SEMARAK ILMU
PUBLISHING
202103268166(003316878-P)

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

A Comparative Study on the Performance of Covariance Functions in Gaussian Process Regression model: Application to Global Wheat Price

Nahamizun Maamor[1,*], Hanita Daud[1], Muhammad Naeim Mohd Aris[2], Nor Izzati Jaini[3], Mahmod Othman[4], Evizal Abdul Kadir[5]

[1] Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia
[2] Centre for American Education, Sunway University, Bandar Sunway 47500, Selangor, Malaysia
[3] Faculty of Industrial Sciences and Technology, Universiti Malaysia Pahang, 26300 Gambang, Pahang
[4] Department of Information System, Universitas Islam Indragiri, Tembilahan 29212, Indonesia
[5] Department of Informatics Engineering, Universitas Islam Riau, Pekanbaru 28284, Indonesia

ARTICLE INFO

Article history:
Received 25 September 2023
Received in revised form 8 November 2023
Accepted 21 February 2024
Available online 26 March 2024

Keywords:
Gaussian Process Regression; Covariance Functions; Global Wheat Price

ABSTRACT

Gaussian Process Regression (GPR) is a nonparametric machine learning model that provides uncertainty quantification in making predictions. GPR utilizes several covariance functions (CFs) in the process of developing models to ensure high accuracy. There are five common CFs in GPR, which are the Radial Basis Function (RBF), Rational Quadratic (RQ), Periodic (Per), Matérn 3/2 (Mat 3/2), and Matérn 5/2 (Mat 5/2), where each covariance function (CF) has different characteristics and behaviors. This paper is to investigate the comparative performances of each CF when applied to the Global Wheat Prices dataset. Error metric measurement like Mean Square Error (MSE) and Root Mean Square Error (RMSE), while Coefficient of Variation (CV) for uncertainty quantification measurement is computed for each CF, and comparisons was made among the CFs to conclude the best CF for this application. The lowest values among them will be the best CF for the data. It should be noted that the CV for each model should be less than 5%, and the CF with the smallest value of CV is considered reliable. The five CFs were fit to the Global Wheat Prices dataset, and it was found that the Mat 3/2 produced the best performances with the lowest values of MSE, RMSE, and CV. Mat 3/2 is the most efficient CF for making predictions since it gives the lowest value of error metric measurement and the lowest value of CV under 5% among the other CFs, making it more reliable for modeling. Overall, the outcome shows that Mat 3/2 is the best CF to be used in developing a GPR model to predict Global Wheat Prices dataset.

1. Introduction

Wheat is one of the earliest cereal grasses that can be consumed by humans and is a part of the Triticum (Poaceae) genus [1,2]. There are six classes of wheat, including hard red winter, hard red spring, soft red winter, hard white, soft white, and durum, each with its own traits and qualities [3]. The unique properties and characteristics of each variety of wheat result in a variety of products, including bread, pasta, and noodles, among others [3]. Each of these categories has a unique

* *Corresponding author.*
E-mail address: mizunmaamor9@gmail.com

https://doi.org/10.37934/araset.42.1.215225

nutritional profile, including carbohydrate, hydration, protein, calories, sugar, and fiber content [4]. All these nutrients are essential for maintaining human health. This has made wheat a primary food in most of the world's nations, including China, Russia, the United States, India, and others [5].

Nonetheless, fluctuations in the global wheat price have a significant effect on the inhabitants of these nations [6]. According to a study conducted using basic statistics, which is a correlation analysis based on price data statistics from international and Hungarian banks, oil prices, population growth, and land destruction are factors that influence the fluctuations in wheat prices and output [7]. Another study that compared the change in global market prices for the five major wheat exporting countries in the world, namely the United States, the European Union, Canada, Australia, and the Former Soviet Union (FSU), from 1980 to 2013, discovered that climate, oil prices, and previous market prices all played a role in wheat price changes but differed by country [8]. Econometric analysis utilizing the Error Correction Model discovered that some events, such as COVID-19 and the Russo-Ukrainian conflict, caused a double increase in the world wheat price and increased volatility for both periods [9]. However, according to the Global Market Analysis published by the Foreign Agricultural Service of the United States Department of Agriculture (USDA), countries such as Nigeria and Indonesia will increase their wheat consumption in 2023–2024 because of the decline in wheat prices during the first few months of 2023 [10].

This demonstrates that the market price of wheat is not fixed and frequently fluctuates based on the circumstances of each period. Indeed, the volatile price fluctuations of wheat cause trends and data patterns to become nonlinear, fluctuating, and dynamic. Non-linear data patterns with unpredictable fluctuations make it difficult to perform predictive analysis on the data. There are several model algorithms that are usually used to perform predictions on time series data that are nonlinear, fluctuating, and dynamic, comprised of statistics models, machine learning models, and deep learning models. The statistical model commonly used in the data series is Autoregressive Integrated Moving Average (ARIMA), while the machine learning model consists of Random Forest, Adaboost, and Gradient Boost [11]. Deep learning models are a subset of machine learning that can be used to predict time series data consisting of Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Multilayer Perceptron (MLP) networks, and Long Short-Term Memory (LSTM) networks [12-16].

However, this study employs Gaussian Process Regression (GPR), a versatile nonparametric machine learning method capable of making predictions on both linear and nonlinear data due to its ability to operate with a variety of covariance functions. This study evaluates the performance of five general covariance functions (CFs) on Global Wheat Price dataset, including the Radial Basis Function (RBF), Rational Quadratic (RQ), Periodic (Per), Matérn 3/2 (Mat 3/2), and Matérn 5/2 (Mat 5/2), with a concentration on prediction by interpolation.

2. Literature on Gaussian Process Regression

The two main subcategories of Gaussian Process (GP) models are regression and classification [17]. Both models are built using the Bayesian probabilistic method. However, this study primarily focuses on GPR, a sort of regression that can handle continuous and numerical data for time series data. Numerous industries use GPR extensively. Particularly in agriculture, GPR is used to anticipate winter wheat yields on a field scale using multi-spatial-type data captured by Unmanned Aerial Vehicles (UAV) [18]. The interpolated GPR approach has also been utilized in controlling air quality (AQ) for fine details in order to reduce air pollution in Beijing, China, and London, UK, using open data [19]. Additionally, in the sphere of sports, GPR has been applied to predict non-linear data with a periodic nature for the Falun Nordic World Ski Championships 2015 data [20]. The investigation of predicting the determination of proteins and small molecules with uncertainty quantification shows that GPR is useful in the pharmaceutical industry [21].

Next, GPR has had a significant impact on energy sustainability, with research being performed specifically for power system design optimization utilizing GPR based on Multi-Objective Bayesian Optimization (GPR-MOBO) [22]. Meanwhile, work on increasing the performance of GPR and Gradient Descent (GD) utilizing the inversion approach has been done to reduce computing time by implementing predictions on hydrocarbon depth in Seabed logging (SBL) [23]. Additionally, research comparing the RBF, RQ, Mat 3/2, and Mat 5/2 covariance functions from GPR has been conducted on NASA's lithium-ion battery data to forecast battery health issues [24]. Finally, in the chemical sector, Reduced GPR (RGPR) based on the Generalized Likelihood Ratio Test (GLRT) has been developed and employed as a flaw detector, and it has been used in simulation of the Tennessee Eastman process [25].

3. Literature Research on Global Wheat Prices

There is a study that has been carried out in early 2023 using global wheat price data from the same source as this study, which is FRED, but differs in terms of the study's goal. The study focuses on the relationship between the global market prices of four agricultural commodities, which are wheat, corn, barley, and sunflower oil, and the Russo-Ukrainian conflict, which involves food security issues around the world [26]. The methods they have used to study the relationship between all the variables are the VAR impulse response function, variance decomposition, Granger causality test, and vector error correction. An additional purpose has also been implemented in the same research paper by implementing 10-month-ahead forecasting using vector autoregressive (VAR) and vector error correction (VECM) for all four agricultural commodities [26]. Another study on the price of wheat has also been carried out, but the price study is specific to a district in Pakistan, namely Faisalabad, Gujaranwala, and Multan, by implementing forecasting using Bagging Tree, GPR, Auto-regressive Integrated Moving Average (ARIMA), and advanced neural network architecture Long Short-Term Memory (LSTM) networks [27]. A wheat price forecasting study specific to India has also been carried out to forecast the future price of wheat in India using an autoregressive moving average (ARMA) and an artificial neural network (ANN) [28]. Finally, a study using stock market price of wheat data specific to the country of Bangladesh, which is not the historical price of global wheat, has also been carried out to study the efficiency of prediction using Support Vector Regression (SVR), Random Forest (RF), Bagging (BG), AdaBoost (AB), Gradient Boosting (GB), LightGBM (LGB), and XGBoost (XGB) [29].

4. Methodology

This section contains five sub-sections consisting of Gaussian Process Regression, Covariance Functions, Hyperparameter Adaptation, Model Validation, and Data Acquisition and Computational Software. This section begins by systematically stating and explaining the equations used in GPR, followed by the CFs of GPR and the hyperparameter adaptation for CFs. The next sub-section discusses model validation, which is used to validate the model. Finally, the fifth sub-section describes the data acquisition from official sources and the software used to perform GPR on the dataset.

4.1 Gaussian Process Regression

Conventionally, a GPR specifies a distribution over functions by acquiring the mean and CFs of the realization of the GPR at $x \in R^d$, denoted by $f(x)$, where $m(x)$ and $k(x, x')$ in Equation (1) are the Mean Function (MF) and Covariance Function (CF), respectively [17].

$$f(x) \sim GP\big(m(x), k(x, x')\big) \tag{1}$$

This study implements the GPR with the presence of noise observation in Equation (2) where ϵ is the Gaussian noise parameterized by σ_n^2 which are the variance values where it is very important in influencing the computation on hyperparameter adaptation for the efficiency in prediction and optimizing the hyperparameter for each CF.

$$y = f(x) + \epsilon,\ \epsilon \sim N(0, \sigma_n^2) \tag{2}$$

In this case, the prior observation is like Equation (3), and the joint distribution of the observation at the test location under prior as Equation (4),

$$K_{\ y}(X, X) = K_f(X, X) + \sigma_n^2 I \tag{3}$$

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K_f(X, X) + \sigma_n^2 I & K_f(X, X_*) \\ K_f(X_*, X) & K_f(X_*, X_*) \end{bmatrix}\right) \tag{4}$$

where $K_y\ (X, X)$ is the covariance matrix for noisy target of y while $K_f\ (X, X)$ is covariance matrix for noise-free latent f where X and X_* matrix training and testing input, respectively. Training data's mean is used in this study as Equation (5) which $m(X_*)$ and is added in equation and the covariance function for the predictive distribution as Equation (6). Both equations are in the presence of noise observation.

$$\overline{f}(X_*) = m(X_*) + K_f(X_*, X)\left[K_f(X, X) + \sigma_n^2 I\right]^{-1}(f - m(X)) \tag{5}$$

$$Cov\big(f(X_*)\big) = K_f(X_*, X_*) - K_f(X_*, X)\left[K_f(X, X) + \sigma_n^2 I\right]^{-1} K_f(X, X_*) \tag{6}$$

4.2 Covariance Functions

GPR works with the CF to make predictions on the data. All the selected CFs in this study consist of Square Exponential, also known as RBF, RQ, Per, Mat 3/2, and Mat 5/2. All the CFs stated use the Euclidean distance metric to measure similarity between data points, as in Equation (7).

$$r = \|x - x'\| \tag{7}$$

The CF also has specific hyperparameters consisting of the signal standard deviation, σ_f and length scale, l for all CF stated. However, RQ and Per has extra hyperparameter on the equation which is positive value of scale-mixture parameter, α while Per, has another extra hyperparameter which is periodicity, p. The equations of RBF, RQ, Per, Mat 3/2 and Mat 5/2 are expressed in Equations (8), (9), (10), (11), and (12).

$$k_f(x,x') = \sigma_f^2 \, exp\left(-\frac{r^2}{2l^2}\right) \tag{8}$$

$$k_f(x,x') = \sigma_f^2 \left(-\frac{r^2}{2\alpha l^2}\right)^{-\alpha} \tag{9}$$

$$k_f(x,x') = \sigma_f^2 \, exp\left(-\frac{2}{l} \, sin^2 \left(\frac{\sqrt{3}\, r}{p}\right)\right)^{-\alpha} \tag{10}$$

$$k_f(x,x') = \sigma_f^2 \left(1 + \frac{\sqrt{3}\, r}{l}\right) exp\left(\frac{\sqrt{3}\, r}{l}\right) \tag{11}$$

$$k_f(x,x') = \sigma_f^2 \left(1 + \frac{\sqrt{5}\, r}{l} + \frac{5\, r^2}{3l^2}\right) exp\left(\frac{\sqrt{3}\, r}{l}\right) \tag{12}$$

4.3 Hyperparameter Adaptation

The hyperparameter of the CF set as $\theta_{f_A} = \{\sigma_f, l,\}$, $\theta_{f_B} = \{\sigma_f, l, \alpha\}$, and $\theta_{f_C} = \{\sigma_f, l, \alpha, p\}$ where $\theta_{f_A}$is the set of the hyperparameter for RBF, Mat 3/2 and Mat 5/2 while θ_{f_B}, and θ_{f_C} are set of hyperparameter for RQ and Per, respectively. All the set of hyperparameter stated needs to be optimized for each CF where it located at posterior as θ_f by using Negative Log Marginal Likelihood (NLML) as shown in Equation (13) where $K_y = K_y(X,X)$ from Equation (3).

$$log\, p(y|X,\theta_f) = -\frac{1}{2} y^T K_y^{-1} y - \frac{1}{2} log|K_y| - \frac{n}{2} log\, 2\pi \tag{13}$$

4.4 Model Validation

All the CFs of the GPR fit to the data have been validated by using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Coefficient of Variation (CV). MSE and RMSE are used to measure the error of the interpolation prediction, while CV is used to check and calibrate the reliability of the model in prediction where the model should achieve CV a less than 5%. The smallest values of MSE, RMSE, and CV are considered the best CFs that can implement interpolation and predict the data nicely. Based on the Equations (14), (15) and (16), y is actual values, y_* is predicted value, and $\mu_{y_i^*}$ is the mean predicted values.

$$MSE = \frac{1}{n} \sum (y - y^*)^2 \tag{14}$$

$$RMSE = \sqrt{\frac{1}{n} \sum (y - y^*)^2} \tag{15}$$

$$CV = \frac{RMSE}{|\mu_{y_i^*}|} \times 100\% \tag{16}$$

4.5 Data Acquisition and Computational Software

The dataset for this study has been retrieved from FRED, the Federal Reserve Bank of St. Louis website, namely the Global Price of Wheat, where the main source is the International Monetary Fund [30]. The dataset as shown in Figure 1 is a monthly time series dataset from January 1, 1990, to April 1, 2023, and is not seasonally adjusted data. Price currency is U.S. dollars per metric ton. The data set in Figure 1 also shows that there are changes in the trend pattern that are uncertain and have large fluctuations.

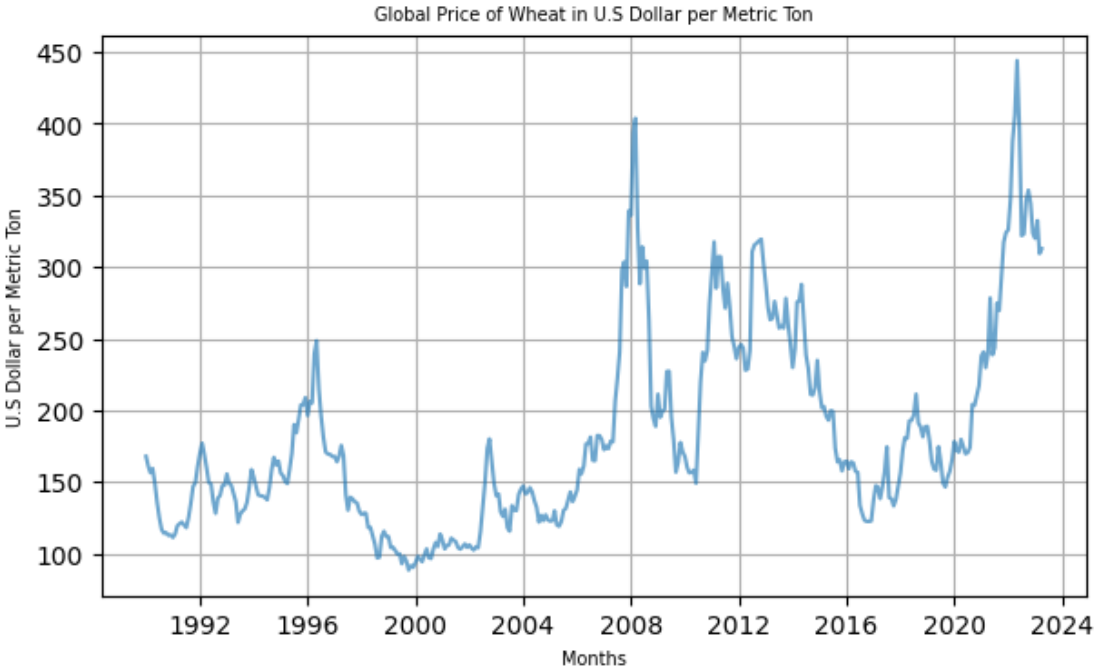


Fig. 1. The plot of global wheat prices in U.S. dollars per metric ton by month every year from 1990 to 2023

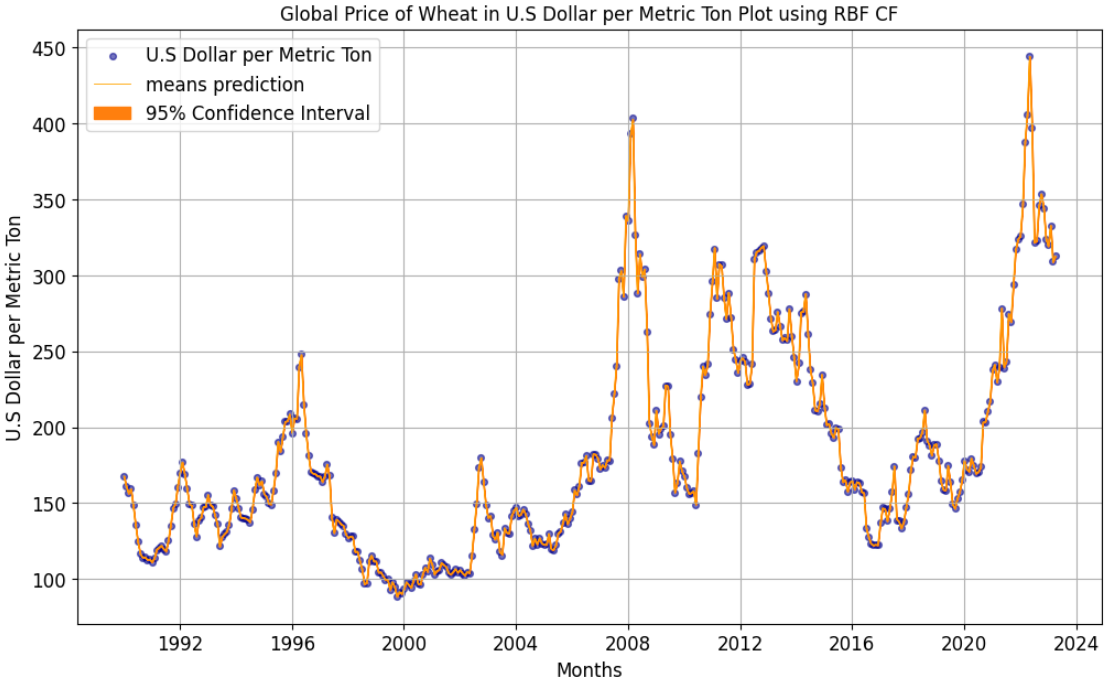
The data has been computed by using Gaussian Process Regression in the Scikit-Learn library on Python using Jupyter Notebook on a laptop with an Intel Core i3-6100 CPU at 2.30 GHz and 12 GB of Random-Access Memory (RAM).

5. Result and Discussion

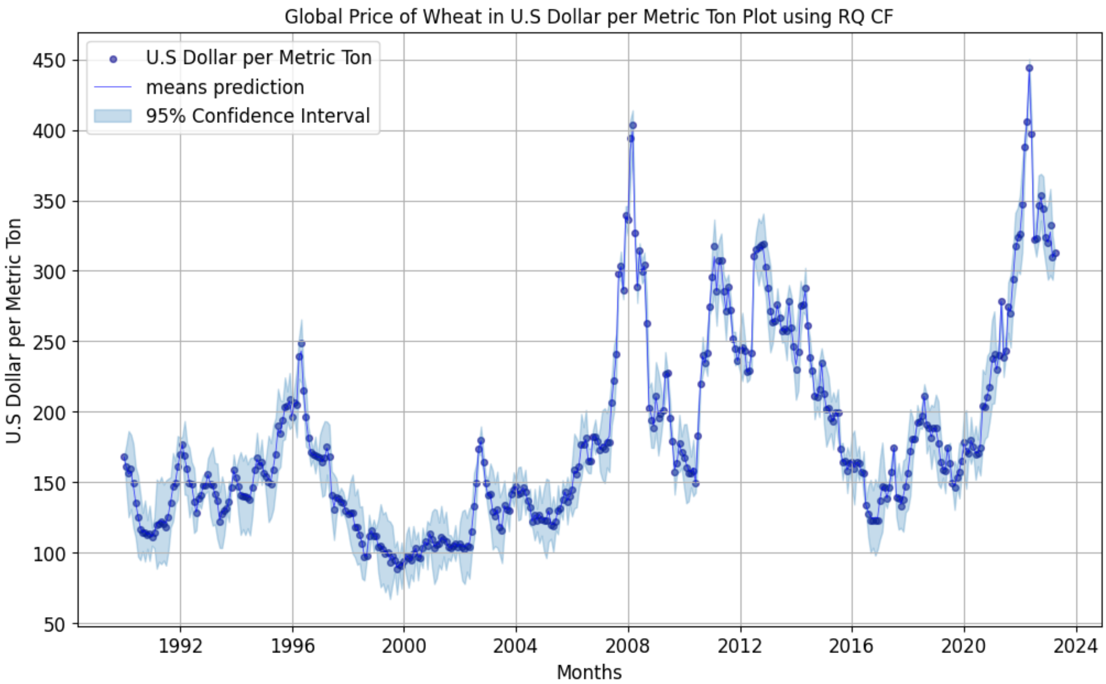
The goal of this study is to assess the efficiency of the CF in performing interpolation prediction in order to acquire the optimal CF for GPR modeling of Global Wheat Price data. The Global Wheat Price data has been fitted to GPR with different CFs, which have then been assessed using the MSE, RMSE, and CV error measurement metrics shown in Table 1. All CFs have also had their plot graphs created, as seen in Figure 2.

Table 1
Result from each fitted Covariance Functions to global wheat prices dataset

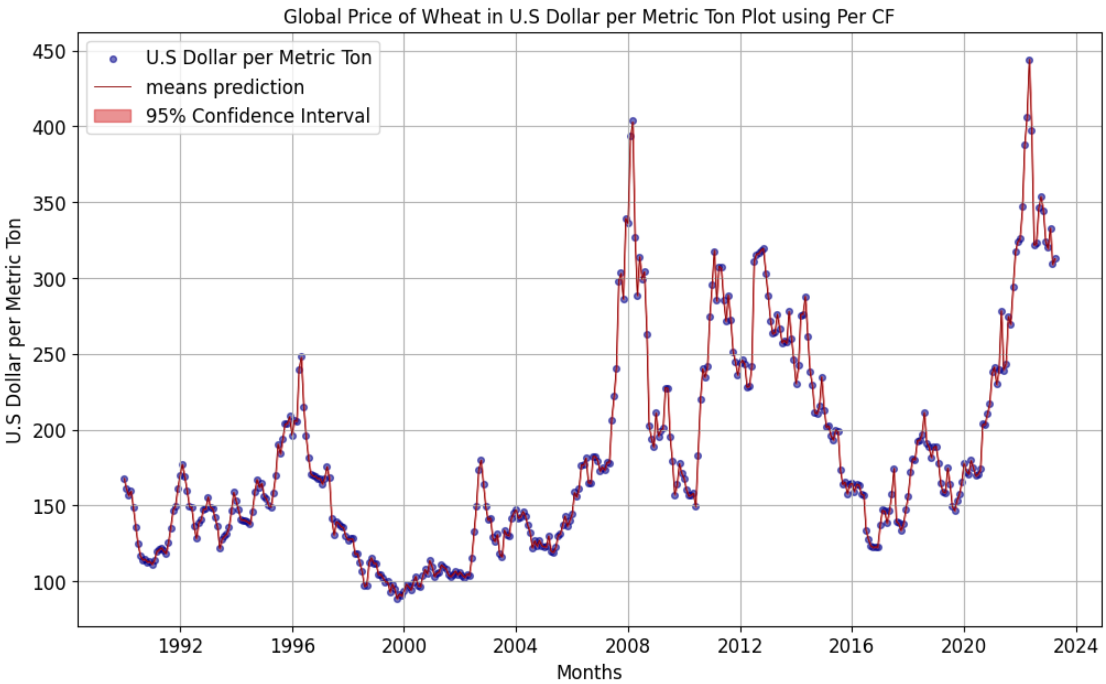
Covariance Function	MSE	RMSE	CV (%)
RBF	0.15369	0.39203	0.21551
RQ	1.54782	1.24411	0.68357
Per	0.15359	0.39191	0.21543
Mat 3/2	0.11714	0.34226	0.18814
Mat 5/2	0.13198	0.36330	0.19971


Global Price of Wheat in U.S Dollar per Metric Ton Plot using RBF CF
U.S Dollar per Metric Ton
means prediction
95% Confidence Interval
U.S Dollar per Metric Ton
Months

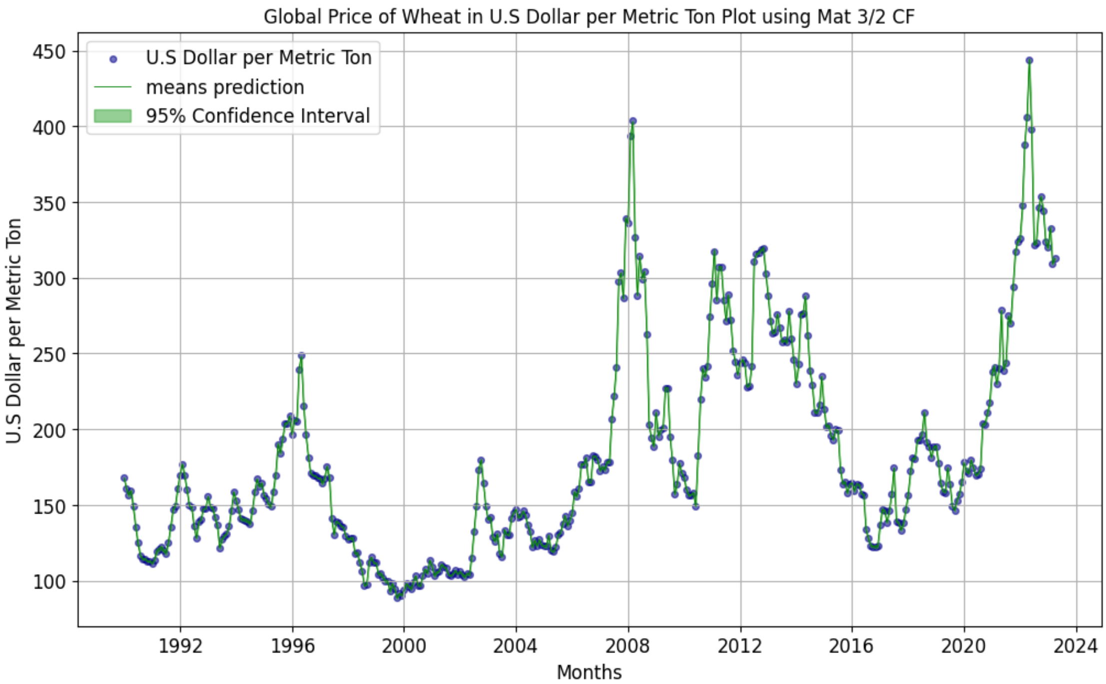
(a)


Global Price of Wheat in U.S Dollar per Metric Ton Plot using RQ CF
U.S Dollar per Metric Ton
means prediction
95% Confidence Interval
U.S Dollar per Metric Ton
Months

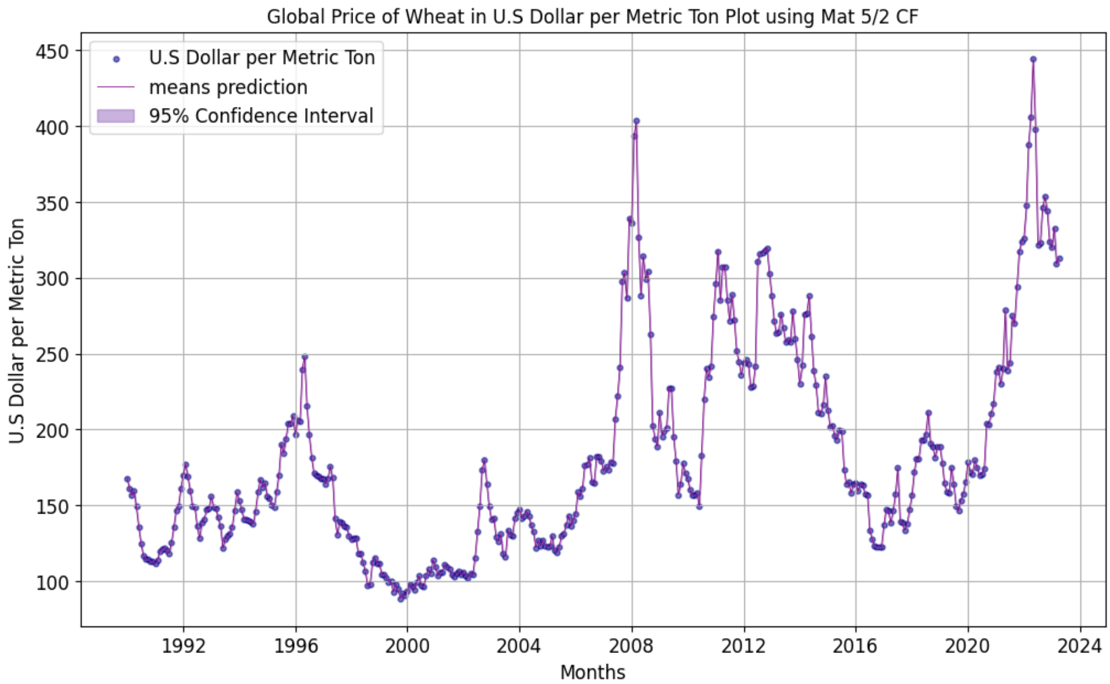
(b)


Global Price of Wheat in U.S Dollar per Metric Ton Plot using Per CF
U.S Dollar per Metric Ton
means prediction
95% Confidence Interval
U.S Dollar per Metric Ton
Months

(c)


Global Price of Wheat in U.S Dollar per Metric Ton Plot using Mat 3/2 CF
U.S Dollar per Metric Ton
means prediction
95% Confidence Interval
U.S Dollar per Metric Ton
Months

(d)



(e)

Fig. 2. Graph plots for each CF from GPR fitted to the global wheat prices dataset consist of (a) RBF, (b) RQ, (c) Per, (d) Mat 3/2, and (e) Mat 5/2

Based on all the graphs plotted in Figure 2, all CFs can perform predictive interpolation for each data point on the studied data, which is shown in Figure 1. However, assessing based on the graph only is not comprehensive in concluding the best CF. In this regard, validation tests such as MSE, RMSE and CV play an important role in determining the accuracy and precision of performance for each CF. According to the value that the CV provided, all covariance functions are capable of modeling and are credible because all the values are less than 5%. However, Mat 3/2 recorded the lowest value compared to other CFs. On the other hand, it has been discovered that there is a variation in the predicted error for MSE and RMSE based on the numerical results that have been generated computationally. Ultimately, Mat3/2 has the lowest MSE and RMSE values in comparison to other CFs, with values of 0.11714 and 0.34226, respectively. However, RQ had the highest number in error for both MSE and RMSE, with 1.54782 and 1.24411, respectively. The values of MSE, RMSE, and CV% at RQ are the highest because there is a large and clearly visible gap at the 95% confidence interval for both the upper bound and the lower bound, which can be referred to in Figure 2 (b). This is very different compared to other CFs such as RBF, Per, Mat 3/2, and Mat 5/2, where the 95% confidence interval gap is not visible based on Figure 2 (a), (c), (d) and (e), respectively. The size of this large confidence interval causes models such as RQ to be less confident in the accuracy of predictions on the studied data because the distance between the predicted point and the actual point is so great. Ultimately, this demonstrates that Mat 3/2's nature and features enable predictions to be made effectively for data of the type of global wheat price from 1st January 1990 until 1st April 2023 that has a dynamic trend with uncertain fluctuations.

6. Conclusion

This study focusses on evaluating the effectiveness performances of the CF from GPR to determine the optimal CF for GPR modeling by performing interpolated predictions on Global Wheat Prices dataset. Graph analysis alone is insufficient to assess a CF's performance. The effectiveness, precision, and reliability of a CF have been assessed and compared in this study using MSE, RMSE, and CV to produce accurate results. Global Wheat Price dataset and each CF from GPR were processed using the Python programming language and Scikit-Learn module to get numerical results. According to the numerical results, Mat 3/2 acquired the MSE and RMSE error levels as well as the reliability value, which is the lowest CV within the 5% threshold when compared to other CFs. Mat 3/2 is hence the most effective CF for interpolation prediction. In the future, investigations combining CFs can be used to predict different data sets with irregular trend patterns and high volatility.

7. Acknowledgements
The authors would like to express our gratitude to the Universiti Teknologi PETRONAS for providing support and funding for the project under National Collaborative Research Fund cost center 015MC0-031.

References
[1] Sousa, Telma de, Miguel Ribeiro, Carolina Sabença, and Gilberto Igrejas. 2021. "The 10,000-Year Success Story of Wheat!" *Foods* 10 (9): 2124. https://doi.org/10.3390/foods10092124.

[2] Hawkesford, Malcolm J., Jose-Luis Araus, Robert Park, Daniel Calderini, Daniel Miralles, Tianmin Shen, Jianping Zhang, and Martin A. J. Parry. 2013. "Prospects of Doubling Global Wheat Yields." *Food and Energy Security* 2 (1): 34–48. https://doi.org/10.1002/fes3.15.

[3] Javid Iqbal, Muhammad, Naureen Shams, and Kalsoom Fatima. 2022. "Nutritional Quality of Wheat." *Wheat*, November. https://doi.org/10.5772/intechopen.104659.

[4] Shewry, Peter R., and Sandra J. Hey. 2015. "The Contribution of Wheat to Human Diet and Health." *Food and Energy Security* 4 (3): 178–202. https://doi.org/10.1002/fes3.64.

[5] Trust Beta, and Corinne Isaak. 2016. "Grain Production and Consumption: Overview." *Elsevier EBooks*, January, 349–58. https://doi.org/10.1016/b978-0-12-394437-5.00051-6.

[6] "Agricultural Production Statistics 2000–2021." 2022, December. https://doi.org/10.4060/cc3751en.

[7] Gyarmati, Gábor. 2017. "On What Factors the Wheat Production and Price Depends." Proceedings- 11th International Conference on Mangement, Enterprise and Benchmarking (MEB 2017), 78–96. https://ideas.repec.org/h/pkk/meb017/78-96.html.

[8] Enghiad, Aliakbar, Danielle Ufer, Amanda M. Countryman, and Dawn D. Thilmany. 2017. "An Overview of Global Wheat Market Fundamentals in an Era of Climate Concerns." *International Journal of Agronomy* 2017: 1–15. https://doi.org/10.1155/2017/3931897.

[9] Martín, Will, and Nicholas Minot. 2023. "The Impact of Price Insulation on World Wheat Markets during Covid-19 and the Ukraine Crisis," January. https://doi.org/10.2499/p15738coll2.136616.

[10] "Grain: World Markets and Trade." 2022. https://apps.fas.usda.gov/psdonline/circulars/grain.pdf.

[11] Amimul, Mohammad, and Wan Hussain. 2023. "Comparison of Machine Learning Models in Forecasting Reservoir Water Level." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 31 (3): 137–44. https://doi.org/10.37934/araset.31.3.137144.

[12] "Analysis of Wind Speed Prediction Using Artificial Neural Network and Multiple Linear Regression Model Using Tinyml on Esp32." 2023. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences* 107 (1): 29–44. https://doi.org/10.37934/arfmts.107.1.2944.

[13] Goparaju, Bhargavi, and Bandla Sreenivasa Rao. 2023. "Distributed Denial-of-Service (DDoS) Attack Detection Using 1D Convolution Neural Network (CNN) and Decision Tree Model." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 32 (2): 30–41. https://doi.org/10.37934/araset.32.2.3041.

[14] Khairul, Wan, None Afiqah Bazlla Md Soom, Aisyah Mat, None Juhaida Ismail, None Aszila Asmat, and Abdul Rahm. 2023. "Sales Forecasting Using Convolution Neural Network" 30 (3): 290–301. https://doi.org/10.37934/araset.30.3.290301.

[15] Muhamad Hadzren Mat, Prakash Nagappan, Fakroul Ridzuan Hashim, Ahmad, M M Saleh, Khairunesa Isa, and Khaleel Ahmad. 2023. "Hybrid Multilayer Perceptron Network for Explosion Blast Prediction." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 30 (3): 265–75. https://doi.org/10.37934/araset.30.3.265275.

[16] Safwan Mahmood Al-Selwi, Mohd Fadzil Hassan, Said Jadid Abdulkadir, and Amgad Muneer. 2023. "LSTM Inefficiency in Long-Term Dependencies Regression Problems." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 30 (3): 16–31. https://doi.org/10.37934/araset.30.3.1631.

[17] Carl Edward Rasmussen, and Christopher Williams. 2005. *Gaussian Processes for Machine Learning*. *The MIT Press EBooks*. The MIT Press. https://doi.org/10.7551/mitpress/3206.001.0001.

[18] Bian, Chaofa, Hongtao Shi, Suqin Wu, Kefei Zhang, Meng Wei, Yindi Zhao, Yaqin Sun, Huifu Zhuang, Xuewei Zhang, and Shuo Chen. 2022. "Prediction of Field-Scale Wheat Yield Using Machine Learning Method and Multi-Spectral UAV Data." *Remote Sensing* 14 (6): 1474. https://doi.org/10.3390/rs14061474.

[19] Sato, Koya. 2022. "Over-The-Air Gaussian Process Regression Based on Product of Experts." *2022 IEEE Globecom Workshops (GC Wkshps)*, December. https://doi.org/10.1109/gcwkshps56602.2022.10008771.

[20] "Gaussian Processes for Flow Modeling and Prediction of Positioned Trajectories Evaluated with Sports Data." n.d. Ieeexplore.ieee.org. Accessed September 30, 2023. https://ieeexplore.ieee.org/document/7528056.

[21] Parkinson, Jonathan, and Wei Wang. 2023. "Linear-Scaling Kernels for Protein Sequences and Small Molecules Outperform Deep Learning While Providing Uncertainty Quantitation and Improved Interpretability." *ArXiv (Cornell University)*, February. https://doi.org/10.48550/arxiv.2302.03294.

[22] Palm, Nicolai, Markus Landerer, and Herbert Palm. 2022. "Gaussian Process Regression Based Multi-Objective Bayesian Optimization for Power System Design." *Sustainability* 14 (19): 12777. https://doi.org/10.3390/su141912777.

[23] Daud, Hanita, Muhammad Naeim Mohd Aris, Khairul Arifin Mohd Noh, and Sarat Chandra Dass. 2021. "A Novel Methodology for Hydrocarbon Depth Prediction in Seabed Logging: Gaussian Process-Based Inverse Modeling of Electromagnetic Data." *Applied Sciences* 11 (4): 1492. https://doi.org/10.3390/app11041492.

[24] Garay, Fernando, William Huaman, and José Muñóz. 2022. "Verification and Validation of the Gaussian Process Regression Model to Predict the State of Health in Lithium-Ion Batteries," August. https://doi.org/10.1109/intercon55795.2022.9870158.

[25] Radhia Fezai, Majdi Mansouri, Nasreddine Bouguila, Hazem Nounou, and Mohamed Nounou. 2019. "Reduced Gaussian Process Regression for Fault Detection of Chemical Processes," December. https://doi.org/10.1109/iintec48298.2019.9112136.

[26] Aliu, Florin, Jiří Kučera, and Simona Hašková. 2023. "Agricultural Commodities in the Context of the Russia-Ukraine War: Evidence from Corn, Wheat, Barley, and Sunflower Oil." *Forecasting* 5 (1): 351–73. https://doi.org/10.3390/forecast5010019.

[27] Rasheed, Ahmed, Muhammad Shahzad Younis, Farooq Ahmad, Junaid Qadir, and Muhammad Kashif. 2021. "District Wise Price Forecasting of Wheat in Pakistan Using Deep Learning." *ArXiv (Cornell University)*, March. https://doi.org/10.48550/arxiv.2103.04781.

[28] Kumar, Raushan. 2020. "Predicting Wheat Futures Prices in India." *Asia-Pacific Financial Markets*, July. https://doi.org/10.1007/s10690-020-09320-6.

[29] Amin, Md Nur. 2020. "Predicting Price of Daily Commodities Using Machine Learning." IEEE Xplore. December 1, 2020. https://doi.org/10.1109/3ICT51146.2020.9312012.

[30] International Monetary Fund. 1980. "Global Price of Wheat." FRED, Federal Reserve Bank of St. Louis. January 1, 1980. https://fred.stlouisfed.org/series/PWHEAMTUSDM.