results are presented in Fig. 2. It can be seen that the Hurst parameter of temperature time series has values between 0.7 and 0.9 under different time lengths, while the Hurst parameter value of the entire time series is relatively smaller, and the value is around 0.72. Therefore, we conclude that the time series of grain temperature is self-similar. Such self-similarity indicates long-range dependence in the temperature time-series data and historical data is useful for predicting future temperature samples due to such long-term correlation.

2) Spatial Correlation: According to the definition of the Pearson correlation coefficient [22], the correlation coefficient between sensor nodes (i.e., the temperature samples) at different locations is expressed as

$$\rho_{pp} = \frac{\sum_{t=1}^{d}(x_{it} - \bar{x}_i)\left(x_{jt} - \bar{x}_j\right)}{\sqrt{\sum_{t=1}^{d}(x_{it} - \bar{x}_i)^2}\sqrt{\sum_{t=1}^{d}\left(x_{jt} - \bar{x}_j\right)^2}} \tag{10}$$

where x_{it} and x_{jt} represent the temperature samples from different locations i and j at time t, respectively; and $\bar{x}_i$ and $\bar{x}_j$ are

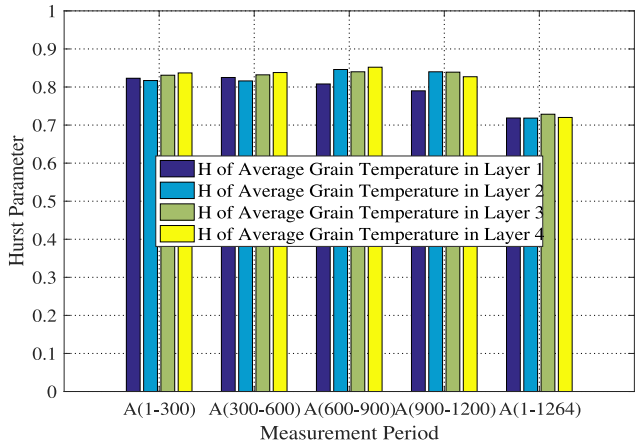

Fig. 2. Hurst parameter of average grain temperature for each layer of sensors in different time periods. A(1-300) denotes X_{a1}^{layer}, X_{a2}^{layer}, ..., X_{a300}^{layer}; A(300-600) denotes X_{a300}^{layer}, X_{a301}^{layer}, ..., X_{a600}^{layer}; A(600-900) denotes X_{a600}^{layer}, X_{a601}^{layer}, ..., X_{a900}^{layer}; A(900-1200) denotes X_{a900}^{layer}, X_{a901}^{layer}, ..., X_{a1200}^{layer}; and A(1-1264) denotes the entire time length of average temperature in each layer.

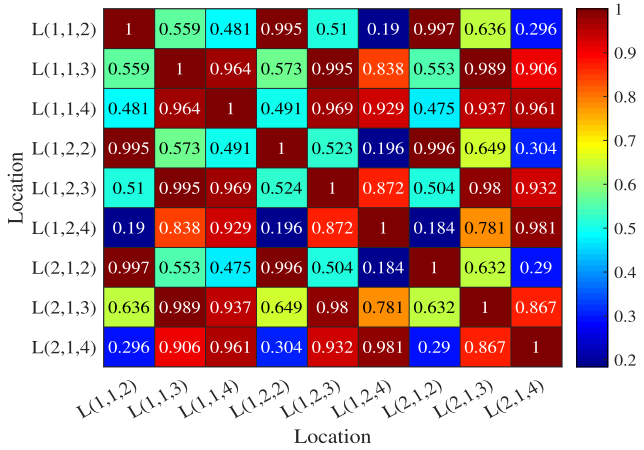

Fig. 3. Correlation matrix showing the correlation between different sampling locations.

the mean values of temperature at the two locations, respectively. The correlation coefficients among different locations are shown in Fig. 3. The three parameters in the bracket, e.g., $L(1, 1, 2)$, indicate the row, column, and layer of the sensor node in the grain pile (see Fig. 1). It can be seen that the correlation coefficient among the different sampling positions is irregular in a certain range. For example, the correlation coefficient between sensor node $L(1, 1, 2)$ and sensor node $L(1, 1, 3)$ is 0.559, However, the correlation coefficient between the sensor node $L(1, 1, 3)$ and sensor node $L(1, 1, 4)$ is 0.964, even if the spatial distance between $L(1, 1, 2)$ and $L(1, 1, 3)$ is the same as that between sensor node $L(1, 1, 3)$ and sensor node $L(1, 1, 4)$. The above results also indicate that the temperature samples in the grain pile have strong spatial correlation.

In fact, stored grain is a kind of active substance, i.e., it has the respiration function, and the granary is an open nonadiabatic system. However, in the spatial distribution of grain temperature, the main function is the heat transfer caused by the spatial temperature difference. Thus, the temperature data of stored grain sampled at the same time exhibit certain spatial correlation.

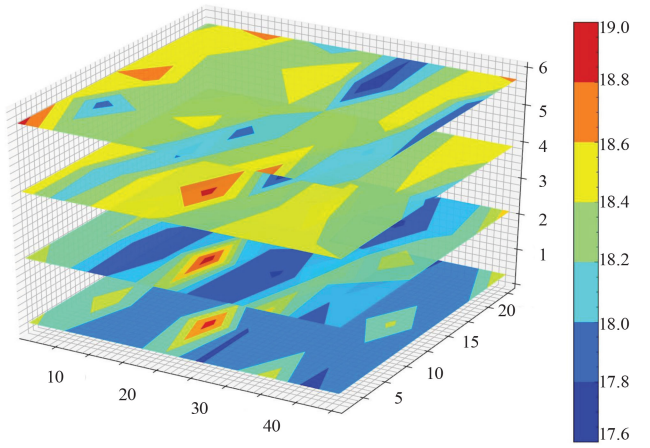

Fig. 4. Isotherm of the average grain temperatures in the four layers of sensor nodes.

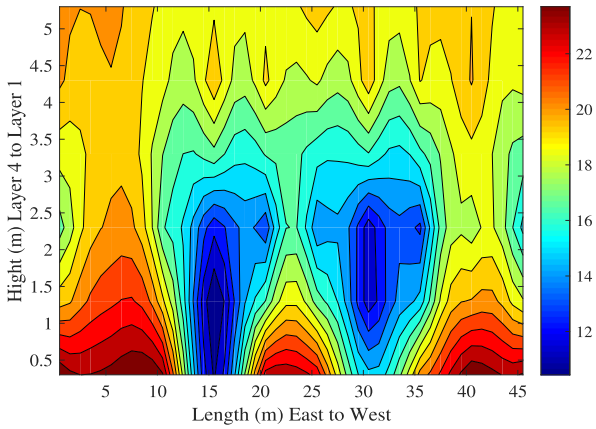

Fig. 5. Isotherm of the third column (i.e., the vertical slice at the center) in the grain pile.

Our correlation analysis of the grain temperature data show that there is considerable mutual influence among the temperature samples over time and at different locations. The overall temperature of the stored grain is a dynamic process. We plot the isotherm of the average temperature of the four sensor node layers in Fig. 4 to show how temperature is changing over space. It can be seen that the average temperature of the grain pile decreases from the first layer (the upper layer) to the fourth layer (the bottom layer). Then, we analyzed the temperature isotherm diagram of the third column [i.e., sensor nodes $L(*, 3, *)$] in the grain pile using data collected in October 2016, which is shown in Fig. 5. It can be seen that there is an obvious hot-skin-and-cold-core phenomenon, which is in line with the temperature change rule of stored grain pile. In the following, we will analyze the grain temperature isotherm map to obtain the local factors that affect the temperature change of the target location.

III. Proposed Temporal and Spatial Prediction Approach

According to the analysis of the time series of grain temperature in the time domain and the spatial domain in Section II, we find that the time series of grain temperature not only

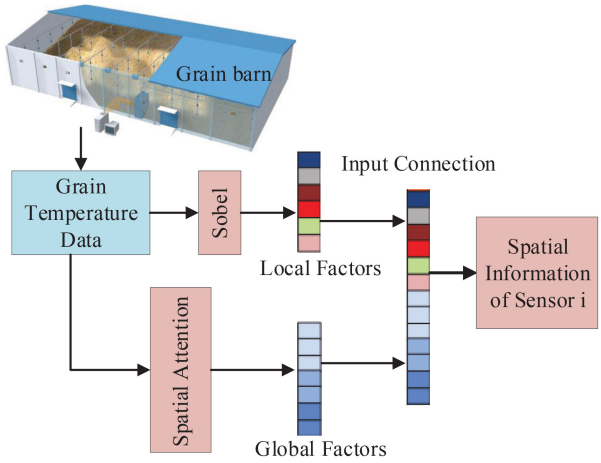

Fig. 6. Process for extracting the spatial features.

has strong correlation and self-similarity in the time domain but also has considerable correlation in the spatial domain. Based on the analysis, a temporal and spatial approach based on the attention mechanism is proposed to predict the temperature of stored grain utilizing both historical grain temperature data and the meteorological data of the corresponding area. By predicting the grain temperature, we can obtain "early knowledge" of the grain situation, so as to take countermeasures in advance to avoid damage of stored grain. The proposed process of grain temperature prediction is presented in this section.

A. Spatial Factor Extraction

To predict the temperature at a target location, the points in the gradient direction are considered to have a bigger impact on the temperature change. However, considering that the stored grain pile is usually a small ecological environment, the temperature samples collected at all the nodes of the entire sensor network interact with each other. Based on this observation, spatial features include local features and global features. The process of extracting the spatial features is illustrated in Fig. 6.

1) Local Features: According to the exploration of the change law of grain temperature observed in Section II, the driving force of heat transfer is the temperature gradient formed in the entire grain pile due to the influence of ambient temperature. The temperature gradient leads to natural convection with different velocity, and in turn transfers heat through convection.

In view of the relatively limited space stored grain, the Sobel operator is used to obtain the approximate gradient of the target location in the grain pile [23]–[25]. The Sobel operator includes two groups of 3×3 matrices, which are in the landscape orientation and portrait orienttion, as illustrated in Fig. 7. The procedure of the Sobel operator is expressed as follows:

$$\begin{aligned} G_x &= (X_7 + 2X_8 + X_9) - (X_1 + 2X_2 + X_3) \\ G_y &= (X_3 + 2X_6 + X_9) - (X_1 + 2X_4 + X_7). \end{aligned} \tag{11}$$

The gradient orientation angle is derived as

$$a(x, y) = \arctan\left(\frac{G_y}{G_x}\right). \tag{12}$$

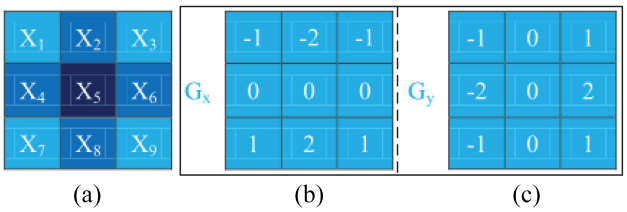

Fig. 7. (a) 3×3 area on the grain heap level. (b) and (c) Convolution templates of the Sobel operators.

The gradient directions of the target location on the three planes are calculated. Then, 10 values are taken in each gradient direction, and the 30 values are obtained as the local characteristics of the target location. The output vector of the local spatial factor is given by

$$\mathbf{l}_t = \left(x_t^{\text{grad},1}, x_t^{\text{grad},2}, \ldots, x_t^{\text{grad},L}\right)^\top \tag{13}$$

where $x_t^{\text{grad},\ell}$ is the ℓth temperature value in the direction of the temperature gradient at the target location (e.g., L is 30 for the temperature sensor network).

2) Global Features: By analyzing the correlation coefficient among temperature samples collected at different locations (see Section II-B), it can be seen that even for some locations with small spatial distance, the correlation coefficient between them could be small and the spatial correlation is weak. Therefore, the performance of prediction may be reduced if all other temperature samples are taken as global characteristics. Based on this rational, we introduce the attention mechanism to capture the key features of target temperature prediction.

Taking the ith temperature sensor as the prediction target, the attention score between the pth sensor and the target sensor i in the sensor network can be calculated as follows [26], [27]:

$$a_t^p = \mathbf{v}_a^\top \tanh\left(\mathbf{W}_a\left[\mathbf{h}_{t-1}; \mathbf{s}_{t-1}\right] + \mathbf{U}_a\mathbf{x}^p + \mathbf{V}_a\mathbf{x}^i + \mathbf{b}_a\right) \tag{14}$$

$$\delta_t^p = \frac{\exp\left(a_t^p\right)}{\sum_{j=1}^{G} \exp\left(a_t^j\right)} \tag{15}$$

where $[\cdot;\cdot]$ represents the concatenation operation, $\mathbf{x}^p$ and $\mathbf{x}^i$ are the temperature time series from sensor p and target sensor i, respectively, and $\mathbf{v}_a, \mathbf{b}_a \in \mathbb{R}^T, \mathbf{W}_a \in \mathbb{R}^{T\times 2m}, \mathbf{U}_a \in \mathbb{R}^{T\times T}$, and $\mathbf{V}_a \in \mathbb{R}^{T\times T}$ are model parameters. According to the target temperature series and the observed input temperature series, the spatial attention mechanism adaptively selects the most suitable input to predict the target sequence. When the attention score is obtained, the output vector of the global spatial factor at time step t is computed as

$$g_t = \left(\delta_t^1 x_t^{i,1}, \delta_t^2 x_t^{i,2}, \ldots, \delta_t^P x_t^{i,P}\right)^\top \tag{16}$$

where $x_t^{i,p}$ is the temperature sample at sensor p in the sensor network (i.e., P is 200 for the temperature sensor network). Note that δ_t^p is calculated as in (15).

After obtaining the local factors $\{\mathbf{l}_t\}_{t=1}^T$ and the global factors $\{\mathbf{g}_t\}_{t=1}^T$ of the target location, they are concatenated as the overall spatial features as: $\mathbf{m}_t = [\mathbf{l}_t; \mathbf{g}_t]$, where $\mathbf{m}_t \in \mathbb{R}^{L+P}$.

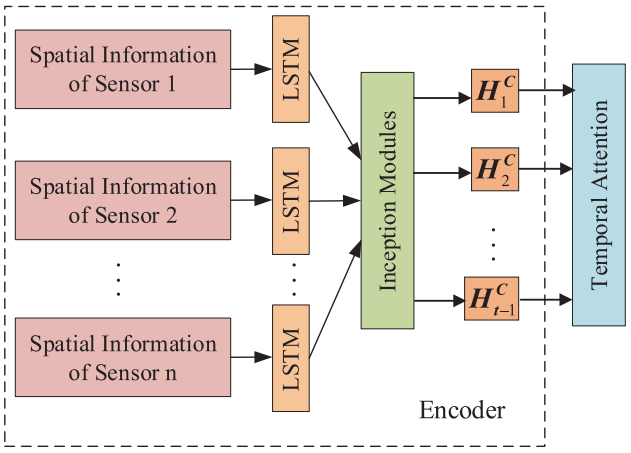

Fig. 8. Encoding process of spatial features.

B. Encoder Design

The encoding process of spatial features is shown in Fig. 8. In the encoder, the aggregate information sequence $\mathbf{m}_t$ of spatial features is taken as input to the encoder LSTM, which usually defines a function $f_e(\cdot)$. According to the spatial data information $\mathbf{m}_t$ of the current time step t and output state $\mathbf{h}_{t-1}$ of the last time step $t-1$, the current output state $\mathbf{h}_t$ is calculated for each time step t as

$$\mathbf{h}_t = f_e(\mathbf{m}_t, \mathbf{h}_{t-1}). \tag{17}$$

The hidden state $\mathbf{H}^i = \{\mathbf{h}_1^i, \mathbf{h}_2^i, \ldots, \mathbf{h}_{t-1}^i\}$ (where $i = 1, 2, \ldots, 200$) of different temperature sensors generated by the LSTM contains all the spatial information of the sensor network. For the hidden state of LSTM in each time step, their features are convoluted by a CNN to extract the features of spatial information.

For the process of using the CNN model to extract spatial features, we refer to the Inception model [28]–[30]. The CNN model contains four Inception modules with a similar structure, each of which has four branches as shown in Fig. 9. The first branch starts with a 1×1 convolution block with 16 output channels, followed by a 3×3 convolution block with 16 output channels. The second branch starts with a 1×1 convolution block with 16 output channels, followed by a 5×5 convolution block with 16 output channels. The third branch begins with a maximum-pooling layer, which is connected to a 1×1 convolution block of ten output channels. The fourth branch is a 1×1 convolution block with 16 output channels. The outputs of the four branches are then combined to produce the output of the first Inception module. The output of the first module, which is connected with a 1×1 convolution of a 16 output channels and a maximum pooling layer, is then used as input to the second Inception module. The structure of the first module is shown in Fig. 9. The structures of the second module, third module, and fourth module are similar to that of the first module except that the number of output channels is different. Finally, there is a fully connected layer whose output is the result of spatial feature extraction and used as input to the decoder.

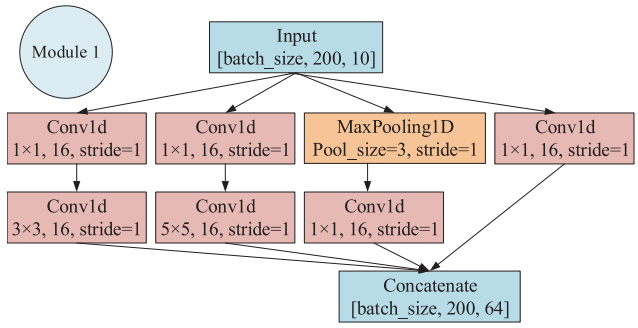

Fig. 9. Structure of the first Inception modules in the CNN model.

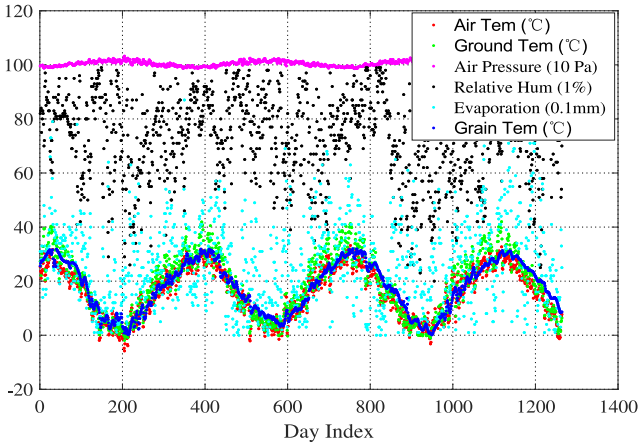

Fig. 10. Grain temperature and meteorological factors over time.

C. External Meteorological Features

In the process of grain storage, the external meteorological factors have a significant impact on the grain temperature. In general, grain temperature changes with external temperatures, which include solar radiation, atmospheric temperature, and ground temperature. In this article, eight meteorological factors are utilized for grain temperature prediction, including daily average air temperature (denoted by Air Tem), air pressure (Air Pressu), 0-cm ground surface temperature (Ground Tem), relative humidity (Relative Hum), daily average evaporation, sunshine duration (Sunshine Dur), wind speed, and precipitation (PRE). The time series of meteorological factors and the time series of mean grain surface temperature are plotted in Fig. 10. It can be seen that the trend of mean gain surface temperature is consistent with that of air temperature and ground temperature, but lags behind with some delay. The correlation coefficients between the mean grain surface temperature and the meteorological factors are plotted in Fig. 11. It can be seen from the figure that there is considerable correlation between grain temperature and meteorological factors, and there is mutual influence among the meteorological factors.

In view of this, in order to avoid using too much redundant information, the CNN composed of Inception modules is used to extract meteorological features. The CNN model contains four Inception modules with a similar structure, each of which has three branches. The first branch has three layers, which are one 1×1 convolution block with 64 output channels and two 3×3 convolution blocks with 64 output channels. The

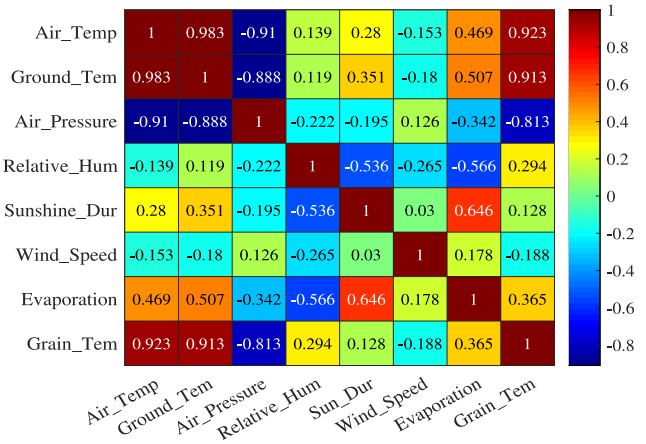

Fig. 11. Correlation coefficients between grain temperature and meteorological factors.

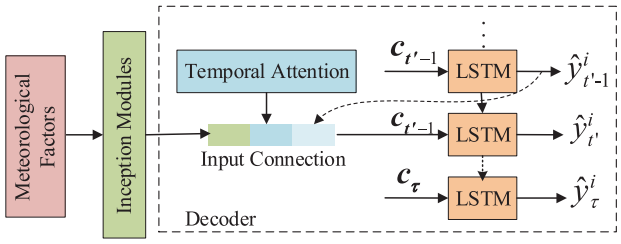

Fig. 12. Decoding process with the temporal attention mechanism.

second branch starts with a 1×1 convolution block with 64 output channels, followed by a 3×3 convolution block with 64 output channels. The third branch is a 1×1 convolution block with 64 output channels. The outputs of the three branches are then combined to produce the output of the first Inception module. The output of the first module, which is connected with a 1×1 convolution block with 64 output channels, is then used as input to the second Inception module. The structure of the second module, third module, and fourth module is similar to that of the first module except that the number of output channels is different. Finally, there is a fully connected layer whose output is used to make the final prediction together with the output of the temporal attention mechanism.

D. Decoder

For each time step t, there is a corresponding space vector $\mathbf{H}_t^C$. In order to capture the temporal correlation between spatial factors and enable the model to select relevant information across multiple time steps, the temporal attention mechanism is introduced in this article. The input temperature time series is usually long. The vector generated by the encoder cannot save all the information of the time series, which limits the prediction ability of the model. The temporal attention mechanism is used to overcome this problem. It is designed to calculate a matching weight between the current input sequence and the output vector, and adaptively select the relevant information to predict the target sequence [31]–[33]. The decoding process with the temporal attention mechanism is shown in Fig. 12.

In order to obtain the attention vector, the context vector $\mathbf{c}_{t'}$ of each output time period t' is computed, which depends on the annotation vector of the encoder, i.e., $(\mathbf{H}_1^C, \mathbf{H}_2^C, \ldots, \mathbf{H}_{t'-1}^C)$. The encoder maps the input spatial factors into these annotation vectors, and each annotation contains the information about the input time sequence. The context vector $\mathbf{c}_{t'}$ is calculated as a weighted sum of the annotation vectors $\mathbf{H}_i^C$, as

$$\mathbf{c}_{t'} = \sum_{i=1}^{T} \alpha_i^{t'} \mathbf{H}_i^C \tag{18}$$

where the weight $\alpha_i^{t'}$ of each annotation vector $\mathbf{H}_i^C$ is calculated as

$$\alpha_i^{t'} = \text{softmax}\left(u_i^{t'}\right) = \frac{\exp\left(u_i^{t'}\right)}{\sum_{j=1}^{T} \exp\left(u_j^{t'}\right)}, \text{ and} \tag{19}$$

$$u_i^{t'} = \mathbf{v}_d^\top \tanh\left(\mathbf{W}_1\left[\mathbf{d}_{t'-1}; \mathbf{s}'_{t'-1}\right] + \mathbf{W}_2 \mathbf{H}_i^C + \mathbf{b}_d\right) \tag{20}$$

where $\mathbf{W}_1 \in \mathbb{R}^{m \times 2n}$, $\mathbf{W}_2 \in \mathbb{R}^{m \times m}$, $\mathbf{v}_d$, and $\mathbf{b}_d \in \mathbb{R}^m$ are model parameters, $\mathbf{d}_{t'-1} \in \mathbb{R}^n$ is the state of the previous time step $t'-1$ in the decoder, and $\mathbf{s}_{t'-1} \in \mathbb{R}^n$ is the cell state of the decoder at time $t'-1$. $\alpha_i^{t'}$ or its associated quantity $u_i^{t'}$ indicates the importance of the annotation vector $\mathbf{H}_i^C$ in determining the next hidden state compared to the previous hidden state. Once the external meteorological information $\mathbf{ex}_t$, the context vector $\mathbf{c}_t$, and the previous temperature value $\hat{y}_{t'-1}^i$ are acquired, they are fed into the decoder LSTM, and the hidden state of the decoder is updated as

$$\mathbf{d}_{t'} = f_d\left(\mathbf{d}_{t'-1}, \left[\hat{y}_{t'-1}^i; \mathbf{ex}_{t'}; \mathbf{c}_{t'}\right]\right) \tag{21}$$

where $f_d(\cdot)$ is the LSTM unit used in the decoder and t' is a future time step. Then, we feed the context vector $\mathbf{c}_{t'}$ and the hidden state $\mathbf{d}_{t'}$ to the LSTM unit. Finally, the decoder is trained to obtain the grain temperature of the target point at the next time, which is given by

$$\hat{y}_{t'}^i = \mathbf{v}_y^\top \left(\mathbf{W}_m[\mathbf{c}_{t'}; \mathbf{d}_{t'}] + \mathbf{b}_m\right) + b_y \tag{22}$$

where $\mathbf{W}_m \in \mathbb{R}^{n \times (m+n)}$ and $\mathbf{b}_m \in \mathbb{R}^n$ map the concatenated $[\mathbf{c}_{t'}; \mathbf{d}_{t'}] \in \mathbb{R}^{m+n}$ to the size of the decoder hidden state; and $\mathbf{v}_y \in \mathbb{R}^n$ and $\mathbf{b}_y \in \mathbb{R}$ are linear transformation parameters, which can predict the final output. The Adam optimizer is used to train the model to minimize the root mean-square error (RMSE) between the predicted temperature value $\hat{\mathbf{y}}^i$ and the ground-truth temperature value $\mathbf{y}^i$.

IV. Performance Evaluation

The grain temperature data set used in our performance evaluation is obtained from a temperature sensor network deployed in a large warehouse in Zhumadian, Henan Province, China. The corresponding meteorological data are obtained from the China meteorological data network. The data set includes daily grain temperature measured by the temperature sensors from July 3, 2015 to December 17, 2018. The frequency of measuring temperature data with a temperature sensor is three times a week. We interpolate the grain temperature data to obtain daily temperature samples, which is then synchronized with the daily meteorological data.

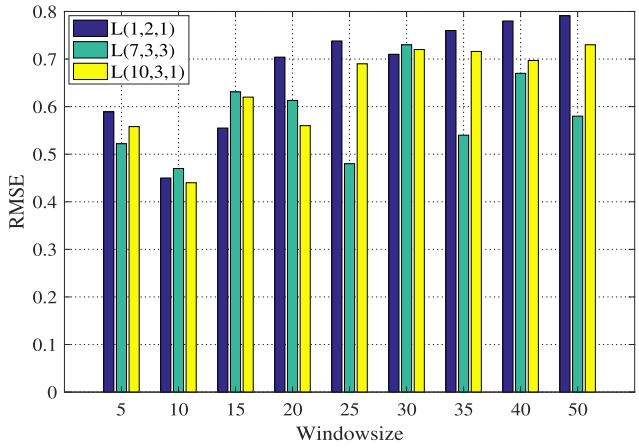

Fig. 13. RMSEs of grain temperature prediction under different time window sizes.

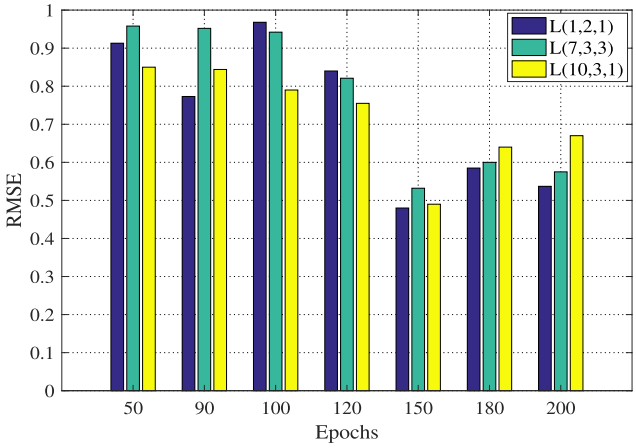

Fig. 14. RMSEs under different epochs.

In this article, model training is implemented using the tensorflow library. In the model training, the initial learning rate is 0.001, and the coefficient of avoiding gradient explosion is 2.5. Data expansion and dropout [34], [35] are used to avoid overfitting, and the ratio of dropout is set to 0.3. We use the RMSE as a performance metric for the accuracy of grain temperature prediction, given by

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(|y_i - \hat{y}_i|^2\right)} \tag{23}$$

where y_i is the ground-truth grain temperature data collected by the sensor network, $\hat{y}_i$ is the grain temperature data predicted by the proposed model, and n represents the number of grain temperature samples.

A. Parameter Tuning Results

First, we train the parameters that affect the performance of the model to select the appropriate model parameters to ensure the accuracy of prediction. We then analyze the influence of time window size, epochs, and batch size on the prediction performance of the model.

1) Impact of Time Window Size: We set the time window size to 5, 10, 15, 20, 25, 30, 35, 40, and 50 to test the effect of time window size on the prediction accuracy. In order to ensure that the results of different time window sizes are not affected by other variables, we assume that the values of other parameters are fixed. The hidden units, epochs, and batch size are set to 128, 150, and 100, respectively. Fig. 13 shows the RMSE of the prediction model under different time window sizes. When the time window size is 10, the prediction errors of three target temperature locations [including two boundary locations $L(1, 2, 1)$ and $L(10, 3, 1)$, and one center location $L(7, 3, 3)$] are all the minimum. In addition, we find that with the increase of time window size, the model training time becomes longer and the training speed becomes slower. Therefore, we choose 10 as the time window size of the prediction model by considering both the prediction accuracy and the training time of the model.

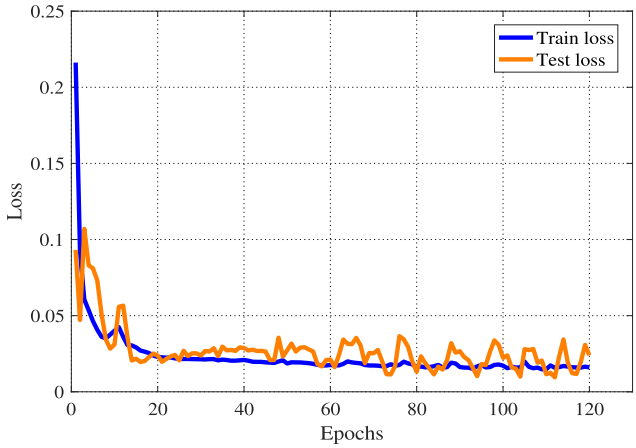

Fig. 15. Training loss and testing loss when epoch is 120.

2) Impact of Epochs and Batch Size: An epoch is when a complete data set passes through the model once and returns once. Batch size refers to the number of samples selected for one training process. For different temperature locations, we train the model with different epoch values and then select the appropriate epoch value by analyzing the resulting prediction errors. Fig. 14 plots the effect of different epoch values on the prediction error. It can be seen from Fig. 14 that the RMSE is the largest when the epoch value is 50. When the value of epoch is greater than 150, the RMSE becomes smaller. However, in addition to the RMSE, the trend of training loss and testing loss should also be considered. As can be seen from Figs. 15–17, the testing loss exhibits a trend of divergence when the value of epoch is less than 150, which indicates that the model is under fitting. However, when the epoch value is greater than 150, the testing loss is greater than the training loss, which indicates that the model is overfitting. As a result, 150 is selected as the optimal value of epoch.

In the training process, we also investigate the influence of batch size on the model performance. We find that the value of batch size has a great influence on the convergence of training. When the value of batch size is small, the training loss curve oscillates and does not converge to a stable value. As the value of batch size is increased, the loss curve becomes smooth and convergent. Finally, we choose 100 as the value of batch size.

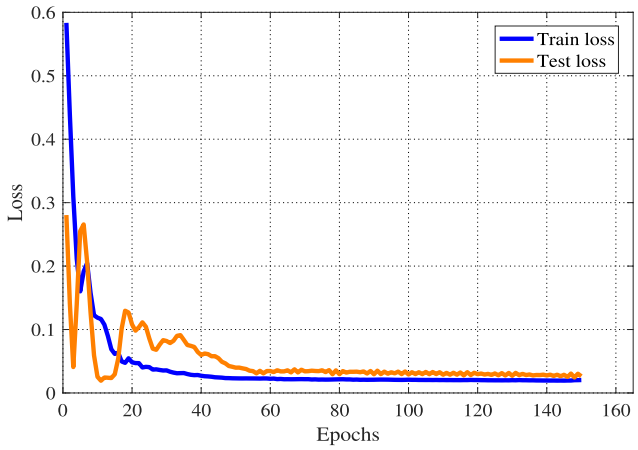

Fig. 16. Training loss and testing loss when epoch is 150.

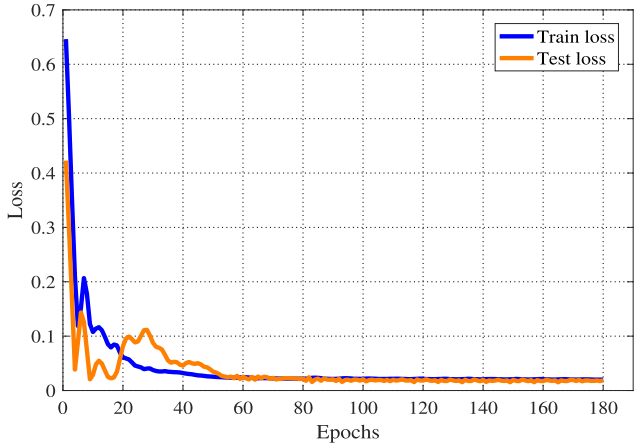

Fig. 17. Training loss and testing loss when epoch is 180.

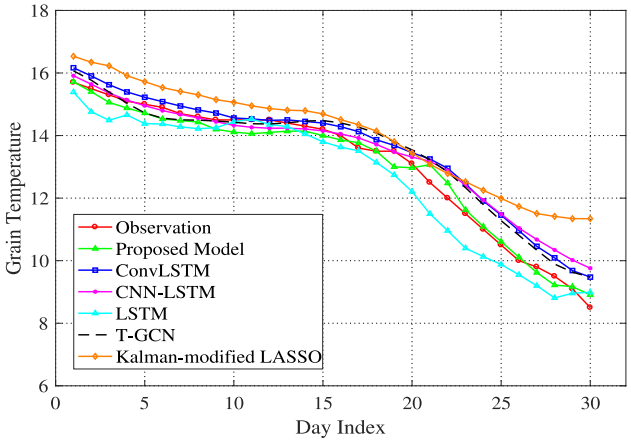

Fig. 18. Grain temperature prediction results at location (row 1, column 1, and layer 1) of stored grain.

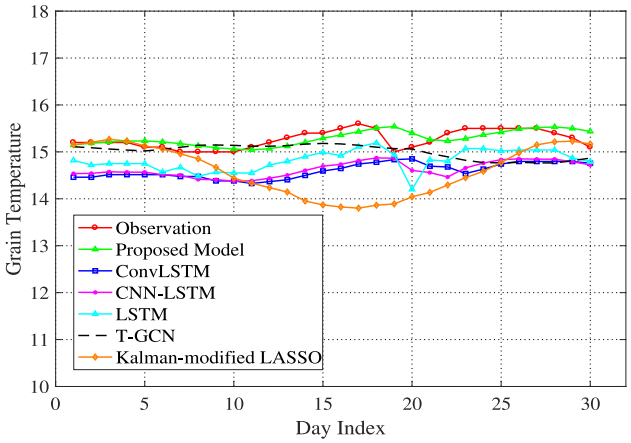

Fig. 19. Grain temperature prediction results at location (row 1, column 5, and layer 4) of stored grain.

B. Comparison Study With State-of-the-Art Schemes

In order to verify the performance of the proposed model, we predict the grain temperature at different target locations, including the boundary of the grain pile (i.e., the the upper, lower, southeast, northeast, southwest, and northwest boundary locations) and near the center of the grain pile. We compare the performance of the proposed model with the following five baseline schemes selected from the recent literature in the following comparison study.

1) Kalman-modified LASSO [36], which first filters the time series, then removes the corresponding variables from the model to perform variable selection, and in this case, sparse solution is realized.
2) A temporal graph convolutional network (T-GCN), which combines graph convolution with gating cycle unit [37].
3) LSTM network [38], which is suitable for dealing with and predicting the problems with relatively long interval and delay in time series.
4) Convolutional LSTM network (ConvLSTM) [39], for the matrix obtained by the sensor network, convolution operation is added to LSTM to extract features effectively.
5) A convolutional neural network (CNN)-long short-term memory (LSTM) combined network (CNN-LSTM) [40], where CNN is used to extract the correlation between multiple sequences recorded by the sensors, and LSTM is used to model and forecast the extracted time information.

The results are presented in Figs. 18–23. In particular, Figs. 18 and 19 present the grain temperature prediction results at two locations close to the eastern boundary of the granary; Fig. 18 is for the sensor close to the southern side of the granary; and Fig. 19 is for the sensor close to the northern side of the granary. Moreover, Figs. 20 and 21 present the grain temperature prediction results for the sensors in the center of the grain pile. Finally, Figs. 22 and 23 present the results for the sensors close to the western boundary of the granary. Fig. 22 is for the sensor close to the northern side of the granary and Fig. 23 is for the sensor close to the southern side of the granary. Figs. 18 and 22 are for the sensors on the upper layer of the grain pile. Figs. 19 and 23 are for the sensor on the lower layer of the grain pile. The results show that the proposed model can accurately predict temperature for all the positions in the grain pile.

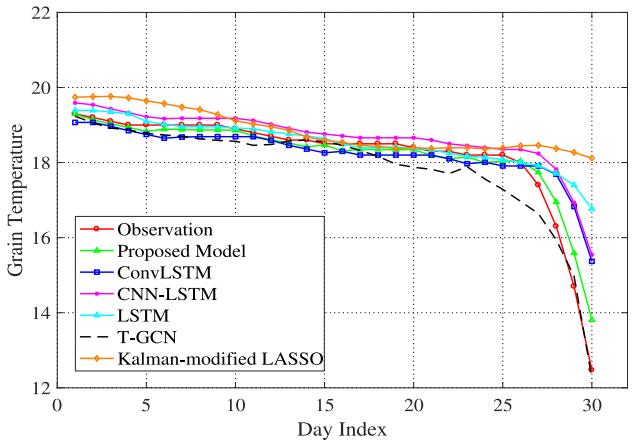

Fig. 20. Grain temperature prediction results at location (row 5, column 3, and layer 3) of stored grain.

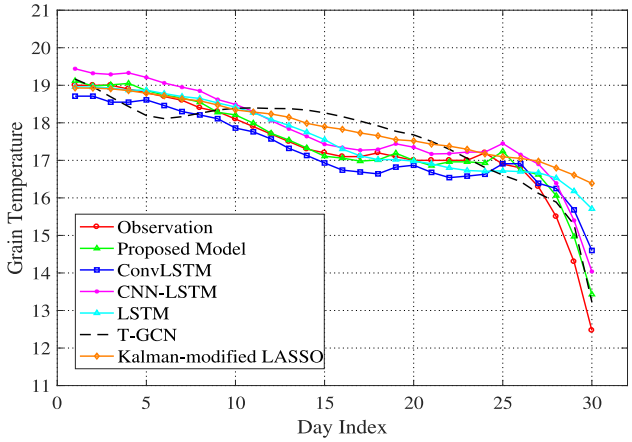

Fig. 21. Grain temperature prediction results at location (row 6, column 3, and layer 3) of stored grain.

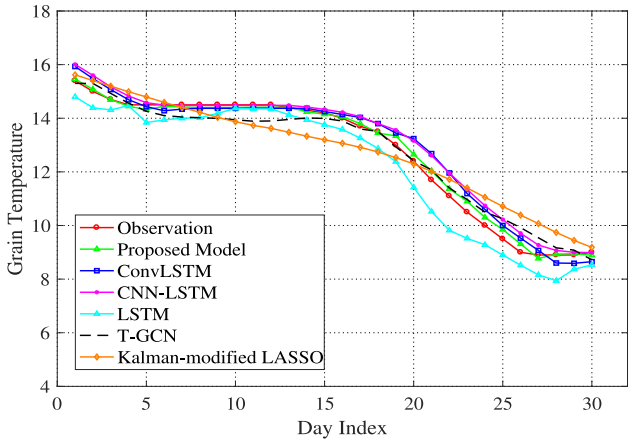

Fig. 22. Grain temperature prediction results at location (row 10, column 5, and layer 1) of stored grain.

The RMSE results of the six prediction models are presented in Table II. For the sensor at (1, 1, 1), i.e., located at the upper layer of the southeast boundary, the RMSE of the proposed model is 0.2794, which achieves a reduction of 48% over the ConvLSTM model, 50% over the CNN-LSTM model, 52%

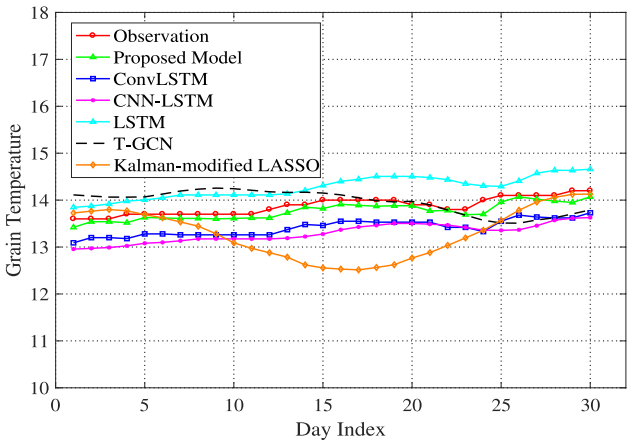

Fig. 23. Grain temperature prediction results at location (row 10, column 1, and layer 4) of stored grain.

TABLE II
RMSEs ACHIEVED BY THE SIX APPROACHES AT DIFFERENT LOCATIONS

RMSEs of different points	(1,1,1)	(1,5,4)	(5,3,3)	(6,3,3)	(10,1,4)	(10,5,1)
Proposed	0.28	0.17	0.34	0.27	0.13	0.18
ConvLSTM	0.54	0.69	0.74	0.56	0.47	0.42
CNN-LSTM	0.56	0.65	0.8	0.51	0.59	0.46
LSTM	0.58	0.47	0.97	0.75	0.41	0.63
T-GCN	0.49	0.38	0.42	0.60	0.39	0.41
Kalman-modified LASSO	0.85	0.9	1.33	0.95	0.81	0.73

over the LSTM model, 43% over the T-GCN model, and 67% over the Kalman-modified LASSO method. For the sensor at (1, 5, 4), i.e., located at the lower layer of the northeast boundary, the RMSE of the proposed model is 0.1736, which achieves a reduction of 74% over the ConvLSTM model, 73% over the CNN-LSTM model, 63% over the LSTM model, 54% over the T-GCN model, and 81% over the Kalman-modified LASSO method. For the sensor at (5, 3, 3), i.e., located in the center of the grain pile, the RMSE of the proposed model is 0.3423, which achieves a reduction of 54% over the ConvLSTM model, 57% over the CNN-LSTM model, 65% over the LSTM model, 19% over the T-GCN model, and 74% over the Kalman-modified LASSO method. For the other sensor at (6, 3, 3), i.e., located in the center of the grain pile, the RMSE of the proposed model is 0.2684, which achieves a reduction of 52% over the ConvLSTM model, 47% over the CNN-LSTM model, 64% over the LSTM model, 55% over the T-GCN model, and 72% over the Kalman-modified LASSO method. For the sensor at (10, 1, 4), i.e., located at the lower layer of the southwest boundary, the RMSE of the proposed model is 0.1307, which achieves a reduction of 72% over the ConvLSTM model, 78% over the CNN-LSTM model, 68% over the LSTM model, 66% over the T-GCN model, and 84% over the Kalman-modified LASSO method. For the sensor at (10, 5, 1), i.e., located at the upper layer of the northwest boundary, the RMSE of the proposed model is 0.1817, which achieves a reduction of 57% over the ConvLSTM model, 60% over the CNN-LSTM model, 71% over the LSTM model, 56% over the T-GCN model, and 75% over the Kalman-modified LASSO method. Among the five other schemes, the statistics method Kalman-modified LASSO has the poorest

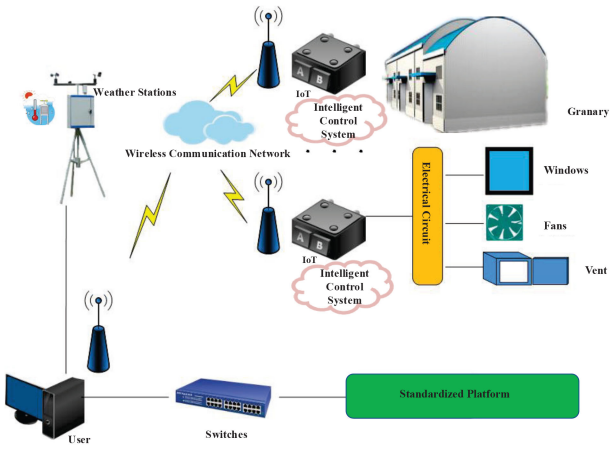

Fig. 24. Intelligent control system of grain situation with IoT technologies integration.

performance, compared to other deep learning-based schemes. The proposed model outperforms all the five other schemes with considerable gains.

C. Application of Developed Prediction Models

The developed prediction models can be used to guide the operation of the intelligent control system of grain storage, as shown in Fig. 24. The intelligent control system of grain depot can acquire temperature and humidity index automatically through the grain condition detection system in real time. Data transmission is carried out by the IoT, network equipment, and communication module. The computer evaluates and intelligently analyzes the information of grain and environment in the granary, and forms the action control strategy or auxiliary suggestions. It is of great importance to realize the temperature reduction, ventilation, dehumidification, and insect control by intelligent control and storage technologies, such as ventilation equipment, air conditioning equipment, low temperature control equipment, and circulation fumigation, and to achieve automatic monitoring of high efficiency, high reliability, and automation of grain warehouse, which is of great importance. for reducing the cost of grain storage, maintaining the quality of stored grain, reducing grain loss, and improving the storage capacity of grain safety.

V. Related Work

The IoT has become a critical technology toward the future precision agriculture, while safe storage of grain is becoming increasingly concerned of the world. In this section, we review related works with respect to deep learning models, spatiotemporal prediction models, and the attention mechanism of deep learning.

CNN is an end-to-end learning model in the field of deep learning, which connects two adjacent layers through global sliding, local connection, and weight sharing, making the network structure simpler and more adaptive [41]. LSTM is an improved model based on RNNs. The basic unit of the LSTM hidden layer is a special cell structure rather than a traditional neuron node. Through the input gate, output gate, and forgetting gate in the cell structure, the inflow and outflow of information and the update of the previous state are realized, respectively. CNN and LSTM were integrated and used in a PM2.5 forecasting system that considers the factors, such as accumulated hours of rain, wind speed, and PM2.5 concentration [42]. LSTM, CNN, and multilayer perceptron (MLP) have been used to predict the electricity demand [43]. A deep long short-time memory (DLSTM) model, which uses the genetic algorithm to optimize its allocation, was developed to accurately predict oil production [44].

Meanwhile, approaches that consider both temporal and spatial impacts have been developed for accurate prediction. A temporal approach that is based on LASSO and LSTM has been integrated for forecasting short-term solar intensity using meteorological data [45]. Lyu *et al.* [46] proposed a smart WiFi access point (AP) management program LAM, which can dynamically control a large amount of APs (i.e., turning on or off) without losing WiFi coverage to achieve the goal of energy saving by mining large-scale spatiotemporal user association data. A spatiotemporal visual question answering model using two dual-layer LSTMs was presented to understand visual content and find the associations of pairs of questions and answers in the natural language form [47]. A deep-metalearning-based spatial and temporal approach was utilized to overcome the diversification and sophistication of urban traffic forecasting [48]. Deep RNNs, such as LSTM and gated recurrent unit (GRU) networks, were utilized for forecasting traffic load of base stations [49]. A two-branch deep learning model was established to predict winter wheat yield in the main producing regions of China at the county level [50]. The deep belief network (DBN) GRU hybrid neural network based on deep learning was used to predict meteorological time series, which avoids the problems of low generalization ability and long training time of RNNs [51]. Furthermore, a deep hybrid spatial and temporal dynamic neural network (DHSTNet) was utilized for forecasting both inflows and outflows at different positions along conurbation [52]. A spatiotemporal position forecasting approach on account of LSTM was used for forecasting the next spatiotemporal track of mobile user [53]. A deep spatial model consisting of a global stacked autoencoder (GSAE) and multiple local SAEs (LSAEs), which can reduce the model size and enable parallel training, was utilized for forecasting the mobile traffic load in cellular networks [54].

Compared with conventional spatiotemporal approaches, the attention mechanism can significantly decrease the number of parameters in the deep spatiotemporal process, and adaptively choose relevant temporal and spatial features to enhance temperature forecasting performance. Furthermore, the computation of the attention mechanism in each step is independent to the results of the previous step, which is amenable for parallel processing in GPU-accelerated deep models. An interpretable spatiotemporal attention LSTM model (STA-LSTM) based on LSTM and the attention mechanism was utilized for flood prediction, where the visualization and interpretation of spatial and temporal attention weights represent the rationality of the attention-based model [55]. A new attention mechanism was used to select the relevant time series, and its frequency-domain information is utilized to make multiple predictions.

It achieved superior performance in many practical tasks, and overcame the shortcoming that the traditional attention mechanism could not capture time patterns across multiple time steps [33]. The attention mechanism has also been effectively used in geography-sensor data prediction [27] and intelligent traffic prediction [56], [57] due to its advantages of using fewer parameters, fast execution, and superior performance.

VI. Conclusion

It is of great practical importance to develop IoT systems of intelligent grain state control for safe grain storage. In this article, a new temporal and spatial approach based on the attention mechanism was proposed to predict the temperature of stored grain. The proposed model captures the temporal and spatial correlation structure in grain temperature data, as well as the influence of external meteorological factors. It overcomes the limitation of traditional prediction methods, which only consider historical temperature data. Our experimental study using real granary sensor network and meteorological data validate the superior prediction performance of the proposed model over five other schemes.

For future work, we will investigate what is a suitable timescale for temperature measurement/forecasting for food storage. The developed grain temperature prediction model will be used in the development of the IoT system for grain intelligent control, so as to realize the automatic notification of dangerous grain situation, optimize the layout of sensors in the grain pile, and reduce the use of sensors and conserve energy.

References

[1] C. L. Mouël and A. Forslund, "How can we feed the world in 2050? a review of the responses from global scenario studies," *Eur. Rev. Agr. Econ.*, vol. 44, no. 4, p. 541–591, 2017.

[2] P. Pinstrup-Andersen, R. Pandya-Lorch, and M. W. Rosegrant, "The world food situation: Recent developments, emerging issues, and long term prospects," Int. Food Policy Res. Inst., Washington, DC, USA, Food Policy Rep., Dec. 1997.

[3] Food and Agriculture Organization of United Nations. (Nov. 2020). *World Food Situation*. [Online]. Available: http://www.fao.org/worldfoodsituation/csdb/en

[4] D. S. Jayas, "Storing grains for food security and sustainability," *Agr. Res.*, vol. 1, no. 1, pp. 21–24, Mar. 2012.

[5] C. B. Singh and J. M. Fielke, "Recent developments in stored grain sensors, monitoring and management technology," *IEEE Instrum. Meas. Mag.*, vol. 20, no. 3, pp. 32–55, Jun. 2017.

[6] W.-L. Chen, Y.-B. Lin, F.-L. Ng, C.-Y. Liu, and Y.-W. Lin, "RiceTalk: Rice blast detection using Internet of Things and artificial intelligence technologies," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1001–1010, Feb. 2020.

[7] A. D. Boursianis *et al.*, "Internet of Things (IoT) and agricultural unmanned aerial vehicles (Uavs) in smart farming: A comprehensive review," *Elsevier Internet Things J.*, to be published.

[8] A. W. Malik, A. U. Rahman, T. Qayyum, and S. D. Ravana, "Leveraging fog computing for sustainable smart farming using distributed simulation," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 3300–3309, Apr. 2020.

[9] W. Yang, X. Wang, A. Song, and S. Mao, "Wi-Wheat: Contact-free wheat moisture detection with commodity WiFi," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kansas City, MO, USA, May 2018, pp. 1–6.

[10] W. Yang, X. Wang, S. Cao, H. Wang, and S. Mao, "Multi-class wheat moisture detection with 5GHz Wi-Fi: A deep LSTM approach," in *Proc. 27th Int. Conf. Comput. Commun. Netw. (ICCCN)*, Hangzhou, China, Jul./Aug. 2018, pp. 1–9.

[11] P. Hu, W. Yang, X. Wang, and S. Mao, "MiFi: Device-free wheat mildew detection using off-the-shelf WiFi devices," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[12] A. K. Mishra *et al.*, "A chain topology for efficient monitoring of food grain storage using smart sensors," in *Proc. 15th Int. Joint Conf. e-Bus. Telecomm.*, Porto, Portugal, Jul. 2018, pp. 89–98.

[13] J. Lawrence, D. E. Maier, and R. L. Stroshine, "Three-dimensional transient heat, mass, momentum, and species transfer in the stored grain ecosystem: Part I. Model development and evaluation," *Trans. ASABE*, vol. 56, no. 1, pp. 179–188, Jan. 2013.

[14] S. S. Panigrahi, C. B. Singh, J. Fielke, and D. Zare, "Modeling of heat and mass transfer within the grain storage ecosystem using numerical methods: A review," *Drying Technol.*, vol. 38, no. 13, pp. 1677–1697, Aug. 2019.

[15] S. Duan, W. Yang, X. Wang, S. Mao, and Y. Zhang, "Grain pile temperature forecasting from weather factors: A support vector regression approach," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Changchun, China, Aug. 2019, pp. 255–260.

[16] S. Duan, W. Yang, X. Wang, S. Mao, and Y. Zhang, "Forecasting of grain pile temperature from meteorological factors using machine learning," *IEEE Access*, vol. 7, pp. 130721–130733, 2019.

[17] S. Duan, W. Yang, X. Wang, S. Mao, and Y. Zhang, "Deep spatiotemporal attention model for grain storage temperature forecasting," in *Proc. IEEE 26th Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Hong Kong, Dec. 2020, pp. 593–600.

[18] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*. New York, NY, USA: Springer, 2016.

[19] B. B. Mandelbrot and J. R. Wallis, "Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence," *Water Resour. Res.*, vol. 5, no. 5, pp. 967–988, Oct. 1969.

[20] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 25, no. 1, pp. 203–213, Oct. 1993.

[21] G. He, Y. Gao, J. C. Hou, and K. Park, "A case for exploiting self-similarity of network traffic in TCP congestion control," *Elsevier Comput. Netw.*, vol. 45, no. 6, pp. 743–766, Aug. 2004.

[22] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise Reduction in Speech Processing*, J. Benesty, J. Chen, Y. Huang, and I. Cohen, Eds. New York, NY, USA: Springer, 2009, pp. 1–4.

[23] R. C. Gonzales and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[24] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. Bengaluru, India: Pearson Educ. India, 2004.

[25] Y. Zhang, X. Han, H. Zhang, and L. Zhao, "Edge detection algorithm of image fusion based on improved Sobel operator," in *Proc. IEEE 3rd Inf. Technol. Mechatronics Eng. Conf.*, Chongqing, China, Oct. 2017, pp. 457–461.

[26] M. Zanfir, E. Marinoiu, and C. Sminchisescu, "Spatio-temporal attention models for grounded video captioning," in *Proc. Asian Conf. Comput. Vis.*, Taipei, Taiwan, Nov. 2016, pp. 104–119.

[27] Y. Liang, S. Ke, J. Zhang, X. Yi, and Y. Zheng, "GeoMAN: Multi-level attention networks for geo-sensory time series prediction," in *Proc. 27th Int. Joint Conf. Artif. Intell. (IJCAI)*, Stockholm, Sweden, Jul. 2018, pp. 3428–3434.

[28] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.

[29] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun./Jul. 2016, pp. 2818–2826.

[30] X. Jin *et al.*, "ILGNet: Inception modules with connected local and global features for efficient image aesthetic quality classification using domain adaptation," *IET Comput. Vis.*, vol. 13, no. 2, pp. 206–212, Mar. 2019.

[31] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," Sep. 2014. [Online]. Available: https://arxiv.org/abs/1409.0473.

[32] K. Xu *et al.*, "Show, attend and tell: Neural image caption generation with visual attention," in *Proc. 32nd Int. Conf. Mech. Learn. (ICML)*, Lille, France, Jul. 2015, pp. 2048–2057.

[33] S.-Y. Shih, F.-K. Sun, and H.-Y. Lee, "Temporal pattern attention for multivariate time series forecasting," *Springer Mach. Learn. J.*, vol. 108, pp. 1421–1441, Jun. 2019.

[34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 56, pp. 1929–1958, Jun. 2014.

[35] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Conf. Neural Inf. Process. Syst. (NIPS)*, Lake Tahoe, NV, USA, Dec. 2012, pp. 1097–1105.

[36] N. Tang, S. Mao, Y. Wang, and R. M. Nelms, "Solar power generation forecasting with a LASSO-based approach," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1090–1099, Apr. 2018.

[37] L. Zhao *et al.*, "T-GCN: A temporal graph convolutional network for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 3848–3858, Sep. 2020.

[38] L. Mei *et al.*, "Realtime mobile bandwidth prediction using LSTM neural network and Bayesian fusion," *Elsevier Comput. Netw.*, vol. 182, no. 9, Dec. 2020, Art. no. 107515.

[39] S. W. Lee and H. Y. Kim, "Stock market forecasting with super-high dimensional time-series data using convlstm, trend sampling, and specialized data augmentation," *Expert Syst. Appl.*, vol. 161, Dec. 2020, Art. no. 113704.

[40] J. Sun, L. Di, Z. Sun, Y. Shen, and Z. Lai, "County-level soybean yield prediction using deep CNN-LSTM model," *MDPI Sens.*, vol. 19, no. 20, p. 4363, Oct. 2019.

[41] P. Robinson, *The CNN Effect: The Myth of News, Foreign Policy and Intervention*. Abingdon, U.K.: Routledge, 2005.

[42] C.-J. Huang and P.-H. Kuo, "A deep CNN-LSTM model for particulate matter ($PM_{2.5}$) forecasting in smart cities," *MDPI Sens. J.*, vol. 18, no. 7, p. 2220, Jul. 2018.

[43] S. A. Adewuyi, S. Aina, and A. I. Oluwaranti, "A deep learning model for electricity demand forecasting based on a tropical data," *Appl. Comput. Sci.*, vol. 16, no. 1, pp. 5–17, Jan. 2020.

[44] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep LSTM recurrent networks," *Elsevier Neurocomput.*, vol. 323, no. 1, pp. 203–213, Jan. 2019.

[45] Y. Wang, Y. Shen, S. Mao, X. Chen, and H. Zou, "LASSO & LSTM integrated temporal model for short-term solar intensity forecasting," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2933–2944, Apr. 2019.

[46] F. Lyu, L. Fang, G. Xue, H. Xue, and M. Li, "Large-scale full WiFi coverage: Deployment and management strategy based on user spatio-temporal association analytics," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9386–9398, Dec. 2019.

[47] Y. Jang, Y. Song, Y. Yu, Y. Kim, and G. Kim, "TGIF-QA: Toward spatio-temporal reasoning in visual question answering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2758–2766.

[48] Z. Pan, Y. Liang, W. Wang, Y. Yu, Y. Zheng, and J. Zhang, "Urban traffic prediction from spatio-temporal data using deep meta learning," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, Anchorage, AK, USA, Aug. 2019, pp. 1720–1730.

[49] B. Mahdy, "Analysis of the utilization of mobile network base stations using traffic load predictions," Ph.D. dissertation, Dept. Elect. Comput. Eng., Queen's Univ., Kingston, ON, Canada, May 2020.

[50] X. Wang, J. Huang, Q. Feng, and D. Yin, "Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches," *MDPI Remote Sens. J.*, vol. 12, no. 11, p. 1744, May 2020.

[51] Y. Cheng, X. Zhou, S. Wan, and K.-K. R. Choo, "Deep belief network for meteorological time series prediction in the Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4369–4376, Jun. 2019.

[52] A. Ali, Y. Zhu, Q. Chen, J. Yu, and H. Cai, "Leveraging spatio-temporal patterns for predicting citywide traffic crowd flows using deep hybrid neural networks," in *Proc. IEEE 25th Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Tianjin, China, Dec. 2019, pp. 125–132.

[53] S. Tian, X. Zhang, Y. Zhang, Z. Cao, and W. Cao, "Spatio-temporal position prediction model for mobile users based on LSTM," in *Proc. IEEE 25th Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Tianjin, China, Dec. 2019, pp. 967–970.

[54] J. Wang *et al.*, "Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Atlanta, GA, USA, May 2017, pp. 1–9.

[55] Y. Ding, Y. Zhu, J. Feng, P. Zhang, and Z. Cheng, "Interpretable spatio-temporal attention LSTM model for flood forecasting," *Elsevier Neurocomput. J.*, vol. 403, pp. 348–359, Aug. 2020.

[56] X. Shi, H. Qi, Y. Shen, G. Wu, and B. Yin, "A spatial-temporal attention approach for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, early access, Apr. 13, 2020, doi: 10.1109/TITS.2020.2983651.

[57] C. Park *et al.*, "ST-GRAT: A novel spatio-temporal graph attention network for accurately forecasting dynamically changing road speed," Nov. 2019. [Online]. Available: https://arxiv.org/abs/1911.13181.