TIME SERIES ANALYSIS AND MODELING

Final Term Project

Liang Gao

2023-12-11

Contents

List of Figures

List of Tables

1 Abstract

In this report, I analyzed hourly temperature in Jena, Germany, from 2018 to 2020, using regression, base models(average, Naive, Drift, Simple Exponential Smoothing), the Holt-Winter method, and SARIMA (Seasonal Autoregressive Integrated Moving Average) techniques. The first 80% data are used as the training set, the left 20% are the test set. Detained analyses with each technique are presented. The results show that the SARIMA model has better performance than the base models.

2 Introduction

The objective is to apply and compare various techniques to accurately model and predict temperature in Celsius. This report meticulously details the application of each analytical technique, elucidating the underlying methodologies and their respective suitability for temperature time series analysis. I chose three-year temperature data from the raw dataset. I first performed regression analysis using 21 numerical variables. The base models, the Holt-Winter method, and the SARIMA model purely use the target temperature time analysis dataset, and this report mainly focuses on comparing performance between base models and SARIMA models by comparing the variance of forecast error and mean square of forecast error. The results are particularly noteworthy, revealing that the SARIMA model demonstrates better performance in comparison to the base models.

3 Dataset Description

The raw dataset, the Jena Weather dataset was recorded every 10 minutes and covers data from January 1st, 2004 to December 31st, 2020. I chose the data from January 1st, 2018 to December 31st, 2021 (3-Year), and averaged the data in each hour. The dataset has 22 columns including "date" and 21 numerical variables [1].

The dependent variable is the temperature in Celsius, the left 20 numerical features are:
p(mbar): atmospheric pressure;
Tpot(K): Temperature in Kelvin;
Tdew(degC): Temperature in Celsius relative to humidity;
rh(%): Relative Humidity;
VPmax(mbar): Saturation vapor pressure;
VPact(mbar): Vapor pressure;
VPdef(mbar): Vapor pressure deficit;
sh(g/kg): Specific humidity(The mass of water vapor per unit of wet air mass);
H2OC(mmol/mol): Water vapor concentration;
rho(g/m**3): Airtight;
max.wv(m/s): Maximum wind speed;
wd(deg): Wind direction in degrees;
rain(mm): The total rainfall depth during a given period, expressed in millimeters (mm);
raining(s): total time of rain;
SWDR(W/m^2): Irradiance of shortwave solar radiation;
PAR($\mu mol/m^2/s$): Photo-synthetically active radiation;
max.PAR($\mu mol/m^2/s$);
Tlog(degC): Temperature log;
CO2(ppm).

I used the drift method to fill in the missing values and changed column names for better identification. After the preliminary clean, the total observation is 17,545. I checked the histogram plots and statistical

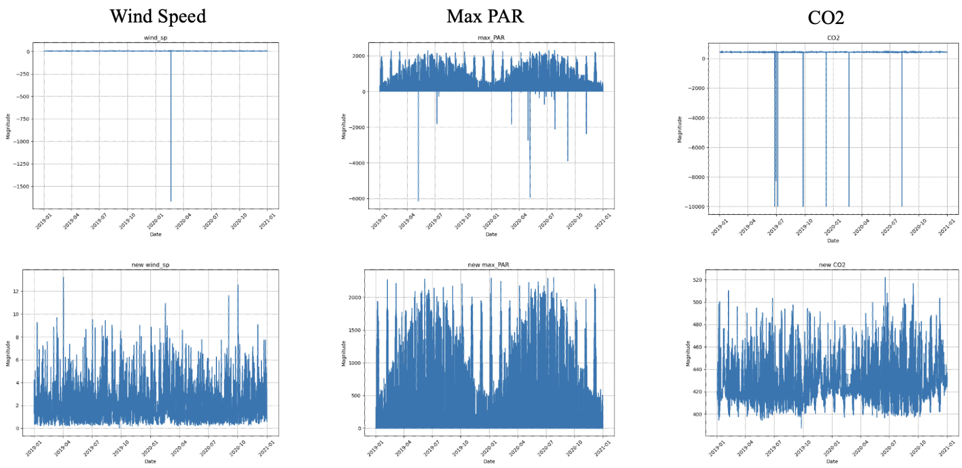

Figure 1: Outlier check

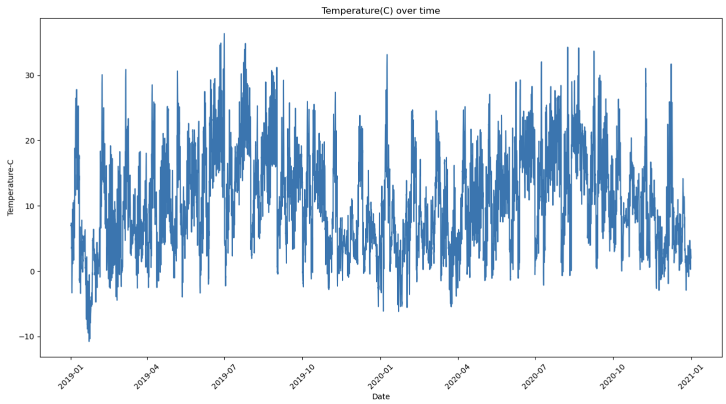

Figure 2: Temperature over time

information for all independent variables. Only three variables, wind speed, max PAR, and CO2, have negative outliers. Figure 1 shows histograms before and after cleaning outliers. The average method was used to fix outliers in wind speed and CO2 and the naive method was used for max PAR.

Figure 2 is the plot of the temperature in Celsius versus time, which shows obvious trends and seasonality.

The seasonality order is 24 from the ACF plot (Figure 3), which is consistent with the characteristics of hourly temperature data.

Figure 4 is the ACF & PACF plot of the dependent variable. The ACF tails off and the PACF cuts off, which indicates the AR process.

Figure 5 is the correlation matrix with Pearson's correlation coefficient. Some variables have a high correlation with the target variable, such as temperature in Kelvin(1), air density(-0.95), and max vapor pressure(0.96). There is also a high correlation between some independent variables, for instance, the correlation coefficient between SWDR and PAR is 1.

The observation is 14,036 in the train set(80%) and 3,509 in the test set(20%).

4 Stationarity

The target variable passes the ADF test with a p-value of 0.00 but fails to pass the KPSS test with a p-value of 0.02(the p-value threshold is 0.05 for both tests). Figure 6 shows the rolling mean and variance of temperature in Celsius, which stabilize once all samples are included. Thus, the dataset is

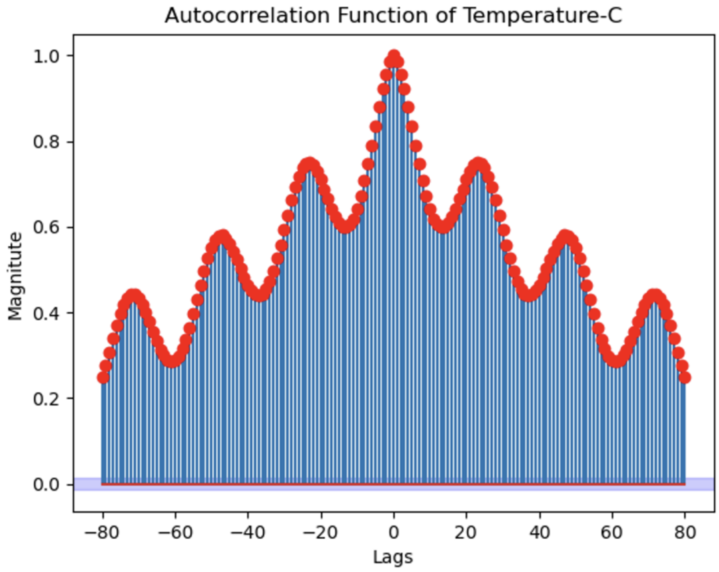

Figure 3: ACF

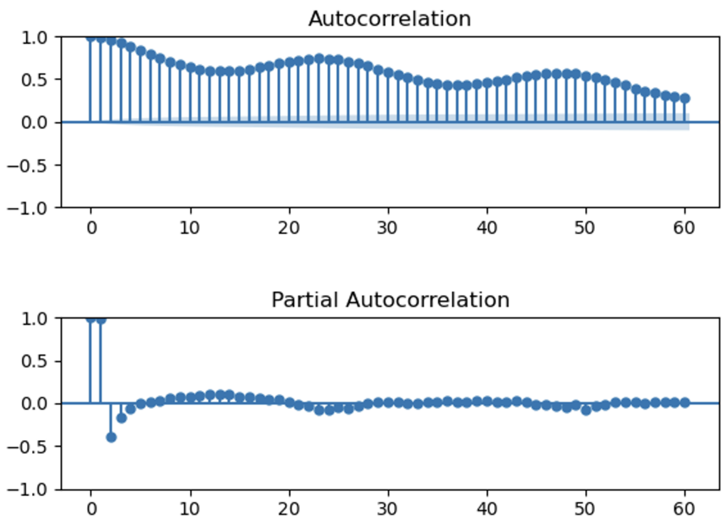

Figure 4: ACF & PACF

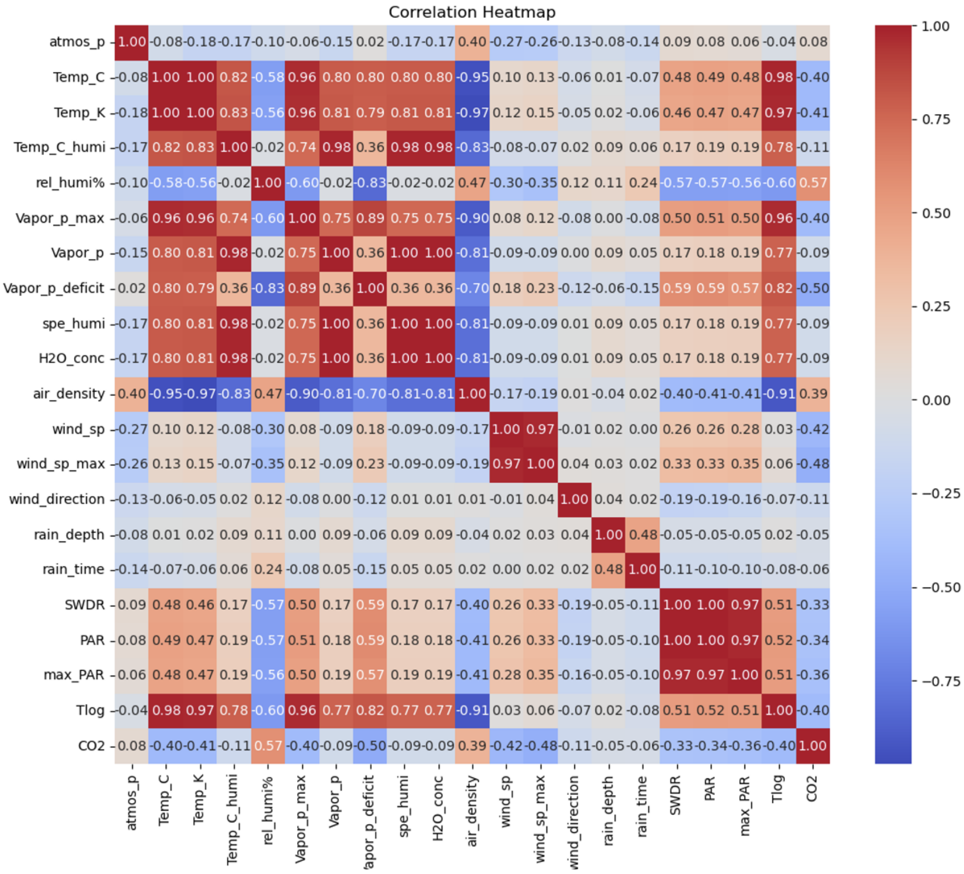

Figure 5: Correlation Matrix

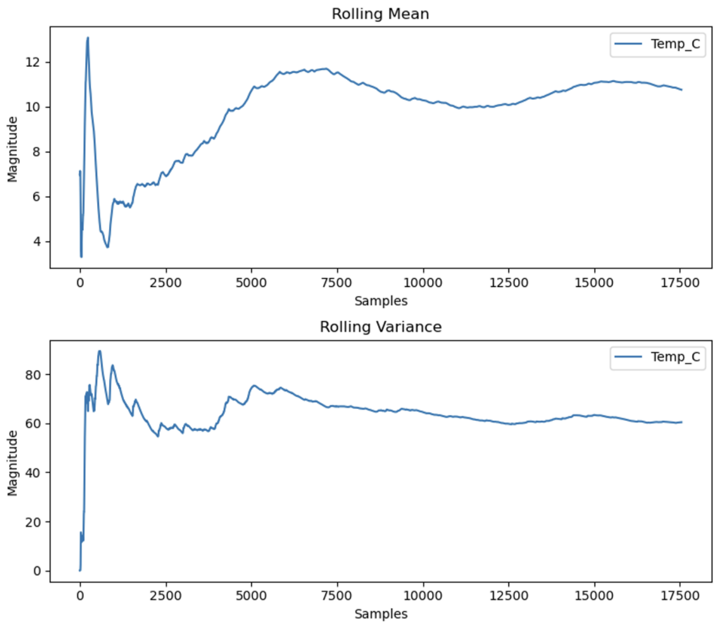

Figure 6: Rolling Mean & Variance

weak-stationary.

5 Time Series Decomposition

Figure 7 is the original, trend, season, and residual data by using the STL method. Figure 8 is the original dataset and de-trend and seasonality-adjusted dataset. The strength of the trend is 94.37%, and the strength of the seasonality is 74.79%.

6 Holt-Winters Method

The Holt-Winters method considers the level, trend, and seasonality of time series data. Figure 9 is the result of the Holt-Winters method. I used 744 (monthly seasonality) as the seasonal period. This method captures most seasonality but not the trend.

7 Feature Selection / Dimensionality Reduction

I normalized the data before running the regression model.

Figure 10 shows all the singular values are greater than 0, but the last few singular values are relatively small compared to the first largest one. The condition number is 1,409,780.69 and is highly greater than 1,000. Both results from singular value and condition number indicate severe co-linearity among some independent variables.

Figure 11 is the calculated variance ratio from the PCA algorithm. The threshold for the PCA feature selection is a variance ratio of less than 0.95. I finally chose 7 features and fitted them into the regression model. The OLS regression model result is in Figure 12. The adjusted R-squared is 0.982, and all the coefficients are statistically significant with p-values less than 0.05. The mean of error is 0.005, the variance of error is 0.017, and the MSE is 0.017. The ACF/PACF for the residual is in Figure 14. The ACF tails off and PACF cuts off. The ACF shows seasonality with order 24, which is the same as the raw target variable. The model's performance is good(Figure 13), but some dynamic information is not captured by the model.

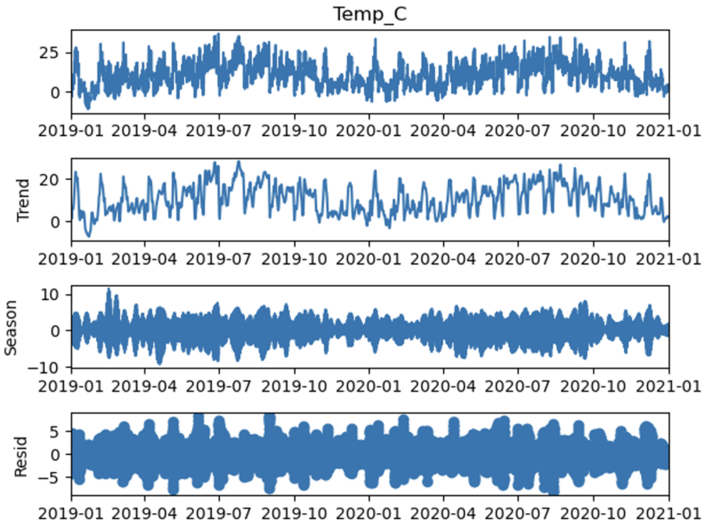

Figure 7: STL method

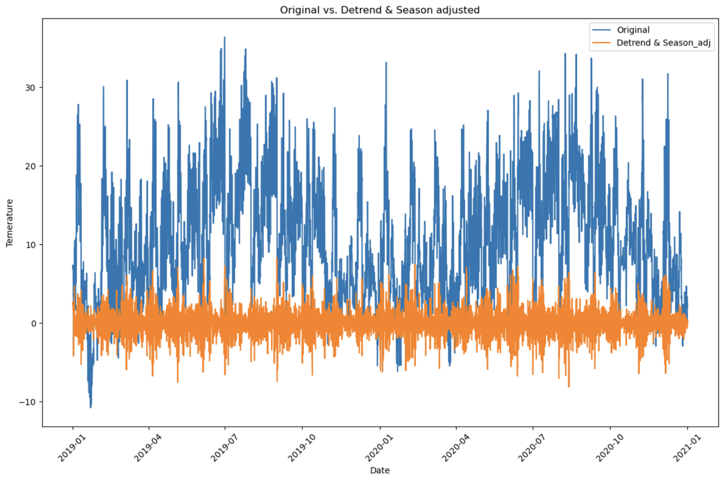

Figure 8: De-trend & Seasonality adjusted data

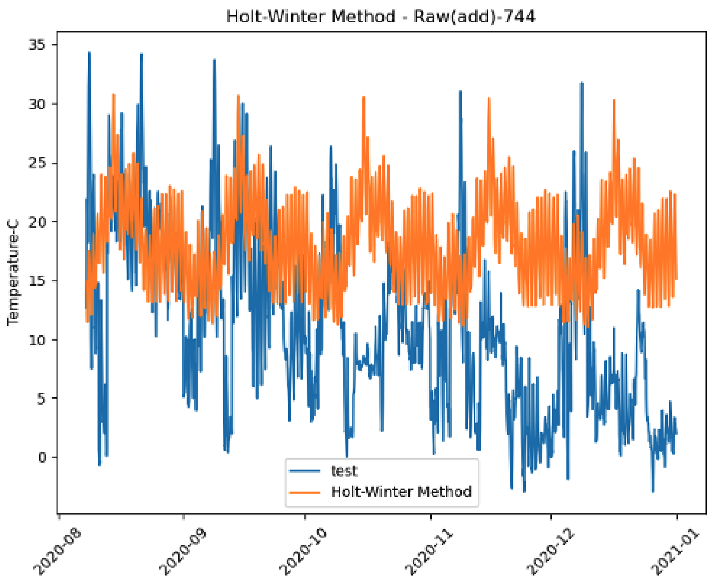

Figure 9: Holt-Winters Method

```
([2.2455168e+05, 8.6037250e+04, 1.3384760e+04, 1.0475040e+04,
  8.4971700e+03, 3.1314400e+03, 2.4113800e+03, 1.3748900e+03,
  8.9651000e+02, 3.9994000e+02, 3.2921000e+02, 2.0236000e+02,
  1.4439000e+02, 5.4280000e+01, 3.5670000e+01, 8.6300000e+00,
  7.7800000e+00, 2.3000000e-01, 1.6000000e-01])
```

Figure 10: Singular values

```
[0.41454358 0.61821955 0.72824184 0.79617275 0.85165992 0.90382252
 0.94367239 0.96735744 0.98691004 0.99324168 0.99589509 0.99785777
 0.99900996 0.99985955 0.99996014 0.99999789 0.99999998 1.
 1.        ]
```

Figure 11: Variance ratio (PCA)

```
                                 OLS Regression Results
=======================================================================================
Dep. Variable:                      y   R-squared (uncentered):                   0.982
Model:                            OLS   Adj. R-squared (uncentered):              0.982
Method:                 Least Squares   F-statistic:                          1.082e+05
Date:                Tue, 05 Dec 2023   Prob (F-statistic):                        0.00
Time:                        21:05:20   Log-Likelihood:                          8067.7
No. Observations:               14036   AIC:                                 -1.612e+04
Df Residuals:                   14029   BIC:                                 -1.607e+04
Df Model:                           7
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1             0.3484      0.000    844.696      0.000       0.348       0.349
x2            -0.0773      0.001   -131.394      0.000      -0.078      -0.076
x3            -0.0249      0.001    -31.090      0.000      -0.026      -0.023
x4            -0.1426      0.001   -139.991      0.000      -0.145      -0.141
x5             0.0363      0.001     32.196      0.000       0.034       0.039
x6             0.0813      0.001     69.928      0.000       0.079       0.084
x7            -0.0210      0.001    -15.767      0.000      -0.024      -0.018
==============================================================================
Omnibus:                     1927.180   Durbin-Watson:                   0.111
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             4810.567
Skew:                          -0.782   Prob(JB):                         0.00
Kurtosis:                       5.404   Cond. No.                         3.23
```

Figure 12: OLS model (PCA)

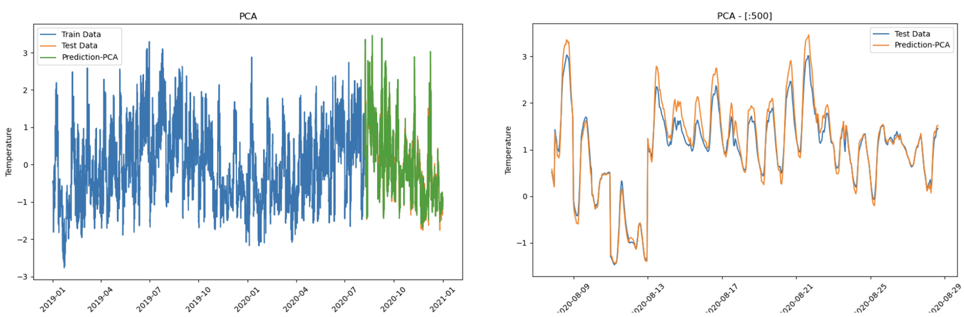

Figure 13: PCA

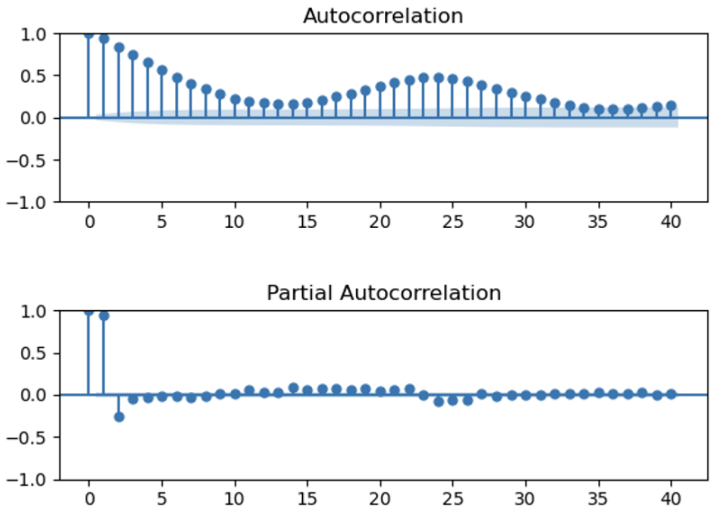

Figure 14: PCA residual

Remove	p-value	Adj_R2
\	\	1.00
PAR	0.79	1.00
Vapor_p_max	0.33	1.00
CO2	0.09	1.00

Figure 15: Insignificant Features

For the Backwards stepwise regression, I started with the model containing all independent variables, removed one predictor with the highest p-value at a time(deleted 3 features - Figure 15), and stopped when the left predictors were all significant. Then I found the confidence interval for some coefficients is very small. For example, the confidence interval for "rain_time" is [-0.001, -0.000]. I decided to remove those variables(Figure 16). 8 features are left and the OLS result is in Figure 17. The adjusted R-squared is 1.000, and all the coefficients are statistically significant with p-values less than 0.05. The plot for model performance is in Figure 18. The mean of error is 0.001, and the variance of error and the MSE are less than 0.00001. However, the condition number of the regression model is 4.1e+03, which indicates strong multi-collinearity or other numerical problems. From Figure 19, the ACF tails off and PACF cuts off. I think the multi-collinearity is one of the reasons for colored residuals.

The threshold for the VIF value is 10. Again, I started with the model containing all independent variables, removed one predictor with the highest VIF value at a time(deleted 9 features - Figure 20), and stopped when the left predictors had VIF values less than 10. Then I deleted one feature with a

wind_sp	0.0014	0.000	5.082	0.000	0.001	0.002
wind_sp_max	-0.0010	0.000	-3.266	0.001	-0.002	-0.000
wind_direction	0.0001	6.3e-05	2.029	0.043	4.31e-06	0.000
rain_depth	-0.0002	8.26e-05	-2.531	0.011	-0.000	-4.72e-05
rain_time	-0.0004	7.43e-05	-4.996	0.000	-0.001	-0.000
SWDR	-0.0019	0.000	-7.350	0.000	-0.002	-0.001
max_PAR	0.0019	0.000	7.652	0.000	0.001	0.002
Tlog	0.0100	0.000	24.268	0.000	0.009	0.011

Figure 16: Features with small coefficient

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Temp_C   R-squared:                       1.000
Model:                            OLS   Adj. R-squared:                  1.000
Method:                 Least Squares   F-statistic:                 3.403e+07
Date:                Wed, 06 Dec 2023   Prob (F-statistic):               0.00
Time:                        11:03:38   Log-Likelihood:                 49237.
No. Observations:               14036   AIC:                        -9.846e+04
Df Residuals:                   14027   BIC:                        -9.839e+04
Df Model:                           8
Covariance Type:            nonrobust
==================================================================================
                     coef    std err          t      P>|t|      [0.025      0.975]
----------------------------------------------------------------------------------
constant       -9.586e-05   6.18e-05     -1.552      0.121      -0.000    2.52e-05
atmos_p            0.3260      0.000    971.741      0.000       0.325       0.327
Temp_C_humi       -0.0051      0.001     -4.966      0.000      -0.007      -0.003
rel_humi%         -0.0121      0.000    -30.375      0.000      -0.013      -0.011
Vapor_p           -0.7306      0.008    -87.197      0.000      -0.747      -0.714
Vapor_p_deficit    0.0806      0.000    444.368      0.000       0.080       0.081
spe_humi          -1.8853      0.079    -23.784      0.000      -2.041      -1.730
H2O_conc           2.6512      0.079     33.355      0.000       2.495       2.807
air_density       -0.9840      0.001  -1052.471      0.000      -0.986      -0.982
==============================================================================
Omnibus:                     2645.156   Durbin-Watson:                   0.138
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            72419.607
Skew:                          -0.147   Prob(JB):                         0.00
Kurtosis:                      14.124   Cond. No.                     4.10e+03
```

Figure 17: OLS model(Backwards)

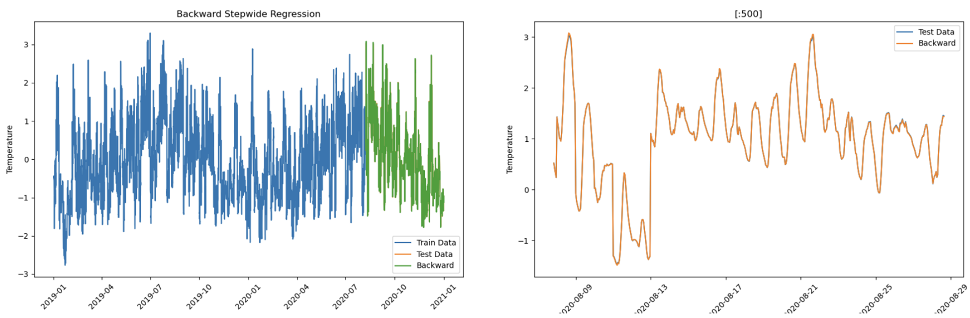

Figure 18: Backwards stepwise regression

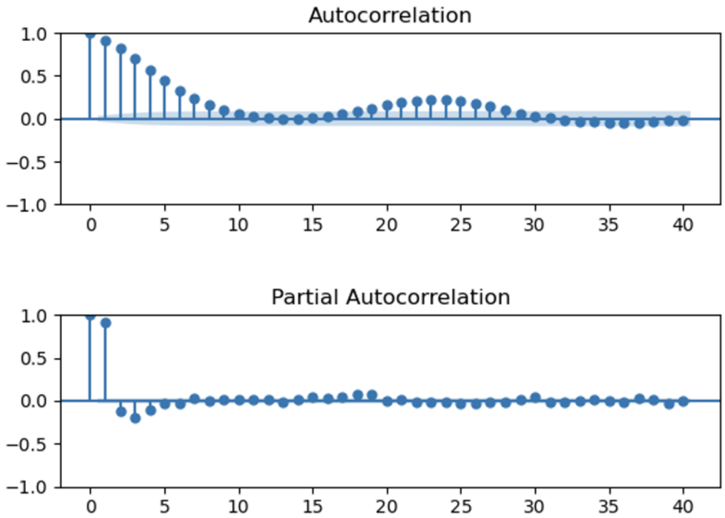

Figure 19: Backwards residual

Method	Mean	Variance	MSE
PCA	0.01	0.02	0.02
Backwards	0.00	0.00	0.00
VIF	0.03	0.03	0.03

Table 1: Regression Residual comparison

p-value of 0.078, which is not significant. I ran the regression model. The adjusted R-squared is 0.971. Some coefficients are less than 0.001. I tried to delete the feature with the lowest coefficient at a time. 6 features were deleted and the adjusted R-squared remains 0.971. So 3 features left and the final model result is in Figure 21. The plot for model performance is in Figure 22. In the second plot (the first 500 observations in the test set), we can see the small difference between forecasts and real test values. The mean of error is 0.031, the variance of error is 0.030 and the MSE is 0.031. The ACF of residuals tails off, and PACF cuts off.(23). The ACF did not show any seasonality information, which means such information is captured by this model.

The residual comparison for the above three feature selection methods is in Table 1.

8 Base Models

The result of h-step prediction based on based models (Average, Naive, Drift, SES) is in Figure 24. The comparison of base models' mean and variance of error, and MSE is in Table 2.

9 Multiple Linear Regression Model

The three regression models in Section 7 all have high adjusted R-squared values. As the model developed in the Backward stepwise regression method has strong multi-collinearity, and the model developed in VIF

Model	Mean	Variance	MSE
Average	0.39	55.47	55.62
Naive	-4.56	55.47	76.23
Dirft	-5.64	61.04	92.82
SES	-4.56	55.47	76.23

Table 2: Base model residual comparison

remove	VIF	Adj_R2
		1.00
Vapor_p_max	14,403,743.53	1.00
H2O_conc	1,664,251.75	1.00
Vapor_p	18,405.61	1.00
PAR	790.68	1.00
air_density	304.34	1.00
Tlog	40.65	1.00
Temp_C_humi	24.89	0.97
wind_sp_max	24.69	0.97
SWDR	19.52	0.97

Figure 20: Features with VIF over 10

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Temp_C   R-squared:                       0.971
Model:                            OLS   Adj. R-squared:                  0.971
Method:                 Least Squares   F-statistic:                 1.565e+05
Date:                Wed, 06 Dec 2023   Prob (F-statistic):               0.00
Time:                        12:58:40   Log-Likelihood:                 4785.9
No. Observations:               14036   AIC:                            -9564.
Df Residuals:                   14032   BIC:                            -9534.
Df Model:                           3
Covariance Type:            nonrobust
==================================================================================
                     coef    std err          t      P>|t|      [0.025      0.975]
----------------------------------------------------------------------------------
constant          -0.0063      0.001     -4.336      0.000      -0.009      -0.003
rel_humi%         -0.3764      0.003   -118.714      0.000      -0.383      -0.370
Vapor_p_deficit    0.2389      0.003     70.797      0.000       0.232       0.246
spe_humi           0.7244      0.002    378.391      0.000       0.721       0.728
==============================================================================
Omnibus:                     3727.518   Durbin-Watson:                   0.054
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             9482.774
Skew:                          -1.451   Prob(JB):                         0.00
Kurtosis:                       5.792   Cond. No.                         4.42
```

Figure 21: OLS model (VIF)

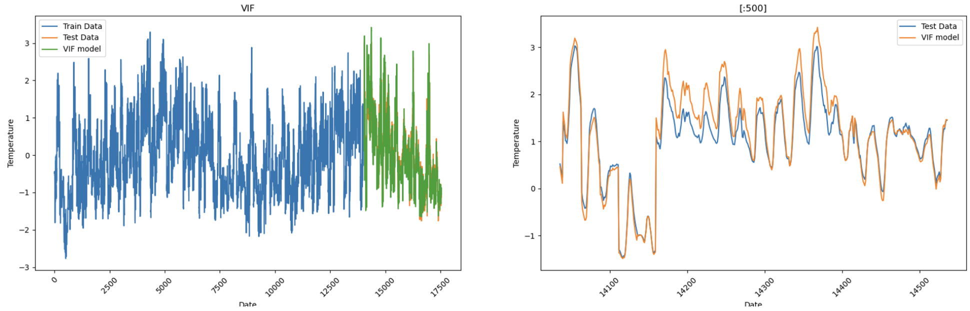

Figure 22: VIF regression

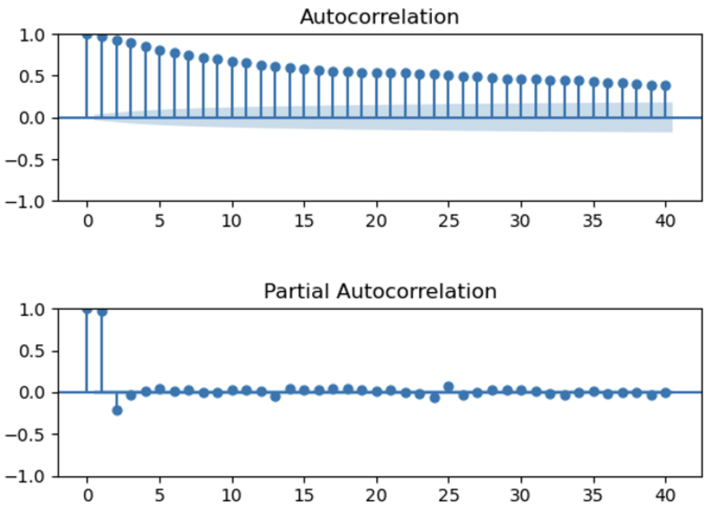

Figure 23: VIF residual

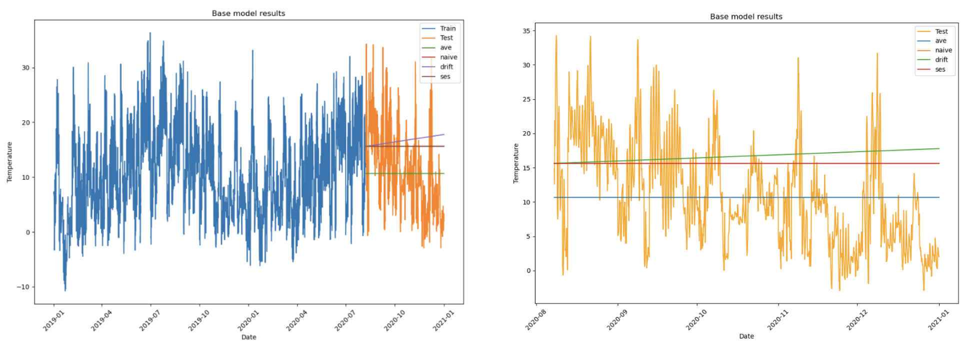

Figure 24: Base models

Subset	MSE	MeanRMSE	R-squared	Adj R-squared
1	0.05	0.22	0.96	0.96
2	0.04	0.19	0.97	0.97
3	0.02	0.15	0.97	0.97
4	0.04	0.19	0.97	0.97
5	0.03	0.16	0.97	0.97

Table 3: Regression Cross Validation

T-test

```
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
c0            -0.0063      0.001     -4.336      0.000      -0.009      -0.003
c1            -0.3764      0.003   -118.714      0.000      -0.383      -0.370
c2             0.2389      0.003     70.797      0.000       0.232       0.246
c3             0.7244      0.002    378.391      0.000       0.721       0.728
==============================================================================
```

F-test

```
F-Test Results:
<F test: F=117357.3159157646, p=0.0, df_denom=1.4e+04, df_num=4>
```

Figure 25: Hypothesis tests: F-test & t-test.

has fewer features(or dimensions) than the model developed in PCA, the final multiple linear regression model is the model developed in the VIF method. The model result is in Figure 21.

The hypothesis tests results including F-test and T-test are in Figure 25. The t-test is used to conduct hypothesis tests on the regression coefficients obtained in multiple regression. The null hypothesis is the coefficient is statistically equal to 0. The alternative hypothesis is not 0. In this case, the coefficients' p-values for all features in the final model chosen are less than 0.05, so we can reject the null hypothesis and support the alternative hypothesis. The F-test compares a model with no predictors (intercept-only) to the model that I specify. The null hypothesis is the fit of the intercept-only model and my model is equal. The alternative hypothesis is the fit of the intercept-only model is significantly reduced compared to my model. The p-value for the F-test here is significantly less than 0.05, so we can reject the null hypothesis and conclude my model provides a better fit than the intercept-only model.

The consistency of the metrics across different folds of cross-validation(Table 3) suggests that the model is stable and generalizes well to different subsets of the data.

For this regression model - MSE: 0.03; RMSE: 0.17; AIC: -9564; BIC: -9534; R-squared: 0.97; Adjusted R-squared: 0.97.

The ACF/PACF of residuals is in Figure 23.

The p-values from the Ljung-Box test(lag equals 20) are 0.0, which is less than 0.05 and indicates a significant auto-correlation that cannot be attributed to chance.

The variance and mean of the residuals are both 0.03.

The plot for the train, test, and predicated values is in Figure 22.

10 SARIMA Model

The temperature dataset is highly trended and seasonal, so I first chose the SARIMA model to perform prediction [2].

I fitted the raw data into GPAC and plotted the ACF/PACF (26). The ACF tails off and PACF cuts off, which indicates the AR process, and with a seasonal pattern at order 24. The GPAC shows a pattern at na equals 1 and nb equals 0 or 1. I tried both nb and then I found nb equals 1 has better performance. The model I fitted is non-seasonal order (1,0,1) and seasonal order (1,0,0,24). The plot for

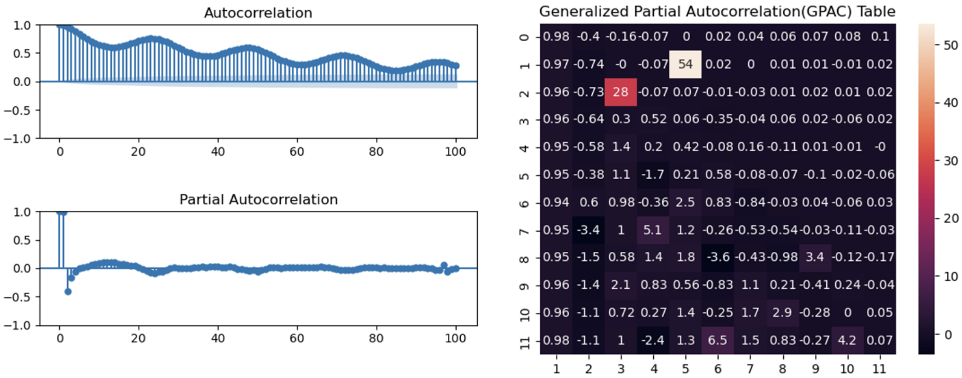

Figure 26: GPAC & ACF/PACF of Raw Dataset

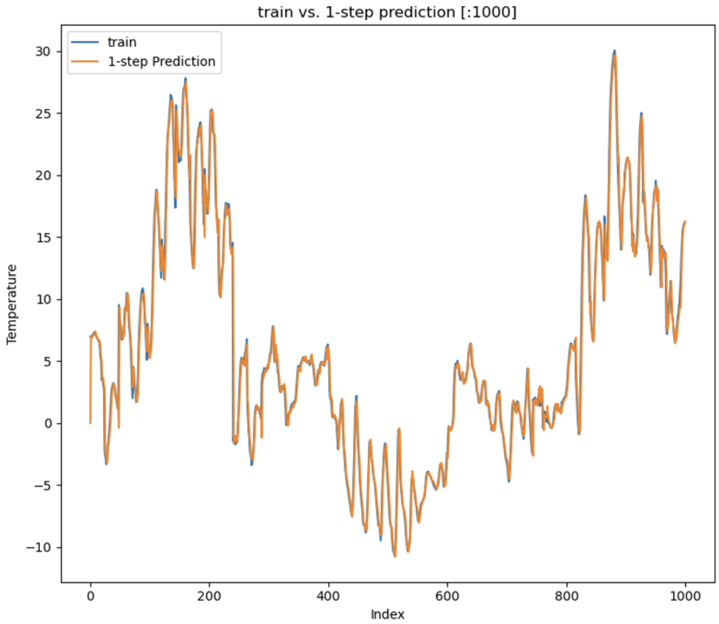

Figure 27: 1-step Prediction I

the train-set versus 1-step prediction is in Figure 27, and for residual's ACF/PACF is in Figure 28. The ACF for residual is in Figure 29. The ACF/PACF of residual shows a large correlation at lag equals 24, so I chose 1 as the seasonal order of AR and MA(1) [3]. The GPAC indicates na equals 1 or 2.

Again, I tried two MA orders and decided to use 2. And add 1 to the seasonal MA order. I fitted non-seasonal order (1,0,3) and seasonal order (1,0,1,24) to the model. The plot for the train-set versus 1-step prediction is in Figure 30, and for residual is in Figure 31. The ACF for residual is in Figure 32. The ACF/PACF shows faint significance in the second seasonal order (48).

I fitted (1,0,3) and seasonal order (2,0,2,24). The plot for the train-set versus 1-step prediction is in Figure 33. The ACF for residual in Figure 32 shows most autocorrelation between residuals is removed. The performance of 1-step prediction is good. I finally chose this model as my final SARIMA model.

11 Parameter Estimation

I used the Levenberg-Marquardt algorithm to estimate the parameters' coefficients, but all confidence intervals include zero, which means the estimated coefficients are not significant. The Levenberg-Marquardt algorithm does not work well in this case.

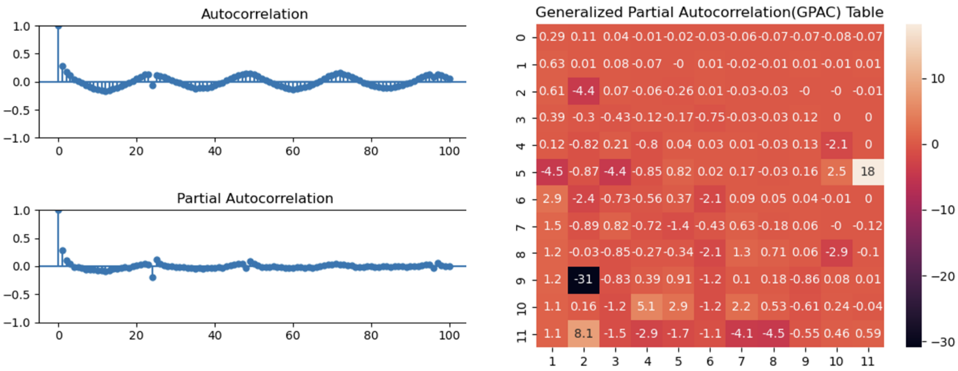

Figure 28: Residual I

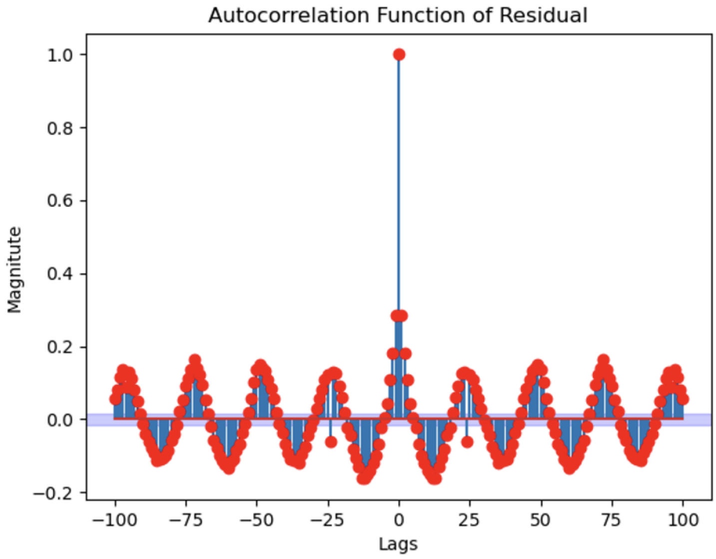

Figure 29: Residual ACF I

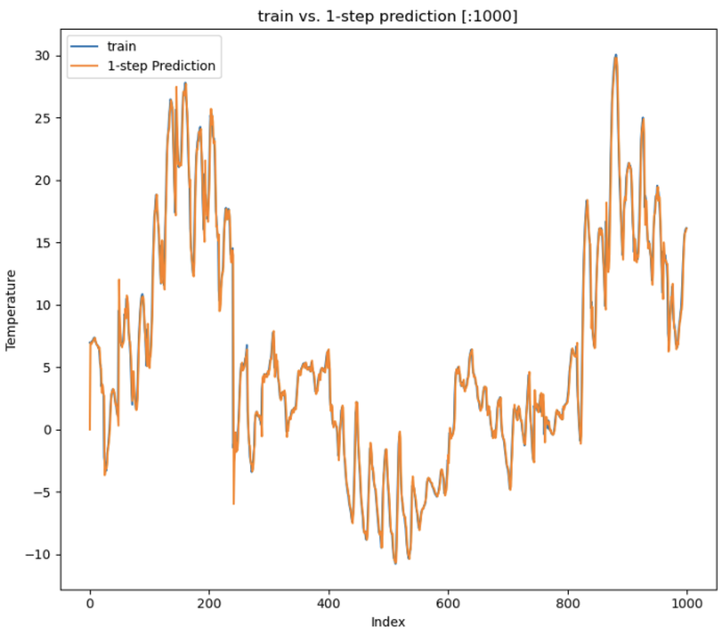

Figure 30: 1-step Prediction II

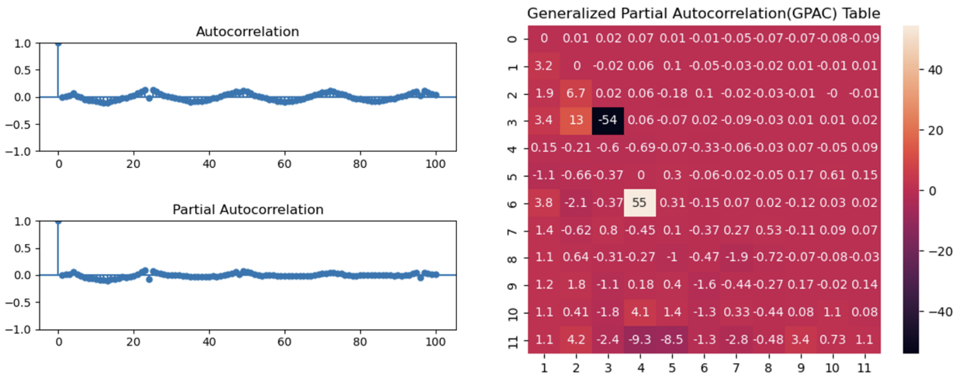

Figure 31: Residual II

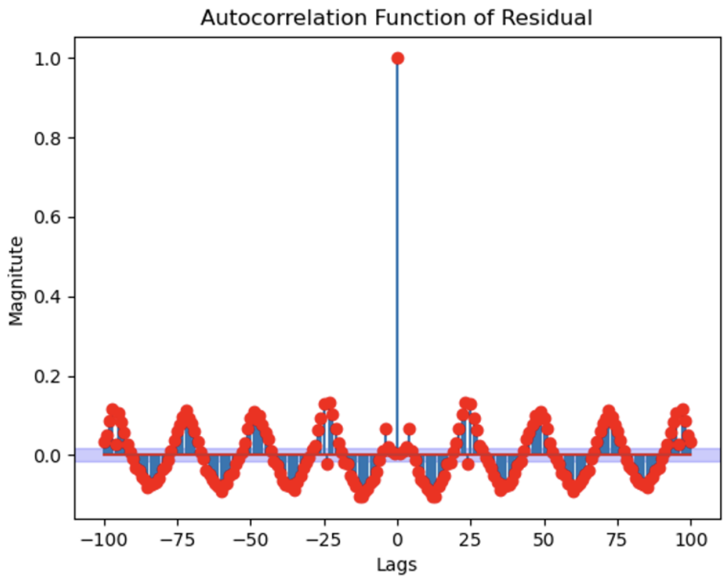

Figure 32: Residual ACF II

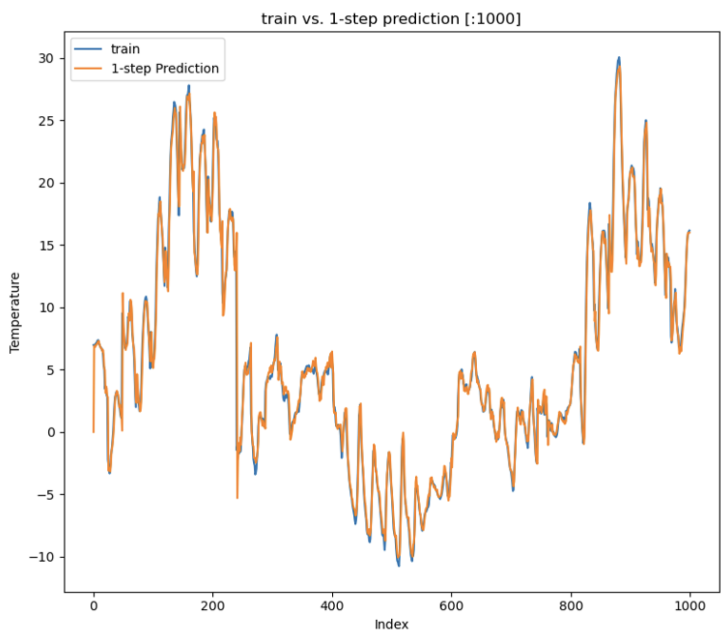

Figure 33: 1-step Prediction III

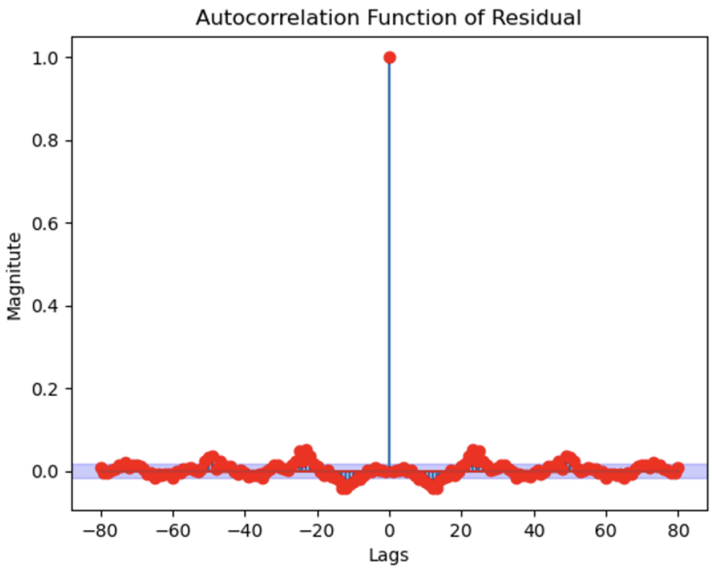

Figure 34: Residual ACF III

Method	Variance	Variance improvement (%) by SARIMA	MSE	MSE improvement (%) by SARIMA
Average	55.47	14.1 %	55.62	9.46%
Naive	55.47	14.1 %	76.23	33.94%
Dirft	61.04	21.94 %	92.82	45.74 %
SES	55.47	14.1%	76.23	33.94%
SARIMA	47.65	-	50.36	-

Table 4: SARIMA vs. Base models

The estimated parameters' coefficients from SARIMAX are in Figure 35. All the p-values are less than 0.05, so the coefficients are significant.

12 Residual Analysis and Diagnostic Test

I performed the Box-Pierce test and the Ljung-Box test to check if the residual error was white noise. The calculated Q value is greater than the Q* value. The p-values from the Ljung-Box test are all less than 0.05. Both tests failed. The residual errors are not white noise.

The estimated variance of the forecast error is 47.65. The estimated covariance of the estimated parameters is in Figure 36.

The estimated mean of the forecast error is -1.64, so the derived mode is biased.

The variance of the residual errors is 1.35, and the variance of forecast errors is 47.65.

I performed a zero-pole cancellation operation and there is no zero cancellation. The final confidence interval is in Figure 37.

13 Final Model Selection

The average improvement of variance of forecast errors by the SARIMA model is 16.06%. The average improvement of MSE is 30.77%. The detailed comparison of MSE and variance of forecast error between base models and SARIMA is in table 4. The final model picked is the SARIMA model.

SARIMAX Results

Dep. Variable:	Temp_C	No. Observations:	14036
Model:	SARIMAX(1, 0, 3)x(2, 0, [1, 2], 24)	Log Likelihood	-22017.440
Date:	Sun, 10 Dec 2023	AIC	44052.881
Time:	22:03:37	BIC	44120.825
Sample:	0	HQIC	44075.497
	- 14036		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.9733	0.002	406.698	0.000	0.969	0.978
ma.L1	0.2057	0.004	45.903	0.000	0.197	0.214
ma.L2	0.1087	0.009	11.817	0.000	0.091	0.127
ma.L3	0.0557	0.013	4.186	0.000	0.030	0.082
ar.S.L24	0.0710	0.010	7.206	0.000	0.052	0.090
ar.S.L48	0.9273	0.010	94.554	0.000	0.908	0.947
ma.S.L24	-0.0361	0.009	-3.938	0.000	-0.054	-0.018
ma.S.L48	-0.9239	0.009	-104.814	0.000	-0.941	-0.907
sigma2	1.3430	0.004	382.838	0.000	1.336	1.350

Ljung-Box (L1) (Q):	0.03	Jarque-Bera (JB):	3071769.19
Prob(Q):	0.86	Prob(JB):	0.00
Heteroskedasticity (H):	0.93	Skew:	-1.09
Prob(H) (two-sided):	0.01	Kurtosis:	75.44

Figure 35: SARIMAX estimated parameters

	ar.L1	ma.L1	ma.L2	ma.L3	ar.S.L24	ar.S.L48	ma.S.L24	ma.S.L48
ar.L1	0.000006	-0.000004	-0.000004	-2.427891e-06	-3.331247e-06	3.265615e-06	2.791145e-06	-0.000003
ma.L1	-0.000004	0.000020	0.000007	2.664634e-06	2.947719e-06	-2.948284e-06	-2.562705e-06	0.000003
ma.L2	-0.000004	0.000007	0.000085	4.574795e-06	7.794026e-06	-7.871232e-06	-7.387701e-06	0.000007
ma.L3	-0.000002	0.000003	0.000005	1.769290e-04	1.442170e-05	-1.444790e-05	-1.417018e-05	0.000013
ar.S.L24	-0.000003	0.000003	0.000008	1.442170e-05	9.709207e-05	-9.657784e-05	-8.952918e-05	0.000085
ar.S.L48	0.000003	-0.000003	-0.000008	-1.444790e-05	-9.657784e-05	9.617582e-05	8.883947e-05	-0.000085
ma.S.L24	0.000003	-0.000003	-0.000007	-1.417018e-05	-8.952918e-05	8.883947e-05	8.426071e-05	-0.000078
ma.S.L48	-0.000003	0.000003	0.000007	1.329496e-05	8.545856e-05	-8.538101e-05	-7.791888e-05	0.000078
sigma2	0.000005	-0.000002	-0.000003	-3.455510e-07	-9.389599e-09	-1.609758e-07	1.074502e-08	0.000001

Figure 36: Covariance of the estimated parameters

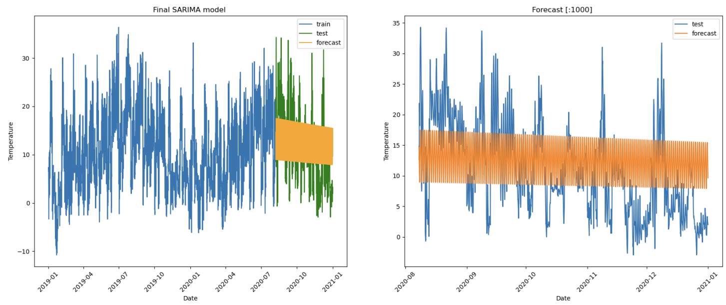

Figure 37: SARIMA h-step prediction

14 H-step Ahead Predictions

The plot for the h-step prediction is in Figure 37. The model captures some seasonality but is unable to capture the trend.

15 Conclusion

According to the above discussion, the selected SARIMA model has a lower variance of forecast error and MSE than base models. Although the residual error of 1-step prediction is not white noise, most autocorrelation between residuals is removed, indicating this model can capture some dynamic information.

This report has several limitations. The biggest limitation is that I down-sampled the dataset, which may remove some patterns between observations. The accuracy of the forecast might be improved by increasing the number of observations, using the raw dataset, and widening the range of parameter combinations.

References

[1] B. Mnassri, “Weather station beutenberg dataset,” 2020, accessed: 2023-12-11. [Online]. Available: https://www.kaggle.com/datasets/mnassrib/jena-weather-dataset/data

[2] P. Chen, A. Niu, D. Liu, W. Jiang, and B. Ma, “Time series forecasting of temperatures using sarima: An example from nanjing,” in *IOP Conference Series: Materials Science and Engineering*, vol. 394. IOP Publishing, 2018, p. 052024.

[3] M. T. Hagan and S. M. Behr, “The time series approach to short term load forecasting,” *IEEE transactions on power systems*, vol. 2, no. 3, pp. 785–791, 1987.