Carbon Credits Price Prediction Model (CCPPM)

Inam Alanazi(✉), Firas AL-Doghman, Abdulrahman Alsubhi, and Farookh Hussain

The University of Sydney Technology, 15 Broadway, Ultimo, NSW 2007, Australia
{Inam.alanazi,Abdulrahmannafeyea.alsubhi}@student.uts.edu.au,
{Firas.al-doghman,Farookh.Hussain}@uts.edu.au

Abstract. The valuation of carbon credits is a multifaceted task and is influenced by a wide range of factors encompassing economic activity, energy prices, weather conditions, policy adjustments, and market expectations. The accurate prediction of carbon credit prices is essential for traders, investors, regulators, and policymakers. To address the various facets of the carbon price prediction challenge, this paper contributes to this evolving field by proposing a solution that employs a machine learning methodology to enhance the accuracy and reliability of carbon credit price predictions. The proposed model is strong, precise and has the potential to guide decision making in the carbon market domain, with its proven accuracy and reliability highlighting its advantages as a valuable tool for stakeholders dealing with the intricacies of the carbon credit landscape.

1 Introduction

Carbon credits play a pivotal role in combating climate change by enabling industries to offset their carbon emissions. The accurate pricing of these credits is essential for fostering effective sustainable practices. This paper examines the realm of carbon credit pricing, with a specific focus on predicting prices through the utilisation of a machine learning model. The objective is to navigate the complexities inherent in this burgeoning market. There are several challenges associated with the accurate prediction of carbon credit prices. A fundamental obstacle is determining the optimal price for carbon credits. The work in [1] notes that research faces formidable hurdles due to the dynamic nature of global markets and complex environmental policies. The work in [2] further underscores the need for market-based instruments, such as carbon pricing mechanisms, to tackle the global scale of climate change and the diversity of emissions.

Additionally, [3] emphasises the pivotal role of precision in predicting carbon prices, asserting its significance in developing a reliable and beneficial carbon pricing mechanism while providing valuable insights for strategic business decisions. These studies collectively emphasise the urgent need for innovative and reliable methodologies in predicting carbon credit prices. To address these challenges, our proposed solution involves leveraging machine learning. This entails

L. Barolli (Ed.): AINA 2024, LNDECT 201, pp. 143–150, 2024.
https://doi.org/10.1007/978-3-031-57870-0_13

deploying advanced algorithms to analyse real-time data and environmental policies. The goal is to construct a predictive model that comprehensively understands patterns in the carbon credit market, ultimately enhancing pricing accuracy. This approach seeks to bring clarity to market participants. The precise forecasting of carbon prices is indispensable in establishing a robust and sustainable carbon pricing system, offering valuable support for informed business decision making. With the implementation of this proposed solution, we anticipate more stable and universally accepted carbon credit pricing mechanisms.

This paper is structured as follows: Sect. 2 presents the related work on carbon price prediction, Sect. 3 details the methodology employed in this research, Sect. 4 presents the results and provides a discussion and Sect. 5 concludes the paper.

2 Related Work

The valuation of carbon credits is subject to a myriad of influencing factors, encom- passing economic activity, energy prices, weather conditions, policy adjustments, and market anticipations.

Consequently, the prediction of carbon credit prices constitutes a formidable and pivotal endeavour for traders, investors, regulators, and policymakers alike. Diverse methodologies have been employed to forecast carbon credit prices, broadly classifiable into three categories: traditional econometric methods, artificial intelligence (AI) algorithms, and hybrid models.

2.1 Statistical Models

Traditional econometric approaches leverage statistical models to scrutinise histor- ical data and discern the variables impacting the supply and demand dynamics of carbon credits. Exemplifying this, [1] and [4] both applied multiple linear regression analysis, with the work in [1] additionally incorporating an auto-regressive integrated moving average model.

2.2 Machine Learning Models

AI algorithms, on the other hand, harness machine learning techniques to glean insights from data, uncovering nonlinear patterns and dependencies in the carbon market. For instance, [5] employed a PSO-RBF model to predict China's carbon trading market prices, showcasing its superior efficacy compared to other neural network models.

The work in [6] introduced an innovative combinatorial optimisation prediction method based on unstructured data, demonstrating its effectiveness in forecasting carbon trading prices in China. The work in [7] developed an ensemble prediction system integrating advanced data feature extraction technology and three sub-models, resulting in enhanced ac- curacy and stability in carbon price forecasting.

The work in [8] utilized six machine learning models, including extreme gradient boosting and support vector machines, to predict daily carbon prices and trading volumes in China, with the CEEMDAN-GWO-KNEA and CEEMDAN-RBFNN models exhibit- ing superior performance. The work in [9] employed a variety of algorithms, including linear regression, decision tree, random forest, extreme gradient boosting, and support vector machines, to predict the trading volume of carbon emissions, with random forest emerging as the most effective.

2.3 Hybrid Models

Hybrid models amalgamate the strengths of both econometric and AI methods, employing diverse techniques to address different facets of the carbon price prediction challenge. Noteworthy examples include the work in [10] which used a hybrid model comprising extreme point symmetric mode decomposition, an extreme learning machine, and a grey wolf optimiser algorithm, which surpassed the benchmark methods in predicting carbon prices in Hubei, Beijing, Shanghai, and Guangdong.

The work in [11] introduced a hybrid model integrating multi-resolution singular value decomposition and an extreme learning machine optimised by the adaptive whale optimisation algorithm, demonstrating superiority over benchmark methods in predicting carbon prices in China and the EU. The work in [12] proposed a hybrid model combining the ICEEMDAN decomposition-reconstruction method with the Sparrow search algorithm-optimised extreme learning machine model, resulting in heightened prediction accuracy, speed, and stability.

Additionally, the work in [13] advocated for a hybrid ARIMA and least squares support vector machine methodology, which proved effective in forecasting carbon prices. The work in [14] identified that a hybrid ARIMA and LSTM deep learning model exhibited the most favourable predictive performance for carbon trading prices in Shenzhen.

3 Method

This section explores the approach employed for carrying out our research. This section overviews the suggested model and the experimental process undertaken to derive results from this model. We use mean absolute error (MAE) which represents the average absolute difference between the predicted values and the actual values [15] to measure the efficiency of our proposed solution.

The level of accuracy impacts the utility or reliability of a model or predictions. A low MAE means that the predicted values are close to the actual values. This precision is crucial, especially in scenarios where accurate predictions are essential for decision making.

3.1 Prepossessing

In the realm of contemporary data-driven research, the significance of meticulous data preprocessing cannot be overstated. High-quality data forms the bedrock upon which accurate and reliable results are built, making the preprocessing phase an indispensable precursor to any analytical endeavour.

The integrity of research outcomes hinges on the quality of the input data, and as such, the process of refining and cleansing raw data emerges as a critical facet in ensuring the robustness and validity of subsequent analyses. Without diligent preprocessing, the potential for distorted or biased results increases, compromising the overall credibility and applicability of the research findings.

In light of these considerations, this section elucidates the various steps under- taken to refine, clean, and prepare the dataset for meaningful analysis, underscoring the pivotal role of data preprocessing in the pursuit of rigorous and dependable research outcomes.

3.1.1 The Selected Database

For the purpose of training and testing, we selected a dataset called "RCPI-data-public-aug2.xlsx," which is accessible via this link here Our choice of this dataset has been driven by its relevance to our intended objectives. We focus only on three columns in the Excel spreadsheet, namely the price, countries (with a specific focus on the USA), and the GHG level (carbon emissions) in line with global policy. These factors play a crucial role in influencing the decision-making process related to the carbon price in our proposed solution. Moreover, once the initial price is determined based on these variables, we further consider the dynamics of supply and demand to establish the ultimate price.

3.2 Algorithm

We used machine learning technology to build the solution and used the dataset to test its performance. Then, we represent our result as MAE.

The pseudo-code:

Function to compute real-time data based on multiple factors function computeRealTimeData(currentPrice, co2EmissionLevel, newRegulation, selectedCountry):

- Step 1: Calculate the impact of current price priceImpact = calculatePriceImpact (currentPrice)
- Step 2: Assess the influence of CO2 emission level emissionImpact = calculateEmissionImpact (co2EmissionLevel)
- Step 3: Evaluate the effect of new regulations regulationImpact = calculateRegulationImpact (newRegulation)
- Step 4: Consider the impact of the selected country countryImpact = calculateCountryImpact (selectedCountry)

- Step 5: Combine individual impacts to derive real-time data realTimeData = combineImpacts (priceImpact, emissionImpact, regulationImpact, countryImpact)
- return realTimeData

Our predictive model for global carbon credit prices is a robust tool that considers a range of influencing factors, ensuring a nuanced and accurate forecast. By incorporating the current price of carbon credits, the model establishes a baseline and identifies trends that can shape future valuations. The analysis of CO2 emission levels allows for an understanding of demand dynamics, reflecting the market's response to environmental commitments. New regulations are meticulously tracked, providing insights into the evolving landscape of carbon trading. Additionally, the model's consideration of selected countries acknowledges the diverse regulatory environments, enabling more precise regional predictions. The real-time data integration further enhances the model's adaptability, ensuring that predictions remain relevant in the face of dynamic market conditions. In essence, our model offers a comprehensive and dynamic approach to predicting carbon credit prices, empowering stakeholders with actionable insights for strategic decision making.

3.3 Implementation

In the development of our proposed solution, we used the Python programming language to craft the solution's code due to its versatility, extensive libraries, and ease of integration. Leveraging Python facilitates efficient coding practices and supports seamless integration with various data processing and analysis tools.

The core of our solution involves the utilisation of the RandomForestRegressor (RFR) machine learning algorithm. This algorithm, a part of the scikit-learn library, is well-suited for regression tasks and achieves robust performance in predicting outcomes based on complex datasets. Random forest is a robust and highly effective tool, demonstrating a level of performance that ranks among the most accurate methodologies to date, as highlighted by [16] The processed database, which forms the foundation of our model, undergoes thorough preprocessing to ensure data quality and relevance. This implementation choice aligns with our goal of creating a robust, adaptable, and accurate solution to address the challenges presented by our problem domain.

4 Result and Discussion

The exceptional MAE result of 0.011507294045069475 obtained from the carbon credits price prediction model underscores its excellence and reliability. This low MAE value demonstrates that, on average, the model's predictions are remarkably close to the actual market prices of carbon credits. The precision achieved by the model is particularly noteworthy in the context of carbon credit markets, where accurate forecasts are pivotal for informed decision making.

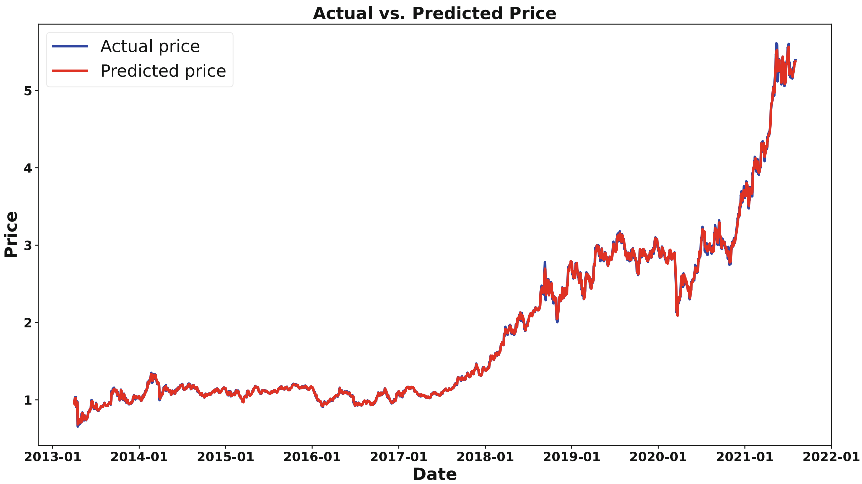

Fig. 1. The performance of the pricing prediction model

The model's ability to consistently provide predictions with such a small margin of error reflects its proficiency in capturing the complex and dynamic factors influencing carbon credit prices, as shown in Fig. 1. This outstanding performance enhances the credibility of the model and positions it as a valuable tool for various stakeholders, including businesses, investors, and policymakers. The reliability of the model, as evidenced by its low MAE, instils confidence in its predictions, making it an excellent resource for strategic planning, financial decision support, and sustainability initiatives. Overall, the noteworthy accuracy demonstrated by the model, as reflected in its low MAE, substantiates its status as an important and dependable tool in the domain of carbon credit price prediction.

5 Conclusion

In conclusion, this research presents a comprehensive exploration of the methodologies employed in predicting carbon credit prices. The proposed solution, underpinned by machine learning, demonstrates exceptional performance, as evidenced by a remarkably low MAE of 0.0115. This outstanding accuracy positions the model as a reliable tool for various stakeholders, offering valuable insights for strategic planning, financial decision support, and sustainability initiatives. The preprocessing phase, highlighted in the methodology section, underscores the significance of high-quality data in ensuring the robustness and validity of predictions. The selected dataset, focusing on price, countries (with a specific emphasis on the USA), and carbon emissions, aligns with global policy considerations, enhancing the relevance of the proposed solution. The implementation of the solution involves a step-by-step algorithm, including real-time data computation and the consideration of factors such as price, CO2 emission levels, new regulations, and selected countries. Overall, the research contributes to the evolving landscape of carbon credit price prediction, providing a robust and accurate model that can inform decision-making in carbon markets. The demonstrated precision and reliability underscore the potential of the proposed solution as a valuable asset for stakeholders navigating the complexities of the carbon credit landscape.

References

1. Kim, H., Kim, Y., Ko, Y., Han, S.: Performance comparison of predictive methodologies for carbon emission credit price in the Korea emission trading system. Sustainability **14**(13), 8177 (2022)
2. Aldy, J.E., Stavins, R.N.: The promise and problems of pricing carbon: theory and experience. J. Environ. Dev. **21**(2), 152–180 (2012)
3. Niu, X., Wang, J., Zhang, L.: Carbon price forecasting system based on error correction and divide-conquer strategies. Appl. Soft Comput. **118**, 107935 (2022)
4. Guðbrandsdóttir, H.N., Haraldsson, H.Ó.: Predicting the price of EU ETS carbon credits. Syst. Eng. Proc. **1**, 481–489 (2011)
5. Huang, Y., Hu, J., Liu, H., Liu, S.: Research on price forecasting method of China's carbon trading market based on PSO-RBF algorithm. Syst. Sci. Control Eng. **7**(2), 40–47 (2019)
6. Huang, Y., He, Z.: Carbon price forecasting with optimization prediction method based on unstructured combination. Sci. Total Environ. **725**, 138350 (2020)
7. Yang, Y., Guo, H., Jin, Y., Song, A.: An ensemble prediction system based on artificial neural networks and deep learning methods for deterministic and probabilistic carbon price forecasting. Front. Environ. Sci. **9**, 740093 (2021)
8. Lu, H., Ma, X., Huang, K., Azimi, M.: Carbon trading volume and price forecasting in China using multiple machine learning models. J. Cleaner Prod. **249**, 119386 (2020)
9. Wong, F.: Carbon emissions allowances trade amount dynamic prediction based on machine learning. In: 2022 International Conference on Machine Learning and Knowledge Engineering (MLKE), pp. 115–120. IEEE (2022)
10. Zhou, J., Huo, X., Xu, X., Li, Y.: Forecasting the carbon price using extreme-point symmetric mode decomposition and extreme learning machine optimized by the grey wolf optimizer algorithm. Energies **12**(5), 950 (2019)
11. Sun, W., Zhang, C.: Analysis and forecasting of the carbon price using multi-resolution singular value decomposition and extreme learning machine optimized by adaptive whale optimization algorithm. Appl. Energy **231**, 1354–1371 (2018)
12. Zhou, J., Chen, D.: Carbon price forecasting based on improved CEEMDAN and extreme learning machine optimized by sparrow search algorithm. Sustainability **13**(9), 4896 (2021)
13. Zhu, B., Chevallier, J.: Carbon price forecasting with a hybrid ARIMA and least squares support vector machines methodology. In: Zhu, B., Chevallier, J. (eds.) Pricing and Forecasting Carbon Markets: Models and Empirical Analyses, pp. 87–107. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57618-3
14. Hu, Y., Xiao, W., He, B., Tang, X.: Carbon trading price forecasting with a hybrid arima and lstm deep learning methodology. In: Proceedings of the 2022 6th International Conference on E-Business and Internet, pp. 289–293 (2022)

15. Hyndman, R., Koehler, A.: Another look at measures of forecast accuracy. Int. J. Forecast. **22**, 679–688 (2006)
16. Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P.: Random forest: a classification and regression tool for compound classification and QSAR modeling. J. Chem. Inf. Comput. Sci. **43**(6), 1947–1958 (2003)