Article

Deep Learning-Based Causal Inference Architecture and Algorithm between Stock Closing Price and Relevant Factors

Wanqi Xing, Chi Chen and Lei Xue *

China School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China; xingwq@shu.edu.cn (W.X.); chenc@shu.edu.cn (C.C.)
* Correspondence: xuelei@shu.edu.cn

Abstract: Numerous studies are based on the correlation among stock factors, which affects the measurement value and interpretability of such studies. Research on the causality among stock factors primarily relies on statistical models and machine learning algorithms, thereby failing to fully exploit the formidable computational capabilities of deep learning models. Moreover, the inference of causal relationships largely depends on the Granger causality test, which is not suitable for non-stationary and non-linear stock factors. Also, most existing studies do not consider the impact of confounding variables or further validation of causal relationships. In response to the current research deficiencies, this paper introduces a deep learning-based algorithm aimed at inferring causal relationships between stock closing prices and relevant factors. To achieve this, causal diagrams from the structural causal model (SCM) were integrated into the analysis of stock data. Subsequently, a sliding window strategy combined with Gated Recurrent Units (GRUs) was employed to predict the potential values of closing prices, and a grouped architecture was constructed inspired by the Potential Outcomes Framework (POF) for controlling confounding variables. The architecture was employed to infer causal relationships between closing price and relevant factors through the non-linear Granger causality test. Finally, comparative experimental results demonstrate a marked enhancement in the accuracy and performance of closing price predictions when causal factors were incorporated into the prediction model. This finding not only validates the correctness of the causal inference, but also strengthens the reliability and validity of the proposed methodology. Consequently, this study has significant practical implications for the analysis of causality in financial time series data and the prediction of stock prices.

Keywords: deep learning; causal inference; causal diagram; granger causality test; potential outcome framework

1. Introduction

As a typical complex system, the stock market exhibits characteristics of non-linearity, non-stationarity, and multi-components [1–3]. In complex real-world scenarios, accurately identifying the specific factors that influence stock market fluctuations becomes challenging. Many traditional approaches are no longer applicable for predictive research in this domain. Therefore, it is necessary to employ more sophisticated methods to uncover the complex features embedded within time series data.

Currently, methods such as logistic regression, gradient boosting models, and deep learning primarily aim to identify correlation information between variables by fitting observed data. Subsequently, variables with high correlations to the target variable are selected as input variables for prediction [4–7]. However, focusing solely on correlations and not considering causality when choosing predictors can affect the accuracy of predictions [8]. Correlations may arise from causal relationships, but it is not equivalent to causality and is further influenced by confounding variables. Confounding variables affect

both the treatment and target variables simultaneously, resulting in correlations between them [9].

In most cases, the Granger causality test is used to infer causal relationships between financial time series [10,11]. Traditional methods for inferring Granger causality mainly include Vector Autoregression (VAR) [12], the Vector Error Correction Model (VECM) [13], and their respective variants [14,15].

Expanding on the ideas discussed above, Xu et al. [16] presented a novel causal decomposition approach and further applied it to investigate information flow between two financial time series on different time scales. The causal decomposition method has three main steps: decomposition, reconstruction, and causality testing. By tracking the driving factors of causal relationships from the perspective of information frequency, the causal decomposition method re-evaluates the causal relationship between stock prices and trading volume from the time-frequency perspective.

Although causal inference methods based on statistical modeling have been extensively studied and have yielded fruitful results, these methods mainly focus on evaluating the interaction between two variables, overlooking the impact of confounding variables [17–24]. Constructing more intricate causal networks, based on the examination and analysis of pairwise causal relationships, still presents significant challenges. This necessitates the exploration of novel methods and theories to tackle issues such as indirect dependence resulting from front-door paths and common driving factors caused by backdoor paths.

Moreover, it has been found that most economic time series are non-stationary according to various unit root test methods, such as the Augmented Dickey-Fuller (ADF) test [25]. By applying preprocessing techniques, like differencing or logarithmic transformation, these time series can be transformed into stationary data. VAR and the VECM are usually more effective when the input data are stationary [26]. Therefore, for the inference of Granger causality in non-stationary time series, the inputs need to be preprocessed to obtain a stationary series in order to avoid forecast distortion. Assuming that the time series becomes stationary after a difference order of 1 or 2, static causal relationships can be inferred. However, economic data often exhibit dynamic properties, and the preprocessing steps may overlook the dynamic causal relationships within the time series. Causal relationships may not only exist between the current and previous period of data. The oversimplification of preprocessing methods can lead to disregarding dynamic causality and losing valuable information from the original data.

Machine learning (ML) models can capture non-linear and complex relationships better than traditional statistical models [27]. Research has been conducted on causal inference utilizing ML models to overcome limitations in conventional approaches. The use of ML models offers a more accurate and comprehensive ability to infer causal relationships. These models can handle complex datasets that involve a substantial number of variables and aim to identify and infer causal relationships within them. This would help to reveal potential causal paths in causal networks [28].

Leng et al. [29] proposed an independent component analysis (ICA) framework, inspired by the decision tree (DT) algorithm in machine learning, for measuring causal relationships by calculating feature importance. The core idea of the framework is to convert time series data into a causal network representation and to make causal inferences based on feature importance. The ICA framework is designed at the network level, serving as a connection and link between traditional mutual causality detection methods and causal network reconstruction.

While the application of machine learning in causal analysis brings forth new perspectives and approaches, it also possesses certain shortcomings. These limitations encompass sensitivity to data bias and causal confounding, as well as restrictions in handling complex non-linear relationships [30,31]. In comparison, deep learning models exhibit more robust expression and pattern learning capabilities, as they can acquire abstract computational methods from intricate and high-dimensional data [32]. For example, Chong et al. [33] explored a deep learning-based model to evaluate the efficacy of three

unsupervised feature extraction methods to predict future market behavior. Similarly, other studies have made significant advancements in the financial domain through the utilization of deep learning models.

Indeed, a dual-stage attention-based recurrent neural network (RNN) model proposed by Qin et al. [34] has shown promising results in predicting stock datasets. By adaptively extracting relevant input features for prediction, the model can effectively capture important information and make accurate predictions. Zahra et al. [35] combined the convolutional neural network (CNN) and long short-term memory (LSTM) with a fundamental analysis. By extracting and synthesizing features at multiple levels and dimensions, the model can capture both local and global patterns and improve prediction accuracy. Rahman et al. [36] applied the GRU model to predict stock data of Coca-Cola and reduced modeling errors. The GRU model is effective in overcoming the problem of vanishing gradients, which can occur in deep learning models and affect their performance. Compared to traditional econometric and ML approaches, the deep learning-based models demonstrate better prediction performance. This highlights the efficiency and effectiveness of deep learning architectures in handling financial time series data.

Wei [37] proposed an interpretable deep learning architecture, known as deep learning inference (DLI), to investigate Granger causality. The main contribution of DLI is to uncover the Granger causality between Bitcoin price and the S&P index, enabling more accurate prediction of Bitcoin prices in relation to the S&P index. However, this method only provides a visual representation of the predicted results for Bitcoin prices and lacks evaluation metrics to quantify the results. Its inference of a Granger causality between the two variables is solely based on the enhanced prediction performance of Bitcoin prices after incorporating historical data from the S&P index. Studies inferring the causal relationship of individual stock-related factors based on deep learning models are gradually stimulating the interest of researchers.

Tank et al. [38] introduced the neural Granger test, enhancing the detection of Granger causality by incorporating non-linear interactions. The researchers proposed a suite of non-linear architectures, wherein each time series is modeled using either Multi-Layer Perceptron (MLP) or RNN. Inputs to this non-linear framework consist of past lags from all series, while the outputs predict the future values of each series. Additionally, a group lasso penalty is applied to effectively reduce the input weights to zero, refining the predictive accuracy of the model.

In order to better comprehend and explicate causal relationships in data, Donald Rubin put forward the POF. The POF seeks to reveal dynamic causal relationships through randomized trials and natural experiments [39,40]. Additionally, Judea Pearl introduced the causal diagrams model as a formalized approach for researchers to describe and infer causal relationships. Combining the potential outcomes framework and causal diagrams can lead to a better understanding and explanation of causal relationships in observed data for causal inference.

Although there has been significant progress in the investigation of causality within stock data, the Granger causality test, designed for linear and stable data, exhibits biases when applied to non-linear and unstable stock data. Moreover, the majority of causal inference methodologies neglect the issue of interference caused by confounding variables, as well as the lack of further validation of the results. To address the aforementioned issue, in this paper, a deep learning-based causal inference architecture and algorithm inspired by the POF, which also integrates causal diagrams and the non-linear Granger test, is proposed. This framework and algorithm are specifically designed to infer the causal relationships between individual stock closing prices and their relevant factors.

The innovation of this paper lies in the utilization of causal diagrams and the establishment of a grouped architecture through deep learning networks. The primary contributions of the proposed methods are summarized as follows:

1. To better understand the impact of confounding variables on causal relationships, causal diagrams are employed in the stock data analysis to explore the relationships

among confounding variables, treatment variables, and target variables. The application of front-door and backdoor adjustments allows for the control of confounding variables and the accurate inference of the relationship between closing prices and relevant factors.

2.　To leverage the computational power of deep learning networks and address the deficiencies of the Granger test, a sliding window strategy is incorporated into the GRU model to achieve precise estimation of closing prices. The enhanced capability of GTU serves to expand the applicability of the Granger test beyond the realm of linear stationary data, facilitating the direct assessment of causal linkages within non-linear time series data.

3.　To control for confounding variables, a grouped architecture structure is built using GRUs combined with a sliding window strategy, inspired by the POF. Additionally, the non-linear Granger test is utilized to infer the causal relationships between individual stock closing prices and relevant factors, thus implementing a deep learning-based causal inference framework and algorithm.

4.　To further validate the accuracy of the inferred causal relationships, different sets of input variables, such as individual closing prices, all related factors, and causal factors, are used when predicting stock closing prices. The results show that including causal factors as input variables significantly enhances prediction accuracy. These findings provide additional validation of the effectiveness and reliability of the proposed algorithm.

The remainder of the paper is organized as follows: In Section 2, we introduce causal diagrams and the Granger test as well as presents a deep learning-based architecture and algorithm for causal inference between stock closing prices and relevant factors. The dataset used as well as the evaluation metrics are also presented in Section 2. The experimental results are presented in Section 3. Section 4 discusses the experimental results to validate the effectiveness of the algorithm. Some conclusions are provided in Section 5.

2. Materials and Methods

2.1. Causal Diagrams

Causal diagrams, which utilize directed acyclic graphs (DAGs), were introduced by Judea Pearl to describe the causal relationships between variables. Causal diagrams are helpful in eliminating estimation bias through conditional distributions [41]. The fundamental idea behind this approach is to estimate and test distributions while minimizing bias introduced by other variables. Due to the presence of confounding variables, three distinct paths arise in causal inference: causal paths, backdoor paths, and front-door paths. Figure 1 illustrates these three paths. Here, Z refers to confounding variable, X represents the treatment variable, and Y represents the target variable.

(**a**)　　　　　　　　　　　　(**b**)　　　　　　　　　　　　(**c**)
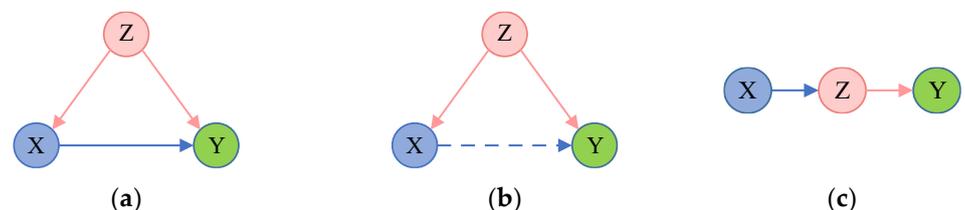
Figure 1. Paths in a causal diagram. (**a**) Causal path: there is a direct causal relationship between the treatment variable X and the target variable Y, and there may be a confounding variable Z. (**b**) Backdoor path: the confounding variable Z affects both the treatment variable X and the target variable Y (not between $X \rightarrow Y$, but both). (**c**) Front-door path: the confounding variable Z affects the path from X to Y (directly affects $X \rightarrow Y$).

The set of variable Z in backdoor path satisfies the backdoor criterion:

1.　Z does not contain any descendant nodes of X.

2. *Z* blocks every path from *Y* to *X* that contains a connection to *X*.

The variable *Z* set in the front-door path satisfies the front-door criterion:

1. *Z* cuts off all directed paths from *X* to *Y*.
2. There is no backdoor path from *X* to *Z*.
3. *X* blocks all backdoor paths from *Z* to *Y*.

The significance of the backdoor criterion and the front-door criterion lies in their ability to estimate certain causal effects using observed data, even when some variables are unobservable. These two criteria are helpful in identifying confounding variables and in designing experimental studies.

The forthcoming experimental design will employ backdoor adjustment and front-door adjustment to truncate the backdoor paths and front-door paths, respectively, based on the backdoor criterion and the front-door criterion. The influence of confounding variables on the causal paths will be eliminated by doing so, allowing the model to correctly identify and assess the causal relationships and effects among stock-related factors.

2.2. Granger Causality Test

The Granger causality test uses statistical techniques to analyze the causality of economic variables [42]. The existence and direction of causal relationships between variables are determined through the assessment of the significance of respective prior period indicators, as reflected by the lagged variables of the economic variables. This assessment helps to explain and influence indicators of each other, leading to conclusive results. The Granger causality test commonly employs a distributional lag model to infer whether the previous level of variable X impacts the subsequent level of variable Y, which is typically represented as follows:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \beta_1 X_{t-1} + \cdots + \beta_p X_{t-p} + \varepsilon_t \tag{1}$$

where X_{t-i} represents the distributed lag term of X, examining whether X has an impact on the current level of Y. The coefficient β_i reflects the magnitude of this impact, indicating the existence of causality. Y_{t-i} is the distributed lag term of Y, and the coefficient α_i represents its impact. ε_t denotes the error term. To infer the causality of X on Y is to examine the following hypotheses:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_p = 0 \tag{2}$$

This hypothesis is generally tested by constructing the *F*-test statistic Equation (3), which is defined as

$$F = \frac{(RSS_0 - RSS_1)/p}{RSS_1/(T - 2p - 1)} \tag{3}$$

where RSS_0 represents the sum of squared errors under the null hypothesis H_0, RSS_1 is the sum of squared errors under the alternative hypothesis H_1 in Equation (4), p denotes the lag length, and T is the sample size. This statistic satisfies the *F*-distribution, when the null hypothesis is that X is not the cause of Y. By referring to the *F*-distribution table, one can determine the statistical significance of the test at a specific confidence level. If the original hypothesis is rejected, this indicates that there is a causal relationship from variable X to variable Y.

The alternative hypothesis H_1 is denoted as

$$H_1: \beta_1 = \beta_2 = \cdots = \beta_p = 1 \tag{4}$$

2.3. Temporal Causal Network

The stock closing price and its relevant factors are time-series data. Figure 2 depicts the causal relationships among these temporal data. In this figure, X_i refers to the treatment

variable, Z_i represents the potential confounding variable, and Y denotes the target variable. The arrows in the figure indicate the direction of influence between the variables.
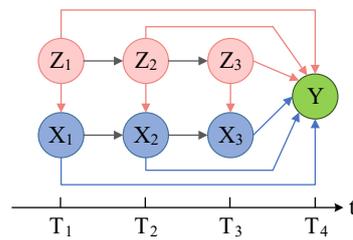
Figure 2. Temporal dynamics of the treatment variable, confounding variable, and target variable in a causal network.

Specific techniques must be employed for these confounding variables that vary over time. Before using these methods, certain generalized premise assumptions must be fulfilled: [43]

1. Sequential Ignorability Assumption: the assumption is that for each time point, if after controlling for a series of values of the confounding factors, the outcome of the treatment at each time point is only affected by its own treatment status and not by the treatment status of other time points, i.e., the effect of accepting the treatment on a single individual is independent of whether other individuals accept the treatment or not. The Assumption can be expressed as

$$Y_{potential}|X_{t-1}, T_0, Z_t = Y_{potential}|X_{t-1}, T_1, Z_t \tag{5}$$

here, for time point t, the values of various confounding variables at each time point prior to the time point, which includes time point t, can be denoted as Z_t. The string of historical values of the treatment variable prior to t, excluding t, is written as X_{t-1}. The various potential values of Y are represented as $Y_{potential}$. T_0 indicates that Y does not accept the treatment and T_1 indicates that Y accepts the treatment. If this assumption is not satisfied, there is in fact a confounding of causal relationships at each time point. And it is impossible to accurately estimate the causal effect at each time point.

2. Consistency Assumption: this assumption requires that the observed value of Y under a specific sequence of treatment variable values is equal to its potential value.

3. Positive Value Assumption: this assumption refers to the probability of an individual receiving a treatment intervention at time point t being between 0 and 1, but not equal to either 0 or 1, after controlling for a series of confounding variables Z_t up to and including time point t, and the sequence of treatment variable values X_{t-1} before time point t.

Introducing a time dimension in deep learning models allows for the creation of temporal structures to better handle time-related confounding variables. For instance, models such as RNN or LSTM are used to capture temporal dependencies. The prediction of stock closing prices using deep learning models can be obtained for the values in potential states, satisfying the consistency assumption. In addition, neural networks can satisfy the positivity assumption with activation functions, such as the Sigmoid function, Tanh function, ReLu function, etc.

However, experimental design is still required to satisfy the sequential ignorability assumption. After fulfilling the three main assumptions, the temporal causal network can be abstracted as a neural network, where the arrows in the neural network are determined by their corresponding weights. If the weight is zero, it implies that the corresponding path does not exist.

2.4. Deep Learning-Based Causal Inference Network Architecture and Algorithms

The causal inference architecture based on deep learning is illustrated in Figure 3. This architecture employs GRU networks to capture temporal dependencies in sequence

data and extract feature representations. Both LSTM and GRUs address the issues of gradient vanishing, gradient explosion, and long-term dependencies in a traditional RNN. However, compared to LSTM, GRUs possess a more concise structure, which makes them computationally faster and suitable for handling large-scale datasets. Meanwhile, GRUs exhibit higher efficiency in memory utilization, reducing the burden of storage and computation.
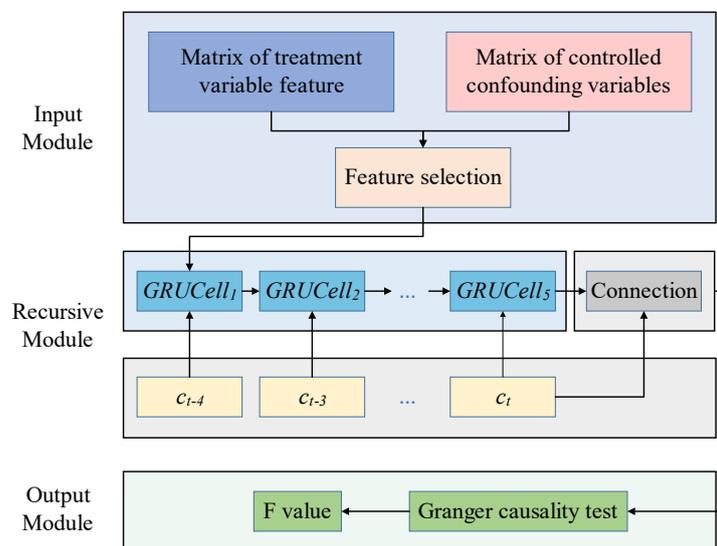
Figure 3. Causal inference network architecture based on deep learning.

A sliding window strategy with a window size of 5 is used in GRUs. This strategy allows the segmentation and processing of time series data in fixed window lengths. By sliding the window, continuous subsequences can be obtained and utilized for further analysis and modeling. This approach proves to be highly effective in capturing local patterns and dynamic features within sequences, thereby enhancing the performance and effectiveness of the model.

The fully connected layer converts the feature extraction and representation from the previous layer into the final output result. By learning the weights of each connection, the fully connected layer can adjust these weights during training to minimize the loss function, allowing the model to make accurate predictions.

To satisfy the assumption of sequential ignorability and eliminate the interference brought by confounding variables in analyzing causal relationships, backdoor adjustment and front-door adjustment were conducted. The input data were divided into two groups: an experimental group and a control group. When one test factor is selected as the treatment variable, the other test factors are the confounding variables, and the stock closing price serves as the target variable.

Control group: the historical information of confounding variables Z and the target variable Y is utilized to predict the target variable Y.

Experimental group: The historical information of the treatment variable X, confounding variables Z, and the target variable Y is utilized to predict the target variable Y.

The distribution of confounding variables Z remains unchanged between the control group and the experimental group, with the target variable consistently being the closing price. According to the backdoor criterion and the front-door criterion, it can be inferred that all the backdoor paths and front-door paths between X and Y are cut off. This experimental design also satisfies the MB-by-MB (Markov Blanket by Markov Blanket) algorithm in the local learning of causal networks [44], and the optimal stepwise intervention design in the active learning of causal networks [45].

The corresponding network model for causal inference architecture is shown in Figure 4. The innovation of this model is to build a grouped architecture using two GRU networks, which achieves the control of confounding variables and the accurate prediction of the closing price of the target variable under different circumstances. It is coupled with a sliding window strategy, so that it satisfies the assumption of sequential ignorability.
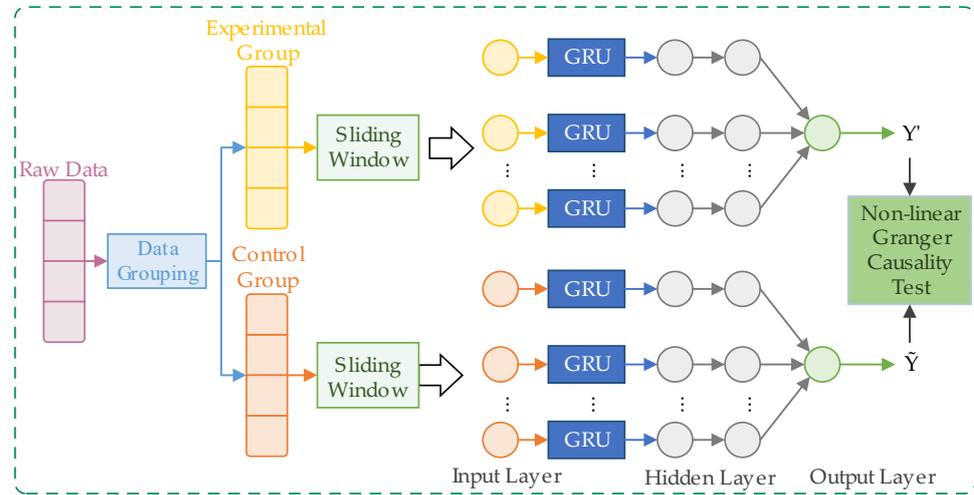
Figure 4. Causal inference network model based on deep learning.

Specifically, when a factor is selected as a treatment variable, other factors are potential confounding variables, and the closing price of an individual stock is the target variable. The data underwent initial processing and grouping to obtain the experimental and control groups. Then the two groups of data were used as inputs, and the potential values of closing prices in different situations were obtained by sliding window and GRU calculations, respectively, and then the causal relationship between the closing prices of individual stocks and the relevant factors was inferred by the non-linear Granger test.

Closing prices were predicted using data from the experimental group and control group. Since the computational process of the GRU network is non-linear, the formula of the lagged distribution model under the H_1 assumption is rewritten as

$$Y_t' = \alpha_0 + f\left(\sum_{i=1}^{p} \beta_i X_{t-i}, \sum_{d=1}^{m}\sum_{i=1}^{p} \gamma_i^d Z_{t-i}^d, \sum_{i=1}^{p} \alpha_i Y_{t-i}\right) + \varepsilon_t \tag{6}$$

where α_0 represents the baseline value; $f(\cdot)$ denotes the composite function; p indicates the lag length; α_i, β_i, and γ_i^d are the coefficients of the corresponding distributional lag terms; X_{t-i}, Z_{t-i}^d, and Y_{t-i} are distributional lags of X, Z, and Y, respectively; and ε_t stands for the error term. The output under the H_0 assumption is obtained as:

$$\tilde{Y}_t = \alpha_0 + f\left(\sum_{d=1}^{m}\sum_{i=1}^{p} \gamma_i^d Z_{t-i}^d, \sum_{i=1}^{p} \alpha_i Y_{t-i}\right) \tag{7}$$

Subsequently, a Granger causality test is conducted by calculating the F-value. The formula is as follows:

$$F = \frac{\sum_{t=1}^{T}\left((\tilde{Y}_t - Y_t)^2 - (\tilde{Y}_t - Y_t)^2\right)/p}{\sum_{t=1}^{T}\left(Y_t' - Y_t\right)^2/(T-2p-1)} \tag{8}$$

where Y_t is the true value of the closing price. If $F > F_a$ (where F_a is obtained by querying the F-distribution table using T and p values), the null hypothesis H_0 is rejected. Thus, it is inferred that there is a causal relationship between the treatment variable and the target variable, i.e., the treatment variable is the cause of the target variable.

Algorithm 1 shows the algorithm for the causal inference network architecture based on deep learning.

Algorithm 1. Causal Inference Algorithm Based on Deep Learning.

Input: The experimental dataset contains T samples, N features $D_I = \{(W_t, Y_t^{'})\}$; the control group dataset contains T samples and M features $D_{II} = \{(H_t, \tilde{Y}_t)\}$, $W_t = \left(X_t^i, Z_t^d, Y_{t-1}, C_t^{'}\right)$, $H_t = (Z_t^d, Y_{t-1}, \tilde{C}_t)$, $Y_t^{'} = c_{t+1}^{'}$, $\tilde{Y}_t = \tilde{c}_t$, $X_t^i = \left[x_{t-p+1}^i, x_{t-p+2}^i, \ldots, x_t^i\right]_{i=1}^N$, $Z_t^d = \left[z_{t-p+1}^d, z_{t-p+2}^d, \ldots, z_t^d\right]_{d=1}^M$, $Y_{t-1} = \left[y_{t-p+1}, y_{t-p+2}, \ldots, y_{t-1}\right]$, $C_t^{'} = \left[c_{t-p+1}^{'}, c_{t-p+2}^{'}, \ldots, c_t^{'}\right]$, $\tilde{C}_t = \left[\tilde{c}_{t-p+1}, \tilde{c}_{t-p+2}, \ldots, \tilde{c}_t\right]$, the lag length p, the number of training cycles E

Output: Granger causality of closing price

```
 1   Initialize the parameters of the GRU model;
 2   result_list = []   # Storage for the results of each iteration.
 3   for e ∈ (1, E) do
 4       for i ∈ (1, N) do
 5           Y_t' _list = []    # Storage for the results y1 of each iteration.
 6           Ỹ_t_list = []     # Storage for the results y2 of each iteration.
 7           for t ∈ (1, T) do
 8               Y_t' = GRU(X_t^i, Z_t^d, Y_{t-1}, C_t');
 9               Ỹ_t = GRU(Z_t^d, Y_{t-1}, C̃_t);
10               Y_t'.append(Y_t');
11               Ỹ_t.append(Ỹ_t);
12           end for
13           F_value = F_calculate(y_{t1}, y_{t2})
14           F_a = F_find(T, p)
15           causal_factor = Granger_test(F_value, F_a)
16           result_list.append(causal_factor)
17       end for
18   end for
19   return causal_factor
```

2.5. Dataset

Using BaoStock to obtain the time series data of a stock, there are 12 factors to consider for evaluating the causal relationship of closing prices of the stock. These factors include opening price, highest price, lowest price, trading volume, trading amount, turnover rate, percentage change, price-earnings (P/E) ratio, price-to-book (P/B) ratio, price-to-sales (P/S) ratio, price-to-cash flow (P/CF) ratio, and the Shanghai Stock Exchange (SSE) Index.

The SHCOMP is a comprehensive stock index that reflects the performance of the overall A-share market. If the SHCOMP undergoes significant fluctuations, the majority of stock prices will be affected. Based on historical data and market performance, certain industries are more affected to changes in the SHCOMP, including the following industries:

- Financial industry: Due to the significant impact of government policies and regulations on the financial market, the volatility of financial stocks is more pronounced relative to other industries. China Taibao (sh.601601) was chosen as a representative. It is a Chinese insurance company with a substantial market capitalization and a particular degree of influence.
- Real estate industry: The real estate market has a large contribution, weighting ratio to the Shanghai stock market. The relaxation or tightening of property market policies has a considerable impact on the volatility of stock prices. Poly Real Estate (sh.600048), a renowned real estate developer in China, involved in diverse sectors, such as residential, commercial real estate, and office buildings, has been selected as a representative.

- Energy and raw materials industry: The profitability of these industries is affected by factors such as changes in the global supply and the demand of raw materials, international oil prices, policy environment, among others. China Petroleum & Chemical Corporation (sh.601857) has been selected. It is one of largest oil and gas producers in China, with abundant energy resources and a significant market share.

There are also some industry stocks that are relatively less affected. Based on historical data and market performance, the following are some of the industries that are less affected by the volatility of the SHCOMP:

- Public utility companies: Public utility companies typically exhibit a relatively stable earnings model and decent cash flow. As a result, they may be relatively less affected by the volatility of the SHCOMP. China Guodian (sh.601985) is chosen because it is one of the largest power companies in China.
- Food & beverage industry: As a general trend, food and beverage enterprises tend to have a stable income and profit model, with a relatively stable market and less volatility in comparison to other industries. Yingjia Gongjiu (sh.603198), a well-known Chinese Baijiu brand, was chosen as a representative.
- Banking industry: Although the financial industry as a whole tends to be volatile, the stock prices of the banking industry are relatively less affected by the SHCOMP because it has substantial cash flows and asset-liability structure. Shanghai Pudong Development Bank (sh.600000) was chosen due to a relatively stable business model and income.

To ensure sufficient data support and incorporate diverse market conditions, the date range of the selected stocks is from 1 July 2017 to 1 July 2022. This approach also aims to minimize the disturbance of structural changes, allowing for a comprehensive observation and analysis of long-term trends and cyclical fluctuations in the stock market, ultimately enhancing the reliability and generalization of the findings.

2.6. Evaluation Parameter

(1) Root mean square error (*RMSE*)

The *RMSE* is the square root of the sum of squared differences between predicted values and true values, which is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - y_i'\right)^2} \tag{9}$$

where y_i denotes the predicted value, y_i' represents the true value, and N is the number of observations. The sum of squared deviations is highly sensitive to errors that are either significantly larger or smaller. Consequently, the resulting error measure provides a reliable assessment of the predictive performance. A smaller *RMSE* value denotes superior prediction performance, while a larger *RMSE* value indicates a greater divergence from the true results.

(2) Mean absolute error (*MAE*)

The MAE is the average of absolute differences between predicted values and true values of the model. As a result, it intuitively captures the discrepancy between predicted and true values. The *MAE* is determined by

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - y_i'\right| \tag{10}$$

A smaller *MAE* value indicates a smaller difference between predicted values and true values, suggesting that the prediction results are closer to true values.

(3) Mean absolute percentage error (*MAPE*)

The *MAPE* diminishes the influence of magnitude in comparison to the previous two metrics, making it well-suited for assessing the efficacy of a model in predicting various stocks, which is formulated as follows:

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{y_i - y_i^{'}}{y_i}\right| \tag{11}$$

The *MAPE* is also an error metric, with smaller values indicating a better performance of the predictive model.

(4)   The R^2 coefficient of determination

The R^2 coefficient of determination is utilized to evaluate the fitting degree of a network model, which can be expressed as

$$R^2 = \frac{\sum_{i=1}^{N}\left(y_i - \overline{y_i}\right)^2}{\sum_{i=1}^{N}\left(y_i^{'} - \overline{y_i}\right)^2} \tag{12}$$

where $\bar{y}_i^{'}$ is the average of the real values. The R^2 coefficient of determination serves as a measure to evaluate the fitting and predictive capabilities of the model. A higher value signifies a greater predictive ability and improved accuracy.

3. Results

3.1. Experimental Hardware and Software Environment

The experiment was conducted using the Python-based TensorFlow framework to build a deep learning network. The central processor used was an Intel® Core™ i7-9750H, and the graphics card was Nvidia GeForce GTX1650 with 4 GB of memory. The learning rate of optimizer Adam was set to 0.001, batch training size was 64, and the total training epochs were set to 80. To accurately examine causality, the causal inference experiments used a fixed seed to initialize weights. The software used for data analysis was Pycharm version 2022.3.2.

3.2. Data Preprocessing and Normalization

After getting the stock data through BaoStock, we checked if there were any missing values in the stock data and, if there were, we filled them with the data of the previous day. After the missing values were processed, the data format was standardized by Equation (13), defined as follows:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{13}$$

where x is the value to be normalized, and x_{min} and x_{max} are the minimum and maximum values in the feature X, respectively. The output of the function x_{norm} is the result of the maximum and minimum normalization of x. This formula scales each value in the data to between [0, 1], eliminating the order of magnitude effect between features while preserving the relative size relationship between the data. The sample size for each statistical analysis was 1215.

3.3. Causal Inference Experiment

The length of the dataset is 1215, denoted as sample size *T*. The GRU model adopts a sliding window strategy with a window size of 5, represented by the lag length *p* as 5. By referring to the *F* distribution table, the critical value F_a corresponding to the values *T* = 1215 and *p* = 5 was determined to be 3.501. If the calculated *F* value surpassed 3.501, it means that the factor passed the Granger test. To enhance the reliability of the results, the average of the results from ten experiments were calculated, shown in Table 1, with a 95% confidence interval.

Table 1. F-test values for each factor in Shanghai Stock Exchange datasets (bolded values are those that pass the Granger test).

	sh.601601	sh.600048	sh.601857	sh.601985	sh.603198	sh.600000
Opening Price	**9.752**	−0.228	−8.304	−31.640	**37.710**	**16.920**
Highest Price	**9.046**	0.837	0.215	−32.979	**20.381**	−19.716
Lowest Price	−4.792	**12.579**	−20.316	−30.295	−6.734	**6.556**
Trading Volume	−2.404	**8.502**	−26.561	**3.524**	−32.866	−0.444
Trading Amount	−3.935	−3.709	**16.142**	**17.580**	**35.538**	**4.309**
Turnover Rate	**5.042**	−0.577	−3.418	**6.490**	**40.980**	**8.669**
Percentage Change	**6.792**	−3.110	−8.044	−53.835	−2.449	**2.660**
P/E Ratio	−0.894	**7.611**	−32.997	**9.709**	**31.841**	−20.402
P/B ratio	**8.782**	**8.408**	**13.859**	**68.545**	−55.818	−22.093
P/S ratio	−19.365	**6.201**	**14.168**	**14.391**	**35.568**	−41.630
P/CF ratio	**10.574**	**8.397**	−8.795	**3.898**	**16.867**	−32.186
SHCOMP	**14.464**	**13.915**	**25.187**	**9.314**	**4.797**	**8.986**

3.4. Prediction Comparison Experiment

(1)   Comparison of different input variables

The baseline RNN, LSTM, and GRU models, which integrated all potentially relevant factors, were compared with the RNN, LSTM, and GRU models that incorporated causal factors in the different datasets. Furthermore, the analysis included comparisons with a GRU model that exclusively utilized closing price data. This GRU model, which focused solely on closing price data, was also part of the comparative evaluation. Figures 5 and 6 illustrate the prediction results of the experiments for China Taibao and Shanghai Pudong Development Bank (SPDB), respectively. Figures 7 and 8 show scatter plots for the experiments conducted on China Taibao and SPDB, respectively. Visual presentations of the remaining datasets, as well as scatterplots, can be viewed in the Supplementary Materials.

Table 2 presents the RMSE, MAE, MAPE, and R^2 for the seven kinds of models. To account for the varying scale ranges of the target variables in different datasets, the evaluation metrics are calculated using standardized data. This standardization process eliminates scale differences, allowing for more intuitive comparisons of errors. We compared the performance of the RNN, LSTM, and GRU models in the stock price prediction task with the same input variables, and provided support for the algorithm to apply the GRU model to calculate the potential value of the target variable. In the case of the same network structure, we compared the performance of models with different input variables to verify the validity and accuracy of the experimental results of causal inference.
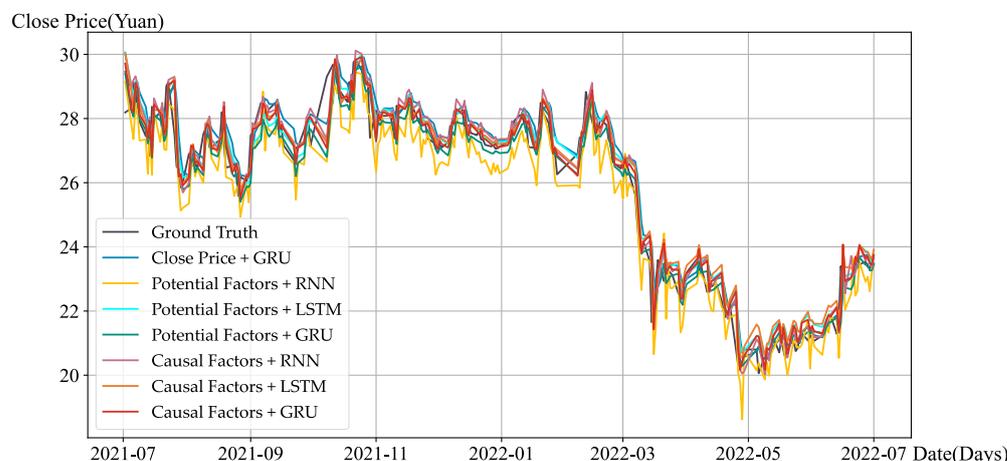
Figure 5. Visualization of prediction results for different models on the China Taibao (sh.601601) dataset.
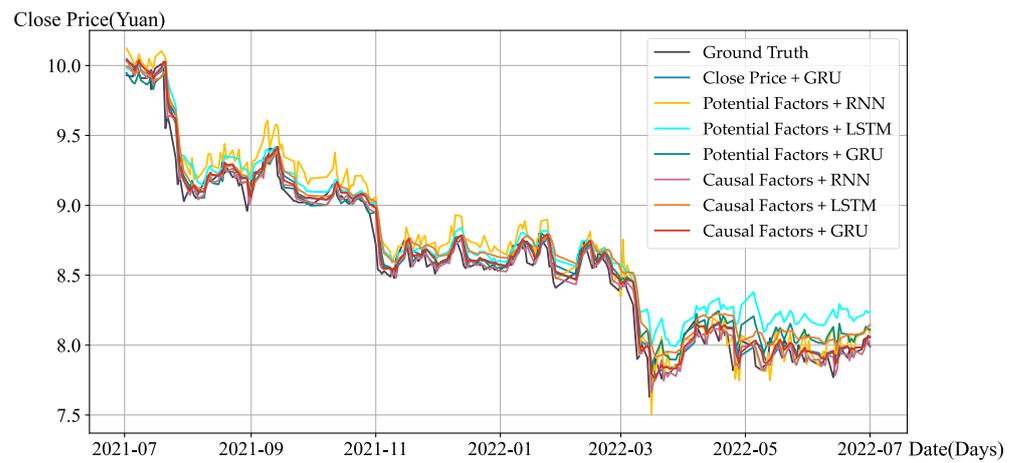
Figure 6. Visualization of prediction results for different models on the SPDB (sh. 600000) dataset.
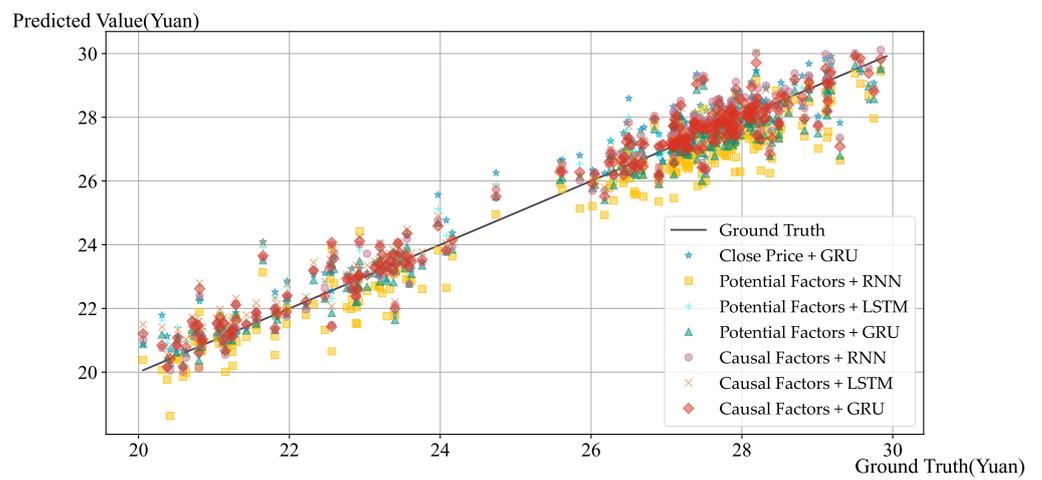
Figure 7. Scatter plot of prediction results for different models on the China Taibao (sh.601601) dataset.
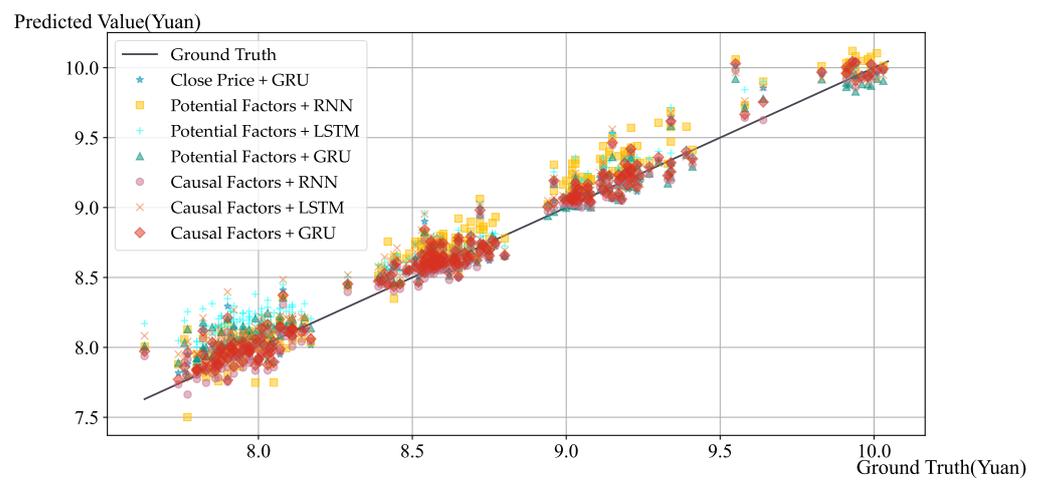
Figure 8. Scatter plot of prediction results for different models on the China SPDB (sh.600000) dataset.

Table 2. Comparison of evaluation metrics for different models in Shanghai Stock Exchange datasets (↑ indicates that larger values are better, ↓ indicates that smaller values are better, the best result among comparative experiments is in bold).

Model	Stock	RMSE↓	MAE↓	MAPE↓	R² ↑
Close Price + GRU	sh.601601	0.0596	0.0453	14.96	0.9544
	sh.600048	0.0514	0.0392	41.43	0.9493
	sh.601857	0.1033	0.0773	21.96	0.7554
	sh.601985	0.0648	0.0488	12.27	0.9356
	sh.603198	0.0522	0.0381	18.69	0.9599
	sh.600000	0.0503	0.0400	17.50	0.9561
Potential Factors + RNN	sh.601601	0.0720	0.0551	17.65	0.9324
	sh.600048	0.0660	0.0510	44.65	0.9349
	sh.601857	0.1133	0.0856	23.63	0.7301
	sh.601985	0.0782	0.0593	15.48	0.9181
	sh.603198	0.0631	0.0468	21.57	0.9453
	sh.600000	0.0572	0.0452	18.99	0.9357
Potential Factors + LSTM	sh.601601	0.0607	0.0485	14.89	0.9476
	sh.600048	0.0534	0.0425	42.78	0.9483
	sh.601857	0.0839	0.0673	19.36	0.8552
	sh.601985	0.0659	0.0504	11.98	0.9361
	sh.603198	0.0548	0.0419	19.91	0.9591
	sh.600000	0.0492	0.0377	16.53	0.9628
Potential Factors + GRU	sh.601601	0.0599	0.0472	14.02	0.9497
	sh.600048	0.0522	0.0412	41.96	0.9501
	sh.601857	0.0806	0.0640	18.91	0.8617
	sh.601985	0.0633	0.0481	11.54	0.9402
	sh.603198	0.0530	0.0412	19.47	0.9610
	sh.600000	0.0475	0.0361	16.21	0.9651
Causal Factors + RNN	sh.601601	0.0628	0.0479	16.74	0.9488
	sh.600048	0.0568	0.0437	43.47	0.9437
	sh.601857	0.1100	0.0830	22.82	0.7431
	sh.601985	0.0714	0.0537	14.19	0.9273
	sh.603198	0.0589	0.0431	20.36	0.9554
	sh.600000	0.0542	0.0425	17.67	0.9449
Causal Factors + LSTM	sh.601601	0.0516	0.0403	14.53	0.9574
	sh.600048	0.0504	0.0395	41.57	0.9518
	sh.601857	0.0758	0.0613	17.42	0.8713
	sh.601985	0.0601	0.0461	11.16	0.9436
	sh.603198	0.0495	0.0371	18.39	0.9649
	sh.600000	0.0417	0.0328	13.47	0.9699
Causal Factors + GRU	sh.601601	**0.0493**	**0.0387**	**12.67**	**0.9608**
	sh.600048	**0.0487**	**0.0377**	**34.60**	**0.9544**
	sh.601857	**0.0667**	**0.0510**	**14.92**	**0.8980**
	sh.601985	**0.0597**	**0.0445**	**10.64**	**0.9452**
	sh.603198	**0.0470**	**0.0354**	**18.07**	**0.9674**
	sh.600000	**0.0383**	**0.0283**	**11.60**	**0.9746**

Figure 9 visualizes the comparative analysis of the GRU model experiment results across different input variables in different datasets.
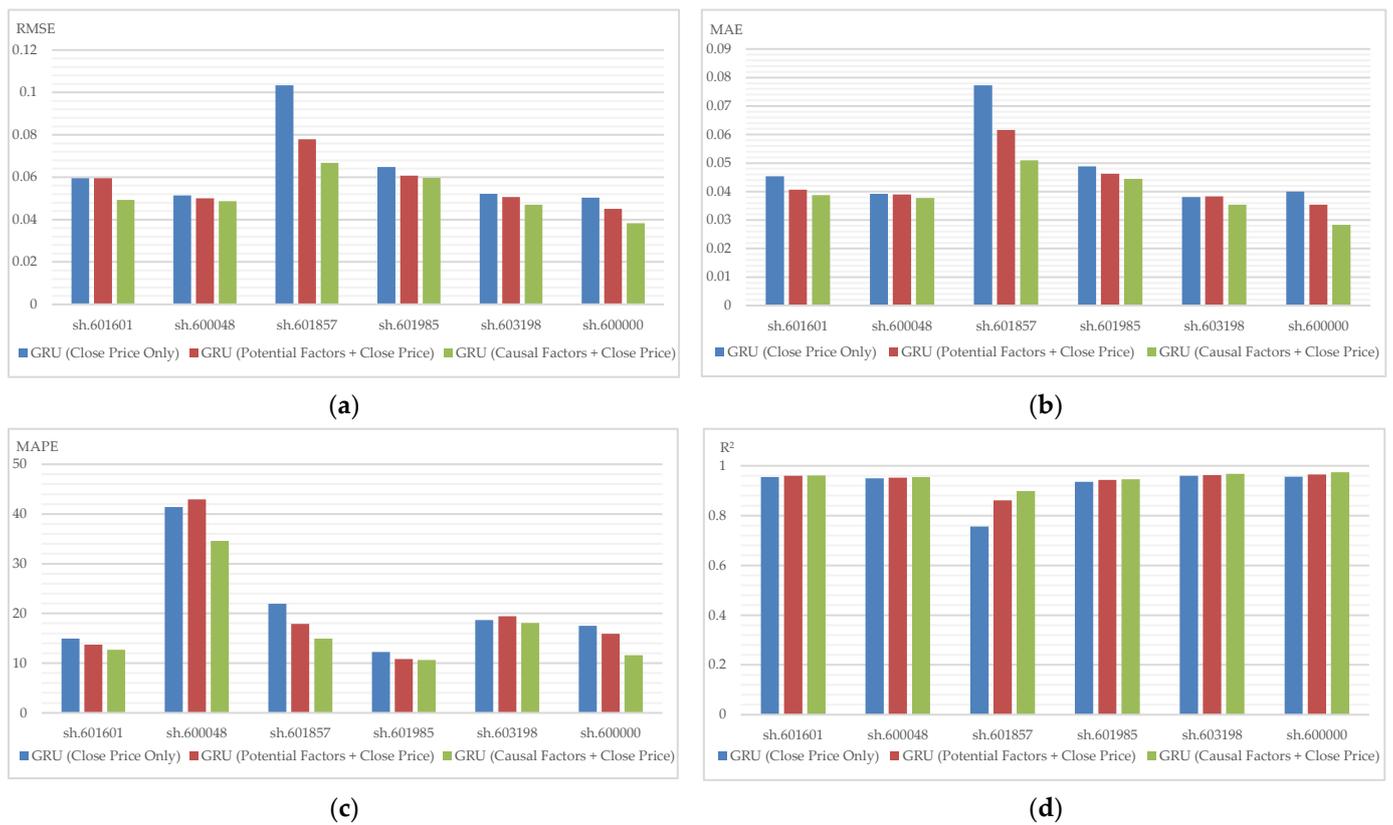
Figure 9. Visualization of comparative experimental results of GRU models with different input variables in different datasets. (**a**) RMSE metric results of various methods. (**b**) MAE metric results of various methods. (**c**) MAPE metric results of various methods. (**d**) R^2 metric results of various methods.

In order to further verify the generality of the model, the data of certain stocks in the Shenzhen Stock Exchange Composite Index (SZCI) from 1 July 2017 to July 2022 were selected; details and the experimental results are in Appendix A.

4. Discussion

The causal inference experiment employs the Granger causality test to determine the causality of factors in industries that are highly and moderately influenced by the SHCOMP. The results show that, in highly influenced industries, causal factors included the open price, high price, low price, trading volume, trading amount, turnover rate, percentage change, P/E ratio, P/B ratio, and the related index itself. In contrast, in less influential industries, the causal relationship between the remaining factors and the closing price is more significant, except for the related index. These findings suggest that, in highly influential industries, individual stock closing prices are more significantly affected by index factors. Meanwhile, the causal relationship between individual stock factors is more pronounced in less influential industries.

The performance of various models was compared using identical input variables, and Table 3 illustrates the percentage improvement in performance of the optimal model compared to other models. From the data presented in the table, it is evident that the GRU model outperformed the RNN and LSTM models in predictive accuracy. This outcome proves that the decision of the proposed algorithm and framework to use GRUs to compute the potential value of the target variable is a reasonable and appropriate choice.

Table 4 shows the percentage improvement in predictive performance of the model, with input causal factors compared to the baseline model with input from all potential factors. The results show that the enhanced model with causal factors input outperformed the corresponding benchmark model with all potential factor input in terms of predictive

performance. For example, in the dataset of stock sh.601601, the inclusion of causal factors resulted in performance improvements for the RNN, LSTM, and GRU models compared to their corresponding baseline models. For the RNN model, there was an enhancement of 12.78% in RMSE, 13.07% in MAE, 5.15% in MAPE, and 1.76% in R2. Regarding the LSTM model, RMSE saw a 14.99% enhancement, MAE improved by 16.91%, MAPE by 2.42%, and R2 by 1.03%. For the GRU model, the RMSE was enhanced by 17.70%, MAE by 18.01%, MAPE by 15.65%, and R2 by 1.16%.

In addition to this, the table also compares the performance between GRU models, predicted using data of closing prices and data predicted using causal factors. The results show that the model using causal factors as input variables performed the best. For example, in the dataset of stock sh.600000, the GRU model using causal factors showed a significant performance improvement compared to the GRU model using only closing price: 23.86% on RMSE, 29.25% on MAE, 33.71% on MAPE, and R^2 1.93%. Furthermore, relative to the GRU model using all latent factors, the causal factor model improved by 15.08% on RMSE, 20.06% on MAE, 26.95% on MAPE, and 1.03% on R^2.

The utilization of causal factors not only reduces the amount of noisy information that the model has to deal with, but also focuses on the most relevant information, thus improving the predictive performance of the model. The causal factors inferred by the method proposed in this paper as inputs to the closing price prediction GRU model outperformed the models using closing price as inputs in terms of prediction accuracy and prediction performance, while the prediction performance of the benchmark model with all potential factors input was inferior to that of the model with causal factors inputs. These results further test the causal inference of correctness and enhance the reliability and validity of the method proposed in this paper.

Table 3. Percentage performance improvement in the optimal model over other models.

Input Variables	Optimal Model	Comparative Model	Stock	Percentage/%			
				RMSE	MAE	MAPE	R²
Potential Factors	GRU	RNN	sh.601601	16.81	14.34	14.9	1.86
			sh.600048	20.91	19.22	6.02	1.63
			sh.601857	28.86	25.23	19.97	18.02
			sh.601985	19.05	18.89	22.22	2.41
			sh.603198	16.01	11.97	9.74	1.66
			sh.600000	16.96	20.13	14.64	3.14
		LSTM	sh.601601	1.32	2.68	5.84	0.22
			sh.600048	2.25	3.06	1.92	0.19
			sh.601857	3.93	4.90	2.32	0.76
			sh.601985	3.95	4.56	3.67	0.44
			sh.603198	3.28	1.67	2.21	0.20
			sh.600000	3.46	4.24	1.94	0.24
Causal Factors	GRU	RNN	sh.601601	21.50	19.21	24.31	1.26
			sh.600048	14.26	13.73	20.40	1.13
			sh.601857	39.36	38.55	34.62	20.85
			sh.601985	16.39	17.13	25.02	1.93
			sh.603198	20.20	17.87	11.25	1.26
			sh.600000	29.34	33.41	34.35	3.14
		LSTM	sh.601601	4.46	3.97	12.80	0.36
			sh.600048	3.37	4.56	16.77	0.27
			sh.601857	12.01	16.80	14.35	3.06
			sh.601985	0.67	3.47	4.66	0.17
			sh.603198	5.05	4.58	1.74	0.26
			sh.600000	8.15	13.72	13.88	0.48

Table 4. Percentage increase in model performance for inputting causal variables versus inputting other variables.

Optimal Model	Comparative Model	Stock	Percentage/%			
			RMSE	MAE	MAPE	R²
Causal Factors + RNN	Potential Factors + RNN	sh.601601	12.78	13.07	5.16	1.76
		sh.600048	13.94	14.31	2.64	0.94
		sh.601857	2.91	3.04	3.43	1.78
		sh.601985	8.70	9.44	8.33	1.00
		sh.603198	6.66	7.91	5.61	1.07
		sh.600000	5.24	5.97	6.95	0.98
Causal Factors + LSTM	Potential Factors + LSTM	sh.601601	14.99	16.91	2.42	1.03
		sh.600048	5.62	7.06	2.83	0.37
		sh.601857	9.65	8.92	10.02	1.88
		sh.601985	8.80	8.53	6.84	0.80
		sh.603198	9.67	11.46	7.63	0.60
		sh.600000	15.24	13.00	18.51	0.74
Causal Factors + GRU	Close Price + GRU	sh.601601	17.28	14.57	15.31	0.67
		sh.600048	5.25	3.83	16.49	0.54
		sh.601857	35.43	34.02	32.06	18.88
		sh.601985	7.87	8.87	13.28	1.03
		sh.603198	9.96	7.09	3.32	0.78
		sh.600000	23.86	29.25	33.71	1.93
	Potential Factors + GRU	sh.601601	17.70	18.01	9.63	1.17
		sh.600048	6.70	8.50	17.54	0.45
		sh.601857	17.25	20.31	21.10	4.21
		sh.601985	5.69	7.48	7.80	0.53
		sh.603198	11.32	14.08	7.19	0.67
		sh.600000	19.37	21.61	28.44	0.98

5. Conclusions

In this study, a causal inference method was applied, combining the GRU model and the Granger causality test, to realize the causality analysis based on stock data. By introducing Granger causality tests, we could identify important factors and determine the degree of influence of index factors on individual stock closing prices. Additionally, the experimental results of incorporating causal factors into the prediction model further validated the correctness of causal inference.

In summary, the deep learning-based causal inference architecture and algorithms proposed in this paper show promising results in analyzing causal relationships in stock data. Future research can further explore the application of causal inference methods in the analysis of other financial time series data. Future research can further explore the application of causal inference methods in the analysis in other financial time series data. In addition, further optimization of the model performance and extension of the application scope can be carried out.

and C.C.; visualization, W.X.; supervision, L.X.; project administration, W.X. and C.C. All authors have read and agreed to the published version of the manuscript.

Appendix A

The SZCI impacts industries and stock prices similarly to the SHCOMP. Historical data show specific industries more influenced by the SZCI:

- Technology industry: Zhong Xing Telecommunication Equipment Corporation (ZTE) (sz.000063) is a leading communications equipment and solutions provider in China, and its position and influence in the technology sector is such that its share price is usually more significantly affected by changes in the SZCI Index.
- Pharmaceutical industry: Aier Ophthalmology (sz.300015) is one of well-known eye medical enterprises of China; its investment and business in the pharmaceutical industry covers a variety of areas, such as ophthalmology diagnosis and treatment, eye surgery, etc.
- New energy industry: BYD (sz.002594) is one of the leading new energy vehicle manufacturers of China, which has strong technical strength and market share in the field of electric vehicles, and its share price is often affected by changes in the Shenzhen Composite Index.

There are some industries whose stock prices are relatively less affected by changes in the SZCI. Listed below are a few industries that may be less affected by the volatility of the SZCI:

- Public service industry: Such as urban infrastructure and domestic waste treatment, etc. These companies are more controlled by the government, and their business is stable and relatively unaffected by industry cyclical factors. China General Nuclear (CGN) Power Corporation (sz.000881) is highly influenced by government policy support and the stability of market demand, and its share price is relatively stable and less affected by fluctuations in the Shenzhen Composite Index.
- Traditional manufacturing industry: Such as machinery, petrochemicals, iron and steel, and other industries related to enterprise. These companies' operating business is relatively stable, profit cycle is more obvious, and they are not affected by fluctuations of the Shenzhen Composite Index too much. The Gree Electric Appliances (sz.000651) sector has a strong market share and brand influence in the traditional manufacturing.
- FMCG industry: Wuliangye (sz.000858) is one of the leading liquor producers, with a stable market share and brand influence in the FMCG industry, and its share price is relatively stable.

Table A1 shows the results of the proposed causal inference method for inferring the causal relationship between correlation factors and stock prices in different datasets.

Table A1. F-test values for each factor in the Shenzhen Stock Exchange datasets (bolded values are those that pass the Granger test).

	sz.000063	sz.300015	sz.002594	sz.000881	sz.000651	sz.000858
Opening Price	−2.015	−5.897	−36.804	**7.329**	**11.129**	−1.516
Highest Price	−7.171	−8.988	−34.948	**17.960**	**34.905**	**12.911**
Lowest Price	−15.864	**5.957**	−17.460	**10.573**	−65.159	−0.072
Trading Volume	−1.259	−5.498	−17.341	**11.056**	−39.112	1.412
Trading Amount	−9.922	**8.055**	**27.314**	−5.350	**5.500**	2.697
Turnover Rate	−8.442	**4.091**	−24.906	−15.264	−24.564	−7.416
Percentage Change	−39.072	−53.889	−24.057	−46.024	**18.947**	−76.086
P/E Ratio	−15.240	−11.801	−3.728	−2.322	−20.583	**6.873**
P/B ratio	−21.259	−3.086	−54.369	**3.817**	−14.869	**6.517**
P/S ratio	−30.907	0.966	0.392	**4.395**	**47.681**	**12.986**
P/CF ratio	**10.837**	**4.083**	−8.094	−49.676	−35.685	**10.580**
SHCOMP	**11.328**	**3.911**	**34.236**	−4.662	−37.955	−1.149

Table A2 shows the results of the experiments comparing different prediction models in different datasets.

Table A2. Comparison of evaluation metrics for different models in Shenzhen Stock Exchange datasets (↑ indicates that larger values are better, ↓ indicates that smaller values are better, the best result among comparative experiments is in bold).

Model	Stock	RMSE ↓	MAE ↓	MAPE ↓	R² ↑
Close Price + GRUs	sz.000063	0.0463	0.0365	31.1451	0.9661
	sz.300015	0.0426	0.0321	13.3153	0.9670
	sz.002594	0.0900	0.0759	22.9663	0.8705
	sz.000881	0.0775	0.0562	12.9659	0.8366
	sz.000651	0.0672	0.0584	21.5701	0.9292
	sz.000858	0.0661	0.0530	21.3026	0.9261
Potential Factors + RNN	sz.000063	0.0565	0.0462	35.5297	0.9497
	sz.300015	0.0455	0.0430	15.1571	0.9672
	sz.002594	0.1086	0.0887	32.3819	0.7870
	sz.000881	0.0964	0.0705	16.7319	0.7473
	sz.000651	0.0654	0.0559	29.6114	0.9347
	sz.000858	0.0734	0.0582	19.4550	0.9088
Potential Factors + LSTM	sz.000063	0.0538	0.0431	35.0179	0.9543
	sz.300015	0.0429	0.0329	13.7151	0.9666
	sz.002594	0.0861	0.0701	27.8958	0.8661
	sz.000881	0.0756	0.0578	13.5359	0.8443
	sz.000651	0.0369	0.0287	14.5130	0.9792
	sz.000858	0.0572	0.0450	16.3179	0.9446
Potential Factors + GRU	sz.000063	0.0420	0.0302	18.8082	0.9722
	sz.300015	0.0411	0.0311	11.8589	0.9712
	sz.002594	0.0719	0.0583	22.5159	0.9065
	sz.000881	0.0747	0.0542	11.7194	0.8480
	sz.000651	0.0344	0.0264	14.0027	0.9819
	sz.000858	0.0547	0.0431	15.8308	0.9494
Causal Factors + RNN	sz.000063	0.0480	0.0367	21.2982	0.9637
	sz.300015	0.0396	0.0352	13.3126	0.9715
	sz.002594	0.0822	0.0633	25.4203	0.8780
	sz.000881	0.0900	0.0670	14.7805	0.7794

	sz.000651	0.0386	0.0306	16.1498	0.9773
	sz.000858	0.0666	0.0525	19.8041	0.9250
Causal Factors + LSTM	sz.000063	0.0381	0.0279	18.3914	0.9771
	sz.300015	0.0410	0.0311	11.2038	0.9695
	sz.002594	0.0754	0.0615	23.0459	0.8972
	sz.000881	0.0749	0.0558	13.1500	0.8474
	sz.000651	0.0347	0.0255	11.7519	0.9816
	sz.000858	0.0543	0.0418	15.7239	0.9501
Causal Factors + GRU	sz.000063	**0.0369**	**0.0266**	**17.8913**	**0.9785**
	sz.300015	**0.0379**	**0.0310**	**11.1971**	**0.9794**
	sz.002594	**0.0680**	**0.0526**	**21.9571**	**0.9163**
	sz.000881	**0.0495**	**0.0393**	**11.0762**	**0.9442**
	sz.000651	**0.0321**	**0.0235**	**11.2748**	**0.9843**
	sz.000858	**0.0538**	**0.0419**	**16.4409**	**0.9509**

Figure A1, Figure A3, Figure A5, Figure A7, Figure A9, and Figure A11 illustrate the experimental predictions in different data sets, respectively. Figure A2, Figure A4, Figure A6, Figure A8, Figure A10, and Figure A12 show scatter plots for the experimental predictions in different data sets, respectively.
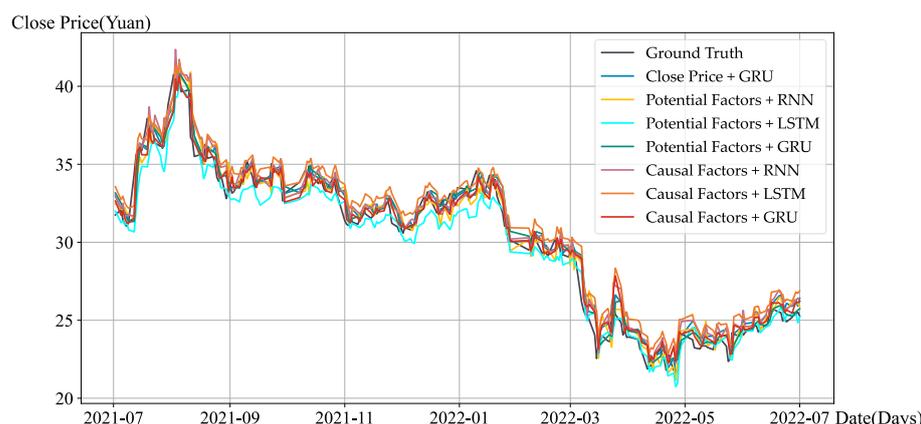
Figure A1. Visualization of prediction results for different models in the ZTE (sz.000063) dataset.
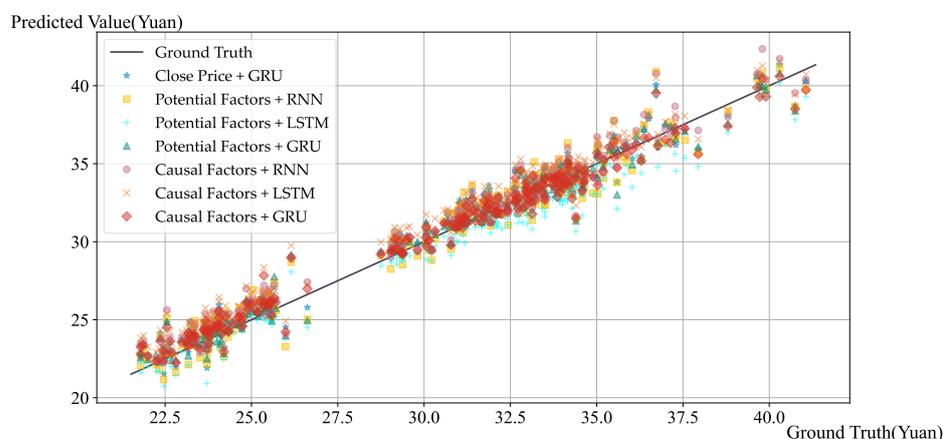
Figure A2. Scatter plot of prediction results for different models in the ZTE (sz.000063) (sz.300676) dataset.
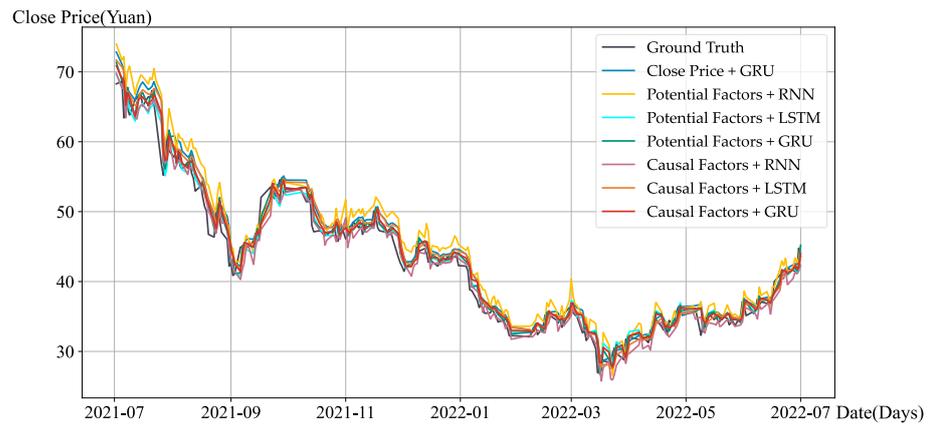
Figure A3. Visualization of prediction results for different models in the Aier Ophthalmology (sz.300015) dataset.
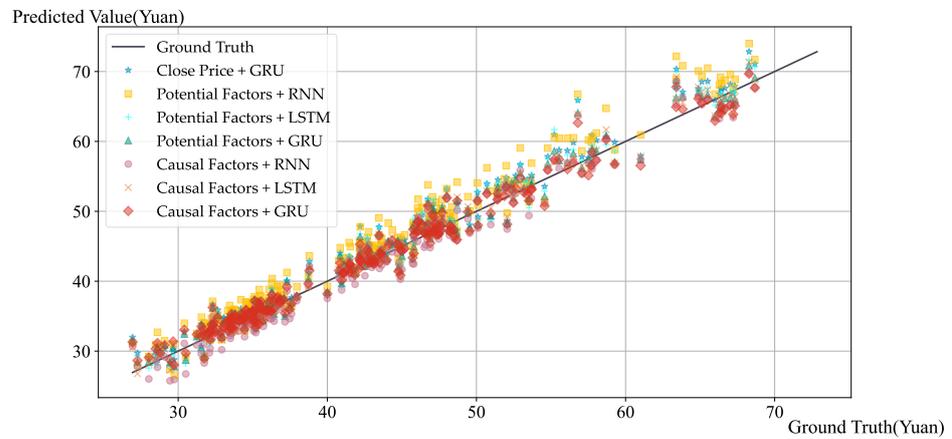
Figure A4. Scatter plot of prediction results for different models in the Aier Ophthalmology (sz.300015) dataset.
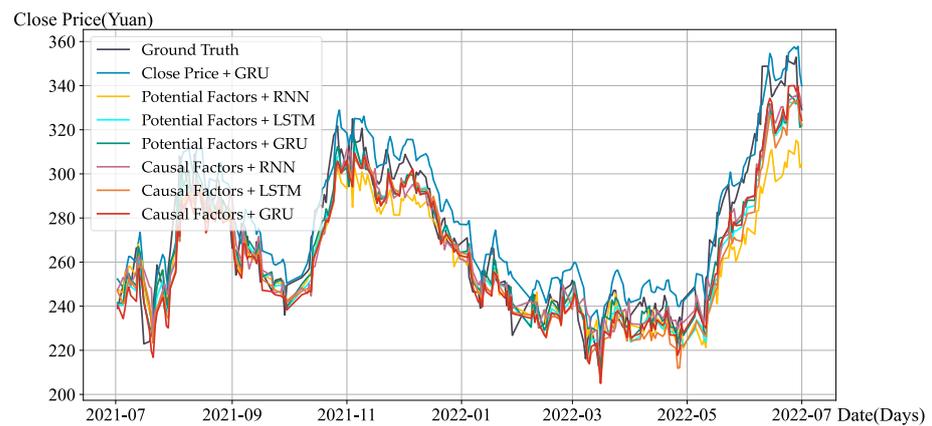
Figure A5. Visualization of prediction results for different models in the BYD (sz.002594) dataset.

Predicted Value(Yuan)
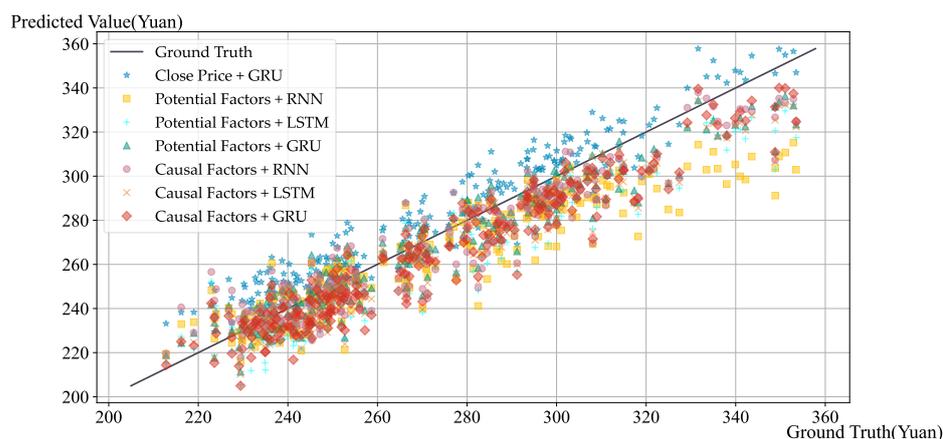
Figure A6. Scatter plot of prediction results for different models in the BYD (sz.002594) dataset.

Close Price(Yuan)
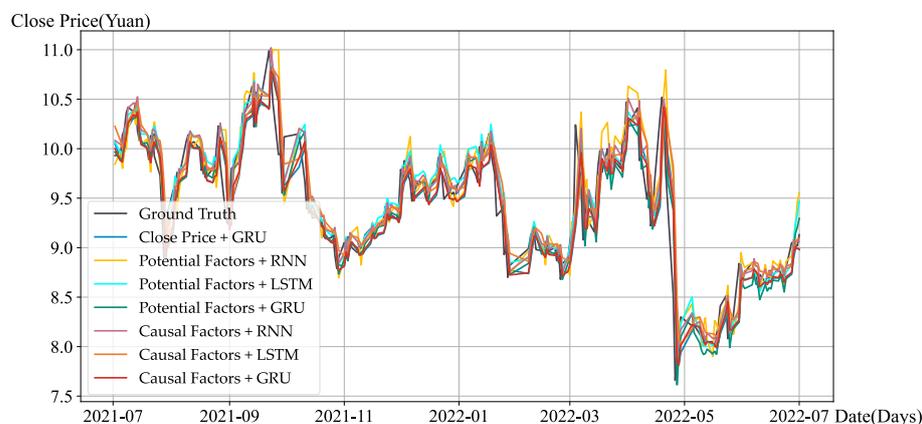
Figure A7. Visualization of prediction results for different models in the CGN Power Corporation (sz.000881) dataset.

Predicted Value(Yuan)
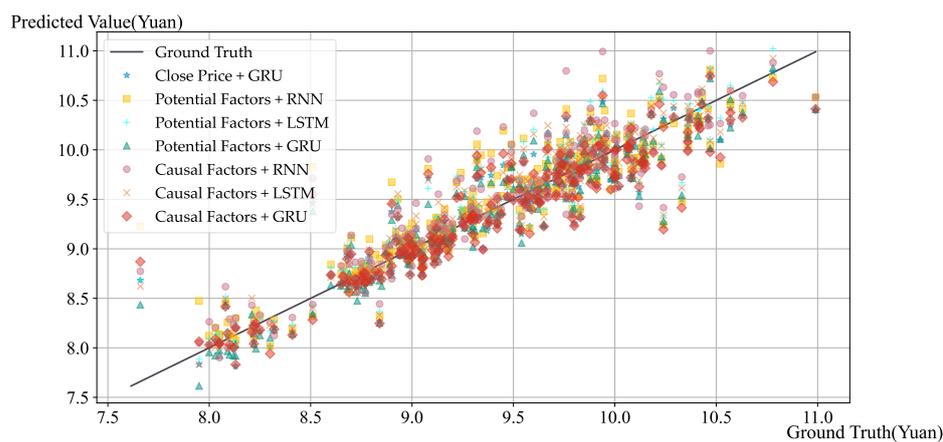
Figure A8. Scatter plot of prediction results for different models in the CGN Power Corporation (sz.000881) dataset.
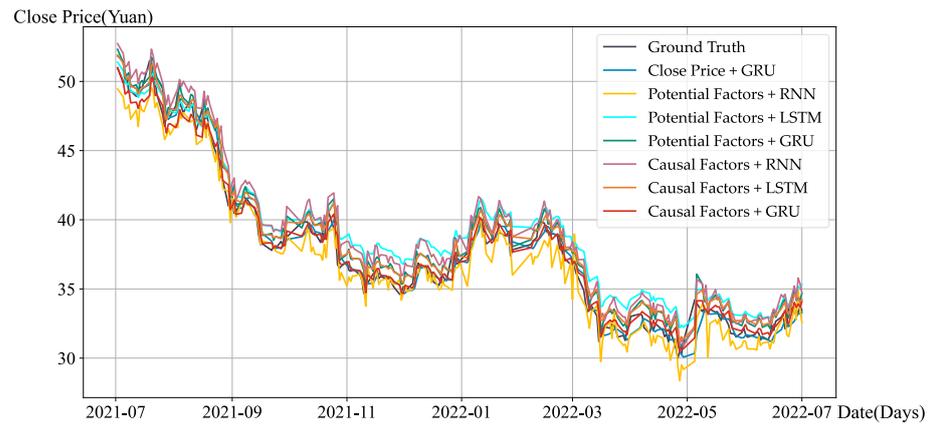
Figure A9. Visualization of prediction results for different models In the Gree Electric Appliances (sz.000651) dataset.
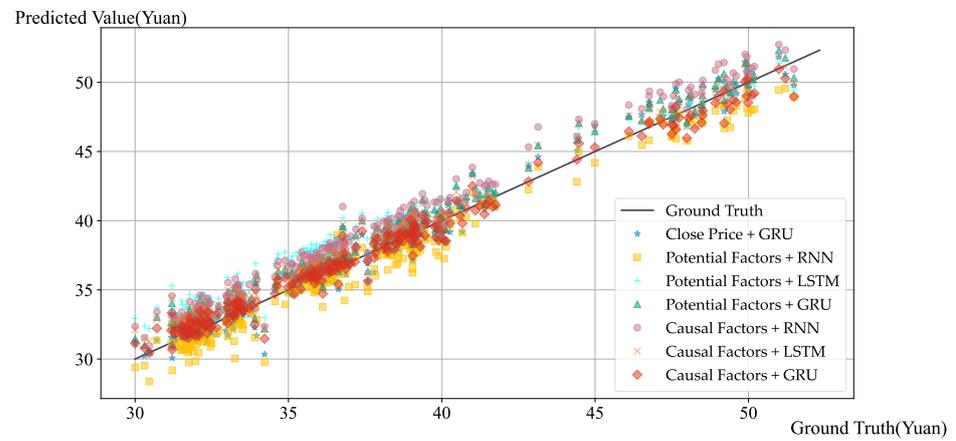
Figure A10. Scatter plot of prediction results for different models in the Gree Electric Appliances (sz.000651) dataset.
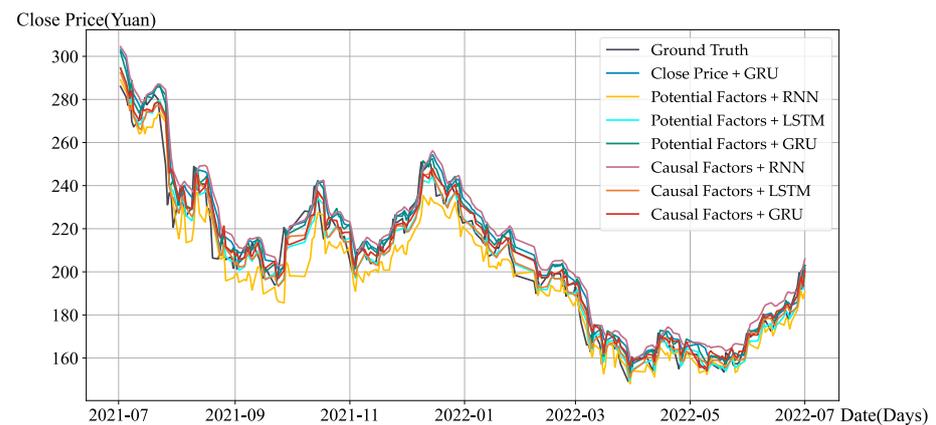
Figure A11. Visualization of prediction results for different models in the Wuliangye (sz.000858) dataset.
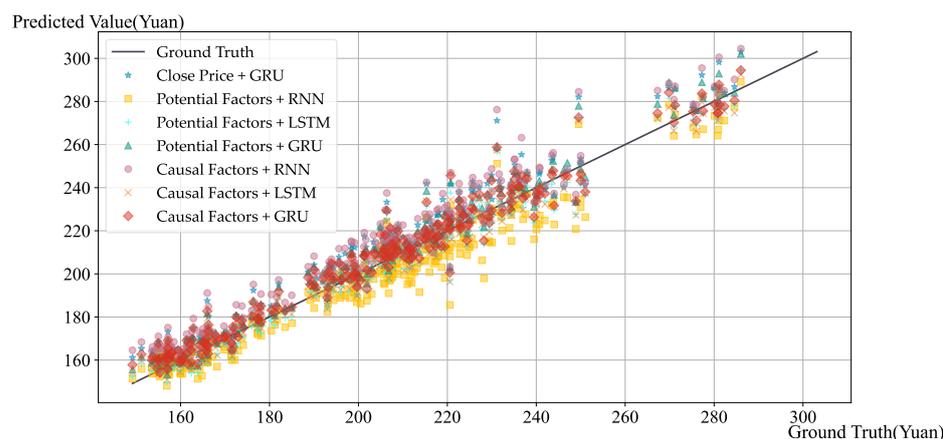
Figure A12. Scatter plot of prediction results for different models in the Wuliangye (sz.000858) dataset.

References

1. Wang, Y.J.; Feng, Q.Y.; Chai, L.H. Structural evolutions of stock markets controlled by generalized entropy principles of complex systems. *Int. J. Mod. Phys. B* **2010**, *24*, 5949–5971.
2. Tiberiu, A.C.; Kumar, T.A.; Phouphet, K. Nonlinearities and Chaos: A New Analysis of CEE Stock Markets. *Mathematics* **2021**, *9*, 707.
3. Olgun, H.; Ozdemir, Z.A. Linkages between the Center and Periphery Stock Prices: Evidence from the Vector ARFIMA Model. *Econ. Model.* **2007**, *25*, 512–519.
4. Moews, B.; Herrmann, J.M.; Ibikunle, G. Lagged Correlation-Based Deep Learning for Directional Trend Change Prediction in Financial Time Series. *Expert Syst. Appl.* **2018**, *120*, 197–206.
5. Zhao, R. Inferring Private Information from Online News and Searches: Correlation and Prediction in Chinese Stock Market. *Phys. A Stat. Mech. Its Appl.* **2019**, *528*, 121450.
6. Liang, M.; Wang, X.; Wu, S. Improving Stock Trend Prediction through Financial Time Series Classification and Temporal Correlation Analysis Based on Aligning Change Point. *Soft Comput.* **2022**, *27*, 3655–3672.
7. Ankit, T.; Dhaval, P.; Preet, S. Pearson Correlation Coefficient-Based Performance Enhancement of Vanilla Neural Network for Stock Trend Prediction. *Neural Comput. Appl.* **2021**, *33*, 16985–17000.
8. Liu, J.; Li, H.; Hai, M.; Zhang, Y. A Study of Factors Influencing Financial Stock Prices Based on Causal Inference. *Procedia Comput. Sci.* **2023**, *221*, 861–869.
9. Shivam, S.; Gyaneshwar, S.K. Effects of Temperature Rise on Clean Energy-Based Capital Market Investments: Neural Network-Based Granger Causality Analysis. *Sustainability* **2022**, *14*, 11163.
10. Wang, J.; Yuan, Y.; Tian, G.; Zheng, Y.; Wang, Z. Micro Input Factors Affecting Industrial Energy Efficiency in Hebei Province Based on Granger Causal Relation Test Model. *J. Comput. Methods Sci. Eng.* **2022**, *22*, 1069–1080.
11. Bachar, M.; Varsakelis, N.C. Causality between International Trade and International Patenting: A Combination of Network Analysis and Granger Causality. *Atl. Econ. J.* **2022**, *50*, 9–26.
12. Hou, X.; Li, S.; Li, W.; Wang, Q. Bank Diversification and Liquidity Creation: Panel Granger-Causality Evidence from China. *Econ. Model.* **2018**, *71*, 87–98.
13. Meng, X.; Han, J. Roads, Economy, Population Density, and CO_2: A City-Scaled Causality Analysis. *Resour. Conserv. Recycl.* **2018**, *128*, 508–515.
14. Zhao, Y.; Billings, S.A.; Wei, H.; He, F.; Sarrigiannis, P.G. A New NARX-Based Granger Linear and Nonlinear Casual Influence Detection Method with Applications to EEG Data. *J. Neurosci. Methods* **2013**, *212*, 79–86.
15. Chang, T.; Gatwabuyege, F.; Gupta, R.; Inglesi-Lotz, R.; Manjezi, N.C.; Simo-Kengne, B.D. Causal Relationship between Nuclear Energy Consumption and Economic Growth in G6 Countries: Evidence from Panel Granger Causality Tests. *Prog. Nucl. Energy* **2014**, *77*, 187–193.
16. Xu, C.; Zhao, X.; Wang, Y. Causal Decomposition on Multiple Time Scales: Evidence from Stock Price-Volume Time Series. *Chaos Solitons Fractals Interdiscip. J. Nonlinear Sci. Nonequilibrium Complex Phenom.* **2022**, *159*, 112137.
17. Baños, R.; Manzano-Agugliaro, F.; Montoya, F.G.; Gil, C.; Alcayde, A.; Gómez, J. Optimization Methods Applied to Renewable and Sustainable Energy: A Review. *Renew. Sustain. Energy Rev.* **2010**, *15*, 1753–1766.
18. Raul, V.; Michael, W.; Michael, L.; Gordon, P. Transfer Entropy—A Model-Free Measure of Effective Connectivity for the Neurosciences. *J. Comput. Neurosci.* **2011**, *30*, 45–67.
19. Brockmann, D.; Helbing, D. The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science* **2013**, *342*, 1337–1342.

20. Deyle, E.R.; Fogarty, M.; Hsieh, C.-H.; Kaufman, L.; MacCall, A.D.; Munch, S.B.; Perretti, C.T.; Ye, H.; Sugihara, G. Predicting Climate Effects on Pacific Sardine. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6430–6435.
21. Tsonis, A.A.; Deyle, E.R.; May, R.M.; Sugihara, G.; Swanson, K.; Verbeten, J.D.; Wang, G. Dynamical Evidence for Causality between Galactic Cosmic Rays and Interannual Variation in Global Temperature. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3253–3256.
22. Hirata, Y.; Amigó, J.M.; Matsuzaka, Y.; Yokota, R.; Mushiake, H.; Aihara, K. Detecting Causality by Combined Use of Multiple Methods: Climate and Brain Examples. *PLoS ONE* **2017**, *11*, e0158572.
23. Joskow, P.L.; Rose, N.L. Chapter 25 The Effects of Economic Regulation. In *Handbook of Industrial Organization*; Elsevier: Amsterdam, The Netherlands, 1989; Volume 2, pp. 1449–1506.
24. Van Nes, E.H.; Scheffer, M.; Brovkin, V.; Lenton, T.M.; Ye, H.; Deyle, E.; Sugihara, G. Causal Feedbacks in Climate Change. *Nat. Clim. Change* **2015**, *5*, 445–448.
25. Sharma, R.K.; Sharma, A. Forecasting Monthly Gold Prices Using ARIMA Model: Evidence from Indian Gold Market. *Int. J. Innov. Technol. Explor. Eng.* **2019**, *8*, 1373–1376.
26. Dickey, D.A.; Fuller, W.A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *J. Am. Stat. Assoc.* **1979**, *74*, 427–431. https://doi.org/10.2307/2286348.
27. Elham, R.S.; Reza, P.H.; Ali, A.; Farshad, S.S.; Clague, J.J. Comparison of Statistical and Machine Learning Approaches in Land Subsidence Modelling. *Geocarto. Int.* **2022**, *37*, 6165–6185.
28. Sun, Z.; Dong, W.; Shi, H.; Ma, H.; Cheng, L.; Huang, Z. Comparing Machine Learning Models and Statistical Models for Predicting Heart Failure Events: A Systematic Review and Meta-Analysis. *Front. Cardiovasc. Med.* **2022**, *9*, 812276.
29. Leng, S.; Xu, Z.; Ma, H. Reconstructing Directional Causal Networks with Random Forest: Causality Meeting Machine Learning. *Chaos* **2019**, *29*, 093130.
30. Ni, W.J.; Shen, Q.L.; Zeng, Q.T.; Wang, H.Q.; Cui, X.Q.; Liu, T. Data-Driven Seeing Prediction for Optics Telescope: From Statistical Modeling, Machine Learning to Deep Learning Techniques. *Res. Astron. Astrophys.* **2022**, *22*, 125003.
31. Ahmed, A.; Ahmed, F.; Ahmad, S.; Mahmoud, B.; Esraa, E. Marine Data Prediction: An Evaluation of Machine Learning, Deep Learning, and Statistical Predictive Models. *Comput. Intell. Neurosci.* **2021**, *2021*, 8551167.
32. Sedai, A.; Dhakal, R.; Gautam, S.; Dhamala, A.; Bilbao, A.; Wang, Q.; Wigington, A.; Pol, S. Performance Analysis of Statistical, Machine Learning and Deep Learning Models in Long-Term Forecasting of Solar Power Production. *Forecasting* **2023**, *5*, 256–284.
33. Chong, E.; Han, C.; Park, F.C. Deep Learning Networks for Stock Market Analysis and Prediction: Methodology, Data Representations, and Case Studies. *Expert Syst. Appl.* **2017**, *83*, 187–205.
34. Qin, Y.; Song, D.; Chen, H.; Cheng, W.; Jiang, G.; Cottrell, G.W. A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction. *arXiv* **2017**, arXiv:1704.02971.
35. Zahra, N.; Narges, H. Combining LSTM and CNN Methods and Fundamental Analysis for Stock Price Trend Prediction. *Multimed. Tools Appl.* **2022**, *82*, 17769–17799.
36. Rahman, M.O.; Hossain, S.; Junaid, T.-S.; Forhad, S.A.; Hossen, M.K. Predicting Prices of Stock Market Using Gated Recurrent Units (GRUs) Neural Networks. *Int. J. Comput. Sci. Netw. Secur.* **2019**, *19*, 213–222.
37. Wei, P. DLI: A Deep Learning-Based Granger Causality Inference. *Complexity* **2020**, *2020*, 5960171.
38. Tank, A.; Covert, I.; Foti, N.; Shojaie, A.; Fox, E.B. Neural Granger Causality. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 4267–4279. https://doi.org/10.1109/TPAMI.2021.3065601.
39. Rubin, D.B. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *J. Educ. Psychol.* **1974**, *66*, 688–701. https://doi.org/10.1037/h0037350.
40. Rosenbaum, P.R.; Rubin, D.B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* **1983**, *70*, 41–55. https://doi.org/10.1093/biomet/70.1.41.
41. Pearl, J. Causal Diagrams for Empirical Research. *Biometrika* **1995**, *82*, 669–688. https://doi.org/10.2307/2337329.
42. Granger, C.W.J. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* **1969**, *37*, 424–438. https://doi.org/10.2307/1912791.
43. Imbens, G.W.; Rubin, D.B. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*; Cambridge University Press: Cambridge, UK, 2015; ISBN 978-0-521-88588-1.
44. Wang, C.; Zhou, Y.; Zhao, Q.; Geng, Z. Discovering and Orienting the Edges Connected to a Target Variable in a DAG via a Sequential Local Learning Approach. *Comput. Stat. Data Anal.* **2014**, *77*, 252–266.
45. He, Y.-B.; Geng, Z. Active Learning of Causal Networks with Intervention Experiments and Optimal Designs. *J. Mach. Learn. Res.* **2008**, *9*, 2523–2547.