| | |
|---|---|
| **Course Name** | **Forecasting and Quantitative Analysis** |
| **Lecturer** | **Dr. Ha Xuan Son** |
| **Student Name** | **Phan Ngan Ha** |
| **Student ID** | **S3926968** |
| **Word count** | **3000, exluding References and Appendix** |

# Table of Contents

## I. Introduction

Germany's temperatures had risen continuously, with 2021 being 11th consecutive warm year (WMO 2022). Germany's cities gradually receive greater warm effect, with temperatures of 24.3°C to 29.8°C (BW 2023). Northernmost regions being exposed to Atlantic Ocean has 3.3°C to 21.11°C (CT 2023). However, northernmost regions bordering Scandinavia is extremely cold, reaching −5°C or -10°C (CLV 2023). Although Germany has temperate, rainy climate zone of mid-latitudes, its overall mean temperatures still get warmer and have increased by 2.1°C between1991-2021 (Worldbank 2023).

Warmer temperatures damage both ecological system and humans. Higher temperature can increase ratio of evaporation, relative humidity, wind speed, wind direction and precipitation patterns. Moreover, Germany's outdoor activities, like agriculture and businesses can receive impact from warmer temperatures. Germany is also highly vulnerable to heat exposure because of large its elderly population (Mücke&Litvinovitch 2020), causing thousands of heat-related deaths annually (Winklmayr et al. 2022). These emphasize importance of monitoring and forecasting temperatures.

This consulting project aims to examine Germany’s temperature conditions for client's request. Germany monthly mean temperatures (°C) between 1991-2021 with 372 observations are collected from World Bank Group, Climate Change Knowledge Portal. Part.A will Section.1 demostrates descriptive measures of monthly temperatures, Section.2 select best ARIMA/SARIMA model to forecast temperatures between 2022-2032.

## II. PART A

### 1. Descriptive statistics

### a) Central tendency

| Method | Mean | Median | Mode |
|---|---|---|---|
| Germany's mean temperature (°C) | 9.548978 | 9.185 | N/A |

***Table.1.Central tendency***

Central tendency of Germany's temperature is considered to choose a value representing the entire distribution. There are no repeated values, so no conclusions are drawn from Mode. Applying IQR, dataset has no outliers, implying assumption observations are not skewed by outliers. However, skewness syntax is applied to test dataset's normal distribution assumption. Skewness syntax's result of –0.0188 implies minorly left-skewed distribution (Figure.1). These 2 methods confirm dataset's non-normal distribution assumption. As Median of 9.185 is unaffected by skewed distributions (Chikkodi&Satyaprasad 2010), Median is selected to represent entire distribution.

```
> skewness(Gts$value)
[1] -0.01881819
```

***Figure.1.Germany's temperatures left-skewed distribution***

| Method | Germany's mean temperature (°C) |
| --- | --- |
| Range | 25.29 |
| IQR | 11.7 |
| Standard Deviation (SD) | 6.59604 squared |
| Sample Variance (SV) | 43.50774 squared |
| Coefficient of Variation (CV) | 0.6907 squared |

***Table.2.Variation***

SD and SV have square root and square in function so they may add more weights to temperatures variation (Appendix.A&Appendix.B). Similar reasons apply to exclude CV as CV's equation comprises SD (Appendix.C). IQR only focuses on central half of data distribution and neglects completely portions below lower-quartile and above upper-quartile, excluding lowest and highest temperatures. If choosing IQR, highest temperatures during sunny season are neglected. Given Germany's unusually warmer temperatures between 2018-2021, highest temperature reached is pivotally considered in methodology, lying above upper-quartile. Thus, range is optimally useful as it can tell spread of entire dataset and know minimum and maximum temperature Germany can reach (Chikkodi&Satyaprasad 2010).

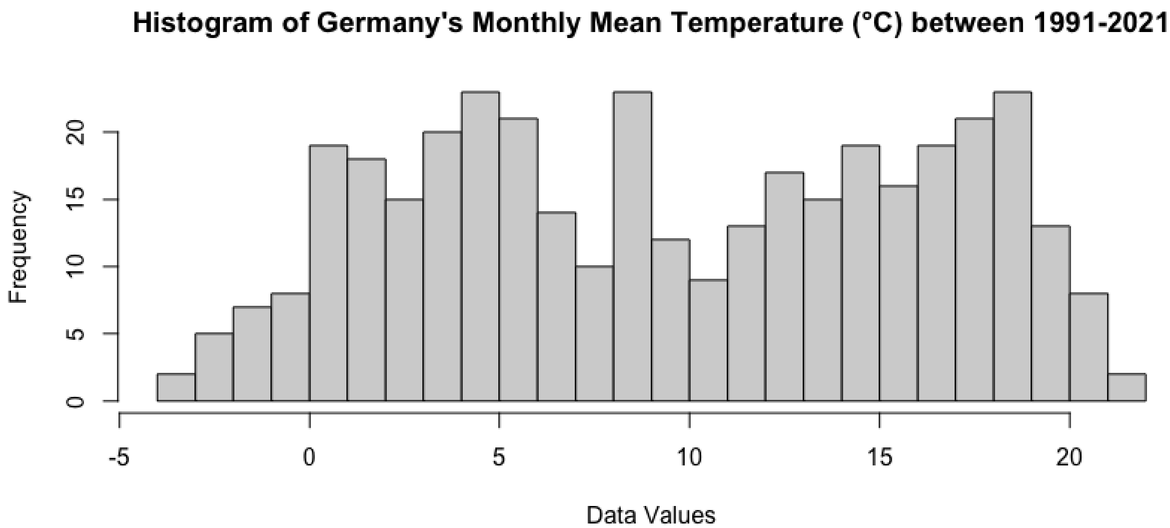


***Figure.2.Histogram***

There are 3 zones where values are mostly lying (0;6],(8;9],(12;19]. Obviously, bin (4;5],(8;9],(18;19] display highest frequency of Germany's temperatures with nearly 25 times each, disclosing Germany's monthly mean temperature usually stood at 4-5°C, 8-9°C, 18-19°C between 1991-2021.

## b) Time series plot and ACF plot

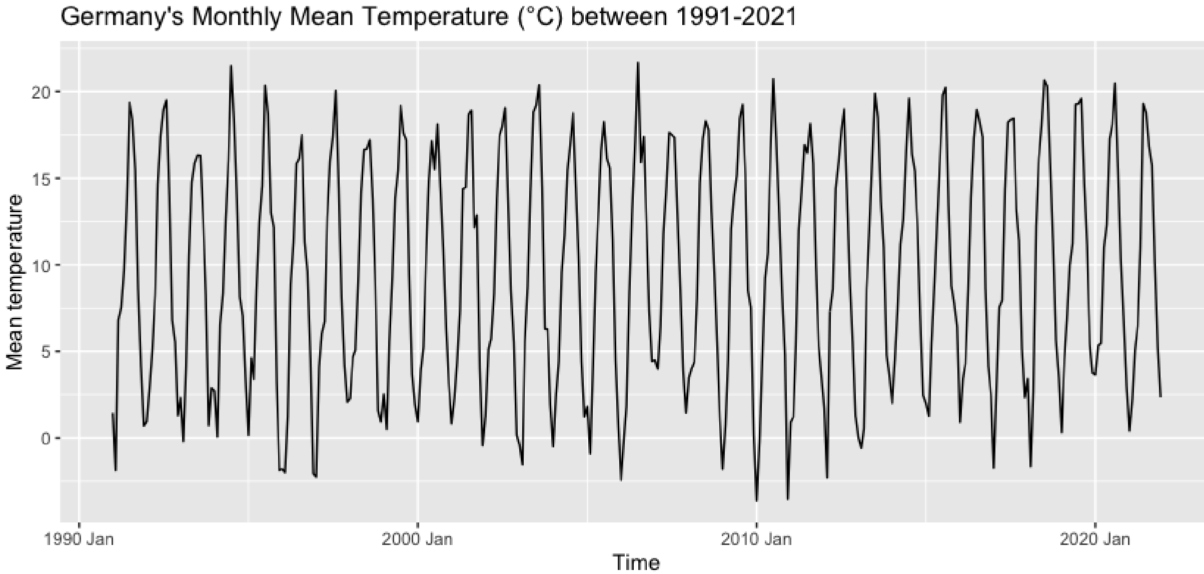


***Figure.3.Germany's monthly mean temperature (°C) time-plot***

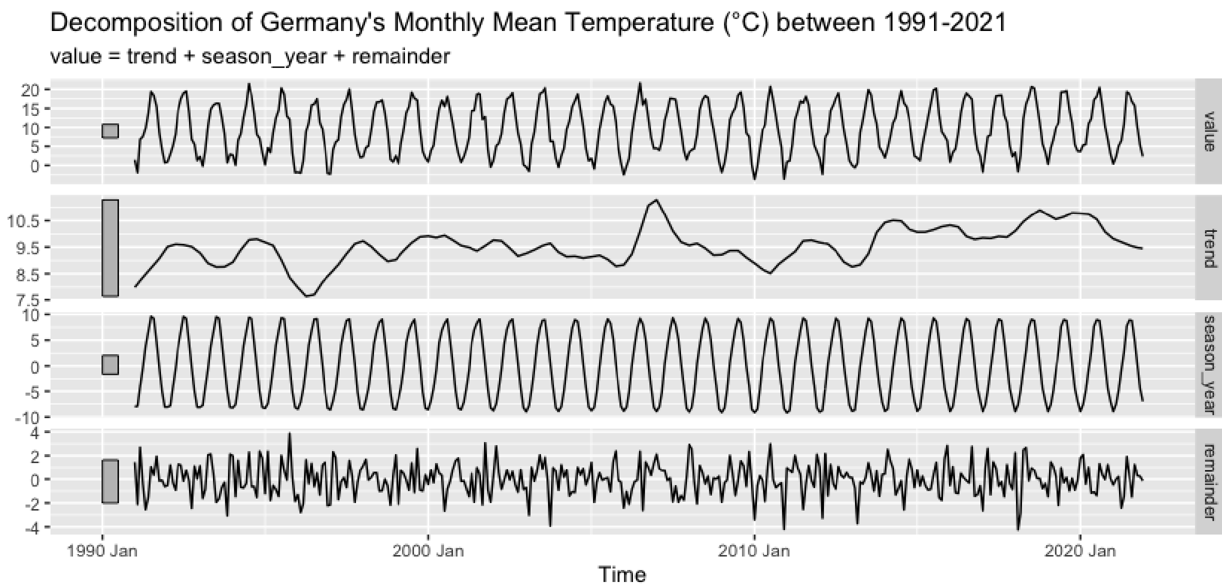


***Figure.4.Germany's 1991-2021 monthly mean temperature (°C) decomposition***

Firstly observing from Figure.3, seasonality can be observed. However, trend is hard to observe. Therefore, time series is decomposed into trend, seasonality, and remainder components (Figure.4).

Mean temperature in 1990 stayed at 8.0°C but gradually increased to 9.5°C in December/2022, signaling an upward trend in Germany's temperature. Additionally, temperature exhibits recurring pattern annually, suggesting seasonality exist.

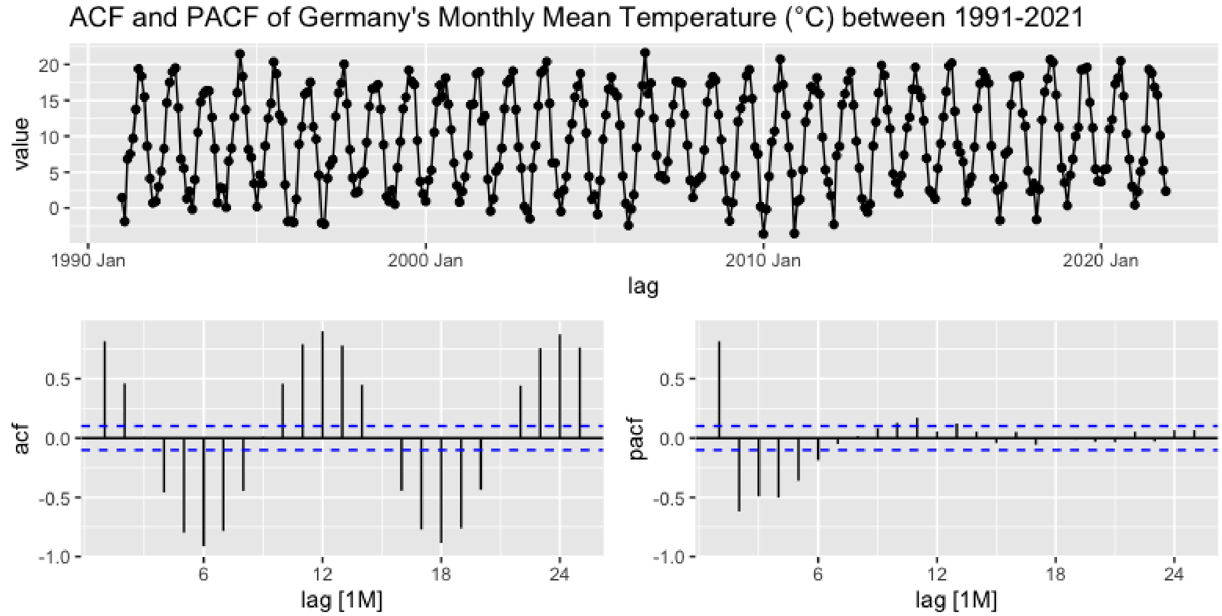


***Figure.5.ACF, PACF***

```
> Gts|> features(value, features = guerrero)
# A tibble: 1 × 1
  lambda_guerrero
            <dbl>
1            1.47
```

***Figure.6.Optimal lambda_guerrero***

Overall, ACF depicts wavelength patterns for monthly Germany' temperatures, swinging at regular frequency that is suspected of being non-stationary owing to seasonality. Excluding lags 3, 9, 15, 21 that are within the bounds in ACF, remaining lags are significant as they exceed bounding lines. More partially, in PACF, there are less significant and more insignificant lags than

ACF but patterns remain in wave, signaling seasonality, validating non-stationary assumption. This can be because mean values of different seasonal patterns vary owing to inconstant temperatures values distribution, hence inconstant mean throughout time-series. Variance is already constant because optimal lambda_guerrero 1.47 suggests denial of box_cox transformation (Figure.6). Accordingly, seasonal differencing of order is applied to remove seasonality to accomplish stationary.

## 2. Model selection

### a) Transformation

Applying unitroot_kpss to check non-stationary dataset assumption, result is 0.0649, above p-value 0.05. Thus, assumption is rejected, implying dataset should be stationary, and no difference is needed. However, limitations of unitroot_kpss concern me as it considers whole data series and might neglect underling patterns, seasonality in this case. Thus, dataset may not require first differencing of order but seasonal differencing is needed. Applying unitroot_ndiffs, results show 0, suggesting no first order differencing whole time series be required. Adopting unitroot_nsdiffs, result shows 1, revealing one more seasonal differencing is required, verifying assumption dataset only needs seasonal differencing.

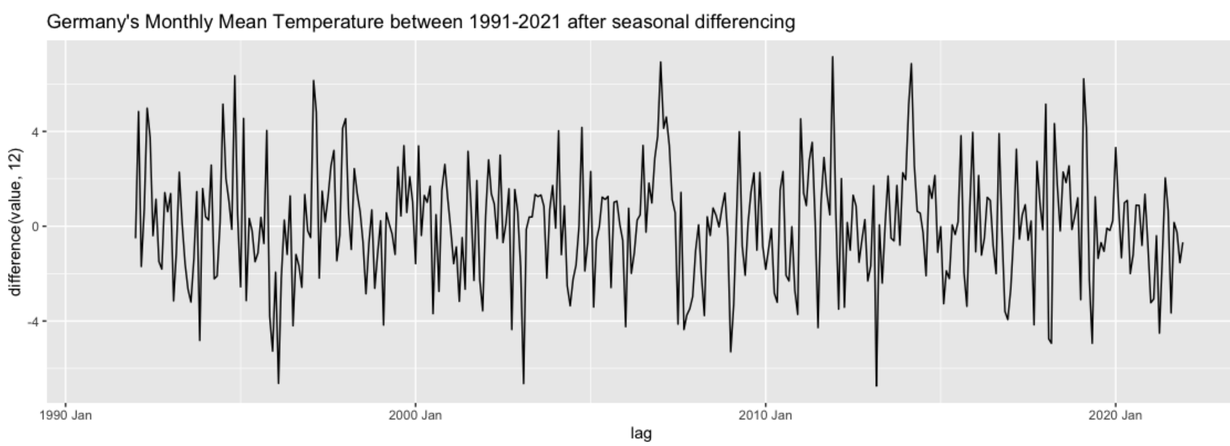


***Figure.7.Germany's 1991-2021 monthly mean temperature (°C) after seasonal differencing***

```
> Gts |> features(difference(value,12), unitroot_kpss)
# A tibble: 1 × 2
  kpss_stat kpss_pvalue
      <dbl>       <dbl>
1    0.0300         0.1
>
> Gts |> features(difference(value,12), unitroot_nsdiffs)
# A tibble: 1 × 1
  nsdiffs
    <int>
1       0
>
> Gts |> features(difference(value,12), unitroot_ndiffs)
# A tibble: 1 × 1
  ndiffs
   <int>
1      0
```

***Figure.8.Unitroot_kpss, unitroot_nsdiffs, unitroot_ndiffs***

Figure.7 shows Germany's monthly temperature between 1991-2021 after seasonal differencing. Checking unitroot_kpss, result is 0.1, above p-value 0.05, suggesting now stationary dataset (Figure.8). Using unitroot_nsdiffs and unitroot_ndiffs on seasonally differenced data, results show 0 (Figure.8), implying no further seasonal differencing or first differencing of order is needed. Accordingly, dataset is now stationary.

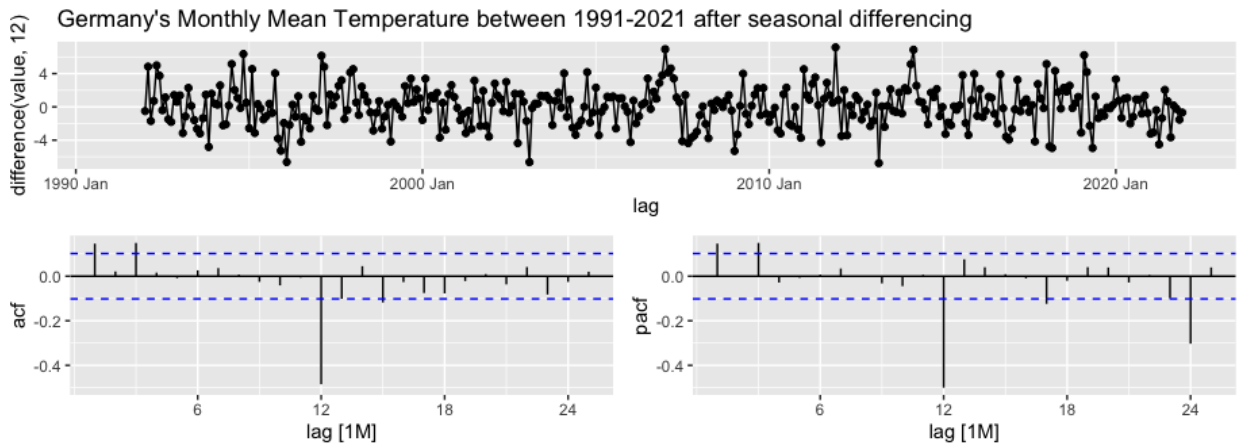


***Figure.9.ACF,PACF after seasonal differencing***

ACF and PACF are plotted to explore whether data after being seasonally differenced meet stationary assumption (Figure.9). Most of ACF's and PACF's lags do not slowly tail off, reflecting weakened correlation of temperature values dependent on previous observations compared to original dataset. As no trend or recurring period is found, dataset is now ready to finish forecasting.

**b) Arima**

My aim now is identifying optimal ARIMA/SARIMA model according to ACF and PACF (Figure.10). Regarding ACF plot (for MA), last significant spike in early lags is at lag 3, and single seasonal spike at lag 12. Moreover, concerning PACF plot (for AR), last significant spike in early lags is at lag 3, and there are 2 seasonal spikes at lag 12 and 24. Lastly, there is only one seasonal differencing.

ACF analyzes correlation between time series and its lagged values at different lags, whereas PACF focuses on measuring direct relationship between time series at specific lag and later lag after eliminating impact of all between lags (Hyndman&Athanasopoulos 2018). In other words, if solely considering ACF to estimate models, underlying relationship like partial correlations between particular lag and later lag can be neglected. Contradictorily, if PACF is solely focused, correlation of whole time series and its lagged values at different lags might be neglected. Both ways affect forecast values' accuracy. This creates assumption applying to both ACF and PACF will diversify estimated model choices and might induce best ARIMA model.

Three recommended models are estimated:

(i) Combine: ARIMA(3,0,3)(2,1,1)[12]

(ii) ACF: ARIMA(0,0,3)(0,1,1)[12]

(iii) PACF: ARIMA(3,0,0)(2,1,0)[12]

```
# A mable: 4 x 2
# Key:     Model name [4]
  `Model name`                    Orders
  <chr>                          <model>
1 arima303211  <ARIMA(3,0,3)(2,1,1)[12]>
2 arima003011               <NULL model>
3 arima300210  <ARIMA(3,0,0)(2,1,0)[12]>
4 auto         <ARIMA(3,0,0)(1,1,2)[12]>
```

***Figure.10.Models***

c) **Diagnostic test**

```
# A tibble: 3 × 6
  .model      sigma2 log_lik   AIC  AICc   BIC
  <chr>        <dbl>   <dbl> <dbl> <dbl> <dbl>
1 auto          2.87   -709. 1432. 1432. 1459.
2 arima303211   2.84   -709. 1438. 1438. 1477.
3 arima300210   3.66   -745. 1503. 1503. 1526.
```

***Figure.11.Accuracy***

Except for model ARIMA(0,0,3)(0,1,1)[12] applying only ACF receiving NULL result, remaining models deserve to forecast upcoming temperatures(Figure.10). Auto model is ARIMA(3,0,0)(1,1,2)[12](Figure.11). Given auto ARIMA(3,0,0)(1,1,2)[12] has lowest AICc and

BIC of 1432 and 1459, it is best model to forecast Germany's monthly temperatures between 2022-2032.

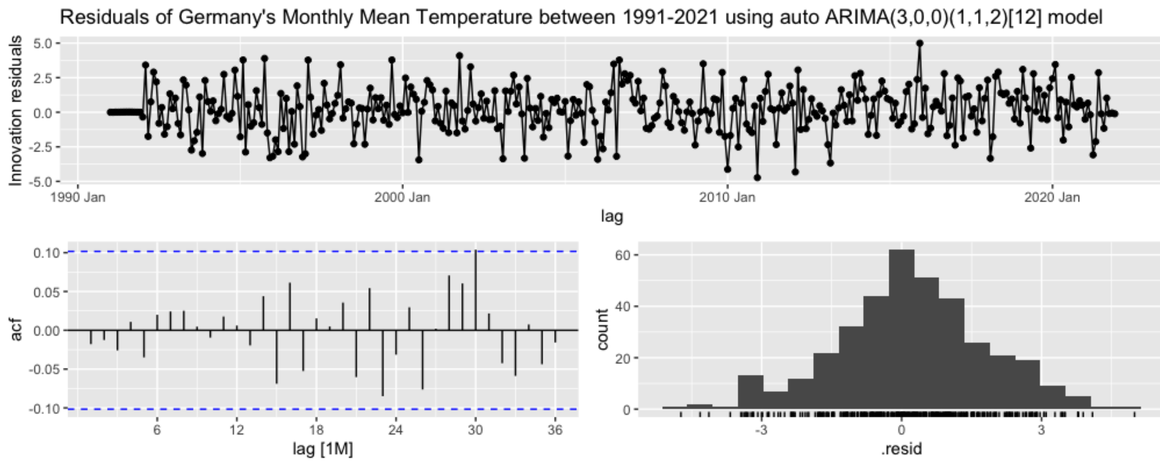


***Figure.12.AutoARIMA(3,0,0)(1,1,2)[12]_residuals***

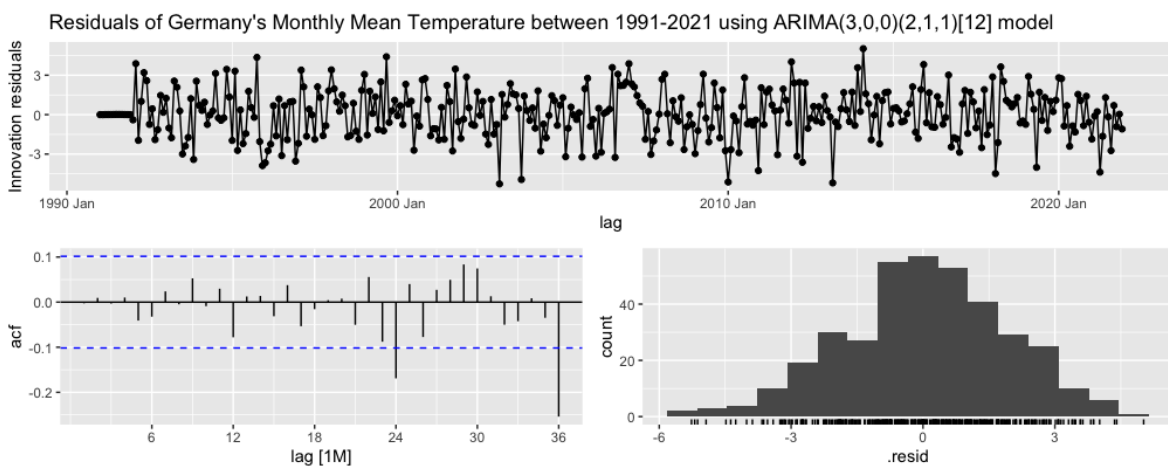


***Figure.13.ARIMA(3,0,3)(2,1,1)[12]_residuals***

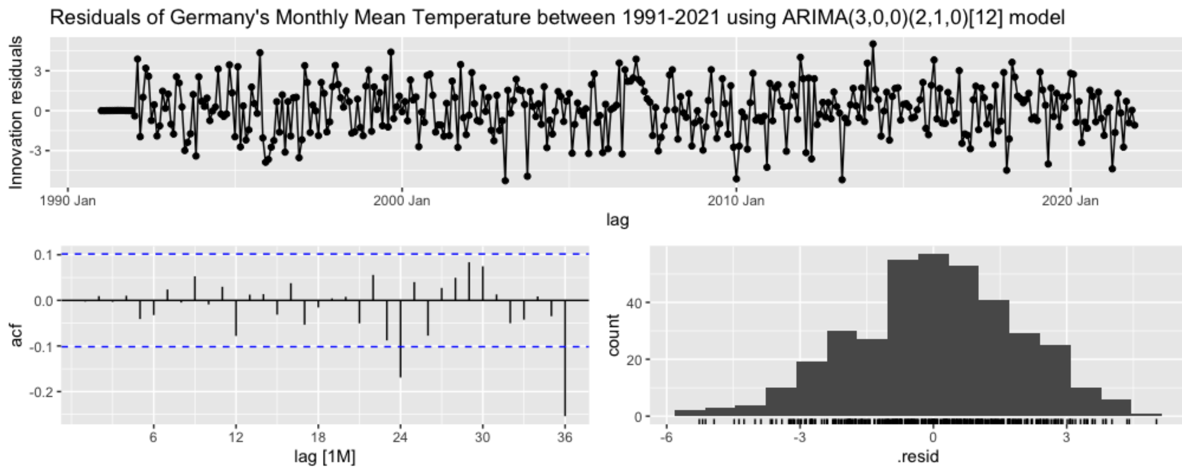


***Figure.14.ARIMA(3,0,0)(2,1,0)[12]_residuals***

| | .model | lb_stat | lb_pvalue |
|---|---|---|---|
| | <chr> | <dbl> | <dbl> |
| 1 | arima003011 | NA | NA |
| 2 | arima300210 | 62.7 | 0.000956 |
| 3 | arima303211 | 34.9 | 0.333 |
| 4 | auto | 27.3 | 0.702 |

***Figure.15.Autocorrelations checked by ljung_box test***

To confirm auto(3,0,0)(1,1,2)[12] best model, both residuals and ljung_box are applied. Regarding residuals, error-terms in ACF are well-defined to have white noise because all lags are within bounds (Figure.13), implying (3,0,0)(1,1,2)[12] neglect no accounted-for trends and seasonality. ACF's error terms of (3,0,3)(2,1,1)[12], (3,0,0)(2,1,0)[12] are not white noise as lags 24 and 36 still exceed bounding-lines. Regarding autocorrelations, only auto(3,0,0)(1,1,2)[12]'s p-value 0.745 above 0.05 confirms white noise assumption while remaining models failed to confirm. Thus, auto(3,0,0)(1,1,2)[12] is best model fitting Germany's monthly temperature.

### d) Final chosen model

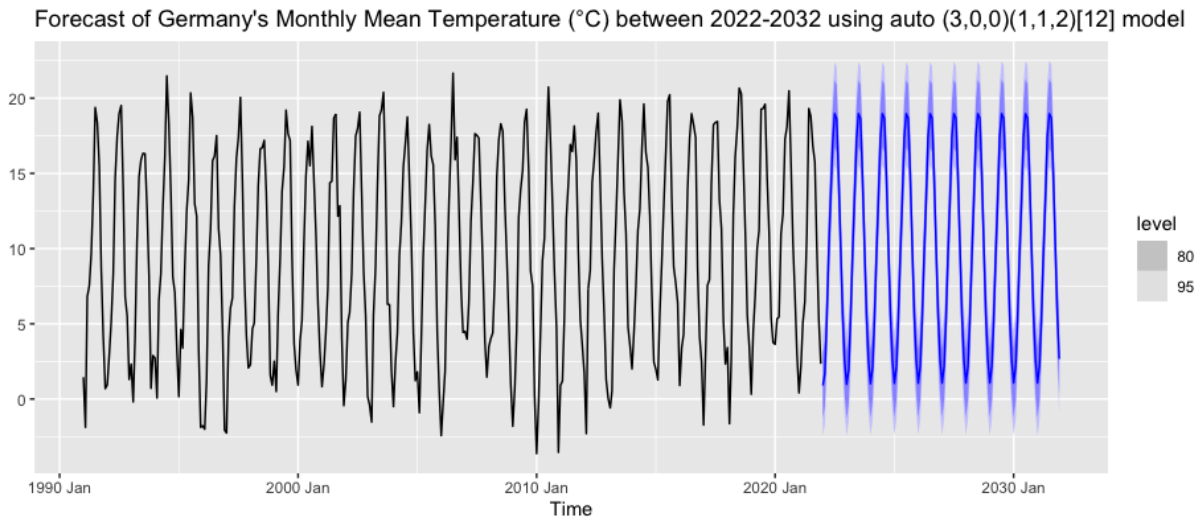


***Figure.16.Forecast 2022-2032 Germany's monthly team temperature(°C)***

By using auto(3,0,0)(1,1,2)[12], Germany's monthly would remain its seasonality. However, the range seems to narrow, with temperatures fluctuating from 1.8°C to 18.9°C, lower than 2021's highest temperature at 19.31°C (Figure.16). This means winter season seems to be warmer between 2022-2032, given Germany's continuously reduced demand for goal imports to heat up during winter (DWD 2022), because human-caused climate change makes every heatwave hotter, accelerating poles' melting (Masson et al.2020;Copernicus2023). Conclusively, Germany experiences rising temperatures with consistent seasonal-patterns where winter will get warmer between 2022-2032.

## III.PART B

### a) Overview

Aluminum, 99.5% minimum purity, is extensive pivotal metal thanks to its affordability and primary construction and industrial component. When aluminum price increased by 22%y-o-y in 2022, such sectors' productions stagnated (IMF 2022). Pincheira&Hardy (2021) discovered economies with moderate or limited base metal production, including UK, New Zealand, South Africa, have some capacities to forecast aluminum prices. This research predicts 2024 monthly aluminum prices based on historical 372 observations between 1992-2022, with forecast being mostly utilized by countries importing aluminum to mitigate production sluggishness.

### b) Time-plot and recommended prediction method

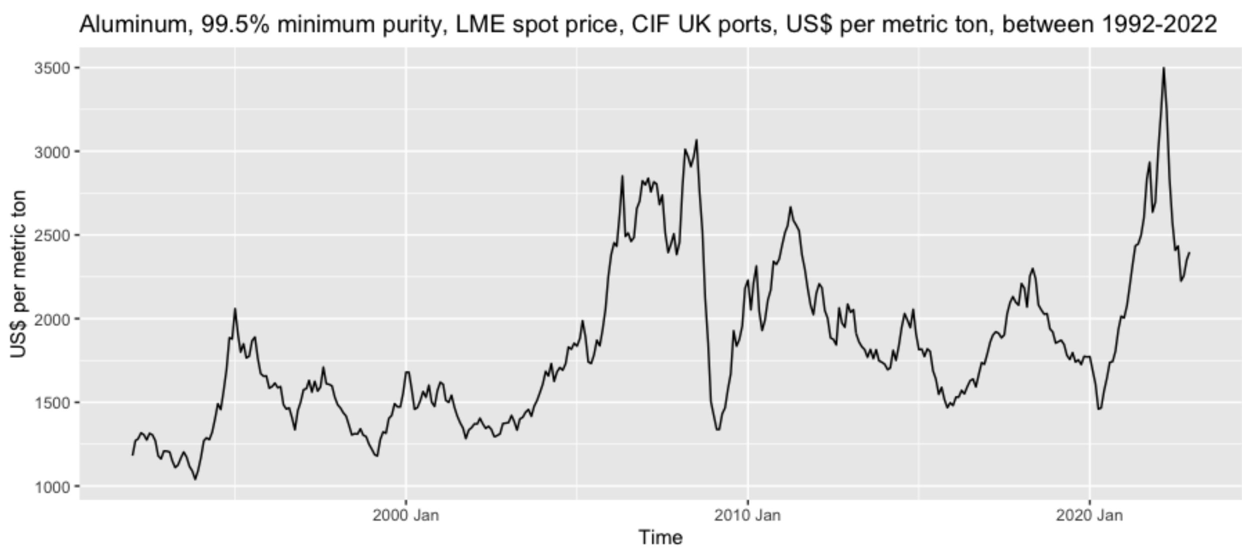


***Figure.17.Time-plot***

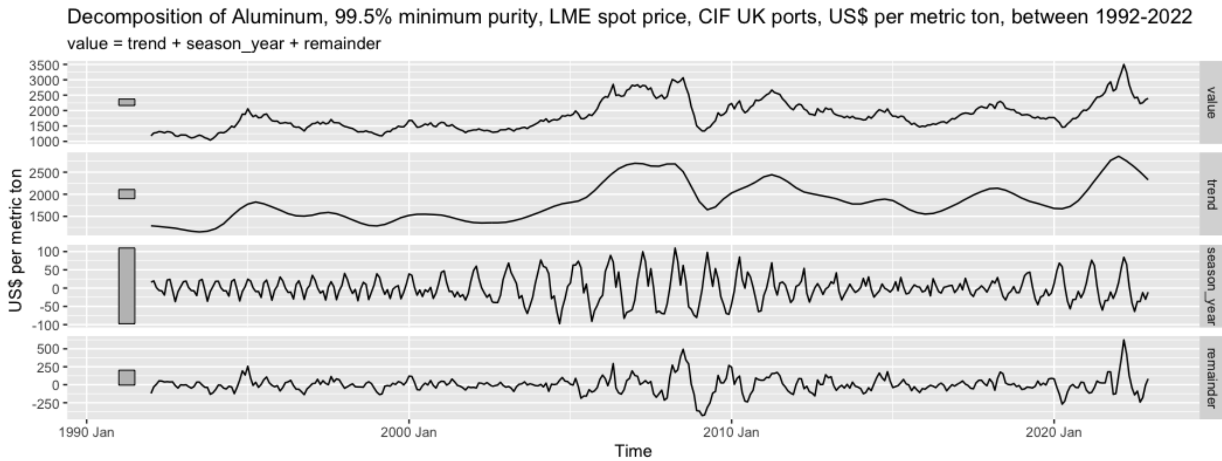


***Figure.18.Decomposition***

Aluminum price has increasing trend from price $1181.227US to $2398.207US (Figure.17). Seasonality is hard identified as time-series does not obviously depict. Moreover, there is spike at $3500US around mid-2022. This is because China, comprising 56% of global supplies, reduced aluminum quantity supplied for global market by 32.4% (S&PGlobal 2023). However, global aluminum demand increased from 93.5 to 96.6 million metric tons in end-2022. Thus, aluminum price was pressurized to increase by 22% (IMF 2022).

Dataset is suspicious of having outliers owing to price surge around mid-2022. Removing outliers can formulate higher precise dataset (Zhang et al.2022). However, removing outlier directly on original dataset can mistakenly remove important values contributing to trend, or seasonal-patterns. Therefore, decomposing dataset and excluding trend and seasonality, there are 3 significant spikes in remainder being suspected of outliers (Figure.18). Applying IQR to detect extreme values in remainder, results validate outliers' existence consumption and that add there are 9 outliers (Appendix.D), which are then removed.

Applying ARIMA to interpolate missing observations, new dataset is ready to finish forecasting(Figure.19).

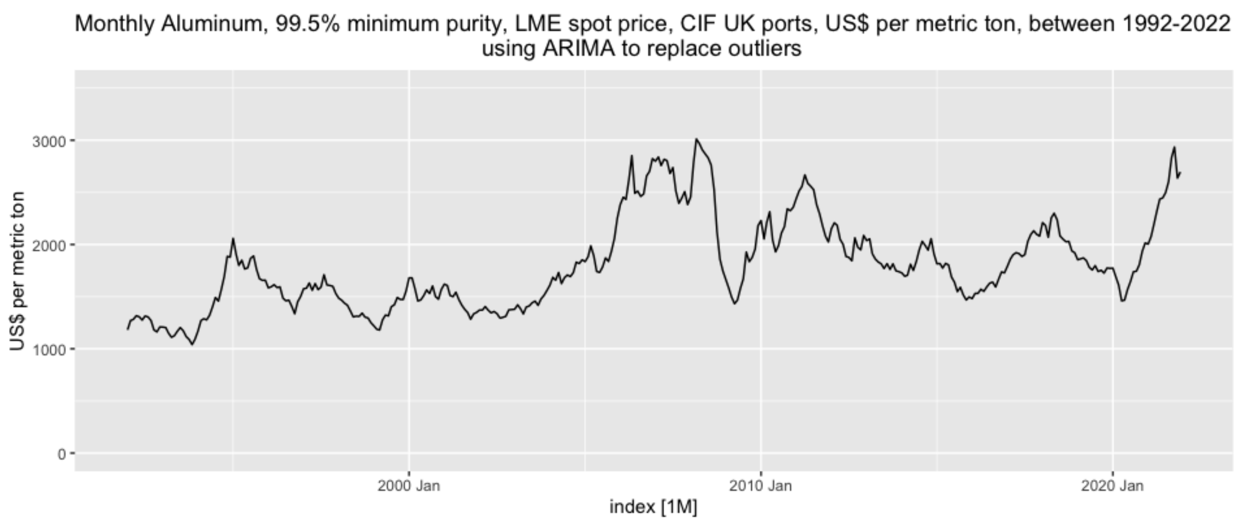


***Figure.19.Using ARIMA to replace outliers***

To forecast 2023-2024 aluminum price, forecast models must fit dataset. Therefore, whole dataset is divided into train and test set. Train set comprises Jan/1992-Dec/2021 to proceed forecasting. Test set comprises Jan/2022-Dec/2022 to check accuracy of train set's forecasting 2022.

Mean, Naïve, Simple exponential smoothing (SES) methods are applied to forecast monthly aluminum between 2023-2024. In Mean, forecast values equal average of all historical values. Naïve forecasts set all future forecasts based on last observation value. As aforementioned, despite removing outliers and replacing them by less insignificant values, trend and seasonality still exist. Therefore, 2 methods with ETS order (Error, Trend, Seasonality), from SES method, are adopted. Regarding trend, first model is AAN, where first A denotes additive errors term with being assumed to be normally distributed white noise (Hyndman&Athanasopoulos 2018). Trend is suspected to exist as prices rose from $1181.227US to $2398.207US between 1992-2022; thus, second A denotes trend with additive values meaning gradually upward linear trendline is expected. Seasonality is found (Figure.18); yet, seasonal patterns' variance is small and insignificant, compared to overall values; thus, seasonality gets no (N). Similar reasons explain ANN model, where middle N denotes no trend. Period 1992-2010 witnesses rising-trend while 2011-2022 experiences decreasing-trend. However, 2011-2022 decreasing level was not significant enough to locate below 1992-2010 rising-trendline. Therefore, 1992-2022 rising-trend is suspected to be denied, which contradicts previous trend assumption in AAN.

Conclusively, 2 trend assumptions are induced, thereby combining AAN and ANN (combination) to leverage and compensate each model's strengths and limitations. Here, combination is recommended to forecast future aluminum prices.

```
FA <- A_train |>
  model(
    Mean = MEAN(value),
    Naive = NAIVE(value),
    ANN = ETS(value ~ error("A") + trend("N") + season("N")),
    AAN = ETS(value ~ error("A") + trend("A") + season("N"))
  ) |>
  mutate(combination = (ANN + AAN) / 2)
```

***Figure.20.Five models&methods***

```
> accuracy(FCA, A_fill) |> arrange(MASE)
# A tibble: 5 × 10
  .model  .type     ME  RMSE   MAE   MPE  MAPE  MASE RMSSE
  <chr>   <chr>  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 ANN     Test   -75.1  311.  295. -4.23  11.6 0.994 0.792
2 Naive   Test   -75.1  311.  295. -4.23  11.6 0.994 0.792
3 combin… Test   -77.4  313.  297. -4.32  11.7 0.999 0.796
4 AAN     Test   -79.7  314.  298. -4.41  11.7 1.00  0.801
5 Mean    Test   819.   873.  819. 30.4   30.4 2.76  2.23
```

***Figure.21.Forecast_accuracy***

```
> fitu <- A_fill |>
+   model(ANN = ETS(value ~ error("A") + trend("N") + season("N")))
> report(fitu)
Series: value
Model: ETS(A,N,N)
  Smoothing parameters:
    alpha = 0.9998998

  Initial states:
     l[0]
 1181.288

  sigma^2:  8388.431

     AIC     AICc      BIC
5566.698 5566.763 5578.454
```

***Figure.22.ANN_alpha***

To improve forecast accuracy, 5 methods and models are applied to forecast aluminum prices between Jan/2022-Dec/2022(Figure.20).

RMSE is chosen to estimate mean error of total projected values(Figure.21). RMSE places more emphasis on errors having larger absolute values as it squares errors before averaging and root squaring, magnifying larger errors impact in overall metric (Chai&Draxler2014). As dataset has some large errors despite having replaced outliers by less significant values, RMSE is optimal to check accuracy.

ANN and Naïve have similar and lowest forecast errors, inducing optimal forecast accuracy among other methods. Both Naïve and ANN do not let past prices intervene Jan/22-Dec/22 forecast. However, regarding Naïve, all monthly prices forecast during 2022 will equal final observation value in train data, given alpha 1. However, ANN imposes weights on final observation value and introduces more varied constraints on recent observations compared to Naïve. This means Naïve still let few past prices receive externals event effect intervene Jan/22-Dec/22 forecast with very small weights, with alpha 0.999 (Figure.22). With such minor disparity, ANN may slightly outperform Naïve. Moreover, with RMSE at 311 (Figure.21), although ANN maybe the best model, gap between forecast values and actual values still can be large.

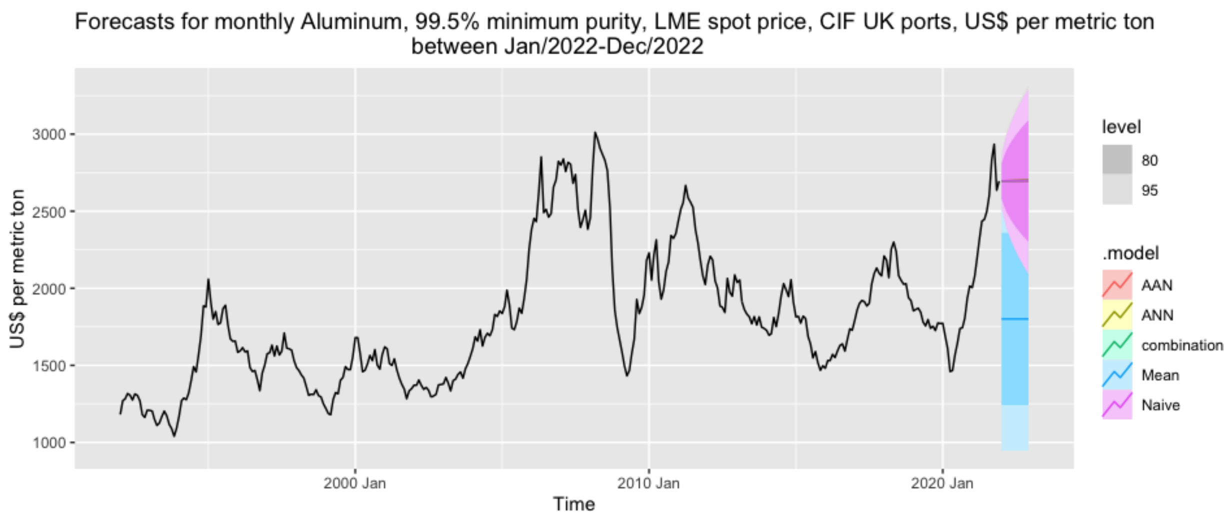


***Figure.23.Forecast Jan/2022-Dec/2022***

During Jan/2022-Dec/2022, Mean forecasts prices will be around $1750US, disqualifying Mean, as projected prices should increase instead of decreasing to that level because demand was rising amid limited supply during 2022(S&PGlobal 2023). Differently, combination forecast highest values for price increasing at nearly $2750US, given green line stay at the top. ANN and Naïve overlap; thus, only pink line appears. ANN and Naïve predict aluminum increases at nearly $2650US compared to combination. Given Russia-Ukraine war still constraints on supply-chain and delivery costs, but such constraints are minorly eased, so prices should continue increasing during 2022 instead of decreasing to level Mean forecasts' or increasing to level combination predicts. Aforementioned-arguments to support ANN selection compared to Naïve add on final decision on ANN even when observing Figure.24.

### c) Using 2022 monthly predicted values and actual values

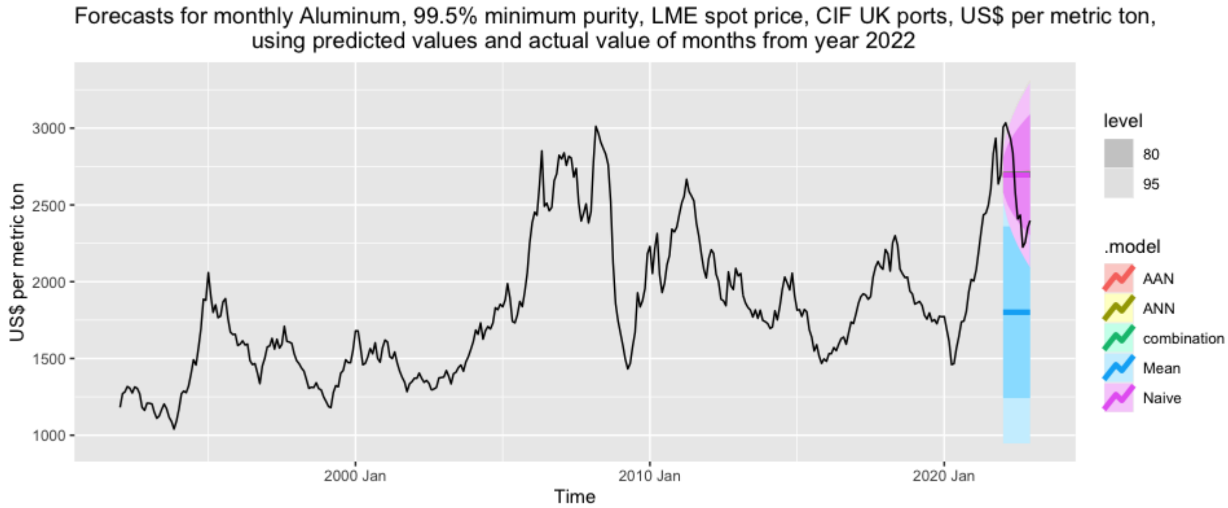


***Figure.24.Accuracy forecast errors***

When using AAN, ANN, combination, Naïve, Mean to predict 2022 monthly temperatures, significant gaps between predicted-values and actual-values arise as above-expected. ANN, Naïve give forecast values most close to actual-values, reflecting lower forecast-errors. ANN adds less bias to final observation to base on that to forecast future-values, compared to Naïve. Conclusively, ANN outperforms.

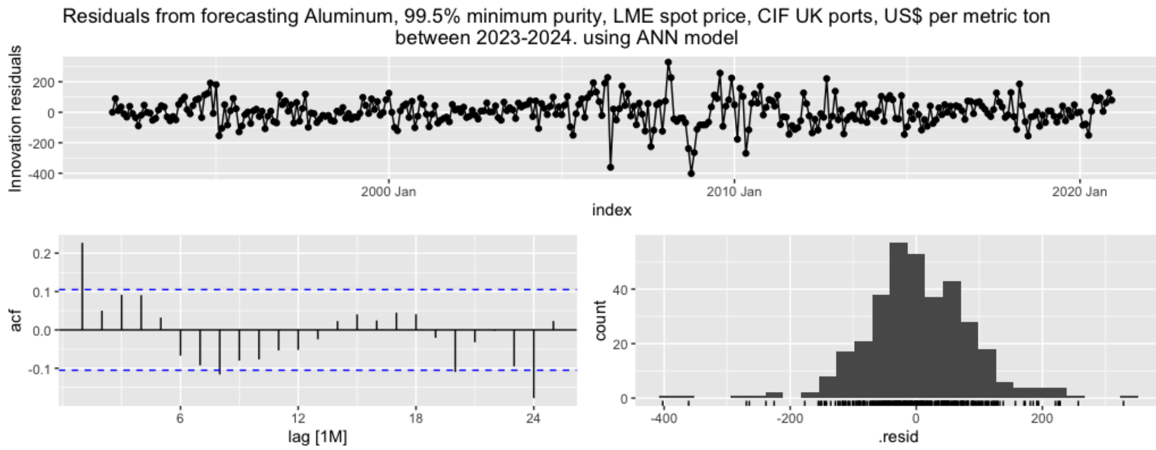


***Figure.25.ANN_residuals***

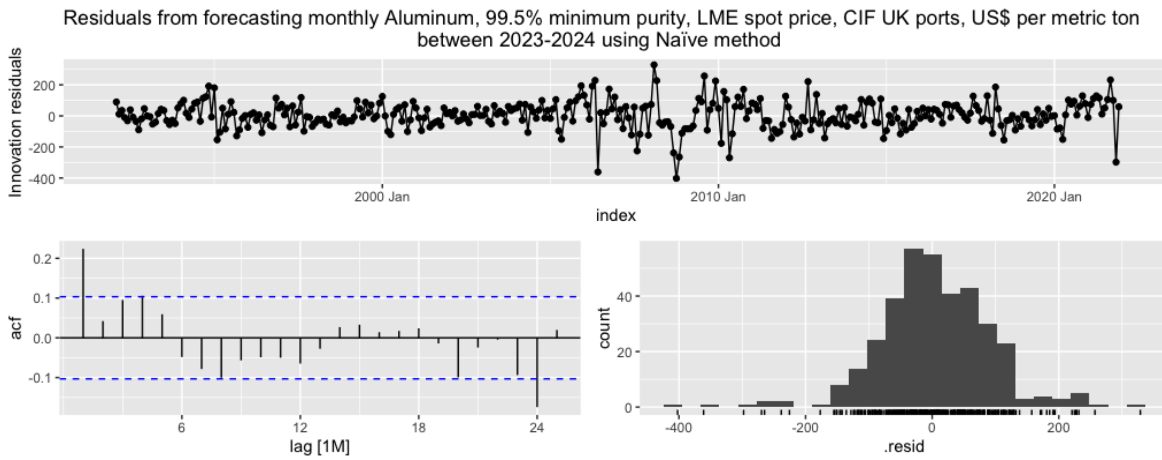


***Figure.26.Naïve_residuals***

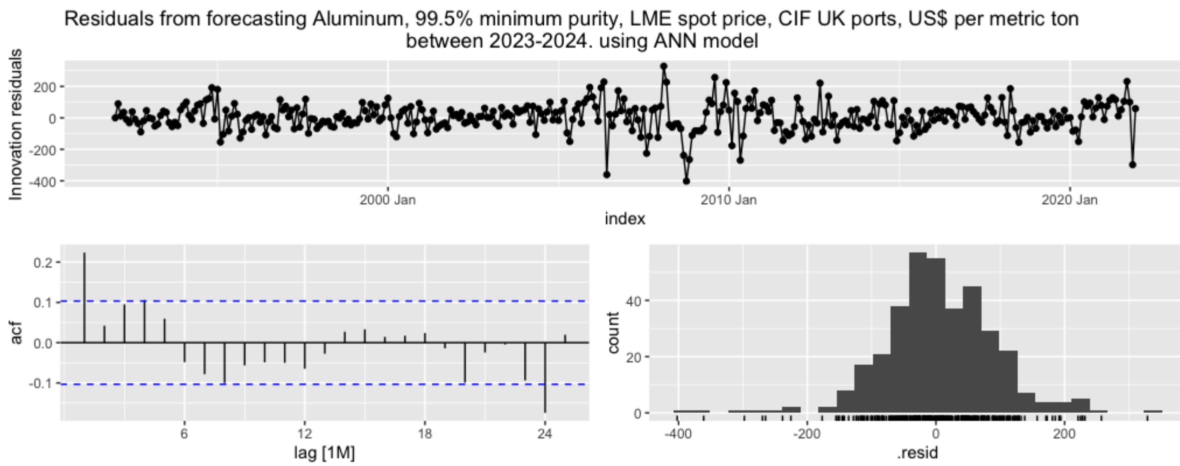


***Figure.27.AAN_residuals***

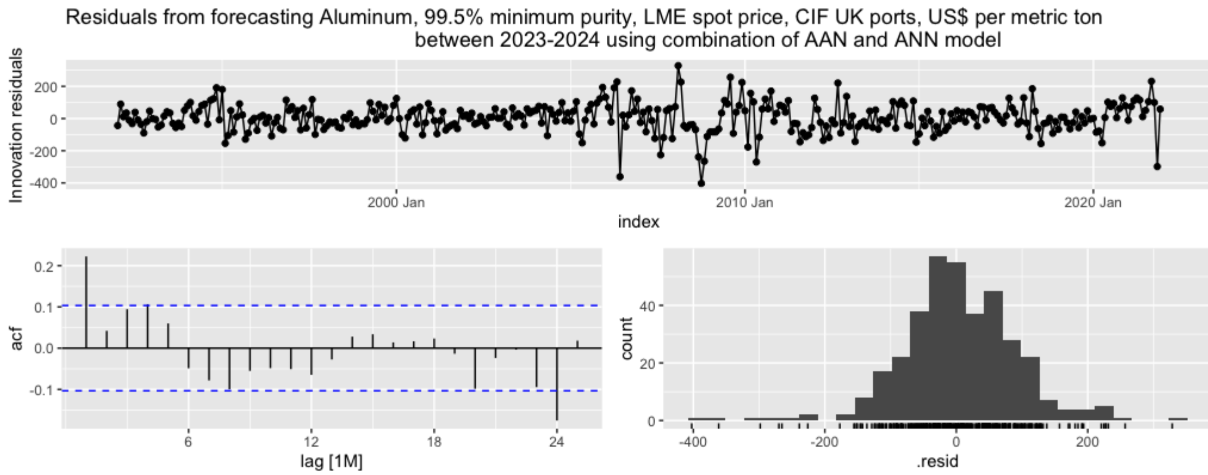


*Figure.28.Combination of AAN&ANN_residuals*

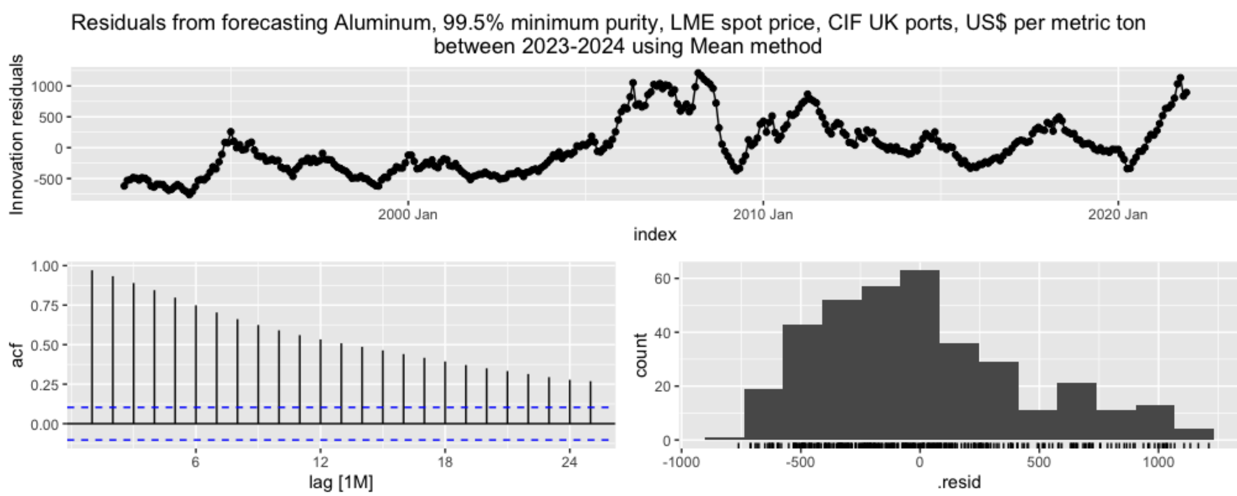


*Figure.29.Mean_residuals*

Five methods and models' residuals, or forecast-errors, are depicted. Mean's ACF shows positive lags tail off and are significant to greatly exceed bounding-lines, signaling neglected trend during forecasting, implying no white noise . Moreover, AAN, ANN, combination, Naïve almost have white noise as only lags 1 and 24 exceed bounding-lines. However, overall, they miss no accounted-for seasonal-patterns and trend by model or strong autoorrelations. With reasons afore-discussed, ANN is still chosen to forecast future prices.

### d) 2024 monthly prices forecast

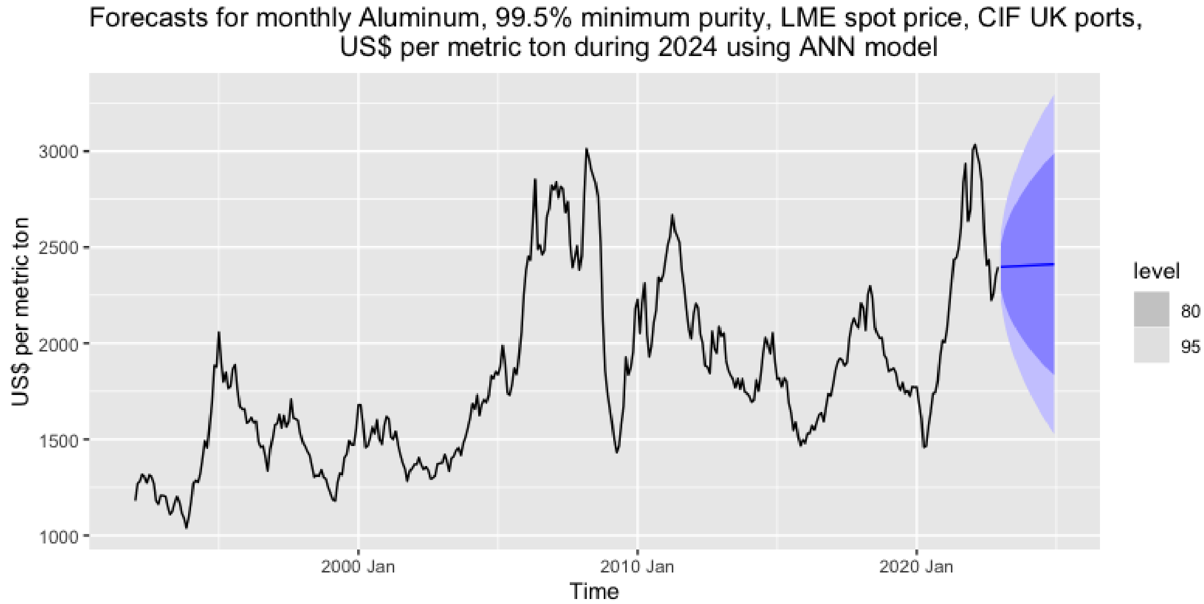


***Figure.30.Monthly aluminum forecast during 2024 using ANN***

To forecast 2024's monthly aluminum prices, ANN model is adopted based on above-mentioned reasons and checking-processes.

Aluminum price reached its peak in Sep/2022 then rebounded till end/2022. Russia-Ukraine effect remains until in 2024, still constraining supply chain (Williams 2023). This is because Russia, 6th largest aluminum exporter, decreases quantity supplied owing to sanctions while other greater China, Canada, India exporters can fill this diminution (Statista 2023). Thus, supply will not change much amid aluminum demand still rising thanks to construction, industrial, and renewable energy demand (Worldbank2023). If the Russia-Ukraine war or unexpected decreased aluminum-production capacity worsens beyond anticipation, aluminum prices may fluctuate beyond ANN's anticipated range, which is blue prediction interval in Figure.30.

Conclusively, aluminum price may increase from $2398,207US in Dec/2022, to the range within $2400US-$2480US regarding monthly prices for Jan/2024-Dec/2024. However, there are still chances that aluminum price will vary between $1500US-$3250US.

## IV. CONCLUSION

### 1. Summary

In part.A, median of 9.185 is chosen to represent Germany's entire monthly temperatures between 1991-2021. Moreover, ARIMA/SRIMA models are presented to predict 2022-2032 Germany's temperatures. Dataset has different historical repeated seasonal patterns, so average every season's temperatures will be different. Correspondingly, seasonal differencing is applied to make those average temperatures become a constant and similar number throughout 1991-2021. Auto model is chosen to finish predicting thanks to lowest AICc, BIC. Germany's 2022-2032 temperatures still have repeated seasonal pattern. However, winter will get warmer, given reduced German coal demand to warm up during winter.

In part.B, nine extreme prices are detected to eliminate based on leftover values having no trend or seasonality. ANN is chosen to predict 2024 monthly aluminum prices because ANN's has least forecasting errors, measured by RMSE. Moreover, ANN's predicted monthly values are nearest to actual monthly values of 2022, validating ANN final selection. Thus, aluminum price is forecast to slightly increase from $2398,207US in Dec/2022, to the range within $2400US-$2480US regarding monthly values for Jan/2024-Dec/2024. If unexpected even worsens than projected, monthly aluminum price can vary between $1500US-$3250US during Jan/2024-Dec/2024.

### 2. Policy

As earlier-mentioned, aluminum importers with moderate or limited base metal production, including UK, New Zealand, South Africa, have better capabilities to forecast aluminum prices than exporters (Pincheira&Hardy2021). However, Russia-Ukraine war, known as instability, will add more errors when forecasting important commodity price like aluminum (Millen 2022). Therefore, National stockpile from such mentioned aluminum-importers should utilize currently moderate price to stockpile more before supply chain restrictions resulting from Russia-Ukraine war once again drive up prices to very high. This can supplement state and local supplies when crisis occurs and mitigate construction and industrial stagnation when unexpected events explode (IEA 2022).

## 3. Limitations

In Part.A, limitation is choosing median as value to represent whole Germany's monthly mean temperatures during 1992-2022. Seasonality is found, implying middle temperatures in warm-seasons are different from cold-seasons. So, choosing best representative temperature value of whole 1992-2022 monthly temperatures becomes unreasonable. However, this step is still applied to help client have snapshot of Germany's temperatures.

Moreover, average temperature (Mean) should be selected to embody whole monthly mean temperatures between 1992-2022, instead of Median. Median takes only average of 2 middle values in categorical order and neglect other seasonal-patterns. Nevertheless, Mean averages all dataset's values, considering all different average point temperature values at different seasonal-periods, implying better representive temperature. Best way to solve this limitation is plotting time series to identify seasonality existence before finding best representative value.

In Part.B, Aluminum prices seem to volatile as there were many upward and downward patterns throughout 1992-2022 with no fixed consitent upward trend. SES helps reduce initial data volatility but lacks capability to address recent fluctuations while SES add more weights to most recent data and base on that to predict (Li et al.2019). Thus, with large sample-dataset like 372 aluminum prices observations, ARIMA can be alternative to better fit data fluctuations, predicting more accurately future aluminum prices (Li et al.2019).

## V. REFERENCES


Basri KI and Sumitra ID (2019) "Comparison of Forecasting the Number of Outpatients Visitors Based on Naïve Method and Exponential Smoothing", *IOP Conference Series: Materials Science and Engineering*, [Online] 662 (4), 42002–.

BW (Better Where) (2023) *What Are the Warmest Cities in Germany?*, betterwhere website, accessed 22 May 2023. https://www.betterwhere.com/warmest-cities-in-germany/

Chai T and Draxler RR (2014) "Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature", *Geoscientific model development*, [Online] 7(3):1247–1250.

Chikkodi CM and Satyaprasad BG (2010) *Business Statistics*, Rev. ed. Mumbai: Himalaya Pub. House, 2010.

CLV (Cedar Lake Venture) (2023) *Climate and Average Weather Year Round in Schleswig*, weatherspark, accessed 22 May 2023. https://weatherspark.com/y/65063/Average-Weather-in-Schleswig-Germany-Year-Round

Copernicus (2023) *Climate change: Europe emerges from second warmest winter on record*, climate.copernicus website, accessed 24 May 2023. https://climate.copernicus.eu/

CT (2023) *Climate – Germany*, Climatestotravel website, accessed 15 May 2023. https://www.climatestotravel.com/climate/germany

DWD (German Weather Service) (2022) *'Clear indication of climate change': Germany logs warmest year on record*, thelocal.de, accessed 21 May 2023. https://www.thelocal.de/20221130/clear-indication-of-climate-change-germany-logs-warmest-year-on-record

Europa (2023) *Climate change: the greenhouse gases causing global warming*, europarl.europa.eu, accessed 15 May 2023. https://www.europarl.europa.eu/news/en/headlines/society/20230316STO77629/climate-change-the-greenhouse-gases-causing-global-warming#:~:text=Greenhouse%20gases%20act%20similarly%20to,be%2C%20supporting%20life%20on%20Earth

GFW (Global Forest Watch) (n.d) Germany, Global Forest Watch website, accessed 15 May 2023. https://www.globalforestwatch.org/dashboards/country/DEU/?category=summary&location=WyJjb3VudHJ5IiwiREVVIl0%3D&map=eyJjYW5Cb3VuZCI6ZmFsc2UsImRhdGFzZXRzIjpbeyJkYXRhc2V0IjoicG9saXRpY2FsLWJvdW5kYXJpZXMiLCJsYXllcnMiOlsiZGlzcHV0ZWQtcG9saXRpY2FsLWJvdW5kYXJpZXMiLCJwb2xpdGljYWwtYm91bmRhcmllcyJdLCJib3VuZGFyeSI6dHJ1ZSwib3BhY2l0eSI6MSwidmlzaWJpbGl0eSI6dHJ1ZX0seyJkYXRhc2V0IjoiTmV0LUNoYW5nZS1TVEFHSU5HIiwibGF5ZXJzIjpbImZvcmVzdC1uZXQtY2hhbmdlIl0sIm9wYWNpdHkiOjEsInZpc2liaWxpdHkiOnRydWUsInBhcmFtcyI6eyJ2aXNpYmlsaXR5Ijp0cnVlLCJhZG1fbGV2ZWwiOiJhZG0wIn19XX0%3D&showMap=true
Gössling S and Humpe A (2020) "The global scale, distribution and growth of aviation: Implications for climate change", *Global environmental change*, 65102194–102194.

Hyndman RJ and Athanasopoulos G (2021) *Forecasting principles and practice*, 3rd edition, Otexts, accessed 20 April 2023. https://otexts.com/fpp3/wn.html

IEA (2022) *International Resource Strategy - National stockpiling system*, IEA website, accessed 23 May 2023. https://www.iea.org/policies/16639-international-resource-strategy-national-stockpiling-system

IMF (2022) *Aluminum Price Forecast: 2021, 2022 and Long Term to 2035*, knoema website, accessed 21 May 2023. https://knoema.com/infographics/ffzioof/aluminum-price-forecast-2021-2022-and-long-term-to-2035

JCG (2023) *Heat in the inner city: a plea for more trees and shade Messner: "We are not at the mercy of the urban heat island effect"*, umweltbundesamt.de, accessed 15 May 2023. https://www.umweltbundesamt.de/en/press/pressinformation/heat-in-the-inner-city-a-plea-for-more-trees-shade#:~:text=This%20phenomenon%20of%20significantly%20higher,city%20neighbourhoods%20and%20their%20buildings

Kallgren J (2022) *Even the trees with deep roots are starting to react to the lack of water*, Euronews website, accessed 15 May 2023.

Li C et al. (2019) "A new method to mitigate data fluctuations for time series prediction", *Applied Mathematical Modelling*, [Online] 65390–407.

Masson V, Lemonsu A, Hidalgo J and Voogt J (2020) "Urban Climates and Climate Change", *Annual Review of Environment and Resources*, 45:411–44. https://doi.org/10.1146/annurev-environ-012320- 083623

Millen RA (2022) *Stability Challenges and Opportunities Regarding the Russo-Ukrainian War*, marshallcenter.org, accessed 24 May 2023. https://www.marshallcenter.org/en/publications/perspectives/stability-challenges-and-opportunities-regarding-russo-ukrainian-war-0

Mitchell TA (2023) *GERMANY WEATHER AND CLIMATE*, internationalliving website, accessed 22 May 2023. https://internationalliving.com/countries/germany/germany-weather-and-climate/#:~:text=Southwestern,-%C2%A9iStock%2FDaLiu&text=The%20southwestern%20region%20is%20comprised,warmest%20part%20of%20the%20country

Mücke HG and Litvinovitch, JM (2020) "Heat Extremes, Public Health Impacts, and Adaptation Policy in Germany", *International journal of environmental research and public health*, [Online] 17 (21), 7862–.

OE (Ocean Explorer) (2023) *How does the ocean affect climate and weather on land?*, oceanexplorer.noaa.gov, accessed 15 May 2023. https://oceanexplorer.noaa.gov/facts/climate.html

Pincheira P and Hardy N (2021) “Forecasting aluminum prices with commodity currencies”, *Resources policy*, [Online] 73102066–.

Rooney G, Lipzig NV and Thiery W (2018) “Estimating the effect of rainfall on the surface temperature of a tropical lake”, *Hydrology and Earth System Sciences*, 22(12):6357–6369. https://doi.org/10.5194/hess-22-6357-2018

S&P Global (2023) *Drought hits hydropower supplies for Chinese aluminum smelting hub*, hydroreview website, accessed 21 May 2023. https://www.hydroreview.com/environmental/drought-hits-hydropower-supplies-for-chinese-aluminum-smelting-hub/#gref

Science (2023) *FOREST FIGHT*, science.org, accessed 15 May 2023. https://www.science.org/content/article/germany-s-trees-are-dying-fierce-debate-has-broken-out-over-how-respond

Statista (2023) *The world's leading exporters of aluminum and aluminum products in 2021, by country(in billion U.S. dollars)* [dataset], Statista website, accessed 21 May 2023. https://www.statista.com/statistics/1113623/global-aluminum-exports-by-country/#:~:text=China%20is%20the%20world's%20leading,around%2018.8%20billion%20U.S.%20dollars

Williams R (2023) *Opinion: Expect Russia’s war in Ukraine to continue into 2024, with higher prices for oil, gas and defense stocks*, marketwatch website, accessed 24 May 2023. https://www.marketwatch.com/story/expect-russias-war-in-ukraine-to-continue-into-2024-with-higher-prices-for-oil-gas-and-defense-stocks-ffaf4bb6

Winklmayr C, Muthers S, Hildegard Niemann H, Mücke HS and Heiden MA (2022) “Heat-Related Mortality in Germany From 1992 to 2021”, *Dtsch Arztebl Int*, 119(26):451-457. doi: 10.3238/arztebl.m2022.0202.

WMO (2022) *Temperatures in Europe increase more than twice global average*, public.wmo.int, accessed 15 May 2023. https://public.wmo.int/en/media/press-release/temperatures-europe-increase-more-twice-global-average

Worldbank (2023) *Competitiveness of Global Aluminum Supply Chains Under Carbon Pricing Scenarios for Solar PV*, worldbank.org, accessed 23 May 2023. https://www.worldbank.org/en/topic/extractiveindustries/publication/competitiveness-of-global-aluminum-supply-chains-under-carbon-pricing-scenarios-for-solar-pv

Worldbank (2023) *Current Climate - Climatology*, worldbank.org, accessed 22 May 2023.https://climateknowledgeportal.worldbank.org/country/germany/climate-data-historical#:~:text=Germany%20is%20part%20of%20the,zone%20of%20the%20mid%2Dlatitudes.

Zhang H, Wang Y, Chen D, Feng D, You X and Wu W (2022) “Temperature Forecasting Correction Based on Operational GRAPES-3km Model Using Machine Learning Methods”, *Atmosphere. 2022*, 13(2):362. https://doi.org/10.3390/atmos13020362

## VI. APPENDICES

$$SD = s = \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i(x_i - \bar{x})^2}$$

SD = Standard Deviation

$x_i$ = Terms Given in the data

$\bar{x}$ = Mean

n = Total number of terms

***Appendix A. Standard deviation calculation, adapted from Chikkodi and Satyaprasad* (2010)**

$$SV = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$s^2$ = Sample variance

$x_i$ = Terms Given in the data

$\bar{x}$ = Mean

n = Total number of terms

***Appendix B. Sample variance calculation, adapted from Chikkodi and Satyaprasad* (2010)**

Coefficient of Variation (CV) = (Standard Deviation/Mean) × 100

***Appendix C. Coefficient of Variation calculation, adapted from Chikkodi and Satyaprasad* (2010)**

| | .model | index | value | trend | remainder | season_adjust |
|---|---|---|---|---|---|---|
| **1** | STL(value ~ season(period = 1), robust = TRUE) | 2009 Feb | 1.338 | 2.029 | -691 | 1.338 |
| **2** | STL(value ~ season(period = 1), robust = TRUE) | 2009 Mar | 1.338 | 2.002 | -664 | 1.338 |
| **3** | STL(value ~ season(period = 1), robust = TRUE) | 2009 Jan | 1.420 | 2.062 | -642 | 1.420 |
| **4** | STL(value ~ season(period = 1), robust = TRUE) | 2008 Dec | 1.504 | 2.100 | -596 | 1.504 |
| **5** | STL(value ~ season(period = 1), robust = TRUE) | 2008 Jun | 2.968 | 2.374 | 594 | 2.968 |
| **6** | STL(value ~ season(period = 1), robust = TRUE) | 2008 Jul | 3.067 | 2.331 | 736 | 3.067 |
| **7** | STL(value ~ season(period = 1), robust = TRUE) | 2022 Feb | 3.246 | 2.606 | 639 | 3.246 |
| **8** | STL(value ~ season(period = 1), robust = TRUE) | 2022 Apr | 3.247 | 2.610 | 637 | 3.247 |
| **9** | STL(value ~ season(period = 1), robust = TRUE) | 2022 Mar | 3.498 | 2.613 | 885 | 3.498 |

***Appendix D. Outliers detected in remainder after decomposing monthly aluminum, 99.5% minimum purity, LME spot price, CIF UK ports, US$ per metric ton***