RESEARCH ARTICLE

JOURNAL OF SUSTAINABLE
AGRICULTURE AND ENVIRONMENT

Climate and environmental data contribute to the prediction of grain commodity prices using deep learning

Zilin Wang[1] | Niamh French[2] | Thomas James[2] | Calogero Schillaci[3] | Faith Chan[4,5,6] | Meili Feng[4] | Aldo Lipani[1]

[1]University College London (UCL), London, UK

[2]Wegaw SA, Trélex, Switzerland

[3]European Commission, Joint Research Centre (JRC), Ispra, VA, Italy

[4]School of Geographical Sciences, University of Nottingham Ningbo China, Ningbo, China

[5]Water@Leeds and School of Geography, University of Leeds, Leeds, UK

[6]Research Centre for Intelligent Management & Innovation Development/Research Base for Shenzhen Municipal Policy & Development, Southern University of Science and Technology, Shenzhen, China

Correspondence
Aldo Lipani, University College London (UCL), London, UK.
Email: aldo.lipani@ucl.ac.uk

Funding information
None

Abstract

Background: Grain commodities are important to people's daily lives and their fluctuations can cause instability for households. Accurate prediction of grain prices can improve food and social security.

Methods & Materials: This study proposes a hybrid Long Short-Term Memory (LSTM)-Convolutional Neural Network (CNN) model to forecast weekly oat, corn, soybean and wheat prices in the United States market. The LSTM-CNN is a multivariate model that uses weather data, macroeconomic data, commodities grain prices and snow factors, including Snow Water Equivalent (SWE), snowfall and snow depth, to make multistep ahead forecasts.

Results: Of all the features, the snow factor is used for the first time for commodity price forecasting. We used the LSTM-CNN model to evaluate the 5, 10, 15 and 20 weeks ahead forecasting and this hybrid model had the lowest Mean Squared Error (MSE) at 5, 10 and 15 weeks ahead of prediction. In addition, Shapley values were calculated to analyse the feature contribution of the LSTM-CNN model when forecasting the testing set. Based on the feature contribution, SWE ranked third, fifth and seventh in feature importance in the 5-week ahead forecast for corn, oats and wheat, respectively, and 7–8 places higher than total precipitation, indicating the potential use of SWE in grain price forecasting.

Conclusion: The hybrid multivariate LSTM-CNN model outperformed other models and the newly involved climate data, SWE, showed the research potential of using snow as an input variable to predict grain prices over a multistep ahead time horizon.

KEYWORDS
commodity prices, LSTM-CNN, multistep ahead prediction, snow water equivalent

1 | INTRODUCTION

Recently, commodities have been considered a core asset class by financial investors. Research has shown that the inclusion of commodities in an investor's portfolio can hedge risk, particularly in turbulent times (e.g., the 2008 financial crisis) (Cheng & Xiong, 2013). Inaccurate forecasting of commodity prices can invite a commodity price collapse with spillover effects to other markets, especially in the current unstable world situation, including pandemics and climate change (Salisu et al., 2020). Of all the categories of commodities,

grain commodities prices play a critical role in everyone's daily life. Fluctuations in grain commodities prices pose a threat to consumers and lead to instability in the incomes and operations of farmers' households (Ayankoya et al., 2016).

Different approaches to agricultural commodity price forecasting have been tried, which can be divided into two major categories, namely univariate models and multivariate models. Univariate models use historical records of prices to make forecasts, such as Auto-Regressive Integrated Moving Average models (ARIMA) (Jadhav et al., 2017; Pujiati et al., 2018) and its variants (Adanacioglu & Yercan, 2012; Divisekara et al., 2021; Li et al., 2012; Mithiya et al., 2019; Naidu et al., 2014), which are statistical modelling technique in commodity price forecasting credited with the ability to capture time series trends, and as such have mostly achieved relatively high levels of accuracy. However, ARIMA and its variants require large datasets and cannot capture the nonlinear relationship between historical and future records, machine learning methods like artificial neural networks (ANN) in commodity price forecasting come into play (Monge & Lazcano, 2022; Wang et al., 2019).

Using univariate models has inherently limited predictability, as the prices are affected by numerous additional input variables (Shu & Gao, 2020). Multivariate variables are necessary to strengthen the performance of prediction models. Commonly used deep learning models include ANN (Ayankoya et al., 2016) and Long Short-Term Memory (LSTM) (Ly et al., 2021; Rasheed et al., 2021). These models have good predictive power in single-step forecasting. For comparison, we choose the LSTM as one of the comparative models. Nevertheless, some studies did not predict multistep ahead prices, and for those that did, the accuracy of multistep ahead predictions would be lower than that of single-step ahead forecasts.

Variables considered in multivariate modelling can be discussed from the supply and demand side. Yield is the main source of supply, which is determined by the area harvested (acreage) and yield per hectare. Acreage, which generally reflects the net return to farmers who grow food, experienced a rising in the US grain market during the 1980s and 1990s (Westcott & Hoffman, 1999), but the acreage in the United States is approaching the maximum capacity recently (USDA, 2022a). Factors such as climate and weather conditions mainly affect agricultural yields. Higher temperatures may not only reduce the time farmers spend in the field (Kjellstrom et al., 2009), but may also cut down the grain yields owing to water shortages and higher evaporation rates (Hertel & Lobell, 2014). Apart from this, for precipitation, Mendelsohn et al. (1994) noted that this factor may affect the price of grains. Hence, temperature and precipitation as easily obtainable weather factors are often taken as predictors in agricultural yields or pricing models (Cammarano et al., 2013; Gu et al., 2022; Nhita et al., 2019; Oktoviany et al., 2021).

Crude oil is another factor that can influence the grain price directly and indirectly from supply side. Farmers need energy to power machinery and fuel transport vehicles, as well as energy-intensive products such as pesticides and fertilisers to help them grow grains (Hitaj & Suttles, 2016). In this way, a major part of the costs of agricultural production stems from crude oil, implying that a significant portion of agricultural production costs is from crude oil. According to the USDA (2022b), the combined costs of chemicals, energy and electricity accounted for 8.8% of the cost of producing corn, 9.8% of soybeans and 10.4% of wheat in the United States (US) from 2010 to 2016. Hence, fluctuations in crude oil prices are often considered in agricultural price models.

Exports are a critical factor affecting the demand for US grains. Grain shortages and grain developments in other countries can increase or decrease the demand for US grains, thereby affecting their prices (Schwartz, 1986). Among the factors that influence export, exchange rates can exert an impact on export demand. If the dollar is weak, buyers of US exports will find the price of US commodities to be lower, and then the demand will increase respectively, leading to higher prices. Nonetheless, if the dollar rises, buyers will turn to the same commodities produced in other countries (Central for Agricultural and Rural Development, 2022). For example, wheat is cheaper in Argentina than in the United States, making cheaper alternatives available to countries such as Indonesia and the Philippines that traditionally import US corn to feed their animals (World Grain, 2016). Moreover, the fluctuations of other countries' currencies can also influence US grain (Chambers & Just, 1981). When the currencies of competitor countries rise, the US export demand rises accordingly even though the US dollar does not change (Central for Agricultural and Rural Development, 2022).

A set of environmental variables have been tested to predict the grain yield (Ayankoya et al., 2016; Ly et al., 2021; Rasheed et al., 2021), however, the snow-related indicators were not yet been included in predictive models. Snow is an essential resource, providing water and climate regulation for communities near mountainous areas (Qin et al., 2020; Sturm et al., 2017). For example, the snow melt, which is the surface runoff generated from melting snowpack, is a vital water resource feeding the downstream daily activities, such as agricultural production, hydroelectricity generation and water supply (Sturm et al., 2017). Grain production in the western US, southern Europe, western China and Central Asia is currently most dependent on snowmelt to support their irrigated agriculture, producing a large fraction of irrigated crops (e.g., wheat, corn and rice). According to a global study by Qin et al. (2020), grains such as wheat, soybeans and corns all rely partly on snowmelt for irrigation, with 50% of irrigation water provided by snowmelt for wheat, 38% for corn and 10%–20% for soybeans. In the United States, 17% of corn, 12% of soybeans and 12% of wheat are irrigated (Lopez et al., 2022). Snowmelt affects not only the production of irrigated crops but rainfed crops, such as oats, can also be affected by snowmelt. SWE can be regarded as a form of delayed precipitation. About 48% of the annual rainfall comes from melting snow, hence the indicator SWE reveals the amount of water that will be available as rainfall in the future, which may inform the weather conditions in a few months or weeks (Field & Heymsfield, 2015). The snowpack also acts as a natural water tower for storing winter precipitation, which melts in springtime. Abnormal low snowpacks can lead to water shortages and groundwater deficits, which can create instability in the production of rain-fed agriculture (Diffenbaugh et al., 2015; Moroizumi et al., 2009; Li et al., 2017). In such cases, taking the snow data into account may improve the performance in predicting commodity prices several weeks or months out.

This study focuses on a hybrid model that is capable of predicting the weekly prices of four grain commodities—oats, corn, soybeans and wheat—in the US market. The hybrid model LSTM-CNN which has not previously been studied in any deep learning predictions of grain prices can predict the prices of the four grain commodities in a multistep ahead horizon, which combines LSTM and CNN architectures. As a multivariate model, it not only involves traditional weather factors such as grain prices and macroeconomic factors to predict grain prices but also considers snow factors. The contributions of various factors for the prediction, in particular the snow factors, are investigated.

2 | METHODS

2.1 | Dataset description

In this study, we considered 17 variables that can be divided into three categories: weather, macroeconomics and the prices of the four crops. Three traditional weather factors, which were mean minimum temperature, mean maximum temperature and total precipitation for the US, were retrieved from the National Oceanic and Atmospheric Association (NOAA) (2022) website. Three novel weather factors, snow-related indicators, have been chosen, namely Snow Water Equivalent (SWE), snowfall and snow depth. SWE indicates the water content of the snowpack when it melts. In other words, SWE represents the amount of water in the snow that can become runoff (Seibert et al., 2015). Snow depth is the depth of the snowpack, and snowfall is the amount of snow that falls on a given day. SWE data were retrieved from the website of the United States Department of Agriculture (USDA), which contains daily records of SNOTEL sensor data covering all western US states (USDA, 2022c). Snowfall and snow depth data were retrieved from a comprehensive database called the Global Historical Climatology Network (GHCNd), which aggregates all daily climate records worldwide (National Oceanic and Atmospheric Administration, 2022). In this study, stations in the US that record snow depth and snowfall were searched to calculate weekly snowfall totals and average snow depth values.

The macroeconomics data included crude oil West Texas Intermediate (WTI) prices, gold prices, and four exchange rates between the US and its four top destinations of grain export (Commodity, 2022), which were the USD and CAD, EUR, CNY and MXN respectively. All the macroeconomics data, spanning from 1990 to 2021, were available at www.Investing.com.

The selected grains were oats, corn, soybeans and wheat (including spring wheat and winter wheat), which are staple foods regularly consumed by humans and are therefore substitutes for each other and compete in the global market.

2.2 | Data preparation

In this study, all variables are time series data, which need to be processed into suitable datasets that can be utilised by deep learning

architectures. First, we address the issue of missing values, since some features are monitoring data such as temperature, precipitation, snowfall, SWE and snow depth are often confronted with this issue. As missing values can affect the training and evaluation process of the model, the forward fill technique is used to replace missing entries in the dataset with the last valid observations. Based on the ADF tests (Supporting Information: Table S1), this data set is nonstationary. In other words, the inherent trends and seasonality of the time series data affect the values of the time series across time. To remove the trend from the time series, data differencing is performed to transform the nonstationary data into stationary data. In this study, first-order differencing is performed, where the previous value is subtracted from each value in the series.

To ensure that all features are in the same range of values, we apply the Min–Max scalar function to normalise all feature vectors in the dataset. MinMax normalisation uses the minimum and maximum values of the observations to convert values in the range of 0 and 1.

After data transformation and normalisation, the dataset is divided into two parts. 70% of the dataset is used for training the model and 30% for testing it. The entire dataset is then regrouped using the sliding window method to create a new series of sample data containing both input and target variables (see Supporting Information: Figure S1). Each sample has 20 input time steps (X) and the target variable can contain 5- or 10-time steps (Y) depending on the prediction target (e.g., 5 weeks ahead or 10 weeks ahead).

To avoid over-fitting, there is a need to apply cross-validation when training the model. Since we are dealing with time series the rolling forecasting origin is used. A small set of data from the origin is utilised to train the model, and the next set of data is used to evaluate it (Hyndman & Athanasopoulos, 2021). The first two sets of data are then combined and used as the training set for the next training round, and another set is used as a validation set and so on, until all the training sets have been used to train the model. The training set has been divided into five folds on a rolling forecasting origin. When using the testing dataset to evaluate the model, the last fold, which is the entire training set, is adopted to train and validate the model. A visual illustration is given in Figure 1.

2.3 | Proposed LSTM-CNN

A schematic illustration of the structure of the proposed hybrid model is provided in Figure 2. Instead of having a CNN accept the input data, the LSTM-CNN model has the LSTM as the initial layer accepting the time series data and extracting valuable patterns from the information in the time-dependent and stored blocks. In their study, He et al. (2019) examined the performance of CNN-LSTM and LSTM-CNN for gold price prediction. It turned out that the LSTM-CNN model performed better than the CNN-LSTM model, probably because using the LSTM layer as the starting layer allows each input unit to have an output unit with the memory/information of all the other units already processed. Afterwards, the one-dimensional (1D) CNN layer receives the output, extracts the local features and makes
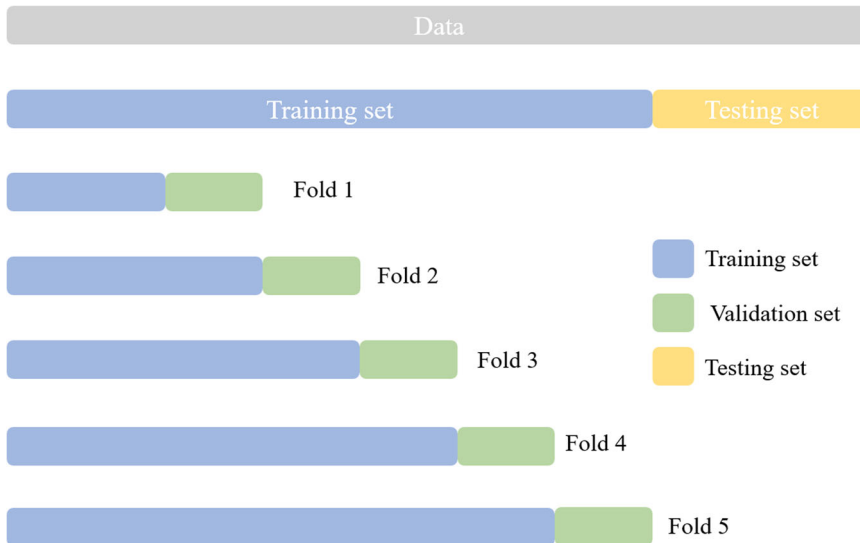
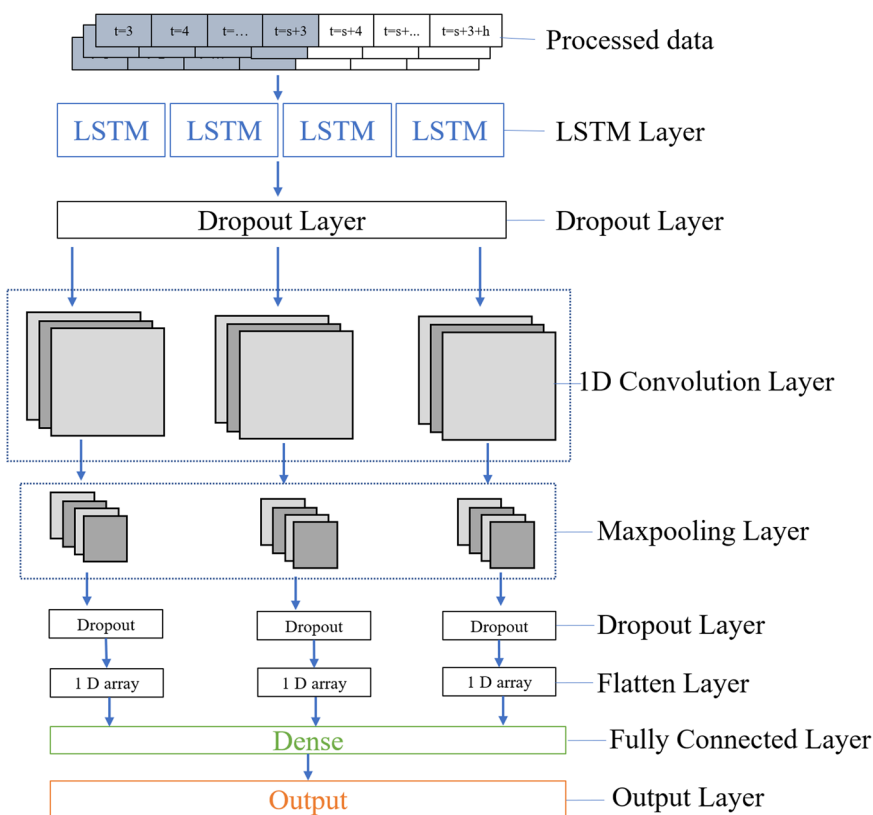**FIGURE 1** The cross-validation on a rolling forecasting origin.

FIGURE 2 The structure of LSTM-CNN model. 1D, one-dimensional; LSTM-CNN, Long Short-Term Memory-Convolutional Neural Network.

predictions (Luft et al., 2022). However, in the CNN-LSTM layer, the CNN layer, as the initial layer, reorganizes the data and extracts only some features (Lu et al., 2022).

After the CNN layer, this research adopts maximum pooling once (=1) and flattens the data. The flattening layer then feeds the output into the fully connected layer to produce predictions. Two dropout layers are added to the LSTM-CNN model. The first dropout layer receives the output of the LSTM model, whilst the second dropout layer precedes the flattening layer, making the data a 1D array. The dropout layer is added to prevent overfitting.

The deep learning model has a series of hyperparameters. To find the optimal combination of hyperparameters, it is desirable to train the model using different sets of hyperparameters, which involve batch size, number of epochs, hidden layers, neurons or filters, and learning rate. A common approach to finding the best combination of hyperparameters is to employ the grid search method proposed by Larochelle et al. (2007), which is an exhaustive search that trains the model in a manually specified hyperparameter space for each combination of hyperparameters. Despite this, sufficiently fine hyperparameter optimization processes are computationally

expensive. Bergstra and Bengio (2012) proposed a random search technique that was proved to be more efficient than the grid search for tuning experiments. Their experimental results reveal that grid search tends to assign too many trials to those dimensions that are not important while covering less of the more critical dimensions. In this study, the random search technique is used to tune the hyperparameters.

Once the best training model is obtained, it is used to predict the testing set. The predicted values were then compared with the observed true values to assess the performance of the developed model. As baselines, three other models were also trained in the same process to make predictions on the testing set, which were LSTM, CNN and ARIMA. As ARIMA is a univariate statistical model, the transformed sliding window dataset was employed to train each feature using ARIMA, and the average performance was used as the comparison value. The Mean Squared Error (MSE) was chosen as the metric to evaluate the models. In statistics, MSE measures the average of the squares of the errors of a prediction model. In other words, it is the calculation of the mean squared difference between the predicted and true values.

2.4 | Feature analysis

To clarify the contribution of each feature in this study, the SHAP tool was used to understand the output of the LSTM-CNN model (see Supporting Information: Tables S2, S3). SHAP is a library developed by Lundberg and Lee (2017), which calculates Shapley values. The Shapley values define the feature contribution of a selected feature. This is done by computing the expected difference between the predicted value of the training model with and without the selected feature for each subset of features (Molnar, 2022). Each input variable (unit of time) has its own Shapley value. Averaging was performed to calculate the mean Shapley value for different weeks and different years. The features were then ranked according to their corresponding years and weeks to obtain how the importance of the feature varies with the year and week.

3 | RESULTS

3.1 | Model performance

The performance in terms of MSE values for the four models: ARIMA, CNN, LSTM and LSTM-CNN, is demonstrated. The hyperparameters of the LSTM-CNN model were tuned for best performance on a 15-week prediction and a 20-week input time step. In the proposed model, LSTM had two layers, but CNN had only one layer. The whole combination of hyperparameters is revealed in Table 1. For a fair comparison, the other models were also trained with a maximum of 80 epochs and a learning rate of 0.001.

In Figure 3a, we showed the MSE trends for the five models when predicting 5, 10, 15 and 20 weeks with the same

TABLE 1 The results of the hyperparameters optimisation for LSTM-CNN model with 15 weeks ahead prediction and all deep learning model use 80 Epochs with the learning rate of 1×10^{-3}.

Model	Hyperparameter	Value
LSTM-CNN	LSTM layers	224
	Dropout rate	0.55
	LSTM layers	96
	Dropout rate	0.15
	Conv 1D	224
	Dropout rate	0.65
	Dense layer	448
ARIMA	ARIMA(p,q,d)	2,1,0
LSTM	LSTM layer	256
CNN	Conv 1D	64

Abbreviations: 1D, one-dimensional; ARIMA, Auto-Regressive Integrated Moving Average models; LSTM-CNN, Long Short-Term Memory-Convolutional Neural Network.

hyperparameter settings. The ARIMA had the highest MSE value at each time horizon and had a low MSE value at 5 weeks, but its performance deteriorated when the number of weeks rose. Like ARIMA, CNN also tended to increase the MSE when the prediction range increased. The LSTM model had a decreasing trend along the changing weeks, and its MSE value was lower than that of the CNN, but in 5 weeks, the MSE values were slightly higher than those of the CNN models. The LSTM-CNN model had the lowest MSE value in 5, 10 and 15 weeks than other models. At 10 weeks, LSTM-CNN achieved the lowest MSE value, being around 0.0086 (Table 2).

The MSE values of the normalised lag difference of the LSTM-CNN model were divided into each grain price to see which grain price is better predicted. Figure 3b indicates the MSE values for four grain products, which are oats, corn, soybeans and wheat. The LSTM-CNN model performed the worst in predicting oat prices and the best in predicting wheat. In addition, the MSE values for corn were close to those of wheat. The MSE values show that corn, soybeans and wheat all had the lowest MSE values in the 20 weeks ahead of the forecast.

3.2 | Feature contribution

3.2.1 | Average contribution

To gain an insight into the variation in the feature contribution of each grain price, this research calculated the Shapley values for all 17 features at 5, 10, 15 and 20 weeks. The mean Shapley value of all samples was calculated to gain a general sense of the feature contribution. 5- and 20-week were chosen to be placed in the result section, as the differences in MSE values between these two steps were the most pronounced. In Figure 4, it can be noted that wheat
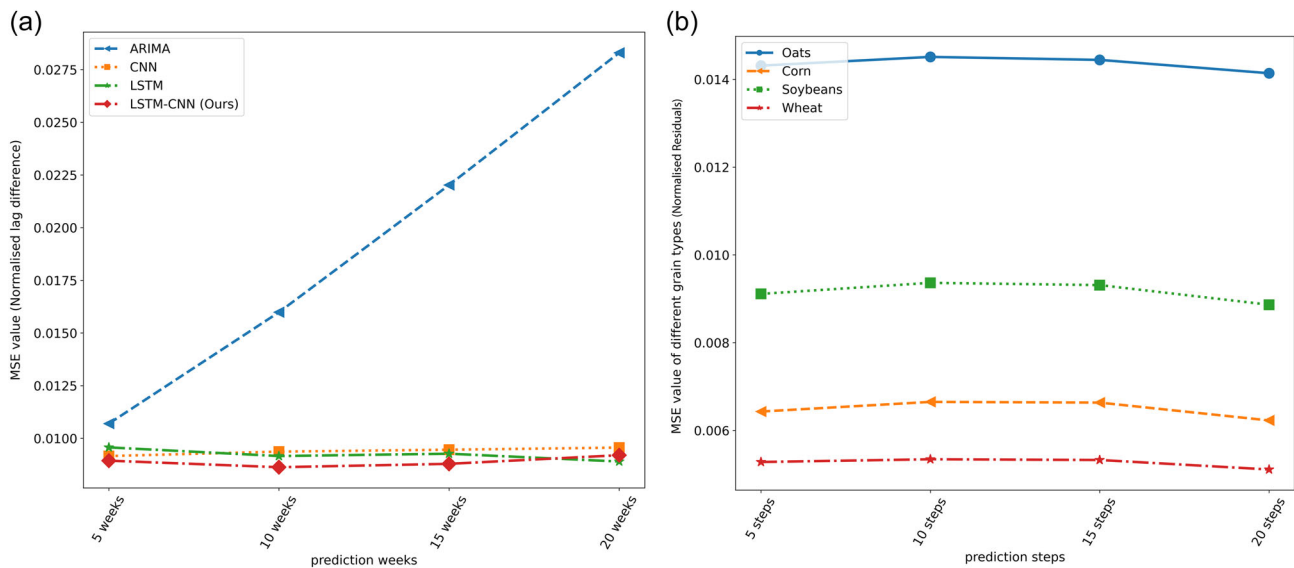
(a)

(b)

FIGURE 3 (a) The MSE distribution of ARIMA, CNN, LSTM and LSTM-CNN model at 5, 10, 15 and 20 weeks. (b) The MSE value of LSTM-CNN in predicting the prices of oats, corn, soybeans and wheat. ARIMA, Auto-Regressive Integrated Moving Average models; LSTM-CNN, Long Short-Term Memory-Convolutional Neural Network; MSE, Mean Squared Error.

TABLE 2 The MSE value of ARIMA, CNN, LSTM and LSTM-CNN model at 5, 10, 15 and 20 weeks.

	ARIMA	CNN	LSTM	LSTM-CNN
5 weeks	0.0107	0.0092	0.0096	0.0090
10 weeks	0.0160	0.0094	0.0092	0.0086
15 weeks	0.0220	0.0095	0.0098	0.0088
20 weeks	0.0283	0.0096	0.0091	0.0092

Abbreviations: ARIMA, Auto-Regressive Integrated Moving Average models; LSTM-CNN, Long Short-Term Memory-Convolutional Neural Network; MSE, Mean Squared Error.

prices made the largest contribution to all grain prices. Of traditional weather factors, the minimum temperature made the greatest contribution to oats and maize, whereas total precipitation made the greatest contribution to soybeans. In addition, the maximum temperature made the greatest contribution to wheat. Of the snow-related features, SWE had the largest contribution, and the contribution of SWE was greater than that of other weather features in predicting oats, corn and wheat. It ranked higher than crude oil WTI in predicting corn prices and wheat prices. Nevertheless, snowfall made a relatively small contribution to grain prices, while snow depth ranked in the middle.

The overall distribution of features in the 20-week prediction was similar to the prediction of the 5-week (Figure 5). Grains' substitutes ranked among the most contributing features. In contrast, the substitute grain feature that contributed most to wheat in the 20-week forecast is corn, while in the 5-week forecast, it is wheat itself. In forecasting oat and corn prices, corn's contribution rose, compared to its ranking in the 5-week forecast. In oats, SWE was also the highest of the weather features (including those related to snow),

but in the other three grain price forecasts, it did not contribute as much as in the 5-week, but still ranked in the middle. Apart from these, snow depth ranked in the middle place in the 5-week ahead of prediction (Figure 4), but its contribution declined in the 20-week ahead of prediction and was among the last few least vital features.

3.2.2 | Yearly variation of feature contribution

The average Shapley value was recalculated based on the year and week of the forecasting period. Figure 6 shows the yearly variation of the feature contribution of wheat for the 10 weeks ahead of the prediction, ranging from 2012 to 2021. Each bar has a base value, which does not involve any feature in the current prediction, that is, the average prediction value of the training set. In the graph, the blue bars indicate that the feature has a negative impact on the prediction, in other words, driving the output to a lower value; conversely, the red bars indicate that the feature has a positive impact, that is, increasing the value. The magnitude of SWE's contribution was significant in 2012 and 2021, whereas in 2012, it showed a negative impact. Then, in the rest of the years, its contribution was very small and hidden from the graph, but it was not until a decade later in 2021 that it made a significant positive contribution.

3.2.3 | Weekly variation of feature contribution

Figure 7 shows how the feature contributions change with the year of the week, reorganised in the same way as the year variation, but according to the weekly values. The x-axis is the weekly value, ranging from the first week to the last week of the
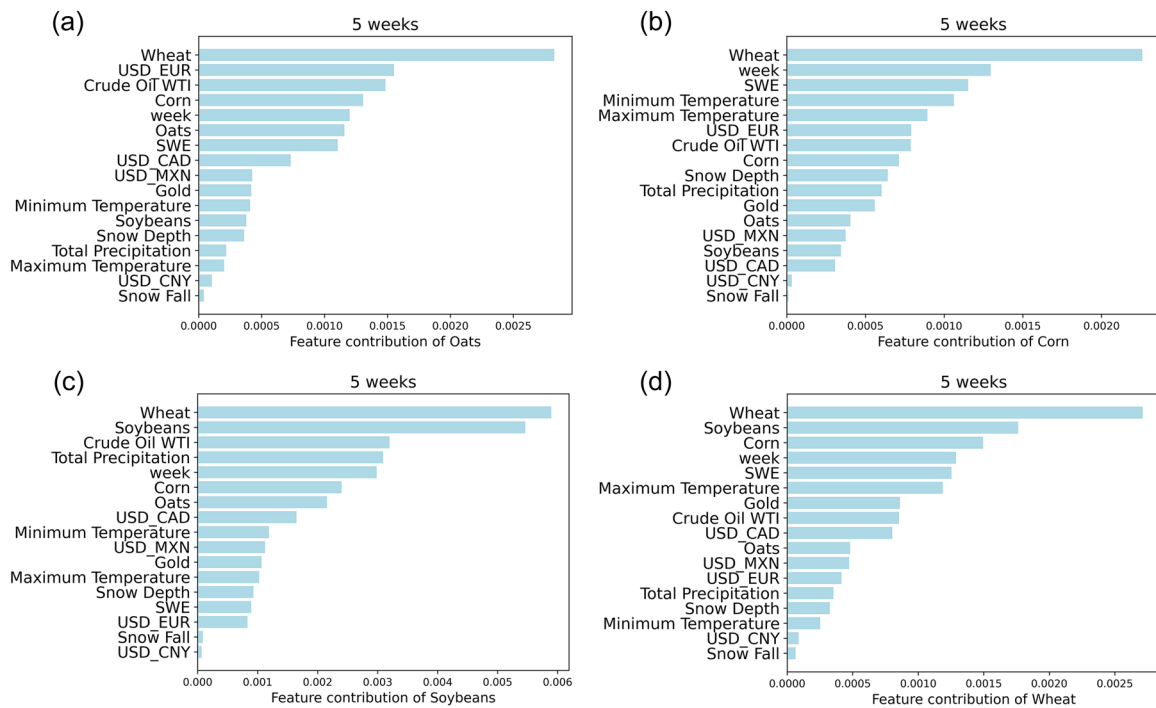
FIGURE 4 Feature contribution of LSTM-CNN model at 5 weeks ahead on (a) Oats, (b) Corn, (c) Soybeans and (d) Wheat. LSTM-CNN, Long Short-Term Memory-Convolutional Neural Network.
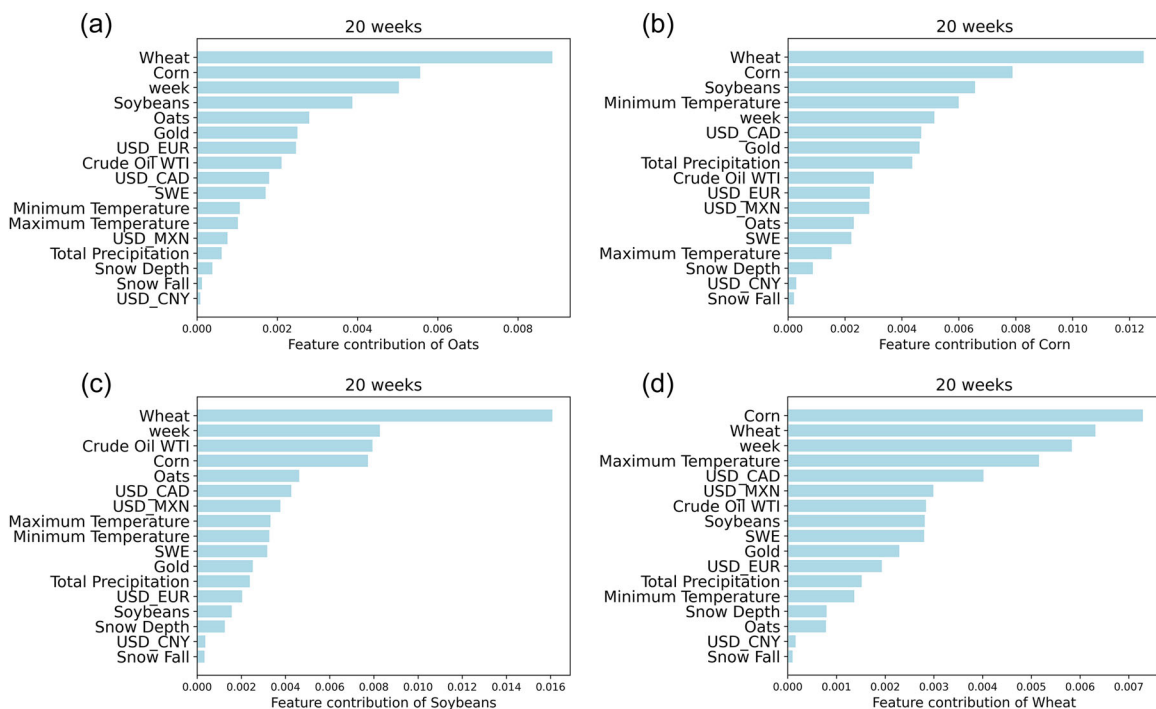

FIGURE 5 Feature contribution of LSTM-CNN model at 20 weeks ahead on (a) Oats, (b) Corn, (c) Soybeans and (d) Wheat. LSTM-CNN, Long Short-Term Memory-Convolutional Neural Network.

year. On the *y*-axis, there is a central point parallel to the *x*-axis that represents the base value. Here, only the SWE, snowfall and snow depth were extracted to understand how snow-related features affect the price of wheat during the year (see Supporting Information: Figures 2–4, for weekly variations in other features). These three snow-related features shared a common characteristic in that they had little impact on prices over a period of approximately 15–20 weeks, but the exact timing did not match.
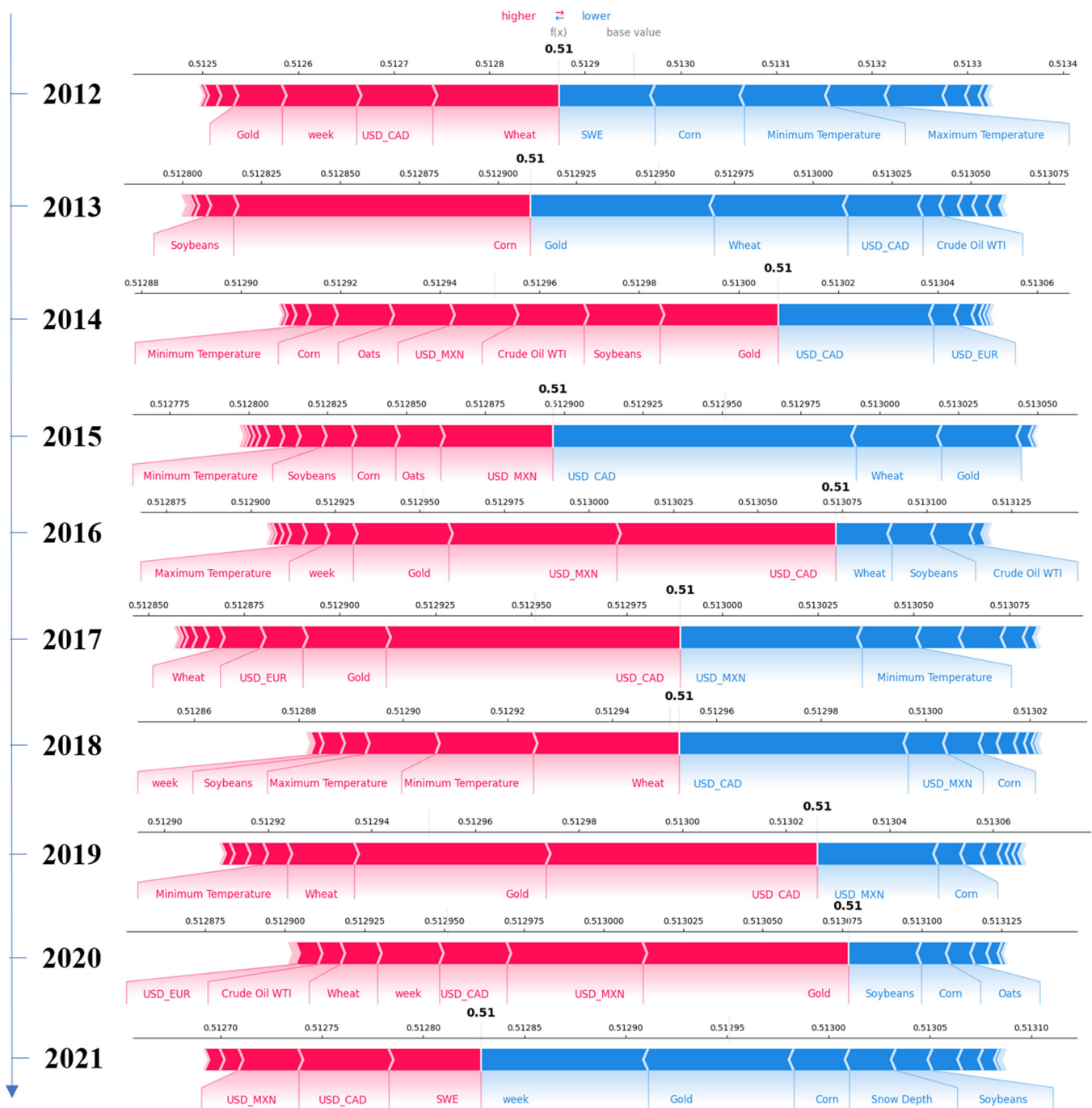
FIGURE 6 The distribution of Shapley values of wheat at 10 weeks ahead of prediction over the years.

For example, SWE occurred between 25 and 40 weeks, snowfall occurred 2–3 weeks earlier, and snow depth started at a similar time to snowfall but ended synchronously as SWE. SWE and snow depth had opposite effects on wheat prices. Specifically, in the first 2 months of the year, snow depth made a positive contribution, but SWE had a negative one. During weeks 8–20, snow depth contributed to explain the prices decrease, but during Weeks 13–25 (where SWE had some delayed patterns over snow depth), SWE was a positive contributor, and its contribution was the greatest during this period. Then, at the end of the year, that is, in

weeks 40–53, SWE contributed negatively, while snow depth contributed positively.

4 | DISCUSSION

4.1 | Performance analysis

In this study, a multivariate hybrid model LSTM-CNN was used to predict grain prices multiple steps ahead. The performance of the
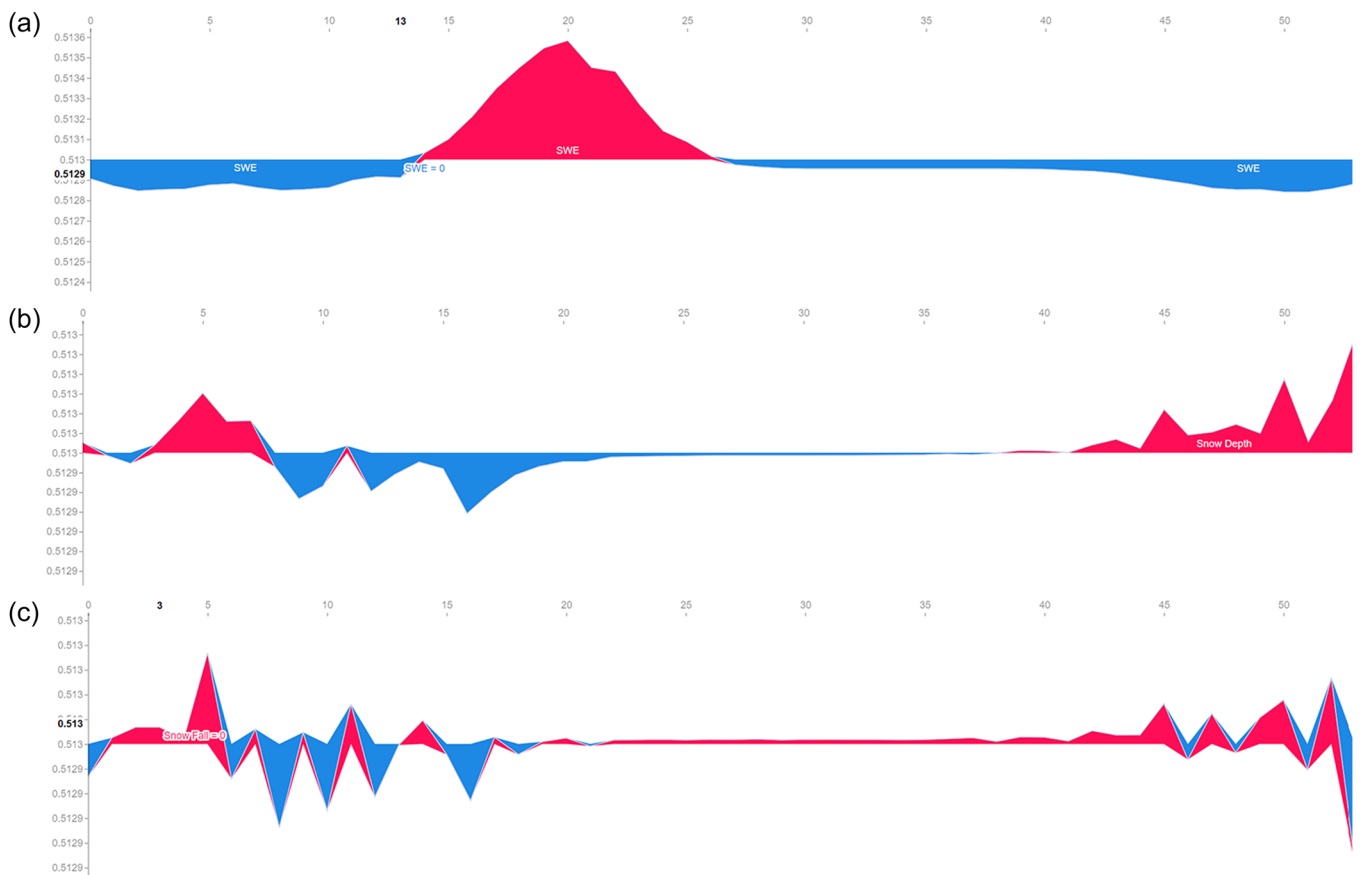
FIGURE 7 Weekly distribution of Shapley values for (a) SWE, (b) snowfall, and (c) snow depth in 10 weeks for wheat price forecasts. SWE, snow water equivalent.

hybrid model achieved the lowest MSE values of normalised lag difference at 10 weeks. However, the prediction performance of the multivariate model did not change significantly with increasing number of prediction time steps. At the short-term prediction, that is, 5 weeks ahead of prediction, the univariate model ARIMA also showed low MSE values, which indicates that ARIMA has good predictive power in short-term forecasting. Such an observation has been found in other studies as well. Ly et al. (2021) also noted in their study that the LSTM did not perform better than the traditional ARIMA model in single-step ahead forecasting the cotton and oil prices. Also, in the study of Sun and Jin (2022), the ARIMA model achieved better results in 1-h ahead to forecast the wind speed. Possible reasons for this are that ARIMA can capture time series correlations within the dataset and is a sophisticated inferential device which can explore the underlying factors of price volatility. These factors are not likely to change significantly in the short term to cause a major change in prices. As a result, ARIMA performs well for short-term forecasting of various time series data (Levenbach, 2017). However, knowing only the underlying factors that cause price volatility is not sufficient to predict a longer range of prices. Hence, as the number of prediction steps increases, the advantages of multivariate models that incorporate more explanatory factors become apparent.

The hybrid model has lower MSE values indicating that the combination of CNN and LSTM performed better in predicting grain prices at 5, 10 and 15 weeks ahead. Such a result could be attributed to the hyperparameter tuning of the LSTM-CNN based on the predictions for 15 weeks ahead. As a result, lower MSE values were observed at 10 and 15 weeks. In the study of Madaeni et al. (2022), they also found that the hybrid model achieved better results than the single CNN and LSTM models. At 5 weeks, CNN had lower MSE values than LSTM, but at 10, 15 and 20 weeks, it performed worse than LSTM. Also, at 20 weeks ahead, the LSTM performed slightly better than other models. This finding coincides with part of the results of Yan et al. (2021), which compared the predictive power of CNN, LSTM and CNN-LSTM for air quality index. Their study shows that the LSTM is the optimal model for multihour prediction, demonstrating its advantages to explicitly modelling the temporal dynamics when predicting longer time series data (Ordóñez & Roggen, 2016).

The better performance of the CNN over the LSTM at 5 weeks may be attributed to the ability of the CNN to partially include the temporal dependence of the dataset (Madaeni et al., 2022). However, when the number of time steps increases, the advantage of incorporating more past information into the LSTM is revealed. The combination of LSTM and CNN may integrate the strengths of both

models in that the CNN is better able to capture correlations between variables, while the LSTM deals with the temporal dynamics of the input variables (Madaeni et al., 2022; Ordóñez & Roggen, 2016).

4.2 | Interpretation of features contribution

Wheat is the main grain source of carbohydrates available to humans, which has a protein content of up to 13%, relatively higher than other grains, and, is grown on more land (Curtis et al., 2002). This significance makes the price of wheat an indicator of food security (Grote et al., 2021) and is often considered in agricultural policy-making. For example, in 2013, the European Union began to consider information on wheat prices in the development of the Common Agricultural Policy (European Union, 2022). Given the importance of wheat, it is reasonable to be the largest contributor to determining grain prices.

Corn also makes a significant contribution to grain prices, especially when the forecast horizon increases. Like oats, corn is mainly used to feed livestock and farmers tend to switch to oats as their feed if the price of corn rises. Alternatively, if corn prices fall, consumption of corn as a feed grain will increase (Zwer, 2016). Despite this, not only does the substitution effect enable corn to influence the prices of other grains, but the US government's ethanol subsidies also enable corn prices to influence other grains' prices. The ethanol subsidies are designed to promote farmers to produce more corn to meet the growing demand for biofuels (Tyner, 2015). Thus, US farmers have increased the area planted to corn at the expense of wheat and soybeans. The reduction in wheat and soybeans production has led to an increase in prices (Babcock, 2012).

Weather factors generally contribute more to the price of corn than other grains, probably since corn yield is sensitive to hail and strong wind damage (Fox et al., 2011). Given this, in the case of corn, temperature and precipitation show a high ranking of contribution than other grains.

4.3 | Snow feature analysis

Three snow-related features were included in this study, namely cumulative SWE, mean snow depth and cumulative snowfall. Of these, SWE contributed most to the four crops at 5- and 20-week ahead of prediction. In addition, SWE ranked higher than the total precipitation when forecasting oat, corn and wheat prices 5 weeks ahead, which also occurred when forecasting oat, soybean and wheat prices 20 weeks ahead. This may be because one of the factors of concern for agricultural development is how much water is available to supply the crops. Precipitation can meet the needs of rainfed crops. However, irrigated crops rely on irrigation techniques to obtain a reliable source of water (Qi et al., 2020). Irrigation can obtain water from a variety of sources including reservoirs, tanks and wells that collect water from snow melt, lakes, and basins (Shah et al., 2019).

SWE refers to the amount of snowpack in terms of water equivalent, which will evaporate as part of precipitation and melt as streams, providing information on water security for agricultural practices in the current year (Biemans et al., 2019; Diffenbaugh et al., 2015). As such, SWE might provide more information on food prices than precipitation in some cases.

Snow depth is a two-sided factor in agricultural practice. For one thing, the snow cover acts as an insulating blanket, protecting the crop from the dynamics of winter minimum temperatures and protecting the soil from deep frosts that can deteriorate soil physical quality and the biological component for the following season (Campbell et al., 2014). Thick snow depth keeps the crop under freezing pressure (Zhu et al., 2022). However, there is less of a direct relationship between snowfall and crop growth. The impact of snowfall on food prices stems in part from the amount of snowfall and the depth of snow (Quante et al., 2021). Another possible relationship between snowfall and grain prices is that when snowfall is extremely heavy, it can give rise to climatic hazards, that is, blizzards, affecting crop growth and local transportation (The New York Times, 2022). Despite this, the spatial and temporal extent of the impact of blizzards is limited compared to the national scale, resulting in its contribution to grain price forecasts being limited.

To understand the specific patterns of snow impacts on grain prices, snow features are discussed in terms of patterns of contribution by year and week, and changes in the contribution of different grain types.

4.4 | Snow features contribution pattern of the year

Among the yearly variations of the contributing features, the contribution of SWE was significant in 2 years, one in 2012 and the other in 2021. In 2012, SWE drove the wheat prices to a lower level, and in 2021, the SWE drove the prices to a higher level, but snow depth contributed negatively. As suggested in the study of McCreight et al. (2014), many areas in the western US had relatively low snow depths and SWE, and depending on the location, they could have been 50% below the normal range in 2012. On 13–17 February 2021, North America experienced a winter storm with temperatures as low as below freezing, extending from southern Texas to the Gulf of Mexico (The New York Times, 2022).

One of the patterns found from the results is that there are sub-weekly variations every year from late spring to early autumn when there is only a little snow. During that time, snow features are hard to make any contribution to price prediction. In spring, the contribution of SWE to wheat prices reached a peak, implying the significance of the amount of SWE in spring for wheat production. However, it would be arbitrary to directly conclude that wheat prices will increase or decrease because of the negative or positive effects from snow-related features. This is because the underlying mechanisms of influence are very complicated, and no linear relationship can be detected. The impact of snow is critical, with both too little and too

much snow causing prices to move in unpredictable directions. However, it hints at the importance of understanding the exact relationship between snow and food, both to better predict price dynamics and to ensure food security and achieve Sustainable Development Goal (SDG) 2.

4.5 | Variation in SWE contribution to different grains

The spatial distribution of oats, corn, soybeans and wheat varies in the US. Oats are mainly grown near the Great Lakes in the northern part of the United States, with South Dakota being the main producer of oats in the US (Government of Alberta, 2022) (Figure 8a). Corn production is concentrated in the heartland of the US (USDA, 2022d), mainly in the Midwestern states of which the states of South Dakota, Nebraska, Minnesota, Iowa, Illinois and Kansas are a part (Figure 8b). States that contribute to soybean production are in the central part of the United States, including Iowa, Illinois, Minnesota, South Dakota, Nebraska, Indiana and Ohio (Figure 8c). Wheat production spreads across the states, with comparable yields of winter and spring wheat (Government of Alberta, 2022). Winter wheat is grown in the western and central regions, whereas spring wheat is grown mainly in the north-western part of the United States, covering Wisconsin, North Dakota, Montana and South Dakota (Figure 8d),

where the share of snowmelt on average runoff is higher than other states according to Li et al. (2017)

A glance at the spatial distribution of grain production in the United States reveals that the main production states for oats, wheat and corn are the states next to the western region, where the SWE monitoring stations are located. The increased contribution of SWE to soybean price prediction from 5 to 20 weeks may be owing to the fact that snowmelt from the western states takes longer to reach the central region, where soybeans are primarily grown. Hence, the SWE contribution is more evident with a longer prediction horizon. Another possible explanation is that oats, corn and spring wheat are sown between April and May, while soybeans are sown between May and June when SWE has the least values and cannot give any indication of price (Government of Alberta, 2022).

Nonetheless, to clarify the relationship between SWE and different grain prices, an in-depth study of streamflow dynamics and the spatial and temporal fluctuations of different grain yields and prices is required. Furthermore, using only 20 weeks of historical data as an input variable to forecast prices 5–20 weeks into the future may lack sufficient information to make longer-term predictions. For predictions of the 5 weeks in the future, snow features may be able to provide sufficient information, but for 20-week forecasts, a longer history should be considered as an input variable to allow the model to learn the full round of weather cycles over the course of a year.
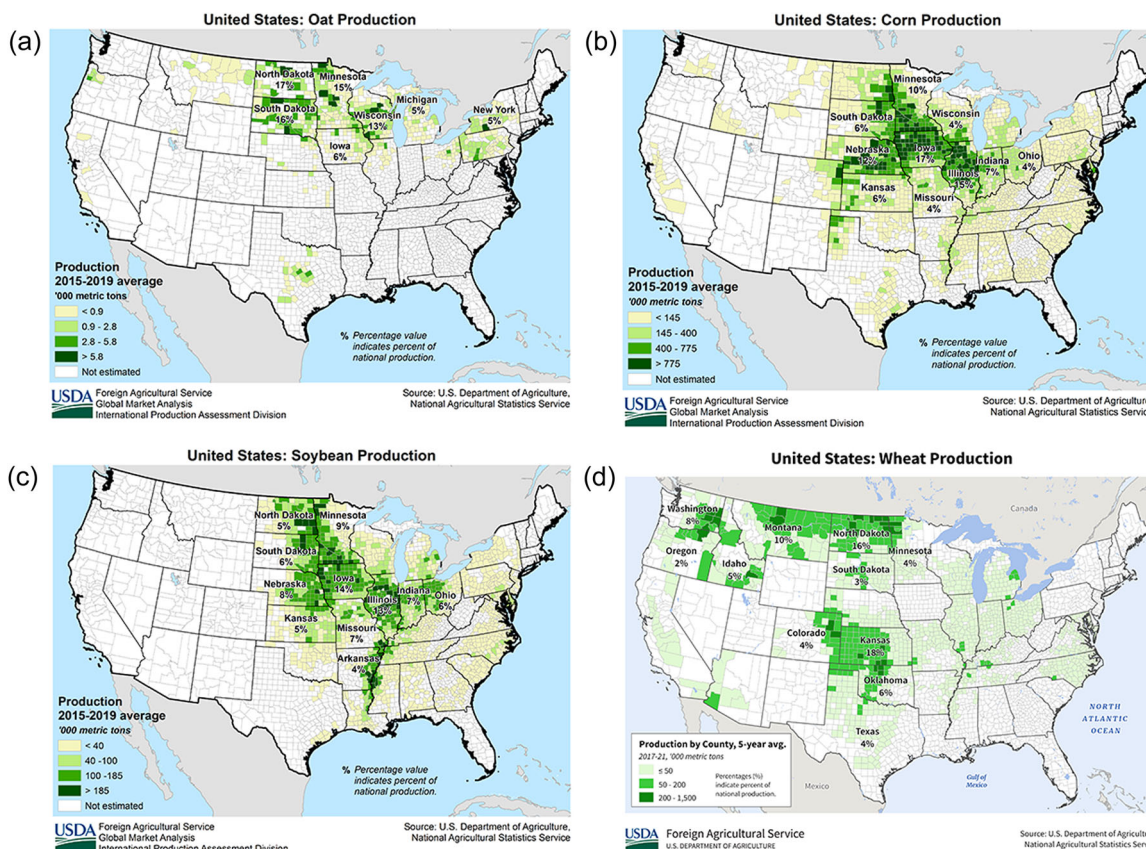

FIGURE 8 The production distribution of (a) Oats, (b) Corn, (c) Soybeans and (d) Wheat in the United States (USDA, 2022e).

262 | JOURNAL OF SUSTAINABLE AGRICULTURE AND ENVIRONMENT — Global Initiative of Crop Microbiome and Sustainable Agriculture — Open Access

WANG ET AL.

4.6 | Application scenarios of the LSTM-CNN model

The study was contributing to lowering uncertainty in the prediction of grain prices almost 5 months in advance, giving farmers enough time to plan the next round of sowing and harvesting. Also, farmers could use the predicted four grain prices, namely oats, corn, soybeans and spring wheat, all of which are sown in the spring, to decide which type of crops they should plant, and how to minimise cultivation and storage costs and maximise benefits.

The model is best suited for predicting spring wheat prices, compared to the other three crops. First, wheat prices are predicted with the highest accuracy. Second, the average time from sowing to germination and then to harvest for spring wheat is approximately 4 months (USDA, 2022f). This is within 20 weeks of the prediction range of the proposed model.

For oats and soybeans, although their predictions are not as accurate as those for wheat prices, they also have a life cycle of around 4–5 months (USDA, 2022f). Thus, the model proposed can also provide some direction for farmers planning to grow oats and soybeans. For corn, however, the life cycle is over 20 weeks. Corn is usually sown in April-May and harvested in October–November, a period of about 6–7 months (USDA, 2022f). Therefore, this model may not provide farmers with future corn prices over a sufficiently long period. In the future, a model with extended prediction time steps could be developed to accommodate the life cycle of corn. A more comprehensive study could also be carried out to examine the extent to which the time series of snow makes a negative and positive contribution to grain prices, considering the accelerated melting rate of glaciers and snowpack due to climate change (Zhu et al., 2022). Considering the importance of snowmelt for the acquirer recharge, we must consider the soil water capacity as a limiting factor in future climates (Zwer, 2016).

4.7 | Limitation

When comparing the LSTM-CNN model with other predictive models, the epochs and learning rate are based on the result of the hyperparameter tuning of the LSTM-CNN in 15 weeks. Consequently, the hyperparameter settings may not be the most optimal combination for the CNN and LSTM models. Alternatively, the better performance of the LSTM-CNN model at Weeks 5, 10 and 15 may be because the hyperparameters were tuned based on the predictions at 15 weeks. This may not be the optimal hyperparameter setting for the LSTM-CNN at 5, 10 and 20 weeks.

To get a thorough understanding of the trends at different input time steps, experiments with different input time steps are necessary to clarify how the output of the proposed model changes and how other models change in the same setup. However, due to time constraints, this experiment was out of the scope of this research. Instead, in this study, the input variable steps were set up fixed to be 20 weeks. For the LSTM-CNN model, the optimal input variable may be represented by other values. Furthermore, the time step of the optimal input variable may differ across the prediction range. Also, it is possible that the current 20-week input step is suitable for the LSTM-CNN model, but when changing to other ranges of input time steps, other comparison models such as CNN and LSTM may perform better than LSTM-CNN.

5 | CONCLUSION

In this study, we propose a hybrid multivariate LSTM-CNN model for grain commodities price prediction that uses weather factors, grain prices, macroeconomic factors and a novel class represented by snow-related time series data—as features to predict multistep price advance. The LSTM-CNN model integrates LSTM and CNN models, with the LSTM responsible for remembering the long-term and short-term memory of grain prices and the CNN is used to learn the dependencies between the input variables and extract key features. We performed hyperparameter optimisation to generate the optimal combination of hyperparameters for the LSTM-CNN model to predict prices 15 weeks in the future. We used the metric MSE to compare the performance of the LSTM-CNN with the ARIMA, CNN and LSTM. The best hyperparameter combination resulted in the LSTM-CNN model having the lowest MSE values at 5, 10 and 15 weeks ahead, compared with other models. The LSTM-CNN model was best suited for predicting wheat prices, with wheat having the best prediction performance than other grains. In addition, the LSTM-CNN model was able to predict prices 20 weeks out, which goes beyond the entire life cycle of wheat, giving farmers enough time to better plan the next round of sowing based on the predicted prices.

To investigate the contribution of features, Shapley values were used to calculate the contribution of each feature to the different grain price predictions. Wheat had the highest contribution in all four grain price forecasts. For the snow-related features, SWE ranked highly in the 5 weeks ahead forecasts and was also the largest contributor to the three snow features, whether 5 or 20 weeks ahead. In some cases, SWE ranked higher than total precipitation, suggesting that SWE contains more hydrological and weather information to support grain price forecasts than total precipitation. Therefore, in the future, SWE could be a potential variable for multistep ahead forecasting of grain prices.

Although SWE has a high ranking of contribution in the 5 weeks, its dropping of ranking in 20 weeks demonstrates the SWE still has exploration potential to contribute to the predictions with a longer input time steps. Hence, studies like increasing the input time steps of variables can be conducted to investigate whether with an increased time range of SWE can improve the prediction outcome of grain prices. Also, the prediction time range can be extended to see how the LSTM-CNN model performed in an even further prediction horizon.

The significant contribution of snow-related features, particularly SWE, suggests the research potential of using snow as an input variable to predict not only grain prices but also other agricultural commodities and even stock and energy prices over a multistep

WANG ET AL.

JOURNAL OF SUSTAINABLE
AGRICULTURE AND ENVIRONMENT

Global Initiative of
Crop Microbiome and
Sustainable Agriculture

263

ahead of the prediction horizon. Furthermore, this study reveals a possible correlation between grain prices and snow time series, but the exact mechanism of correlation cannot be observed from the feature contribution analysis alone. These studies could help policy stakeholders to develop timely policies to ensure better food and social security.

AUTHOR CONTRIBUTIONS

Conceptualization: Niamh French, Thomas James and Aldo Lipani. *Methodology*: Aldo Lipani and Zilin Wang. *Validation*: Zilin Wang. *Formal analysis*: Niamh French, Zilin Wang and Aldo Lipani. *Investigation*: Niamh French, Thomas James, Calogero Schillaci and Zilin Wang; *Resources*: Niamh French, Thomas James and Aldo Lipani; *Writing—original draft preparation*: Zilin Wang. *Writing—review and editing*: Aldo Lipani, Calogero Schillaci, Meili Feng and Faith Chan. *Visualization*: Niamh French, Thomas James and Zilin Wang. *Supervision*: Aldo Lipani, Niamh French and Thomas James; all authors have read and agreed to the published version of the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The temperature, precipitation, snow fall and snow depth data are available in National Oceanic and Atmospheric Association website at https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily. The snow water equivalent data were retrieved from the website of the United States Department of Agriculture (USDA) at https://www.nrcs.usda.gov/wps/portal/wcc/home/snowClimateMonitoring/30YearNormals. The macro-economics data are available on the Investing website at www.Investing.com.

ETHICS STATEMENT

The authors declare that they followed the ethics policies of the journal.

ORCID

Zilin Wang http://orcid.org/0000-0002-2582-3076
Calogero Schillaci http://orcid.org/0000-0001-7689-5697
Faith Chan http://orcid.org/0000-0001-6091-6596

REFERENCES

Adanacioglu H, Yercan M. An analysis of tomato prices at wholesale level in Turkey: an application of sarima model. Custos e Agronegocio. 2012;8(4):52–75.

Ayankoya K, Calitz AP, Greyling JH. Real-time grain commodities price predictions in South Africa: a big data and neural networks approach. Agrekon. 2016;55(4):483–508.

Babcock BA. The impact of us biofuel policies on agricultural price levels and volatility. China Agric Econ Rev. 2012;4:407–26.

Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res. 2012;13(2):281–305.

Biemans H, Siderius C, Lutz AF, Nepal S, Ahmad B, Hassan T, et al. Importance of snow and glacier meltwater for agriculture on the indo-gangetic plain. Nature Sustainability. 2019;2(7):594–601.

Cammarano D, Stefanova L, Ortiz BV, Ramirez-Rodrigues M, Asseng S, Misra V, et al. Evaluating the fidelity of downscaled climate data on simulated wheat and maize production in the southeastern US. Reg Environ Change. 2013;13(1):101–10.

Campbell JL, Socci AM, Templer PH. Increased nitrogen leaching following soil freezing is due to decreased root uptake in a Northern hardwood forest. Global Change Biol. 2014;20(8):2663–73.

Central for Agricultural and Rural Development. (2022). *Exchange rates and agricultural commodity prices*. Available from https://www.card.iastate.edu/iowaagreview/fall07/article4.aspx, Accessed 16 Jul 2022.

Chambers RG, Just RE. Effects of exchange rate changes on us agriculture: a dynamic analysis. Am J Agric Econ. 1981;63(1):32–46.

Cheng H, Xiong W. *The financialization of commodity markets* (Tech. Rep.). National Bureau of Economic Research; 2013.

Commodity. (2022). *US Economy: What are the exports and imports it relies on most?* Available from https://commodity.com/data/usa/, Accessed 4 Jul 2022.

Curtis BC, Rajaram S, Gómez Macpherson H. Bread wheat: improvement and production, Rome: Food and Agriculture Organization of the United Nations (FAO); 2002.

Diffenbaugh NS, Swain DL, Touma D. Anthropogenic warming has increased drought risk in California. Proc Natl Acad Sci. 2015; 112(13):3931–6.

Divisekara RW, Jayasinghe G, Kumari K. Forecasting the red lentils commodity market price using SARIMA models. SN Bus Econ. 2021;1(1):1–13.

European Union. (2022). *The common agricultural policy at a glance*. Available from https://agriculture.ec.europa.eu/common-agricultural-policy/capoverview/cap-glanceen, Accessed 16 Aug 2022.

Field PR, Heymsfield AJ. Importance of snow to global precipitation. Geophys Res Lett. 2015;42(21):9512–20.

Fox JF, Fishback PV, Rhode PW. The effects of weather shocks on crop prices in unfettered markets: The united states prior to the farm programs, 1895–1932. In: Libecap GD, Steckel RH, editors. The economics of climate change: adaptations past and present. University of Chicago Press; 2011. p. 99–130.

Government of Alberta. (2022). *Us crops—where are they grown?* Available from https://www.alberta.ca/us-crops-where-arethey-grown.aspx, Accessed 30 Aug 2022.

Grote U, Fasse A, Nguyen TT, Erenstein O. Food security and the dynamics of wheat and maize value chains in Africa and Asia. Front Sustain Food Syst. 2021;4:617009.

Gu YH, Jin D, Yin H, Zheng R, Piao X, Yoo SJ. Forecasting agricultural commodity prices using dual input attention LSTM. Agriculture. 2022;12(2):256. https://doi.org/10.3390/agriculture12020256

He Z, Zhou J, Dai H-N, Wang H (2019). Gold price forecast based on LSTM-CNN model. *2019 IEEE Int conf on dependable, autonomic and secure computing, intl conf on pervasive intelligence and computing, intl conf on cloud and big data computing, intl conf on cyber science and technology congress (dasc/picom/cbdcom/cyberscitech)* (pp. 1046–1053).

Hertel TW, Lobell DB. Agricultural adaptation to climate change in rich and poor countries: current modeling practice and potential for empirical contributions. Energy Econ. 2014;46:562–75.

Hitaj C, Suttles S (2016). *Trends in us agriculture's consumption and production of energy: Renewable power, shale energy, and cellulosic biomass* (Tech. Rep.).

Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. 3rd ed. Melbourne, Australia: OTexts; 2021. OTexts.com/fpp3

Jadhav V, Chinnappa RB, Gaddi G. Application of ARIMA model for forecasting agricultural prices. J Agric Sci Technol. 2017;19:981–92.

Kjellstrom T, Kovats RS, Lloyd SJ, Holt T, Tol RS. The direct impact of climate change on regional labor productivity. Arch Environ Occup Health. 2009;64(4):217–27.

Larochelle H, Erhan D, Courville A, Bergstra J, Bengio Y. An empirical evaluation of deep architectures on problems with many factors of variation. *Proceedings of the 24th International Conference on Machine Learning* 2007. pp. 473–80.

Levenbach H. Change & chance embraced: achieving agility with smarter forecasting in the supply chain. Delphus Publishing; 2017.

Li D, Wrzesien ML, Durand M, Adam J, Lettenmaier DP. How much runoff originates as snow in the Western United States, and how will that change in the future? Geophys Res Lett. 2017;44:6163–72.

Li Y, Hu N, Li G, Yao X (2012). Forecasting mineral commodity prices with ARIMA-Markov chain. *2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics* (Vol. 1, p. 49–52). https://doi.org/10.1109/IHMSC.2012.18

Lopez JR, Winter JM, Elliott J, Ruane AC, Porter C, Hoogenboom G, et al. Sustainable use of groundwater may dramatically reduce irrigated production of maize, soybean, and wheat. Earth's Fut. 2022;10(1): e2021EF002018. https://doi.org/10.1029/2021EF002018

Lu Y, James T, Schillaci C, Lipani A. Snow detection in alpine regions with convolutional neural networks: discriminating snow from cold clouds and water body. GISci Remote Sens. 2022;59(1):1321–43. https://doi.org/10.1080/15481603.2022.2112391

Luft H, Schillaci C, Ceccherini G, Vieira D, Lipani A. Deep learning based burnt area mapping using sentinel 1 for the Santa Cruz mountains lightning complex (CZU) and Creek Fires 2020. Fire. 2022;5(5):163. https://doi.org/10.3390/fire5050163

Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30:4768–77.

Ly R, Traore F, Dia K. *Forecasting commodity prices using long short-term memory neural networks* (Vol. 2000). Intl Food Policy Res Inst. 2021; 1–26.

Madaeni F, Chokmani K, Lhissou R, Homayouni S, Gauthier Y, Tolszczuk-Leclerc S. Convolutional neural network and long short-term memory models for ice-jam predictions. The Cryosphere. 2022;16(4):1447–68.

McCreight JL, Small EE, Larson KM. Snow depth, density, and swe estimates derived from gps reflection data: validation in the Western US. Water Resour Res. 2014;50(8):6892–909.

Mendelsohn R, Nordhaus WD, Shaw D. The impact of global warming on agriculture: a ricardian analysis. Am Econ Rev. 1994;84(4):753–71.

Mithiya D. Forecasting of potato prices of Hooghly in West Bengal: time series analysis using SARIMA model. Int J Agric Econ. 2019;4(3):101.

Molnar C. (2022). Interpretable machine learning: a guide for making black box models explainable. 2nd ed. christophm.github.io/interpretable-ml-book/

Monge M, Lazcano A. Commodity prices after covid-19: persistence and time trends. Risks. 2022;10(6):128.

Moroizumi T, Hamada H, Sukchan S, Ikemoto M. Soil water content and water balance in rainfed fields in northeast Thailand. Agricult Water Manag. 2009;96(1):160–6. https://doi.org/10.1016/j.agwat.2008.07.007

Naidu GM, Sumathi P, Reddy BR, Kumari VM. Forecasting monthly prices of onion in Kurnool market of Andhra pradesh employing seasonal ARIMA model. Bioinfolet—A Q J Life Sci. 2014;11(2b):518–20.

National Oceanic and Atmospheric Administration. (2022). *Global historical climatology network daily (ghcnd)*. Available from https://www.ncei.noaa.gov/data/global-historical-climatology-network-daily/, Accessed 18 Aug 2022.

Nhita F, Saepudin D, Paramita A, Marliani S, Wisesty UN. Price prediction for agricultural commodities in Bandung regency based on functional link neural network and artifical bee colony algorithms. J Comput Sci. 2019;15(10):1390–95. https://doi.org/10.3844/jcssp.2019.1390.1395

Oktoviany P, Knobloch R, Korn R. A machine learning-based price state prediction model for agricultural commodities using external factors. Decis Econ Finance. 2021;44(2):1063–85. https://doi.org/10.1007/s10203-021-00354-7

Ordóñez F, Roggen D. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. Sensors. 2016;16(1):115.

Pujiati A, Oktavilia S, Damayanti N. The projection model as an early warning of food price commodity fluctuation. KnE Social Sciences. 2018;3:860–8.

Qi J, Brantley S, Golladay S. Simulated irrigation reduction improves low flow in streams—a case study in the lower flint river basin. J Hydrol Reg Stud. 2020;28:100665.

Qin Y, Abatzoglou JT, Siebert S, Huning LS, AghaKouchak A, Mankin JS, et al. Agricultural risks from changing snowmelt. Nat Clim Change. 2020;10(5):459–65.

Quante L, Willner SN, Middelanis R, Levermann A. Regions of intensification of extreme snowfall under future warming. Sci Rep. 2021;11(1):16621.

Rasheed A, Younis MS, Ahmad F, Qadir J, Kashif M. District wise price forecasting of wheat in Pakistan using deep learning. arXiv:2103.04781. 2021:1–9.

Salisu AA, Akanni L, Raheem I. The covid-19 global fear index and the predictability of commodity price returns. J Behav Exp Finance. 2020;27:100383.

Schwartz NE. The consequences of a floating exchange rate for the us wheat market. Am J Agric Econ. 1986;68(2):428–33.

Seibert J, Jenicek M, Huss M, Ewen T. Chapter 4 snow and ice in the hydrosphere. In: Shroder JF, Haeberli W, Whiteman C (Eds.) Snow and ice-related hazards, risks, and disasters. Boston: Academic Press; 2015. p. 99–137 https://doi.org/10.1016/B978-0-12-394849-6.00004-4

Shah HL, Zhou T, Sun N, Huang M, Mishra V. Roles of irrigation and reservoir operations in modulating terrestrial water and energy budgets in the Indian subcontinental river basins. J Geophys Res Atmos. 2019;124(23):12915–36.

Shu W, Gao Q. Forecasting stock price based on frequency components by EMD and neural networks. IEEE Access. 2020;8:206388–95.

Sturm M, Goldstein MA, Parr C. Water and life from snow: a trillion dollar science question. Water Resour Res. 2017;53(5):3534–44.

Sun F, Jin T. A hybrid approach to multi-step, short-term wind speed forecasting using correlated features. Renew Energy. 2022;186: 742–54.

The New York Times. (2022). *Texas farmers tally up the damage from a winter storm 'massacre'*. Available from https://www.nytimes.com/2021/03/04/dining/texas-farms-stormdamage.html, Accessed 17 Aug 2022.

Tyner WE. *Us ethanol policy—possibilities for the future* (Tech. Rep.). Historical Documents of the Purdue Cooperative Extension Service; 2015.

USDA. (2022a). *Crop progress report*. Available from https://usda.library.cornell.edu/concern/publications/8336h188j, Accessed 15 Jun 2022.

USDA. (2022b). *Commodity costs and returns*. Available from https://www.ers.usda.gov/data-products/commodity-costs-andreturns/, Accessed 30 Jun 2022.

USDA. (2022c). *1991-2020 climatic and hydrologic normals*. Available from https://www.nrcs.usda.gov/wps/portal/wcc/home/snowClimateMonitoring/30Year, Accessed 4 Jun 2022.

WANG ET AL.

JOURNAL OF SUSTAINABLE
AGRICULTURE AND ENVIRONMENT
Open Access

Global Initiative of
Crop Microbiome and
Sustainable Agriculture

265

USDA. (2022d). *Feedgrains sector at a glance*. Available from https://www.ers.usda.gov/topics/crops/corn-and-otherfeedgrains/feedgrains-sector-at-a-glance/, Accessed 18 Aug 2022.

USDA. (2022e). *United States—crop production maps*. Available from https://ipad.fas.usda.gov/rssiws/al/uscropprod.aspx, Accessed 18 Aug 2022.

USDA. (2022f). Crop production 2019 summary. Available from https://www.nass.usda.gov/Publications/TodaysReports/reports/cropan20.pdf, Accessed 30 Jul 2022.

Wang C, Zhang X, Wang M, Lim MK, Ghadimi P. Predictive analytics of the copper spot price by utilizing complex network and artificial neural network techniques. Resour Policy. 2019;63:101414.

Westcott P, Hoffman L (1999). Price determination for corn and wheat: the role of market factors and government programs (Tech. Rep.).

World Grain (2016). Corn-wheat price relationship complicates wheat market. Available from https://www.world-grain.com/articles/6084-corn-wheatprice-relationship-complicates-wheat-market, Accessed 16 Jul 2022.

Yan R, Liao J, Yang J, Sun W, Nong M, Li F. Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNNLSTM, and spatiotemporal clustering. Expert Syst Appl. 2021;169:114513.

Zhu P, Kim T, Jin Z, Lin C, Wang X, Ciais P, et al. The critical benefits of snowpack insulation and snowmelt for winter wheat productivity. Nat Clim Change. 2022;12(5):485–90.

Zwer P. Oats: overview. In: Wrigley C, Corke H, Seetharaman K, Faubion J, editors. Encyclopedia of food grains. second edition. Oxford: Academic Press; 2016. p. 173–83. Retrieved from https://doi.org/10.1016/B978-0-12-394437-5.00013-9

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.