



# Comparison Of Jenkins Box Method And Multiple Linier Regression In Predicting The Noble Metal Price

Bayu Gunara<sup>1</sup>, Yennimar<sup>2</sup>

<sup>1,2</sup>Teknik Informatika,

<sup>1,2</sup>Universitas Prima Indonesia, Jl. Sekip Sei Kambing Medan 20111, Indonesia

Email : bgunara1@gmail.com

## ARTICLE INFO

### Article history:

Received: 23 -06- 2019

Revised: 22 -08- 2019

Accepted: 01 -09- 2019

### Keywords:

gold investment, forecasting, Multiple Linear Regression, Box-Jenkins, ARIMA, hybrid methods.

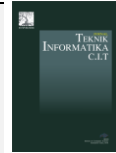
## ABSTRACT

*Nowadays, gold investment practitioners generally use instinct and guess in investing in gold. This is certainly a problem because it has a high error margin. To solve these problems, the forecasting process can be carried out. To be able to forecast gold prices with low error rates, various studies have been conducted. The Box-Jenkins method performs better than other methods in predicting the price of gold, because the Box-Jenkins method applies forecasting by relying on the historical statistics of gold prices beforehand. The Box-Jenkins method is an iterative of choosing the best model for the stationary series of a group of linear time series models called the ARIMA (Autoregressive Integrated Moving Average) model. However, the ARIMA method is a complex method and is not easy to use and requires a long execution time to obtain forecasting results with a high degree of accuracy. To improve the accuracy of prediction results from ARIMA, the ARIMA method can be combined with the multiple regression method into a hybrid method. The Multiple Linear Regression Method is a mathematical technique that minimizes the difference between the actual value and the predicted value. The results of this study are an application of forecasting the price of gold using the ARIMA method and Multiple Linear Regression. The application also provides a facility to test the results of the methods used. Based on the results of testing the accuracy of the prediction results from the hybrid method with 30 data = 48%, 60 data = 40%, and 118 data = 40.81%.*

© 2019 JTI C.I.T. All rights reserved.

## 1. Introduction

Gold is a precious metal that is often used as a medium of exchange in trade and as a financial standard in various countries. To get optimal profits, people who want to invest in gold certainly expect to get a low price at the time of purchase and an expensive price at the time of sale. Between 1975 and 1979, for example, gold price fluctuations ranged from \$ 121.00 - \$ 236.10. In 1982 to 2005, the price range of gold was between \$ 200 - \$ 400. In 2016, the highest selling price of gold reached \$ 1365 [3]. In the current conditions, gold investment practitioners generally use instinct and guess in making gold investments [4]. Investors in gold usually buy gold when the price of gold is considered low enough in the hope of reselling it when the price of gold is high. This is certainly a problem because this method has a high margin of error, where the price of gold that is falling can continue to fall to a much lower point, this can actually cause large losses to gold investors if they have to sell the gold when the price gold is much lower than the price when it was bought [4]. The economic factors that affect the price of gold, include. To solve these problems, the forecasting process can be done (forecasting). Forecasting is a function of management to support the decision making process. The forecasting process is also described as an estimation process for a situation that is not yet known [2]. To be able to forecast gold prices with low error rates, various studies have been conducted, Yuan (2012) uses an algorithm Projection Pursuit to find non-linear variables and construct a Backpropagation (BP) Neural Network to predict the price of gold [2]. Other research on correlation analysis and concludes that international business, politics, market conditions, especially commodity markets, consumer spending trends, and inflation are factors that have a major influence on the movement of gold prices. Meanwhile, the Box-Jenkins method has better



performance than other methods of forecasting prices gold, This is because the Box-Jenkins method applies forecasting by relying on historical gold price statistics, the more historical gold price data available, the forecast results will also be better and have a lower error rate. The Box-Jenkins method is an iterative of choosing the best model for series the stationary of a group of models time series linear called the ARIMA model (Autoregressive Integrated Moving Average). This method assumes that the value series produced by the process stochastic (random) with a form that can be explained [5]. According to Madhur Srivastava, et. al (2010), the ARIMA method is a complex method and is not easy to use and requires a long execution time to obtain forecasting results with a high degree of accuracy.

According to [6], to improve the accuracy of forecasting results from ARIMA, the ARIMA method can be combined with the method multiple regression. Madhur Srivastava, et. al. (2010) put forward a hybrid method between methods Regression and ARIMA in predicting the number of errors (bug) which will appear in the making software. Multiple Linear Regression can be used for interactions between variables dependent with several variables independent. Metode Multiple Linear Regression is a mathematical technique that minimizes the difference between the actual value and the estimated value [1]. Based on testing conducted by PT Madhur Srivastava, et. al. (2010), obtained information that the intermediate hybrid method ARIMA dan regression has a higher level of accuracy when compared to the ARIMA method only.

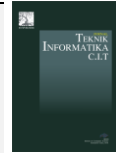
Based on the description above, this research will make an application of forecasting the price of gold with a hybrid method that combines the methods Box Jenkins Approach and Multiple Linear Regression to help people who want to invest in gold know the movement of gold prices in the future. Gold investment will be better if the investment knows the right time to buy gold when prices are low and sell gold when prices are high so that it can minimize losses and optimize the benefits of gold investment. With the application of this deepening is expected to overcome the problems that exist in gold investment.

## 2. Method

The system development method used is SDLC (System Development Life Cycle) or often referred to as the waterfall method.

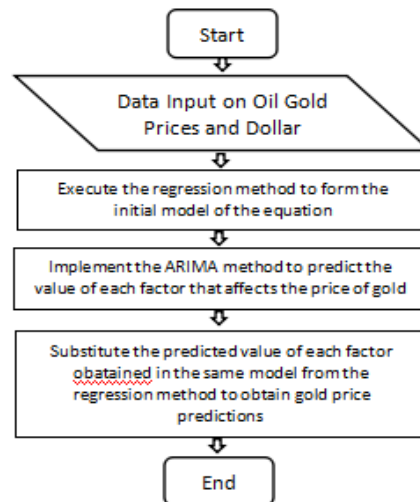
The steps in this methodology are:

- a) Needs Analysis  
All needs will be defined individually. The analysis is done by asking the user what is needed in the development of the system, especially historical data on gold prices. Analysis of functional needs of the system created using the help of use case diagrams and analysis of non-functional requirements of the system using the framework of Performance, Information, Economics, Control, Efficiency, Service (PIECES)
- b) Process Analysis  
At this stage, the work process of the Hybrid method will be described using a flowchart. After that, the calculation process of the Hybrid method will be elaborated in the process of forecasting the price of gold using a simple example.
- c) System design  
At this stage, a comprehensive design of the system will be built. This design includes system modeling using use case diagrams, database design which includes table design and inter-table relations using Microsoft SQL Server 2014 applications, user interface design that includes output design and input design. The application used to design is Balsamiq Mockup.
- d) Writing programs or coding (coding)  
After all needs are defined, the next step is to translate the results of the design process into a form of computer program that is understood by computer machines. The algorithm used to predict is the algorithm Box Jenkins Approach and Multiple Linear Regression. Making coding of this forecasting application will use the Microsoft Visual C # 2015 programming language.
- e) Testing the program (testing)  
At this stage a trial is carried out on the program to look for abnormalities (bugs) that can still occur in the program. The testing process will use a data set of gold prices (closing price per day in the afternoon, taken from the website [price-emas.org](http://price-emas.org)) starting in 2017 hingga 2018 dan will be compared to the forecast results obtained with the actual price of gold in a certain period. From the results of this test, it will be known the accuracy of the results of forecasting the price of gold from the hybrid model between the Box Jenkins Approach method and Multiple Linear Regression. The



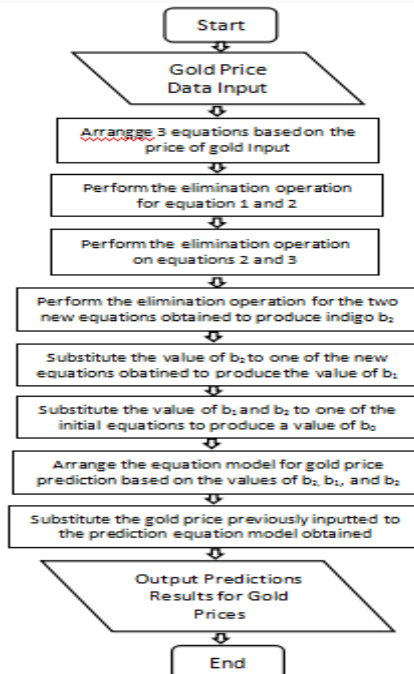
testing process of the accuracy of the hybrid model between the Box Jenkins Approach algorithm and the Multiple Linear Regression will use the confusion matrix method for the values of  $p$ ,  $q$  and  $d = 1$ . This testing process will use real gold price data as a comparison with the forecast price.

To improve the accuracy of the prediction results from the Box Jenkins (ARIMA) method, the ARIMA method will be combined with the Regression method into a hybrid model. The work process of this Hybrid method can be described in the form of a flowchart as shown in the following figure:

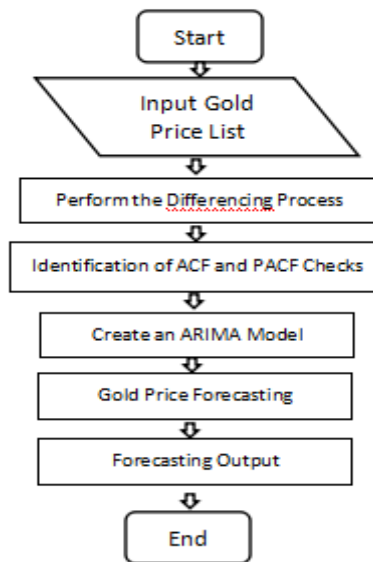


**Fig 1.** Flowchart of the Hybrid Method Process.

The work process of the linear regression method can be described in the form of a flowchart as shown in the following figure:



**Fig 2.** Flowchart of the Multiple Linear Regression Process Process



**Fig 3.** Flowchart of the ARIMA Method.

The working procedure of the ARIMA method can be described in the form of a flowchart as shown in the following figure :

### 3. Result And Discussion

#### 3.1. Result

When you first run the gold price forecasting application with a hybrid model between the Box Jenkins Approach method and Multiple Linear Regression, the form that will appear first is the 'Login' form, which can be seen in Figure 4 below:

**Fig 4.** Login Form.

If the user wants to fill in historical data on gold prices, oil prices and dollar exchange rates, the user must log in as an admin user. The user must enter a valid user ID and password, in order to use the admin section of the software.

If the login process is successful, then the system will display the Main form as shown in the following image:





**Fig 5.** Display Main Form.

In this Main form, there are several menus that can be accessed, such as:

- a) Main Menu, which serves to provide an interface to fill in the main data required by the system. This main menu has several sub menus, namely:
  - a. Gold Price Input sub menu, which functions to display the Gold Price Input form
  - b. Dollar Exchange Input sub menu, which functions to display the Dollar Exchange Price Input form.
  - c. Oil Price Input sub menu, which functions to display the Oil Price Input form..
  - d. Gold Price Forecast menu, which functions to forecast gold prices by displaying the Gold Price Input form.
- b) Report Menu, which functions to display the forecasting information using the Hybrid method.
- c) Testing Menu, which functions to carry out the testing process by displaying the Test Results form.
- d) Add New User Menu, which functions to display the Add User form.

In the 'Main Menu', there are several sub menus that can be accessed, such as:

- a) Gold Price Input Sub Menu, which functions to display the Gold Price Input form as shown in the following image:

**Fig 6.** Form Input Gold Prices.

In this Gold Price Input form, users can enter historical details of the gold price, such as period and price (USD). After completing the historical details of the gold price, the user can click the Save button to save the data to the Kurs\_Emas table in the database. Data stored in that database will be displayed in the Gold Price Data table.

- b) Dollar Exchange Input Sub Menu, which is performed to display the Dollar Exchange Data Input form as shown in the following figure:

**Fig 7.** Input Dollar Exchange Forms.

In this Dollar Exchange Input form, users can enter historical details of the dollar exchange rate, such as the date and price of the exchange rate conversion. After completing the details of the dollar exchange rate data, the user can click the Save button to save the data to the Kurs\_Dollar table in the database. Data stored in that database will be displayed in the Dollar Exchange Data table.

- c) Oil Price Input Data Sub Menu, which functions to display the Oil Price Input Data form as shown in the following image:





## Jurnal Teknik Informatika C.I.T

journal homepage: [www.medikom.iocspublisher.org/index.php/JTI](http://www.medikom.iocspublisher.org/index.php/JTI)



**Fig 8. Oil Price Input Forms.**

On this Oil Price Input Data form, users can enter data on the historical details of oil prices, such as the date and price of oil. After completing the detailed oil price data, the user can click the Save button to save the data to the Oil table in the database. Data stored in that database will be displayed in the Oil Data table.

- d) Gold Price Prediction Sub Menu, which functions to display the Gold Price Forecasting form as shown in the following image:

**PENGANTARAN HARGA EMAS (per Oz)**

Pilih tanggal: 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792

**Fig 9. Form Forecasting Gold Prices.**

In this Gold Price Forecasting form, users can enter the initial data needed in the forecasting process, such as the data set period that will be used in the forecasting process and the period that will be predicted. In addition, the user can also determine the parameter values  $p$ ,  $d$  and  $q$  manually. These three parameter values are used in working procedures of the ARIMA method. After the user enters all the initial data needed, the user can click the Prediction button to start the forecasting process.

In this Forecasting Form, there are two main buttons that can be accessed namely:

- Forecasting Results button used to display all the forecasting results steps.
- The Previous Step button is used to display the previous calculation step.
- The Next Step button is used to display the next calculation step.

Add User Sub Menu, which functions to display the Add User form as shown in the following image:

[illegible]

**Fig 11. Form Add User**

User data entered will be used when you want to log into the system. The details of user data entered include data user names, passwords.

The testing process can be done by clicking on the Test Results menu, so the system will display the Test Results form as shown in the following image:

[illegible]

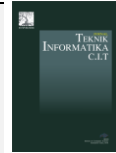
**Fig 12. Form Testing Results**

In this form the average error will be calculated from the results of tests conducted by the user. The testing process will use the mean absolute error (MAE) method.

As shown in the picture above, based on testing using the MAE method performed on 30 data, 60 data and 118 data, obtained information that the average value of the error from the MAE method is 135.2479.

### 3.2. Discussion

Meanwhile, to test the accuracy of the predicted results of gold prices rising or falling, the Confusion



Matrix method will be used, along with the details:

Based on the results of tests conducted on the software, the following information can be obtained:

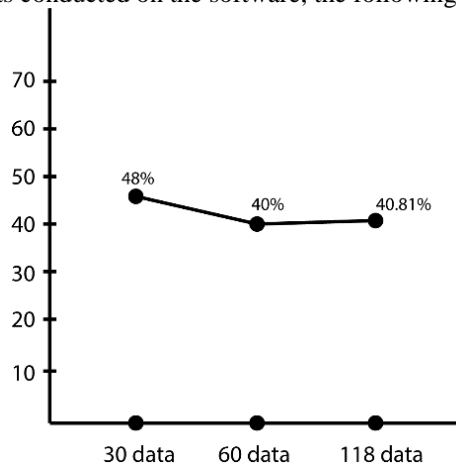


Fig 14. Outline of Test

- By using the MAE method to calculate the difference in the price of gold between the real value and the predicted value, information is obtained that the average error for 118 data obtained is equal to 135.2479
- Dengan By using the confusion matrix method to calculate the accuracy of the predicted results up or down, obtained information that the accuracy of the prediction results is equal to 40.81% for the dataset totaling 118 pieces of data.
- Hybrid method can be used to predict gold prices increase or decrease, but the level of accuracy is less satisfactory. To improve the accuracy of the predicted results, other supporting factors must be included that affect the price of gold.
- Hybrid method has a higher level of accuracy in predicting gold prices in the short term, but the level of accuracy is less good in predicting gold prices in the long run.

#### 4. Conclusion

From the discussion in the previous chapters, finally the research in this final project can be drawn several conclusions, among others.

- Based on the results of tests conducted, information is obtained that the hybrid method can be used to forecast gold prices with accurate gold price forecasting which reaches 40.81% from 118 test data. This means that the gold price forecasting results from the Hybrid method can be used as a reference only when investing in gold.
- he hybrid method is suitable for use in forecasting the price of gold for the short term with an accuracy of 48% from 30 test data.

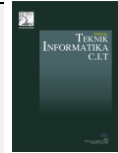
#### 5. Reference

- [1] Bodie Z, Kane A, Marcus A. J. (2009). Investments: eight edition. McGraw Hill. USA.
- [2] Ismail, Z., A. Yahya dan A. Shabri. (2009). Forecasting Gold Prices Using Multiple Linear Regression Method, American Journal of Applied Sciences 6 (8): 1509-1514, ISSN 1546-9239, Science Publications.
- [3] Kitco. (2017). Kitco.com, Gold Prices History.
- [4] Mohamad Heykal, Erlin. (2011). Analisis Kinerja Perdagangan Kontrak Berjangka Logam Mulia pada Periode Oktober 2009 - Desember 2009, BINUS BUSINESS REVIEW Vol. 2 No. 1 Mei 2011: 181-191.
- [5] Puji Catur Siswipraptini, Sri Rahayu. (2016). Aplikasi Simulasi dan Pemodelan Menggunakan Metode Linier Trend pada CV. Bina Multi Barokah, P-ISSN 2085-4315 / E-ISSN 2502-8332.
- [6] Srivastava, M., D. Badal, R. K. Jain. (2010). Regression and ARIMA Hybrid Model for New Bug Prediction, Madhur Srivastava et al. / (IJCSSE) International Journal on Computer Science and Engineering Vol. 02, No. 08, 2010, 2622-2628.
- [7] Yuan, J., F. Z. Guang, Q. M . Fang. (2012). Coal Calorific Value Prediction Based on Projection Pursuit Principle, TELKOMNIKA Indonesian Journal of Electrical Engineering Vol.10, No.6, October 2012, pp. 1287~1292.



## Jurnal Teknik Informatika C.I.T

journal homepage: [www.medikom.iocspublisher.org/index.php/JTI](http://www.medikom.iocspublisher.org/index.php/JTI)



- [8] Anita. (2015). Analisis Komparasi Investasi Logam Mulia Emas Dengan Saham Perusahaan Pertambangan Di Bursa Efek Indonesia 2010-2014. Jurnal Bisnis dan Manajemen. Volume 5 No.2.
- [9] Tanti Octavia, Yulia, Lydia, Aplikasi Simulasi dan Pemodelan Menggunakan Metode Linier Trend pada CV. Bina Multi Barokah, Seminar Nasional Informatika 2013 (semnasIF 2013), UPN "Veteran" Yogyakarta 2013.
- [10] Chandra, A. (2017). Estimasi Trip Frequency dengan Menggunakan Model Multiple Linear Regression, Jurnal Metris 18 (2017) 25–28,