Climate Data Forecasting -

Atmospheric CO_2 Concentration / Temperature / Precipitation

Wolfgang Vollmer

2026-01-08

Contents

1 Forecasting of Hohenpeissenberg - Temperature and Precipitation Climate Analysis

1.1 Stationarity and differencing

Stationary time series is one whose properties do not depend on the time at which the series is observed. Thus, time series with trends, or with seasonality, are not stationary — the trend and seasonality will affect the value of the time series at different times. On the other hand, a white noise series is stationary — it does not matter when you observe it, it should look much the same at any point in time.

Stationary time series will have no predictable patterns in the long-term. Time plots will show the series to be roughly horizontal (although some cyclic behaviour is possible), with constant variance.

If y_t is a *stationary* time series, then for all s, the distribution of $(y_t, ..., y_{t+s})$ does not depend on t.

If Time Series data with seasonality are non-stationary

- => first take a seasonal difference
- if seasonally differenced data appear are still non-stationary
- => take an additional first seasonal difference

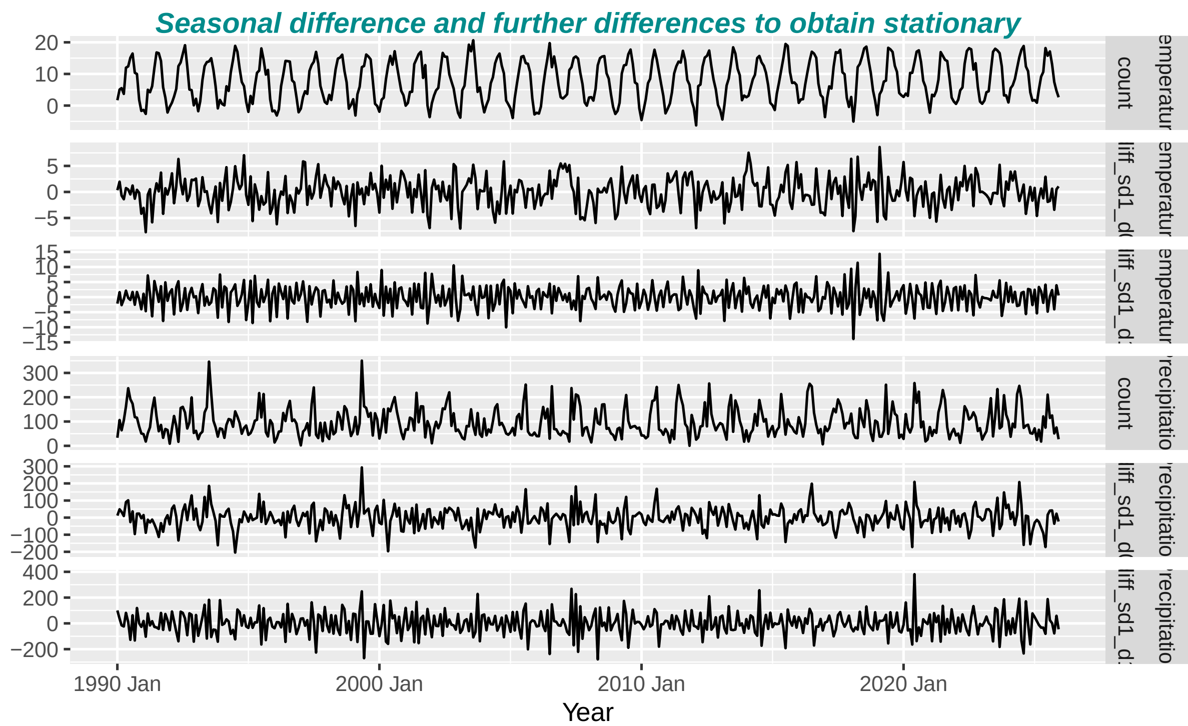
The model fit residuals have to be stationary. For good forecasting this has to be verified with residual diagnostics.

Essential:

- Residuals are uncorrelated
- The residuals have zero mean

Useful (but not necessary):

- The residuals have constant variance.
- The residuals are normally distributed.



1.1.1 Unitroot KPSS Test - fix number of seasonal differences/differences required

- For **ARIMA model** use Unitroot KPSS Test to fix number of differences
  - unitroot_nsdiffs() to determine D (the number of seasonal differences to use)
  - unitroot_ndiffs() to determine d (the number of ordinary differences to use)
  - The selection of the other model parameters (p,q,P and Q) are all determined by minimizing the AICc
- kpss test of stationary of the differentiated data (with seasonal and ordinary differences) used in the **ARIMA model**
  - stationary times series: the distribution of $(y_t, ..., y_{t-s})$ does not depend on t.

  - *Null Hypothesis* H_0: stationary is given in the time series: data are stationary and non seasonal
    * for $p < \alpha = 0.05$: reject H_0
    * for $p >> \alpha = 0.05$: conclude: differentiated data are stationary and non seasonal
- kpss test of stationary w/ unitroot_nsdiffs & unitroot_ndiff provides
  - minimum number of seasonal & ordinariel differences required for a stationary series
  - first fix required seasonal differences and then apply ndiffs to the seasonally differenced data
  - returns 1 => for stationarity one seasonal difference rsp. difference is required

```
#>   ndiffs gives the number of differences required and
#>  nsdiffs gives the number of seasonal differences required
#>  - to make a series stationary (test is based on the KPSS test
#> unitroot_kpss test to define seasonal (nsdiffs) and ordinary (ndiffs) differences
#> # A tibble: 2 x 5
#>   Measure       kpss_stat kpss_pvalue nsdiffs ndiffs
#>   <fct>             <dbl>       <dbl>   <int>  <int>
#> 1 Temperature        4.07        0.01       1      1
#> 2 Precipitation     17.0         0.01       0      1
#> unitroot_kpss test on (difference(count, 12) + difference())
#> # A tibble: 2 x 3
#>   Measure       kpss_stat kpss_pvalue
#>   <fct>             <dbl>       <dbl>
#> 1 Temperature     0.00184         0.1
#> 2 Precipitation   0.00181         0.1
```

1.1.2 Ljung-Box Test - independence/white noise of the time series

The Ljung-Box Test becomes important when checking independence/white noise of the forecasts residuals of the fitted ETS rsp. ARIMA models. There we have to check whether the forecast errors are normally distributed with mean zero

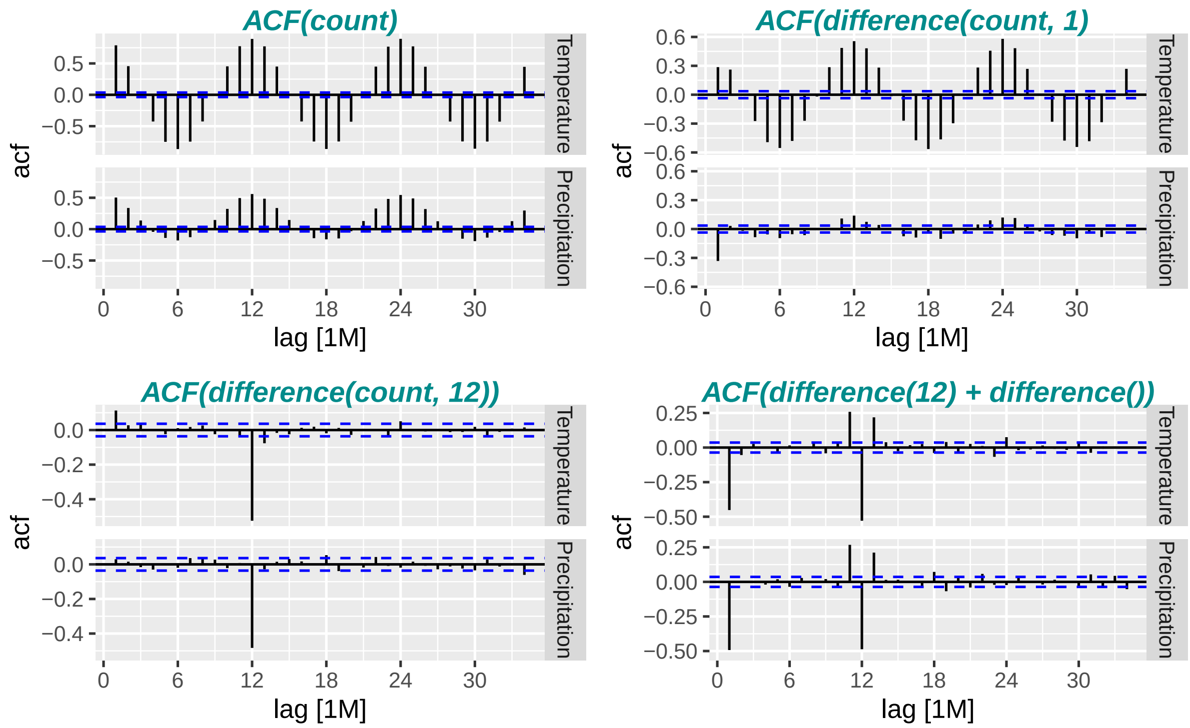
- augment(fit) |> features(.innov, ljung_box, lag=x, dof=y) (Ch. 5.4 Residaul diagnostics)
  - portmanteau test suggesting that the residuals are white noise
  - *Null Hypothesis* H_0: independence/white noise for residuals, i.e. each autocorrelation in the time series residuals of the model for lag l is close to zero.
    * for $p < \alpha = 0.05$: reject H_0
    * for $p >> \alpha = 0.05$: conclude: the residuals are not distinguishable from a white noise series
  - lag = 2*m (period of season, e.g. m=12 for monthly season) | no season: lag=10
  - dof = p + q + P + Q (for ARIMA models only, degree of freedom)

Time series with trend and/or seasonality is not stationary. Tests are to be taken on the residuals of the fitted models.

```
#> Ljung-Box test with (count), w/o differences
#> # A tibble: 2 x 3
#>   Measure       lb_stat lb_pvalue
#>   <fct>           <dbl>     <dbl>
#> 1 Temperature     9619.         0
#> 2 Precipitation   1715.         0
#> Ljung-Box test on (difference(count, 12))
```

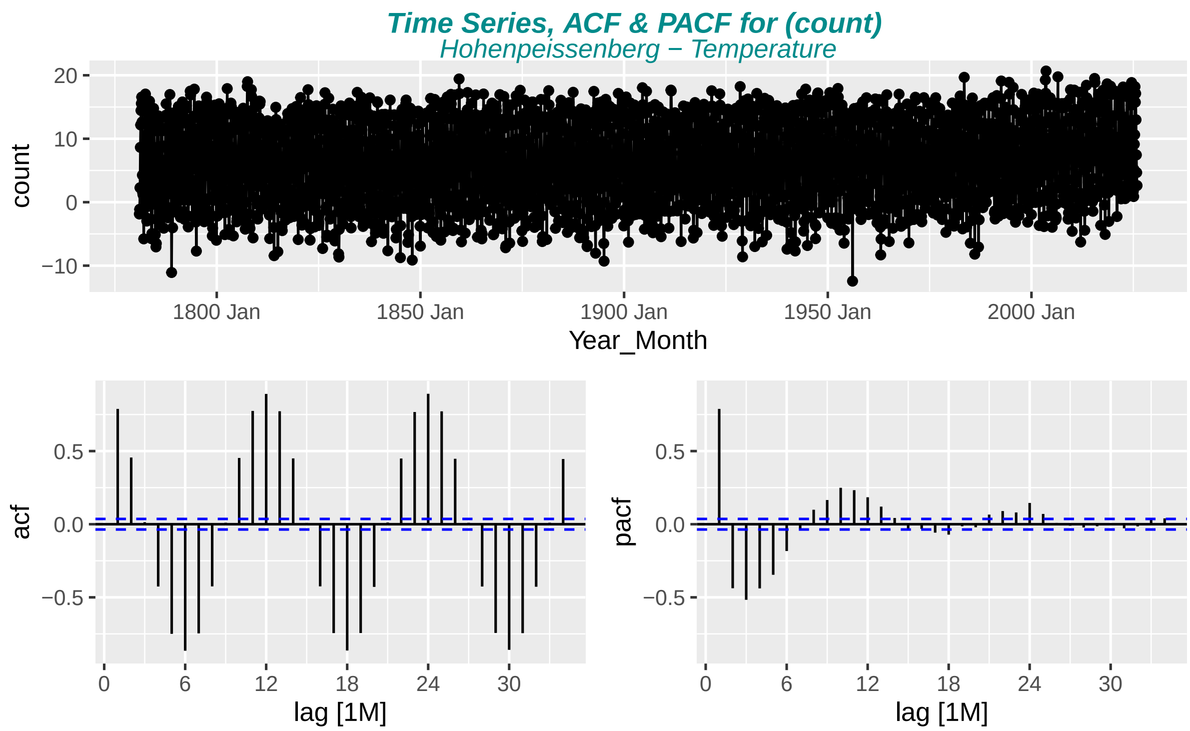
```
#> # A tibble: 2 x 3
#>   Measure       lb_stat   lb_pvalue
#>   <fct>           <dbl>       <dbl>
#> 1 Temperature      50.4 0.000000227
#> 2 Precipitation    19.3 0.0367
#> Ljung-Box test on (difference(count, 12) + difference())
#> # A tibble: 2 x 3
#>   Measure       lb_stat lb_pvalue
#>   <fct>           <dbl>     <dbl>
#> 1 Temperature      624.         0
#> 2 Precipitation    726.         0
```

1.1.3 ACF (Autocorrelation Function) Plots of Differences

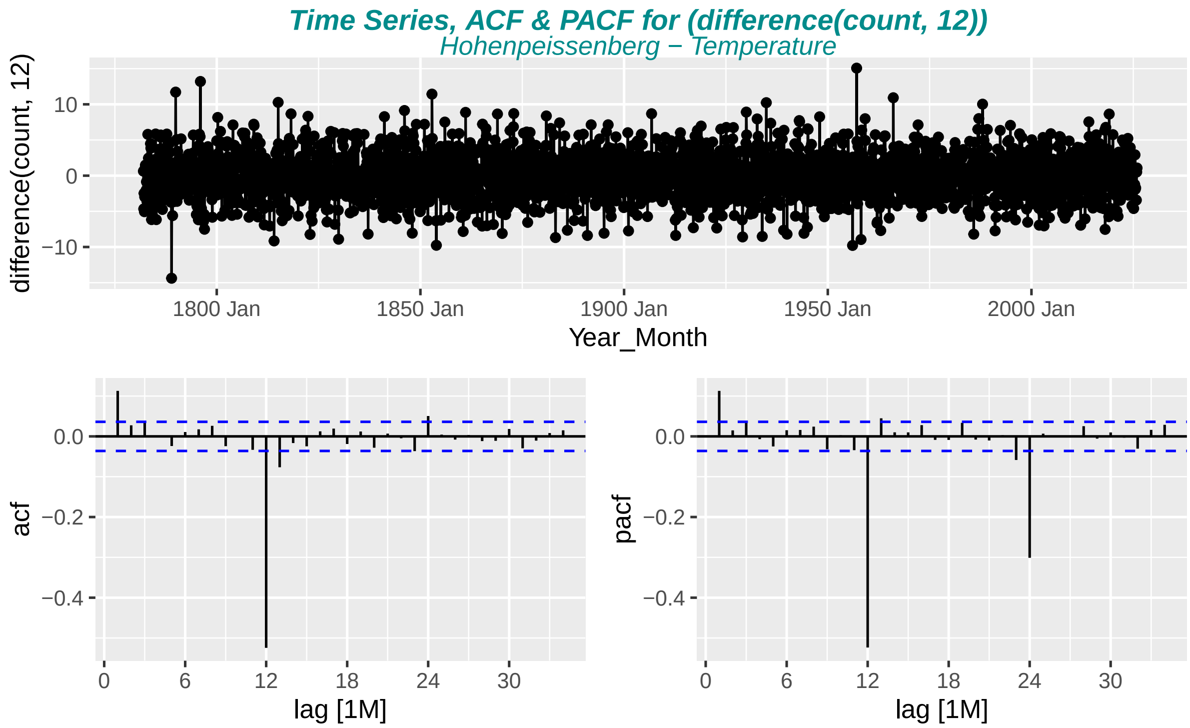


1.1.4 Time Series, ACF and PACF (Partial) Plots of Differences - for ARIMA p, q check

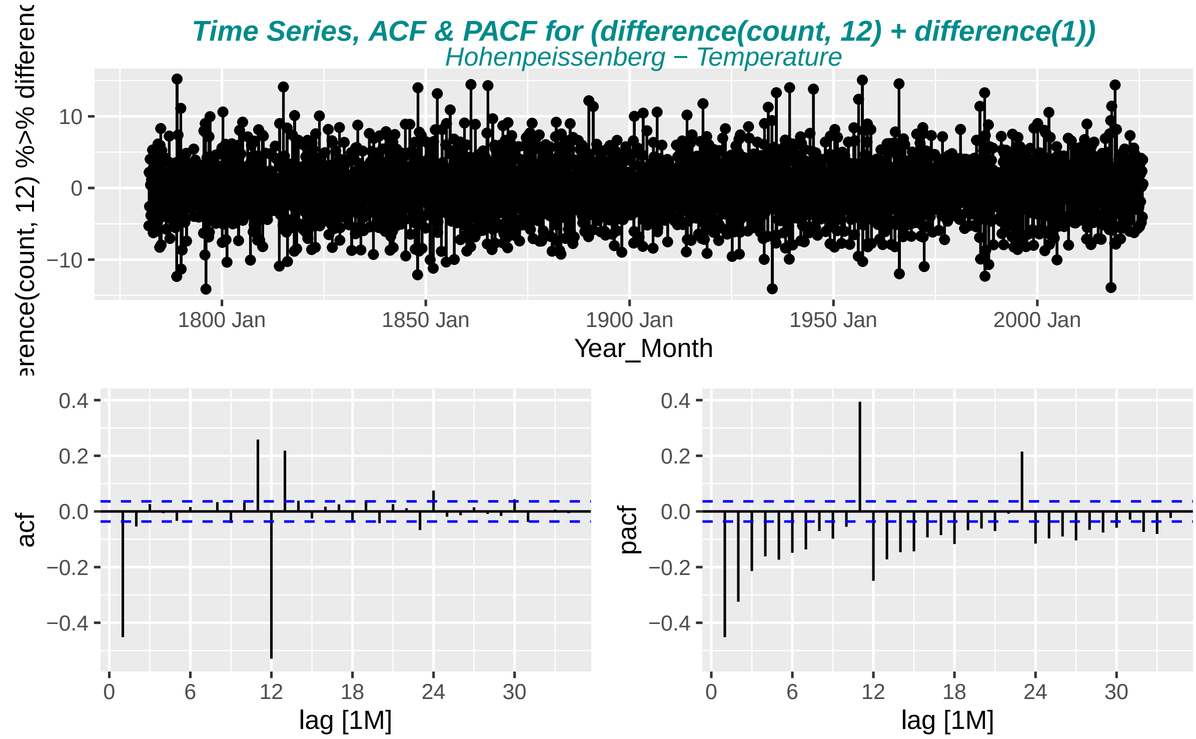
```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City             Measure         Sum  Mean
#>   <chr>            <fct>         <dbl> <dbl>
#> 1 Hohenpeissenberg Temperature 18597.   6.33
```



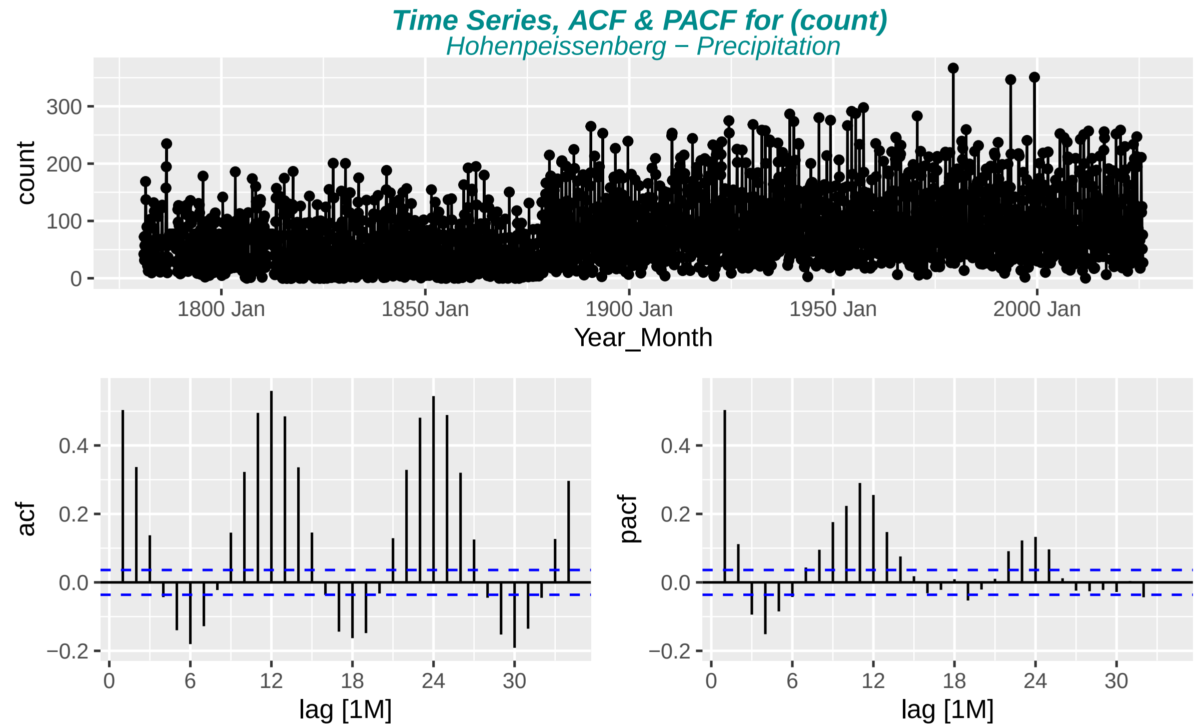

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City             Measure       Sum    Mean
#>   <chr>            <fct>       <dbl>   <dbl>
#> 1 Hohenpeissenberg Temperature  20.2 0.00689
```



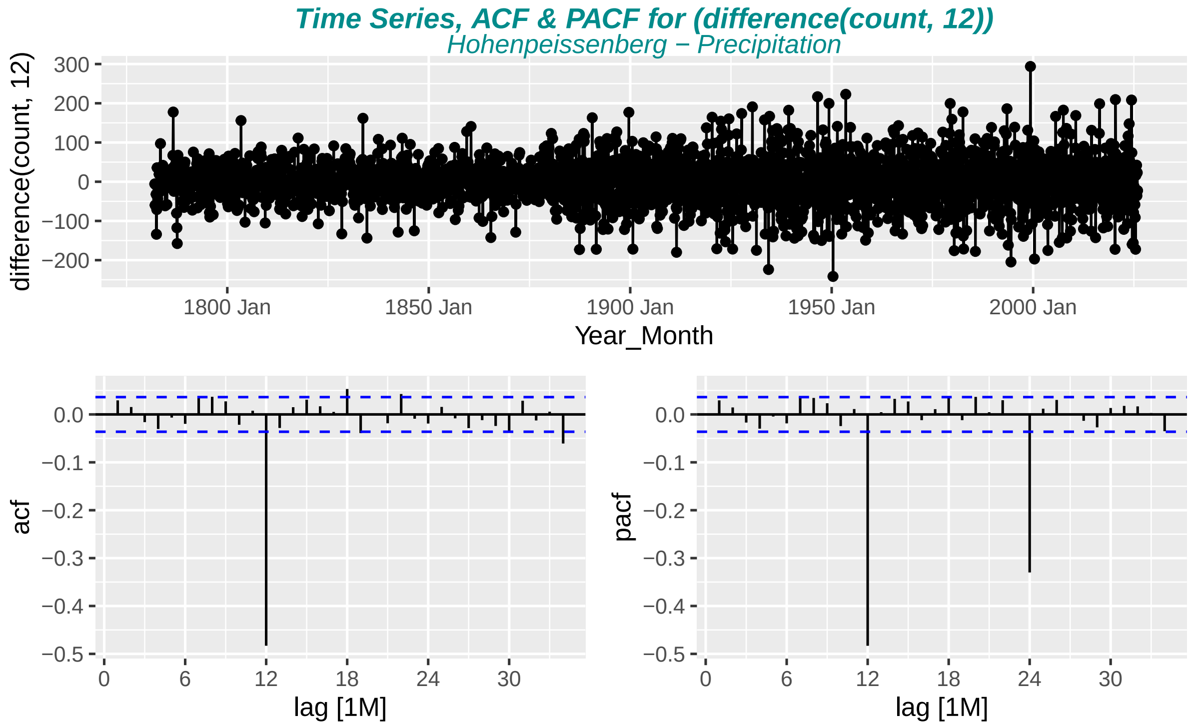

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City             Measure       Sum     Mean
#>   <chr>            <fct>       <dbl>    <dbl>
#> 1 Hohenpeissenberg Temperature 0.440 0.000150
```



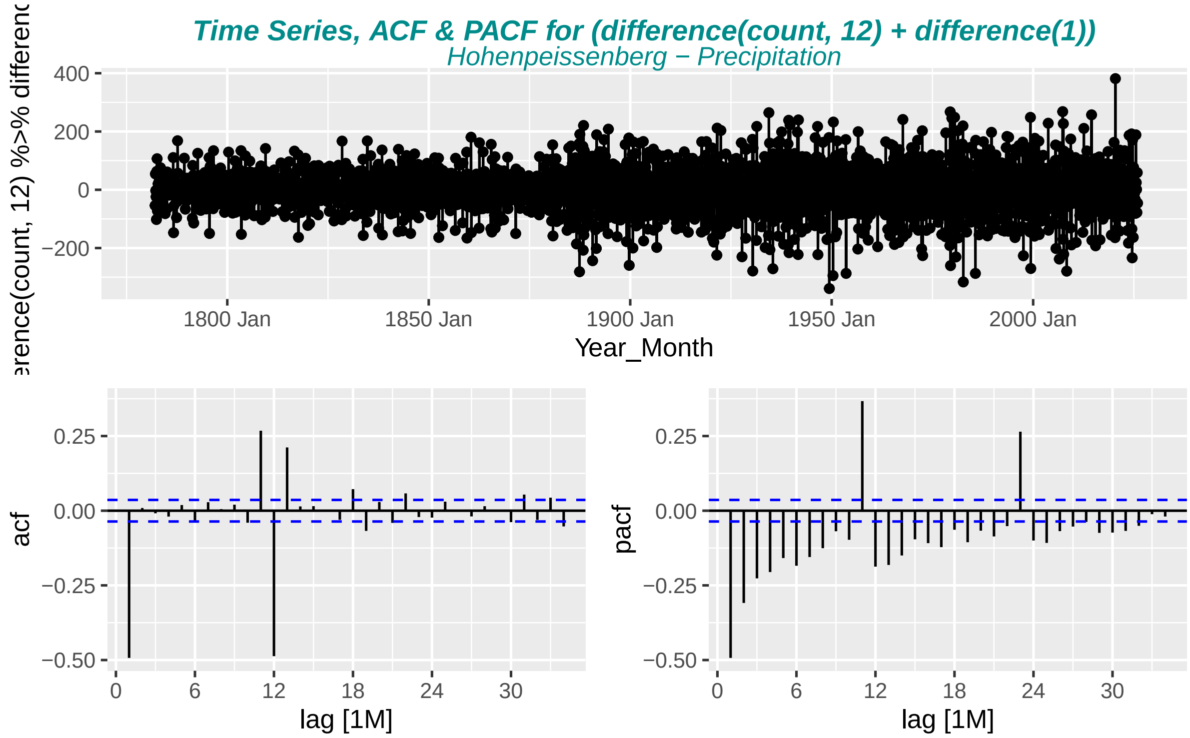

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City             Measure            Sum  Mean
#>   <chr>            <fct>            <dbl> <dbl>
#> 1 Hohenpeissenberg Precipitation 225987.   76.9
```




```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City             Measure          Sum   Mean
#>   <chr>            <fct>          <dbl>  <dbl>
#> 1 Hohenpeissenberg Precipitation  84.6 0.0289
```




```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City             Measure          Sum     Mean
#>   <chr>            <fct>          <dbl>    <dbl>
#> 1 Hohenpeissenberg Precipitation -17.0 -0.00581
```



2 ExponenTial Smoothing (ETS) Forecasting Models

Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older.

The parameters are estimated by maximising the "likelihood". The likelihood is the probability of the data arising from the specified model. AIC, AICc and BIC can be used here to determine which of the ETS models is most appropriate for a given time series (see output glance(fit_ets)).

The model selection is based on recognising key components of the time series (trend and seasonal) and the way in which these enter the smoothing method (e.g., in an additive, damped or multiplicative manner).

- Mauna Loa CO_2 data best Models: ETS(M,A,A) & ETS(A,A,A)
- Basel Temperature data best Models: ETS(A,N,A), ETS(A,A,A), ETS(A,Ad,A) (close togehter). Best Forecast accuracy is with ETS(A,A,A), ETS(A,Ad,A).
- Basel Precipitation data best Models: ETS(A,N,A), ETS(A,Ad,A), ETS(A,A,A) (close togehter). Best Forecast accuracy is with ETS(A,A,A), ETS(A,Ad,A), ETS(A,N,A),

Trend term "N" for Basel Temperature/Precipitation correspondends to a "pure" exponential smooothing which results in a slope $\beta = 0$. This results in a forecast predicting a constant level. This does not fit to the result of the STL decomposition. Therefore best model choice is **ETS(A,A,A)**.

Method Selection

Error term: either additive ("A") or multiplicative ("M").

Both methods provide identical point forecasts, but different prediction intervals and different likelihoods. AIC & BIC are able to select between the error types because they are based on likelihood.

Nevertheless, difference is for

- Mauna Loa CO_2 not relevant and AIC/AICc/BIC values are only a little bit smaller for multiplicative errors. The prediction intervall plots are fully overlapping.
- Basel Temperature AIC/AICc/BIC of additive error types are much better than the multiplicative ones.
- Basel Precipitation AIC/AICc/BIC of additive error types are much better than the multiplicative ones.

Note: For Basel Temperature and Precipitation Forecast plots the models ETS_MAdA, ETS_MMA, ETS_MMA, ETS_MNA are to be taken out since forecasts with multiplicative errors are exploding (forecast $>$ 3 years impossible !!)

Therefore finally **Error term = "A"** is chosen in general.

Trend term: either none ("N"), additive ("A"), multiplicative ("M") or damped variants ("Ad", "Md").

Note: Mauna Loa CO_2 model ETS(A,Ad,A) fit plot shows to strong damping. For Basel Temperature model ETS(A,N,A) and ETS(A,Ad,A) are providing more or less the same forecast. This means that forecast remains on constant level since Trend "N" means "pure" exponentiell smoothing without trend (see above).

Therefore finally **Trend term = "A"** is chosen in general.

Seasonal term: either none ("N"), additive ("A") or multiplicative ("M").

For CO2 and Temperature Data we have a clear seasonal pattern and seasonal term adds always a (more or less) fix amount on level and trend component. Therefore "A" additve term is chosen. For Precipitation the seasonal patttern is only slight. Indead, a multiplicative seasonal term results in "exploding" forecasts.

Since monthly data are strongly seasonal **seasonal term "A"** is chosen.

2.1 ETS Models and their componentes

ETS model with automatically selected $ETS(A|M,N|A|M,N|A|M)$

```
#> [1] "model(ETS(count)) => provides default / best automatically chosen model"
#> # A mable: 2 x 3
#> # Key:     City, Measure [2]
#>   City             Measure                ETS
#>   <chr>            <fct>              <model>
#> 1 Hohenpeissenberg Temperature   <ETS(A,N,A)>
#> 2 Hohenpeissenberg Precipitation <ETS(M,N,M)>
#> [1] "Hohenpeissenberg Temperature"
#> Series: count
#> Model: ETS(A,N,A)
#>   Smoothing parameters:
#>     alpha = 0.02275054
#>     gamma = 0.0001000033
#>
#>   Initial states:
#>      l[0]      s[0]    s[-1]    s[-2]    s[-3]    s[-4]    s[-5]    s[-6]
#>  7.248062 -6.847594 -4.16179 1.008964 4.686353 8.417777 8.507015 6.738792
#>     s[-7]     s[-8]     s[-9]    s[-10]    s[-11]
#>  3.218513 -1.250983 -4.580917 -7.595268 -8.140861
#>
#>   sigma^2:  4.109
#>
#>      AIC     AICc      BIC
#> 5770.405 5771.087 5839.094
#> [1] "Hohenpeissenberg Precipitation"
#> Series: count
#> Model: ETS(M,N,M)
#>   Smoothing parameters:
#>     alpha = 0.01455169
#>     gamma = 0.0001004975
#>
#>   Initial states:
#>      l[0]      s[0]   s[-1]     s[-2]    s[-3]    s[-4]    s[-5]    s[-6]
#>  114.4223 0.6580902 0.79261 0.7729174 1.038345 1.629733 1.534146 1.533975
#>     s[-7]     s[-8]     s[-9]    s[-10]    s[-11]
#>  1.340229 0.8028894 0.6359576 0.6092697 0.6518372
#>
#>   sigma^2:  0.2165
#>
#>      AIC     AICc      BIC
#> 10175.86 10176.54 10244.55
#> # A tibble: 2 x 8
#>   City             Measure       .model    AIC   AICc    BIC     MSE   MAE
#>   <chr>            <fct>         <chr>   <dbl>  <dbl>  <dbl>   <dbl> <dbl>
#> 1 Hohenpeissenberg Temperature   ETS     5770.  5771.  5839.    4.03 1.60
#> 2 Hohenpeissenberg Precipitation ETS    10176. 10177. 10245. 2051.   0.360
```

Fit of different pre-defined $ETS(A|M,N|A|M,N|A|M)$ models

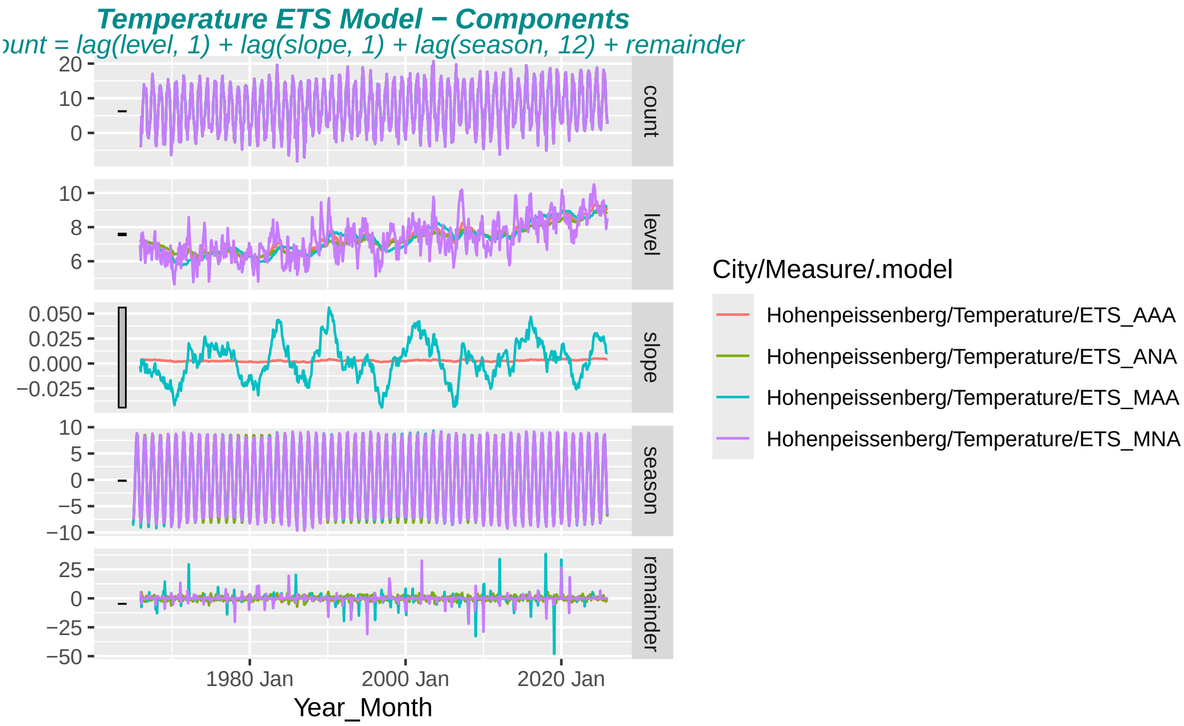
Model Selection by Information Criterion - lowest AIC, AICc, BIC

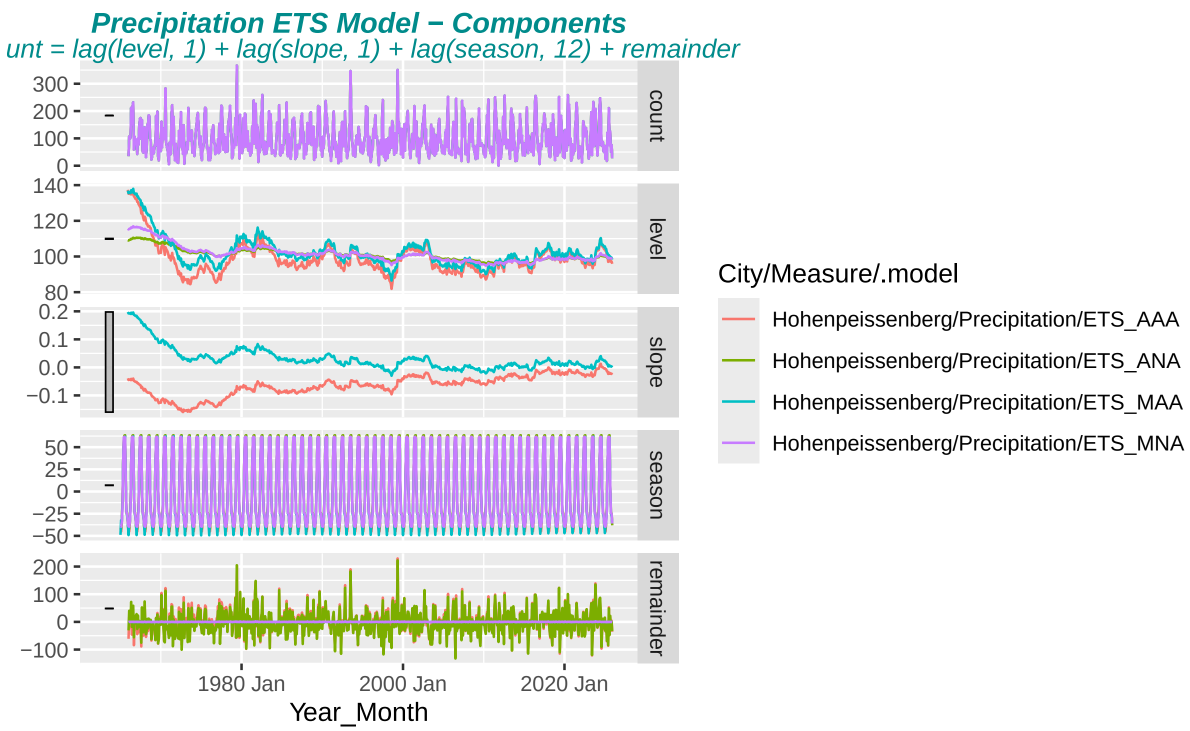
Best model fit with

- CV, AIC, AICc and BIC with the lowest values
- Adjusted R^2 the model with the highest value.

```
#> # A tibble: 16 x 9
#>    City             Measure     .model    AIC   AICc    BIC    MSE   AMSE    MAE
```

```
#>    <chr>           <fct>       <chr>   <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
#>  1 Hohenpeissenberg Temperature ETS_A~  5768.  5768.  5845. 3.99e0 3.98e0  1.58
#>  2 Hohenpeissenberg Temperature ETS_A~  5770.  5771.  5839. 4.03e0 4.02e0  1.60
#>  3 Hohenpeissenberg Temperature ETS_A~  5771.  5772.  5854. 4.00e0 3.99e0  1.59
#>  4 Hohenpeissenberg Temperature ETS_A~  5774.  5775.  5852. 4.02e0 4.02e0  1.60
#>  5 Hohenpeissenberg Temperature ETS_M~  7397.  7398.  7475. 1.37e1 1.42e1  0.486
#>  6 Hohenpeissenberg Temperature ETS_M~  8769.  8770.  8838. 5.01e0 5.06e0  1.29
#>  7 Hohenpeissenberg Temperature ETS_M~  9021.  9022.  9099. 4.54e0 4.54e0  1.53
#>  8 Hohenpeissenberg Temperature ETS_M~  9160.  9161.  9243. 4.78e0 4.92e0  1.33
#>  9 Hohenpeissenberg Precipitat~ ETS_M~ 10182. 10182. 10250. 2.04e3 2.04e3  0.356
#> 10 Hohenpeissenberg Precipitat~ ETS_M~ 10198. 10199. 10280. 2.03e3 2.03e3  0.360
#> 11 Hohenpeissenberg Precipitat~ ETS_M~ 10226. 10227. 10304. 2.10e3 2.10e3  0.365
#> 12 Hohenpeissenberg Precipitat~ ETS_M~ 10226. 10227. 10304. 2.09e3 2.08e3  0.367
#> 13 Hohenpeissenberg Precipitat~ ETS_A~ 10249. 10250. 10331. 2.01e3 2.01e3 33.9
#> 14 Hohenpeissenberg Precipitat~ ETS_A~ 10251. 10251. 10319. 2.03e3 2.03e3 34.3
#> 15 Hohenpeissenberg Precipitat~ ETS_A~ 10267. 10267. 10344. 2.06e3 2.06e3 34.4
#> 16 Hohenpeissenberg Precipitat~ ETS_A~ 10268. 10269. 10346. 2.07e3 2.06e3 34.4
```





2.1.1 Residual Accuracy with one-step-ahead fitted residuals - check RMSE, MAE

Residual accuracy can be computed directly from models as the one-step-ahead fitted residuals are available. Select forecast models that minimises for lowest

- MAE (Mean absolute error, will lead to forecasts of the median) and
- RMSE (Root mean squared error, lead to forecasts of the mean)

```
#> # A tibble: 16 x 7
#>    City             Measure       .model   .type          ME  RMSE   MAE
#>    <chr>            <fct>         <chr>    <chr>       <dbl> <dbl> <dbl>
#>  1 Hohenpeissenberg Temperature   ETS_AMA  Training  0.0516   2.00  1.58
#>  2 Hohenpeissenberg Temperature   ETS_AAdA Training  0.140    2.00  1.59
#>  3 Hohenpeissenberg Temperature   ETS_AAA  Training  0.00629  2.01  1.60
#>  4 Hohenpeissenberg Temperature   ETS_ANA  Training  0.0982   2.01  1.60
#>  5 Hohenpeissenberg Temperature   ETS_MAA  Training  0.00880  2.13  1.69
#>  6 Hohenpeissenberg Temperature   ETS_MAdA Training  0.00312  2.19  1.75
#>  7 Hohenpeissenberg Temperature   ETS_MNA  Training  0.00830  2.24  1.80
#>  8 Hohenpeissenberg Temperature   ETS_MMA  Training -2.45     3.70  2.88
#>  9 Hohenpeissenberg Precipitation ETS_AAdA Training -0.575   44.8  33.9
#> 10 Hohenpeissenberg Precipitation ETS_ANA  Training -2.12    45.1  34.3
#> 11 Hohenpeissenberg Precipitation ETS_MAdA Training -1.56    45.1  34.3
#> 12 Hohenpeissenberg Precipitation ETS_MNA  Training -2.53    45.2  34.5
#> 13 Hohenpeissenberg Precipitation ETS_AMA  Training  0.574   45.4  34.4
#> 14 Hohenpeissenberg Precipitation ETS_AAA  Training  0.181   45.5  34.4
#> 15 Hohenpeissenberg Precipitation ETS_MMA  Training -1.75    45.7  34.7
#> 16 Hohenpeissenberg Precipitation ETS_MAA  Training -2.74    45.8  35.1
```

2.1.2 Ljung-Box Test - independence/white noise of the forecasts residuals

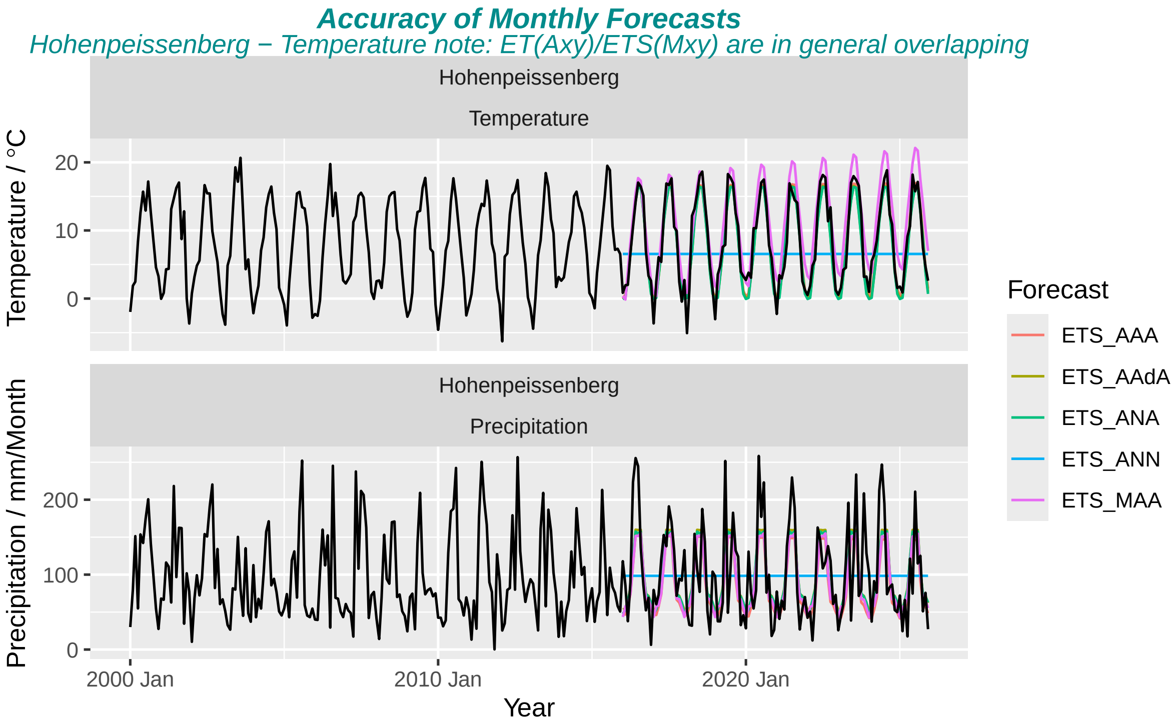
```
#> Null Hypothesis of independence/white noise for residuals - for p < 0.05: reject H_0
#> # A tibble: 16 x 5
#>    City             Measure       .model   lb_stat lb_pvalue
#>    <chr>            <fct>         <chr>      <dbl>     <dbl>
#>  1 Hohenpeissenberg Temperature   ETS_ANA     21.5 0.429
```

```
#>  2 Hohenpeissenberg Temperature   ETS_AMA     22.3 0.384
#>  3 Hohenpeissenberg Temperature   ETS_AAdA    23.1 0.339
#>  4 Hohenpeissenberg Temperature   ETS_AAA     24.5 0.268
#>  5 Hohenpeissenberg Precipitation ETS_MMA     25.5 0.227
#>  6 Hohenpeissenberg Precipitation ETS_MAA     25.9 0.210
#>  7 Hohenpeissenberg Precipitation ETS_ANA     26.5 0.188
#>  8 Hohenpeissenberg Precipitation ETS_MNA     26.6 0.185
#>  9 Hohenpeissenberg Precipitation ETS_MAdA    26.8 0.179
#> 10 Hohenpeissenberg Precipitation ETS_AAdA    27.0 0.172
#> 11 Hohenpeissenberg Precipitation ETS_AMA     27.2 0.165
#> 12 Hohenpeissenberg Precipitation ETS_AAA     27.2 0.164
#> 13 Hohenpeissenberg Temperature   ETS_MAA     33.3 0.0432
#> 14 Hohenpeissenberg Temperature   ETS_MAdA    44.9 0.00178
#> 15 Hohenpeissenberg Temperature   ETS_MNA     56.3 0.0000449
#> 16 Hohenpeissenberg Temperature   ETS_MMA    377.  0
```

2.1.3 ETS Models - components of ETS(A,N,A), ETS(A,A,A), ETS(A,Ad,A), models

2.1.4 Forecast Accuracy with Training/Test Data

```
#> # A tibble: 10 x 7
#>    .model   City             Measure       .type     ME  RMSE   MAE
#>    <chr>    <chr>            <fct>         <chr>  <dbl> <dbl> <dbl>
#>  1 ETS_AAA  Hohenpeissenberg Temperature   Test   0.374  1.89  1.51
#>  2 ETS_AAdA Hohenpeissenberg Temperature   Test   0.715  2.00  1.63
#>  3 ETS_ANA  Hohenpeissenberg Temperature   Test   0.830  2.05  1.69
#>  4 ETS_MAA  Hohenpeissenberg Temperature   Test  -2.33   3.24  2.73
#>  5 ETS_ANN  Hohenpeissenberg Temperature   Test   2.16   6.67  5.69
#>  6 ETS_AAdA Hohenpeissenberg Precipitation Test   4.37  47.3  37.3
#>  7 ETS_ANA  Hohenpeissenberg Precipitation Test   5.79  47.5  37.3
#>  8 ETS_MAA  Hohenpeissenberg Precipitation Test  10.1   48.5  38.0
#>  9 ETS_AAA  Hohenpeissenberg Precipitation Test  14.7   49.2  38.0
#> 10 ETS_ANN  Hohenpeissenberg Precipitation Test   3.21  62.9  50.7
```

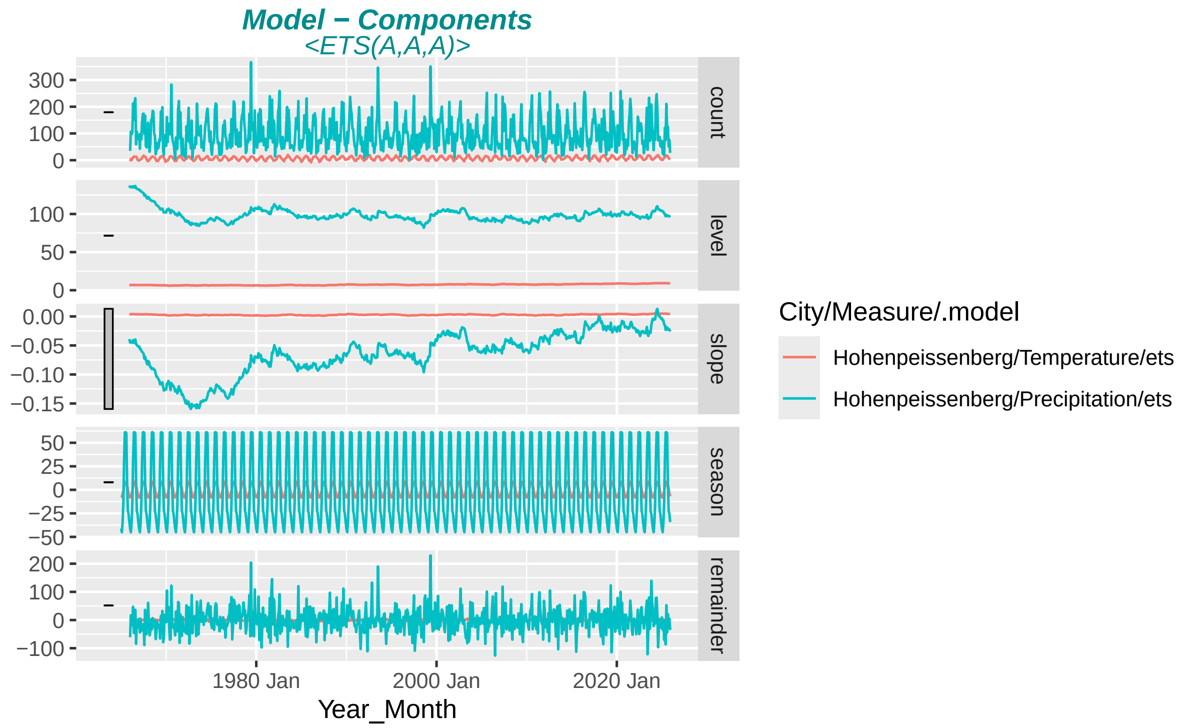


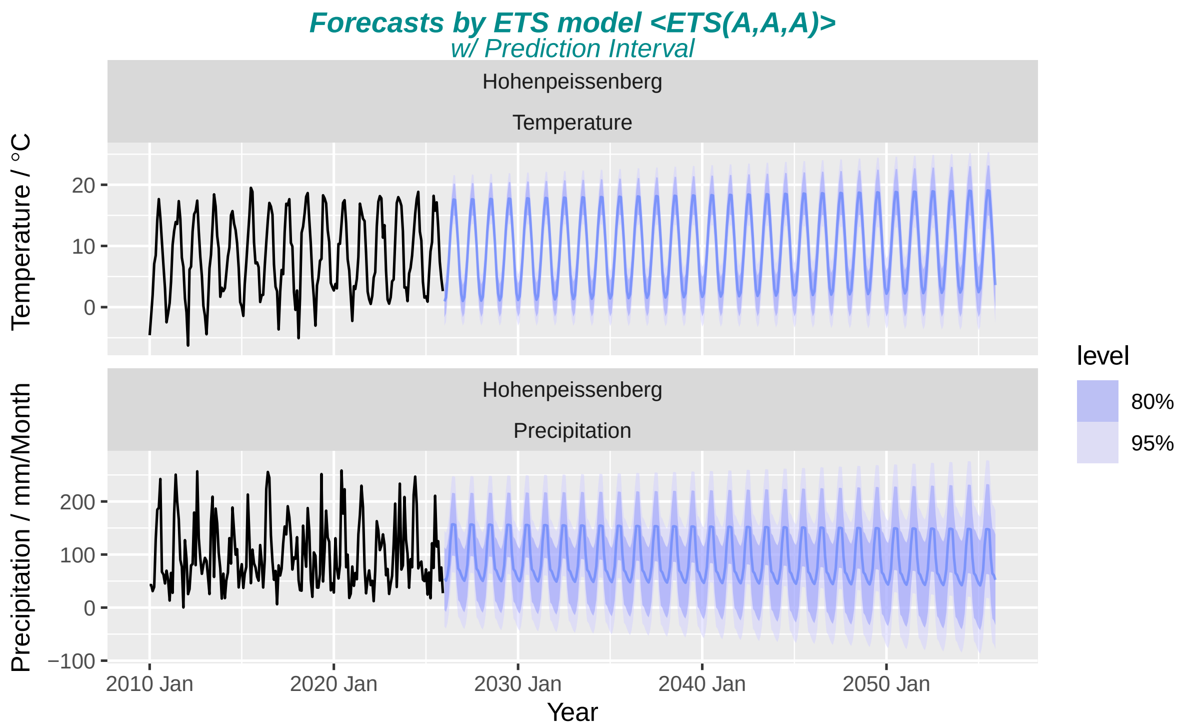
Accuracy of Monthly Forecasts
Hohenpeissenberg − Temperature note: ET(Axy)/ETS(Mxy) are in general overlapping
Hohenpeissenberg
Temperature
Hohenpeissenberg
Precipitation
Temperature / °C
Precipitation / mm/Month
20
10
0
200
100
0
2000 Jan
2010 Jan
2020 Jan
Year
Forecast
ETS_AAA
ETS_AAdA
ETS_ANA
ETS_ANN
ETS_MAA

2.2 Forecasting with selected ETS model <ETS(A,A,A)>, <ETS(A,A,A)>

2.2.1 Forecast Plot of selected ETS model

```
#> Provide model coefficients by report(fit_model)
#> # A tibble: 2 x 11
#>   City     Measure .model sigma2 log_lik    AIC   AICc    BIC    MSE   AMSE   MAE
#>   <chr>    <fct>   <chr>   <dbl>   <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl> <dbl>
#> 1 Hohenp~ Temper~ ets    4.12e0  -2870.  5774.  5775.  5852. 4.02e0 4.02e0  1.60
#> 2 Hohenp~ Precip~ ets    2.12e3  -5117. 10268. 10269. 10346. 2.07e3 2.06e3 34.4
```

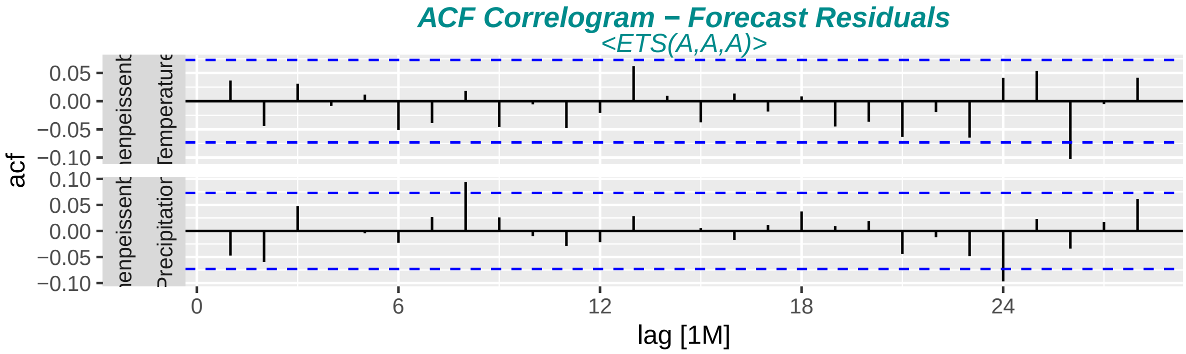


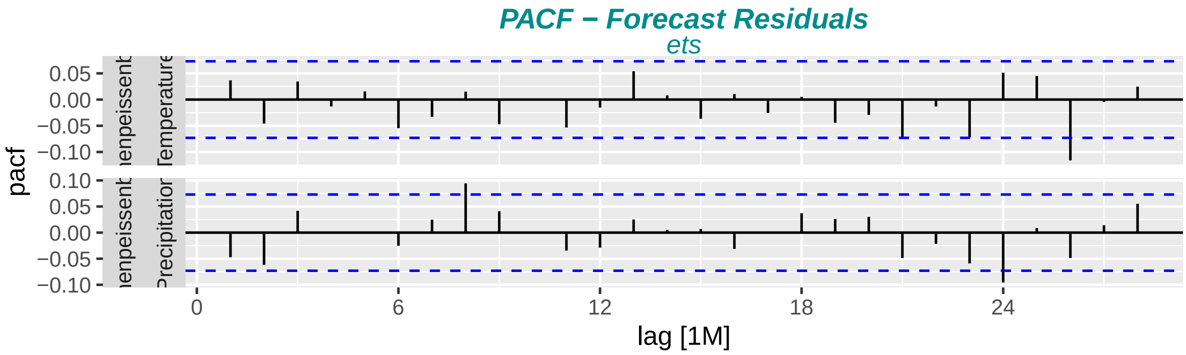


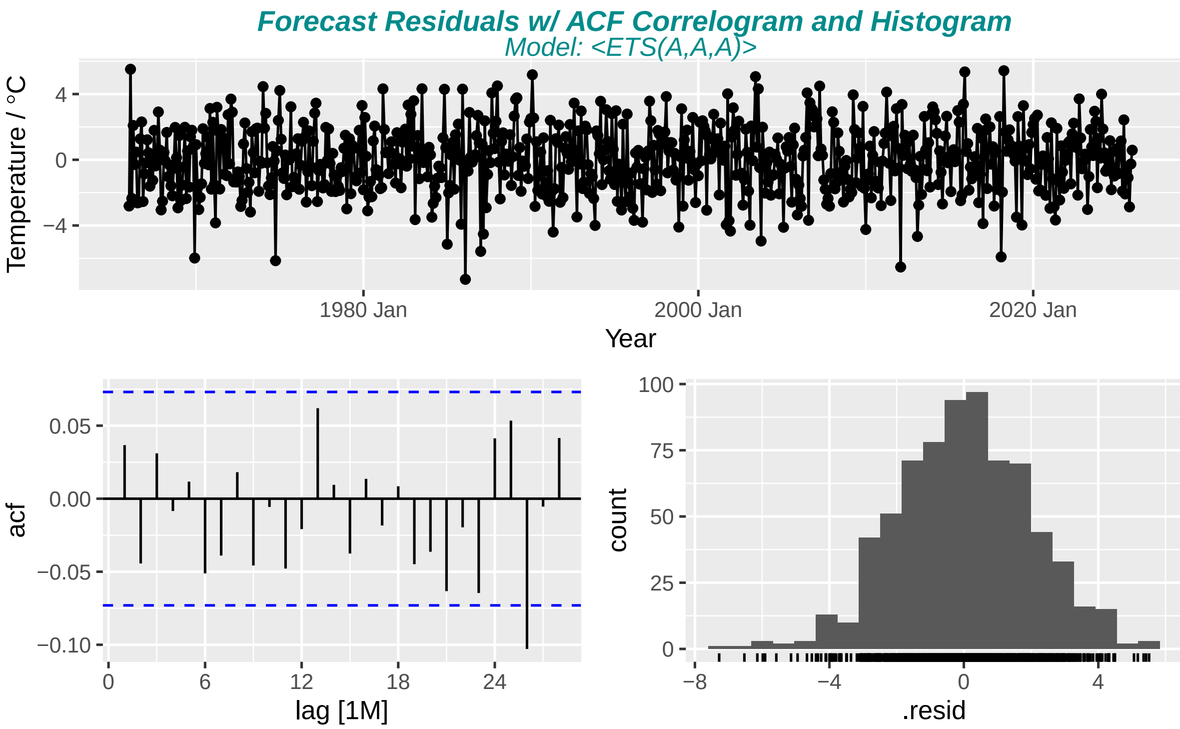
2.2.2 Residual Stationarity

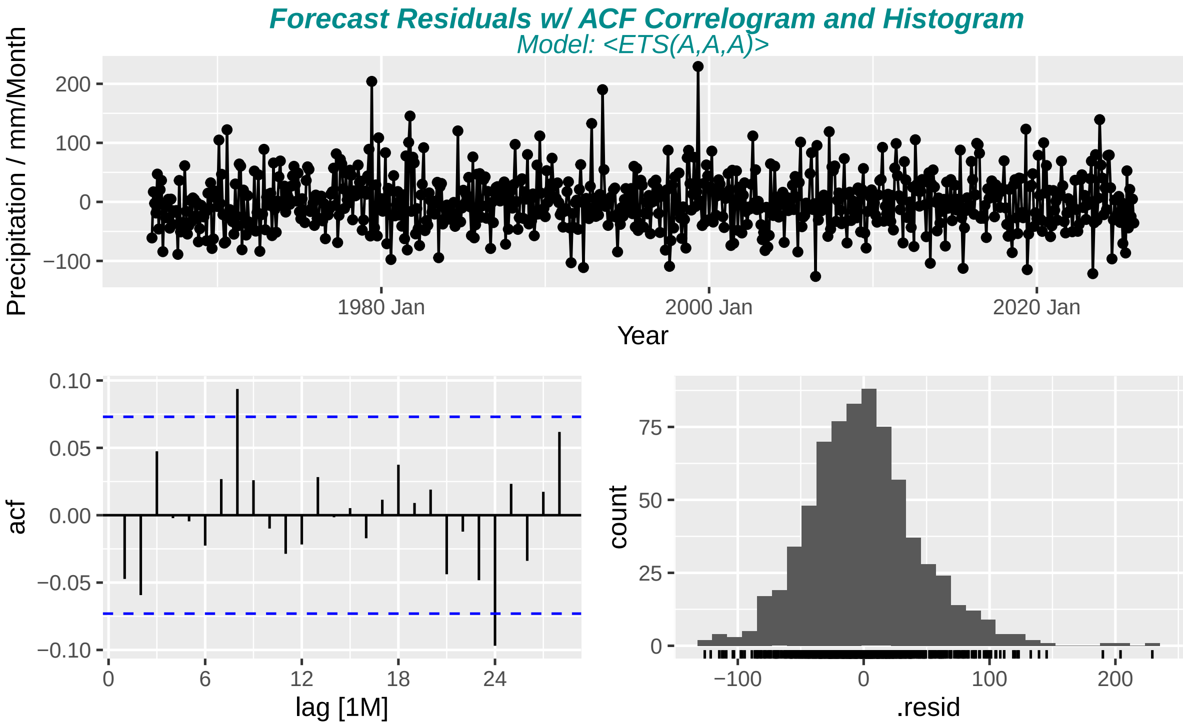
Required checks to be ready for forecasting:

- ACF Forecast Residual: all spikes are within the significance limits, so the residuals appear to be white noise
- The Ljung-Box test also shows that the residuals have no remaining autocorrelations
- Forecast Residuals are more or less normally distributed with roughly centred on zero



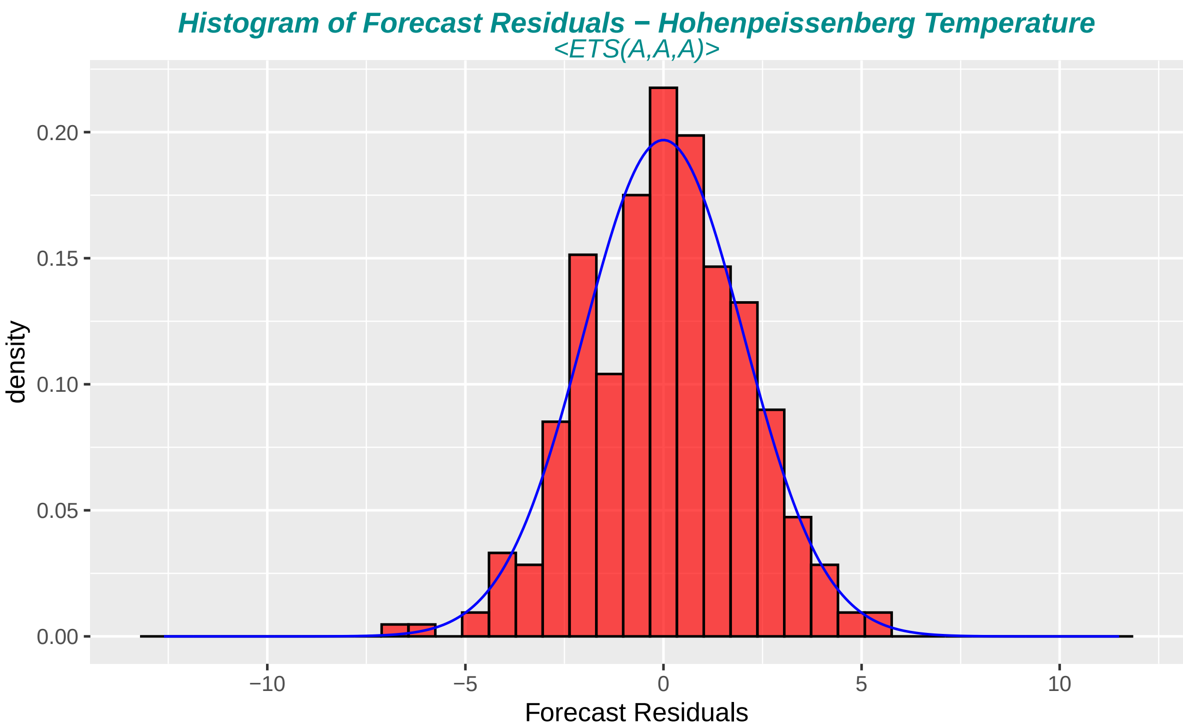


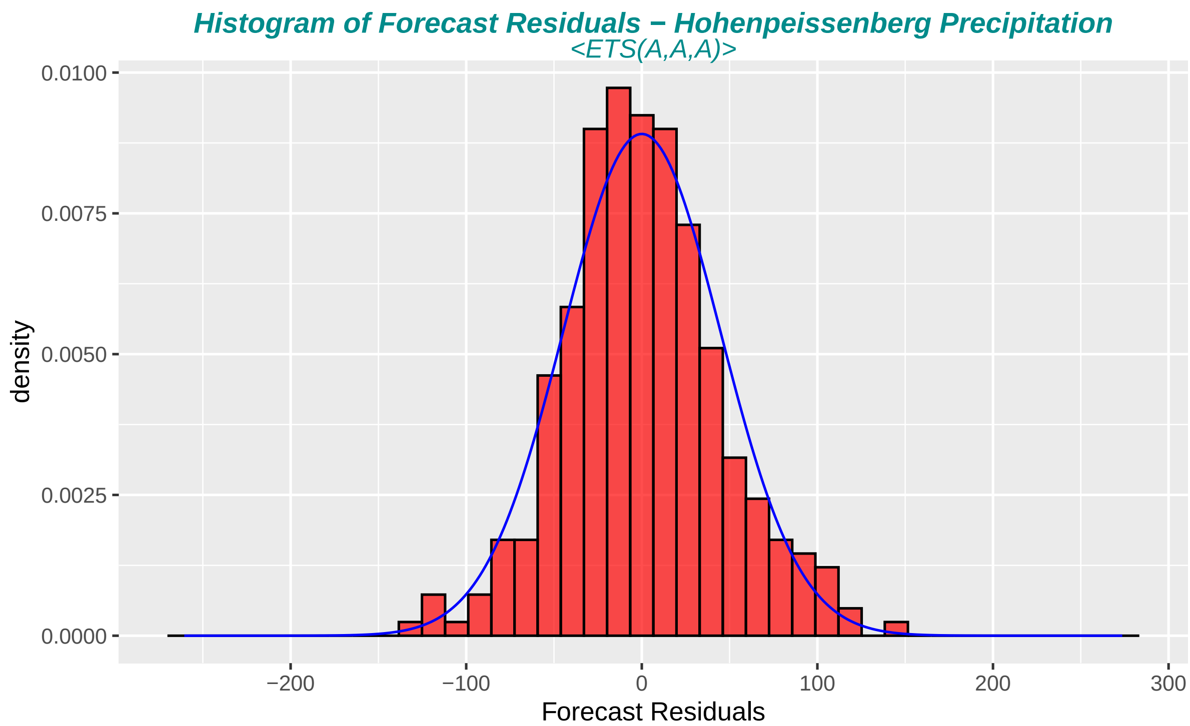




2.2.3 Ljung-Box Test and Histogram of forecast residuals with overlaid normal curve

```
#> Null Hypothesis of independence/white noise for residuals - for p < 0.05: reject H_0
#> # A tibble: 2 x 5
#>   City             Measure       .model lb_stat lb_pvalue
#>   <chr>            <fct>         <chr>    <dbl>     <dbl>
#> 1 Hohenpeissenberg Temperature   ets       23.0     0.519
#> 2 Hohenpeissenberg Precipitation ets       23.1     0.515
```




Histogram of Forecast Residuals − Hohenpeissenberg Precipitation
<ETS(A,A,A)>
density
0.0100
0.0075
0.0050
0.0025
0.0000
−200
−100
0
100
200
300
Forecast Residuals

3 ARIMA Forecasting Models - AutoRegressive-Integrated Moving Average

Exponential smoothing and ARIMA (AutoRegressive-Integrated Moving Average)models are the two most widely used approaches to time series forecasting, and provide complementary approaches to the problem.

While exponential smoothing models are based on a description of the trend and seasonality in the data, ARIMA models aim to describe the autocorrelations in the data.

3.1 Seasonal ARIMA models

Non-seasonal ARIMA models are generally denoted ARIMA(p,d,q) where parameters p, d, and q are non-negative integers, * p is the order (number of time lags) of the autoregressive model * d is the degree of differencing (number of times the data have had past values subtracted) * q is the order of the moving-average model of past forecast errors .

The value of d has an effect on the prediction intervals — the higher the value of d, the more rapidly the prediction intervals increase in size. For d=0, the point forecasts are equal to the mean of the data and the long-term forecast standard deviation will go to the standard deviation of the historical data, so the prediction intervals will all be essentially the same.

Seasonal ARIMA models are usually denoted ARIMA(p,d,q)(P,D,Q)m, where m refers to the number of periods in each season, and the uppercase P,D,Q refer to the autoregressive, differencing, and moving average terms for the seasonal part of the ARIMA model.

$ARIMA(pdq)(PDQ)$ model with automatically selected $(pdq)(PDQ)$ values

Fit of different pre-defined $ARIMA(pdq)(PDQ)$ models

```
#> Model Selection by Information Criterion - lowest AIC, AICc, BIC
#>  choose p, q parameter accordingly - but only for same d, D values
#> # A tibble: 8 x 8
#>   City             Measure     .model          sigma2 log_lik   AIC  AICc   BIC
#>   <chr>            <fct>       <chr>            <dbl>   <dbl> <dbl> <dbl> <dbl>
#> 1 Hohenpeissenberg Temperature arima_012_011     4.11  -1523. 3055. 3055. 3073.
#> 2 Hohenpeissenberg Temperature arima_111_011     4.11  -1523. 3055. 3055. 3073.
#> 3 Hohenpeissenberg Temperature arima_211_011     4.11  -1523. 3056. 3056. 3079.
#> 4 Hohenpeissenberg Temperature arima_111_012     4.11  -1523. 3057. 3057. 3080.
#> 5 Hohenpeissenberg Temperature arima_012_112     4.12  -1523. 3058. 3058. 3086.
#> 6 Hohenpeissenberg Temperature arima_100_210     5.26  -1594. 3196. 3196. 3214.
#> 7 Hohenpeissenberg Temperature arima_200_011     5.74  -1624. 3255. 3255. 3274.
#> 8 Hohenpeissenberg Temperature arima_100_110_c   5.75  -1623. 3257. 3257. 3280.
#> # A tibble: 8 x 8
#>   City             Measure       .model          sigma2 log_lik   AIC  AICc   BIC
#>   <chr>            <fct>         <chr>            <dbl>   <dbl> <dbl> <dbl> <dbl>
#> 1 Hohenpeissenberg Precipitation arima_012_011   2087.  -3727. 7462. 7463. 7481.
#> 2 Hohenpeissenberg Precipitation arima_111_011   2087.  -3727. 7463. 7463. 7481.
#> 3 Hohenpeissenberg Precipitation arima_211_011   2083.  -3726. 7463. 7463. 7485.
#> 4 Hohenpeissenberg Precipitation arima_111_012   2092.  -3727. 7464. 7465. 7487.
#> 5 Hohenpeissenberg Precipitation arima_012_112   2090.  -3727. 7466. 7466. 7494.
#> 6 Hohenpeissenberg Precipitation arima_200_011   3255.  -3868. 7743. 7743. 7762.
#> 7 Hohenpeissenberg Precipitation arima_100_110~  3260.  -3868. 7745. 7745. 7768.
#> 8 Hohenpeissenberg Precipitation arima_200_110~  3260.  -3868. 7745. 7745. 7768.
```

Good models are obtained by minimising the AIC, AICc or BIC (see glance(fit_arima) output). The preference is to use the AICc to selec p and q.

These information criteria tend not to be good guides to selecting the appropriate order of differencing (d) of a model, but only for selecting the values of p and q. This is because the differencing changes the data on which the likelihood is computed, making the AIC values between models with different orders of differencing not comparable.

3.1.1 Residual Accuracy with one-step-ahead fitted residuals - check RMSE, MAE

Residual accuracy can be computed directly from models as the one-step-ahead fitted residuals are available. Select forecast models that minimises for lowest

- MAE (Mean absolute error, will lead to forecasts of the median) and
- RMSE (Root mean squared error, lead to forecasts of the mean)

```
#> # A tibble: 8 x 7
#>   City              Measure     .model          .type        ME  RMSE   MAE
#>   <chr>             <fct>       <chr>           <chr>     <dbl> <dbl> <dbl>
#> 1 Hohenpeissenberg Temperature arima_012_011   Training 0.127    2.00  1.57
#> 2 Hohenpeissenberg Temperature arima_111_012   Training 0.127    2.00  1.57
#> 3 Hohenpeissenberg Temperature arima_211_011   Training 0.127    2.00  1.57
#> 4 Hohenpeissenberg Temperature arima_111_011   Training 0.127    2.00  1.57
#> 5 Hohenpeissenberg Temperature arima_012_112   Training 0.127    2.00  1.57
#> 6 Hohenpeissenberg Temperature arima_100_210   Training 0.0756   2.27  1.81
#> 7 Hohenpeissenberg Temperature arima_100_110_c Training 0.00110  2.37  1.87
#> 8 Hohenpeissenberg Temperature arima_200_110_c Training 0.00110  2.37  1.87
#> # A tibble: 8 x 7
#>   City              Measure       .model          .type        ME  RMSE   MAE
#>   <chr>             <fct>         <chr>           <chr>     <dbl> <dbl> <dbl>
#> 1 Hohenpeissenberg Precipitation arima_211_011   Training  1.98   45.1  33.8
#> 2 Hohenpeissenberg Precipitation arima_012_112   Training  1.88   45.1  33.8
#> 3 Hohenpeissenberg Precipitation arima_012_011   Training  1.88   45.2  33.8
#> 4 Hohenpeissenberg Precipitation arima_111_011   Training  1.87   45.2  33.9
#> 5 Hohenpeissenberg Precipitation arima_111_012   Training  1.88   45.2  33.9
#> 6 Hohenpeissenberg Precipitation arima_001_002   Training -0.113  54.3  42.3
#> 7 Hohenpeissenberg Precipitation arima_100_110_c Training -0.128  56.5  42.4
#> 8 Hohenpeissenberg Precipitation arima_200_110_c Training -0.128  56.5  42.4
```

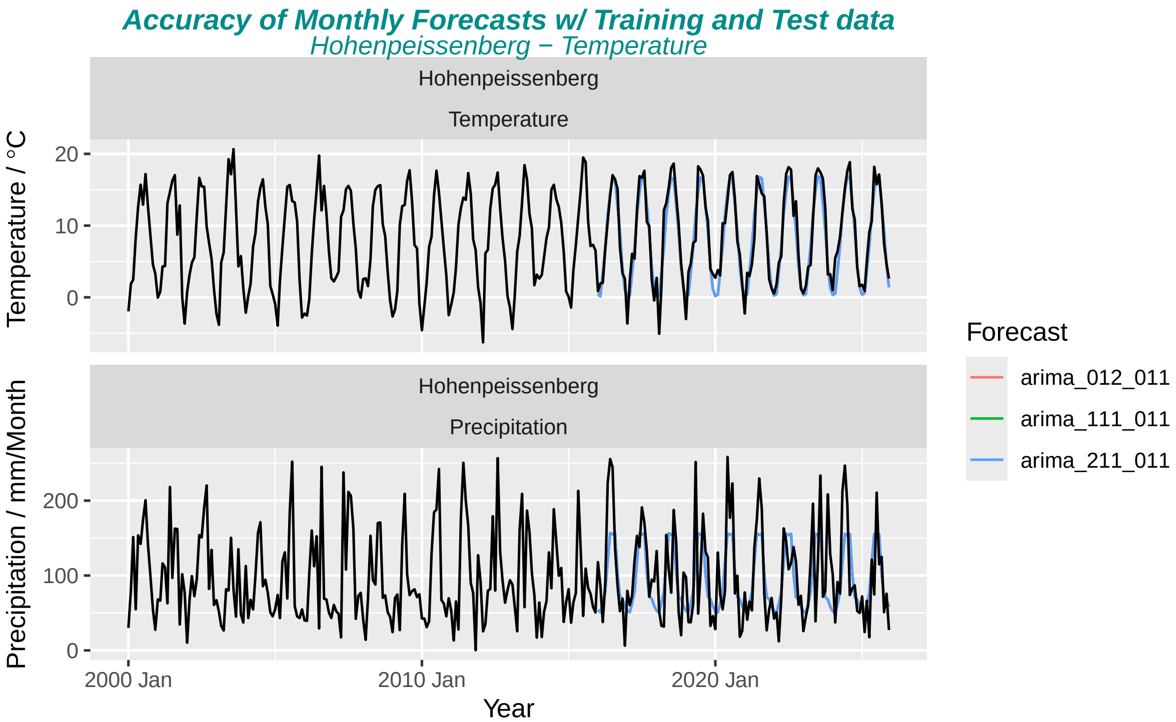
3.1.2 Ljung-Box Test - independence/white noise of the forecasts residuals

```
#> Null Hypothesis of independence/white noise for residuals - for p < 0.05: reject H_0
#> # A tibble: 8 x 5
#>   City              Measure     .model          lb_stat lb_pvalue
#>   <chr>             <fct>       <chr>             <dbl>     <dbl>
#> 1 Hohenpeissenberg Temperature arima_211_011      23.7  3.10e- 1
#> 2 Hohenpeissenberg Temperature arima_012_112      24.0  2.94e- 1
#> 3 Hohenpeissenberg Temperature arima_012_011      24.2  2.81e- 1
#> 4 Hohenpeissenberg Temperature arima_111_011      24.4  2.76e- 1
#> 5 Hohenpeissenberg Temperature arima_111_012      24.4  2.76e- 1
#> 6 Hohenpeissenberg Temperature arima_100_210      47.5  8.01e- 4
#> 7 Hohenpeissenberg Temperature arima_200_011      97.7  7.41e-12
#> 8 Hohenpeissenberg Temperature arima_100_110_c    97.8  6.96e-12
#> # A tibble: 8 x 5
#>   City              Measure       .model        lb_stat lb_pvalue
#>   <chr>             <fct>         <chr>           <dbl>     <dbl>
#> 1 Hohenpeissenberg Precipitation arima_211_011    24.0     0.292
#> 2 Hohenpeissenberg Precipitation arima_012_112    24.6     0.266
#> 3 Hohenpeissenberg Precipitation arima_012_011    25.7     0.216
#> 4 Hohenpeissenberg Precipitation arima_111_011    25.8     0.215
#> 5 Hohenpeissenberg Precipitation arima_111_012    26.0     0.208
#> 6 Hohenpeissenberg Precipitation arima_001_002   236.      0
#> 7 Hohenpeissenberg Precipitation arima_010_110   358.      0
#> 8 Hohenpeissenberg Precipitation arima_012_010   182.      0
```

3.1.3 Forecast Accuracy with Training/Test Data

```
#> # A tibble: 6 x 7
```

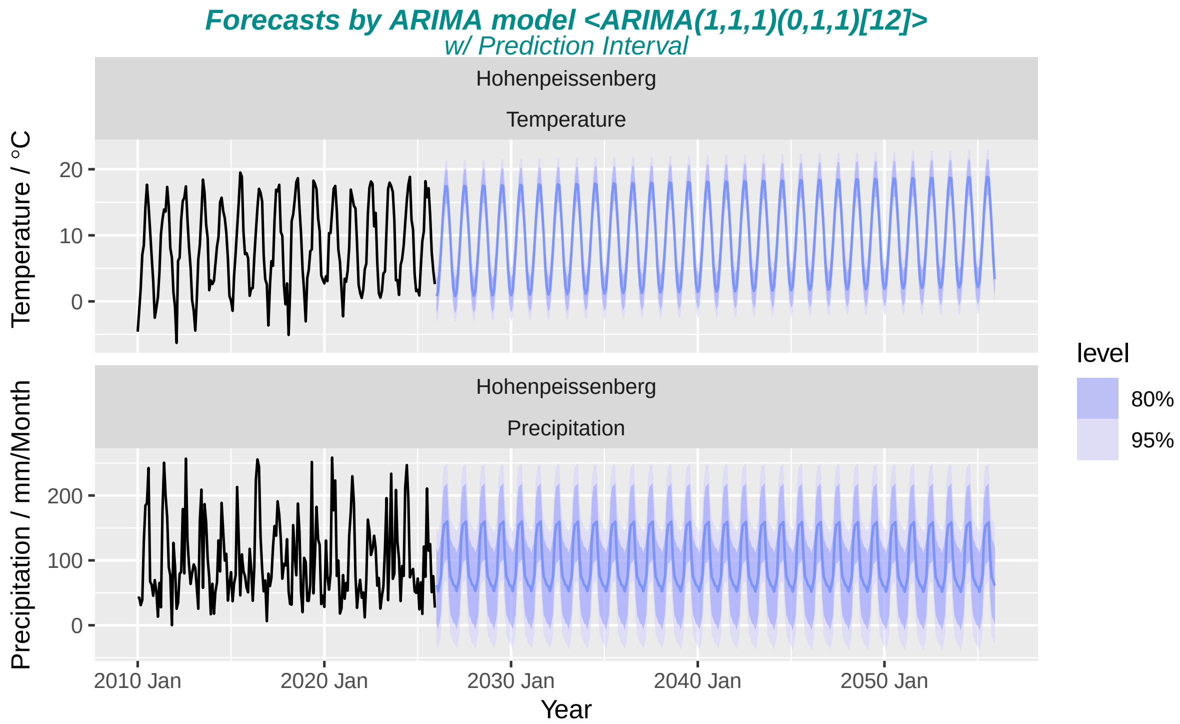
```
#>   .model        City             Measure       .type    ME  RMSE   MAE
#>   <chr>         <chr>            <fct>         <chr> <dbl> <dbl> <dbl>
#> 1 arima_111_011 Hohenpeissenberg Temperature   Test  0.501  1.92  1.54
#> 2 arima_012_011 Hohenpeissenberg Temperature   Test  0.502  1.92  1.54
#> 3 arima_211_011 Hohenpeissenberg Temperature   Test  0.505  1.92  1.55
#> 4 arima_211_011 Hohenpeissenberg Precipitation Test  7.53  47.5  37.1
#> 5 arima_012_011 Hohenpeissenberg Precipitation Test  7.55  47.5  37.2
#> 6 arima_111_011 Hohenpeissenberg Precipitation Test  7.55  47.5  37.2
```



3.2 Temperature, Precipitation - Forecasting with selected ARIMA model <ARIMA(1,1,1)(0,1,1)[12]>, <ARIMA(1,1,1)(0,1,1)[12]>

3.2.1 Forecast Plot of selected ARIMA model

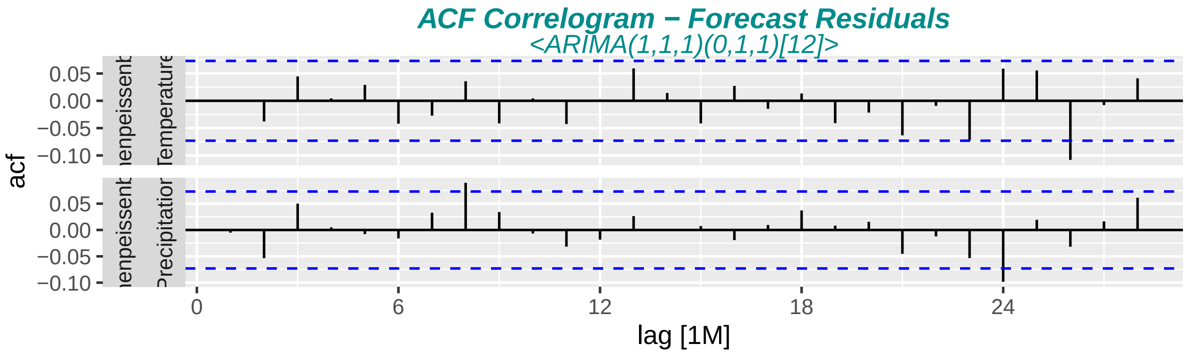
```
#> Provide model coefficients by report(fit_model)
#> # A tibble: 2 x 10
#>   City         Measure .model sigma2 log_lik   AIC  AICc   BIC ar_roots ma_roots
#>   <chr>        <fct>   <chr>   <dbl>   <dbl> <dbl> <dbl> <dbl> <list>   <list>
#> 1 Hohenpeisse~ Temper~ arima  4.11e0  -1523. 3055. 3055. 3073. <cpl>    <cpl>
#> 2 Hohenpeisse~ Precip~ arima  2.09e3  -3727. 7463. 7463. 7481. <cpl>    <cpl>
```

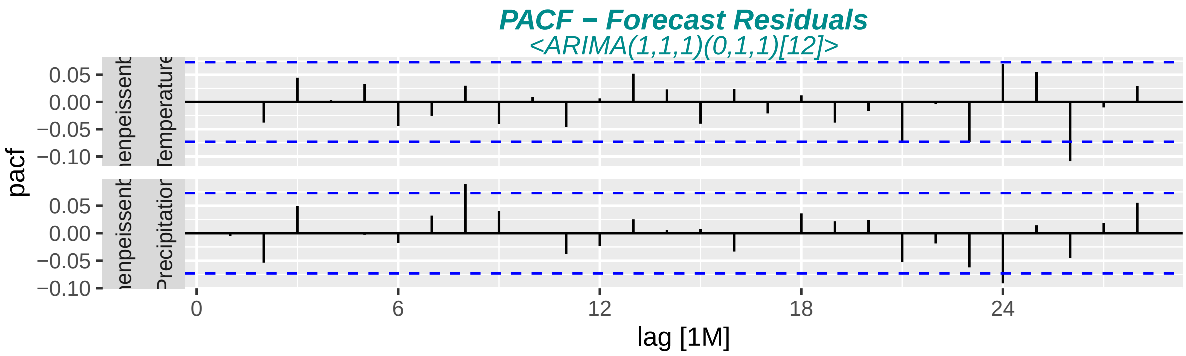


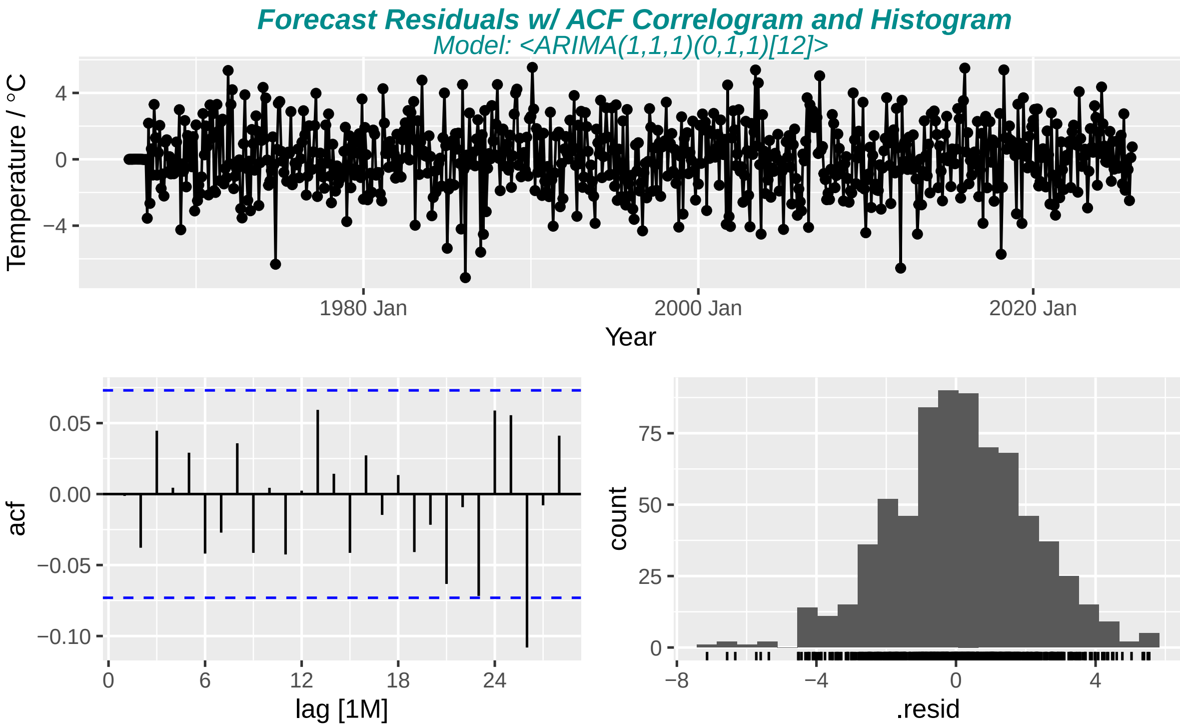
3.2.2 Residual Stationarity

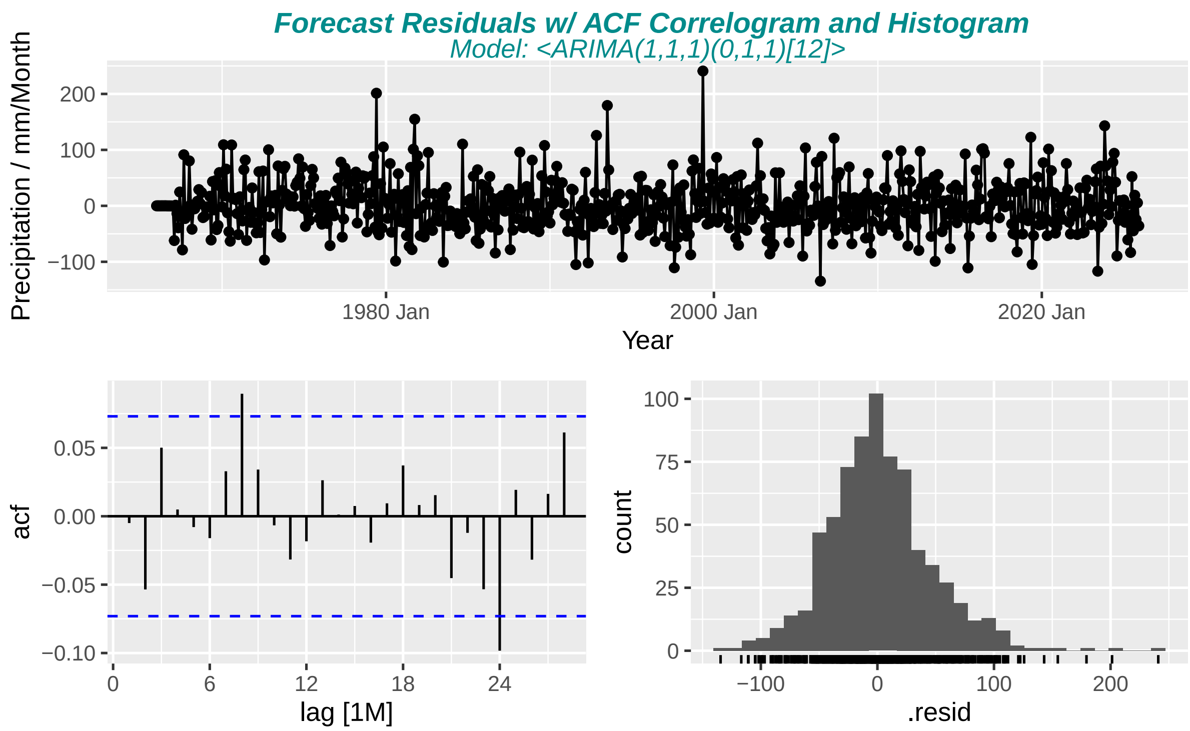
Required checks to be ready for forecasting:

- ACF Forecast Residual: all spikes are within the significance limits, so the residuals appear to be white noise
- The Ljung-Box test also shows that the residuals have no remaining autocorrelations
- Forecast Residuals are more or less normally distributed with roughly centred on zero


ACF Correlogram – Forecast Residuals
<ARIMA(1,1,1)(0,1,1)[12]>
acf
Temperature
Precipitation
lag [1M]

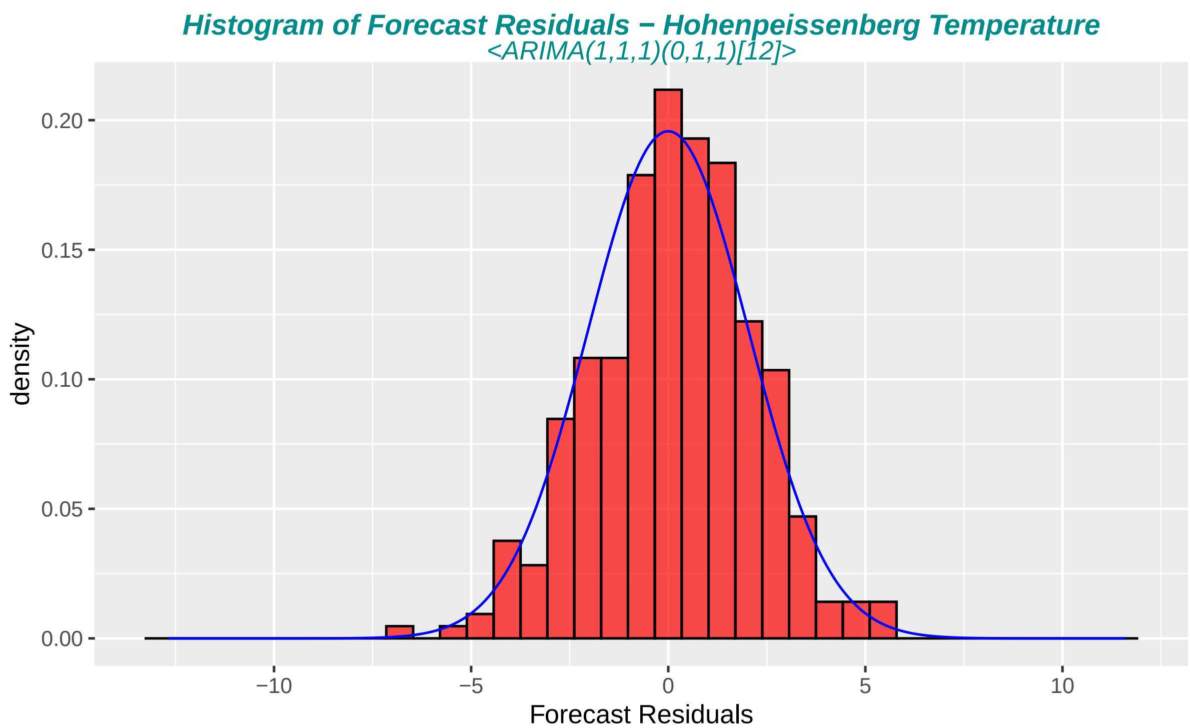

PACF – Forecast Residuals
<ARIMA(1,1,1)(0,1,1)[12]>
pacf
Temperature
Precipitation
lag [1M]

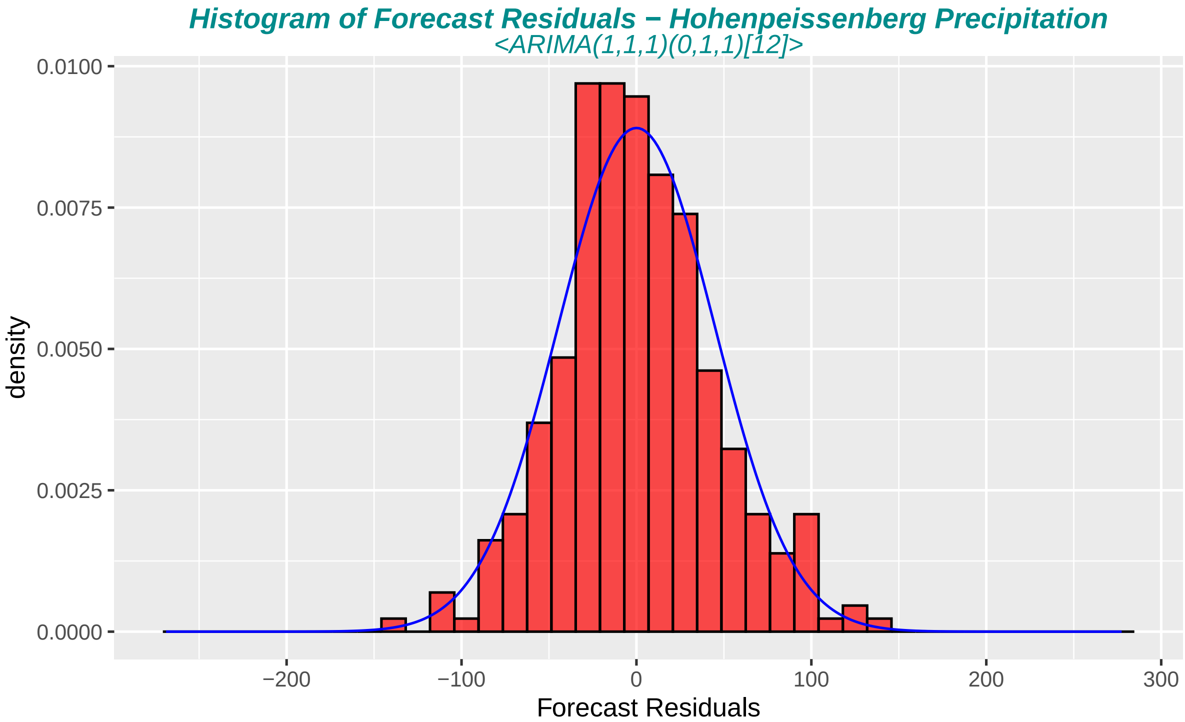

Forecast Residuals w/ ACF Correlogram and Histogram
Model: <ARIMA(1,1,1)(0,1,1)[12]>
Temperature / °C
Year
1980 Jan
2000 Jan
2020 Jan
acf
lag [1M]
count
.resid



3.2.3 Ljung-Box Test and Histogram of forecast residuals with overlaid normal curve

```
#> Null Hypothesis of independence/white noise for residuals - for p < 0.05: reject H_0
#> # A tibble: 2 x 5
#>   City            Measure       .model lb_stat lb_pvalue
#>   <chr>           <fct>         <chr>    <dbl>     <dbl>
#> 1 Hohenpeissenberg Temperature   arima     19.2     0.572
#> 2 Hohenpeissenberg Precipitation arima     23.3     0.330
```




Histogram of Forecast Residuals − Hohenpeissenberg Precipitation
<ARIMA(1,1,1)(0,1,1)[12]>
density
0.0100
0.0075
0.0050
0.0025
0.0000
−200
−100
0
100
200
300
Forecast Residuals

4 ARIMA vs ETS

In particular, all ETS models are non-stationary, while some ARIMA models are stationary.

The ETS models with seasonality or non-damped trend or both have two unit roots (i.e., they need two levels of differencing to make them stationary). All other ETS models have one unit root (they need one level of differencing to make them stationary).

We compare for the chosen ETS rsp. ARIMA model the RMSE / MAE values. Lower values indicate a more accurate model based on the test set RMSE, . . . , MASE.

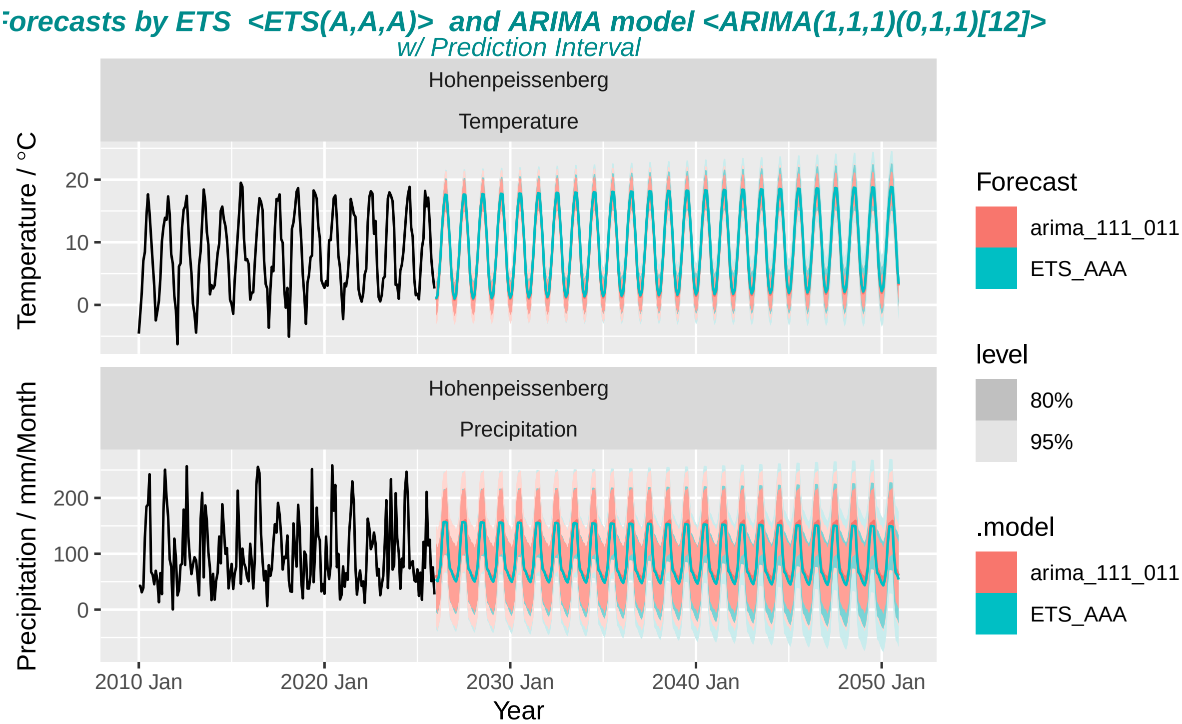
- Residual Accuracy with one-step-ahead fitted residuals
- Forecast Accuracy with Training/Test Data

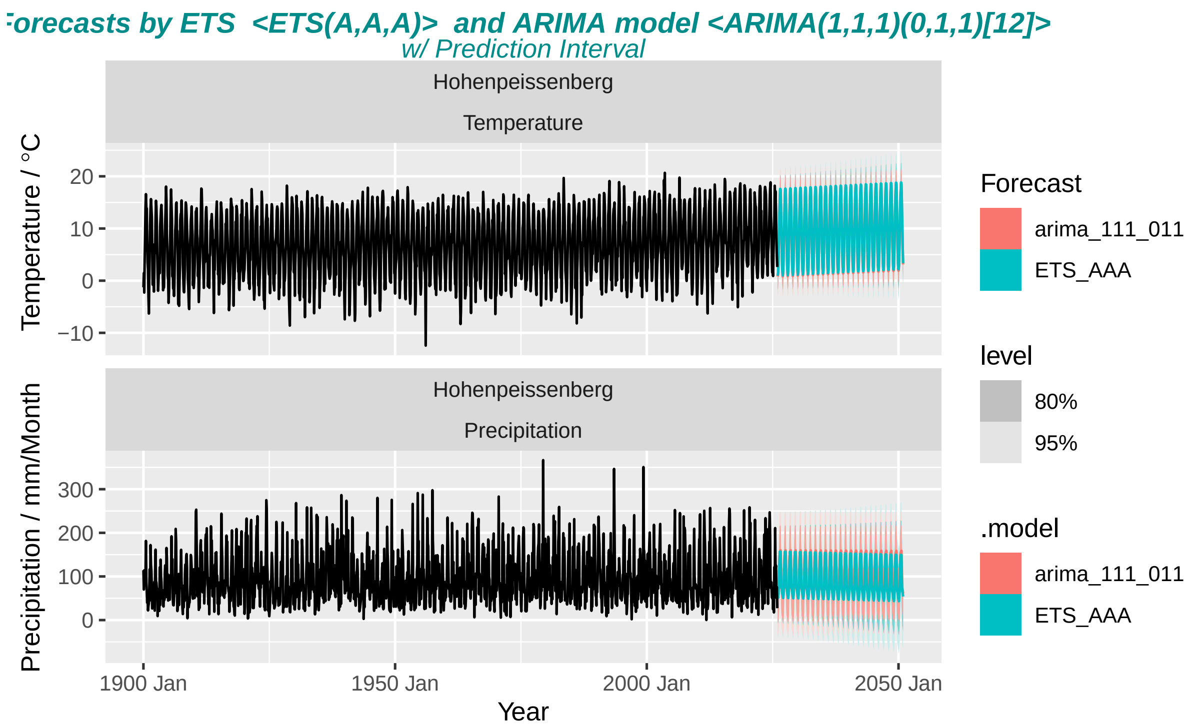
Note: a good fit to training data is never an indication that the model will forecast well. Therefore the values of the Forecast Accuracy are the more relevant one.

4.0.1 Comparing Residual and Forecast Accuracy of selected ETS and ARIMA model

```
#> # A tibble: 8 x 9
#>   City             Measure       .model       .type  RMSE   MAE  MAPE  MASE RMSSE
#>   <chr>            <fct>         <chr>        <chr> <dbl> <dbl> <dbl> <dbl> <dbl>
#> 1 Hohenpeissenberg Temperature   ETS_AAA      Test   1.89  1.51  31.8 0.684 0.675
#> 2 Hohenpeissenberg Temperature   arima_111_~  Test   1.92  1.54  32.7 0.699 0.686
#> 3 Hohenpeissenberg Temperature   arima        Trai~  2.00  1.57 119.  0.710 0.712
#> 4 Hohenpeissenberg Temperature   ets          Trai~  2.01  1.60 133.  0.720 0.713
#> 5 Hohenpeissenberg Precipitation arima        Trai~ 45.2  33.9  106.  0.695 0.704
#> 6 Hohenpeissenberg Precipitation ets          Trai~ 45.5  34.4  107.  0.707 0.709
#> 7 Hohenpeissenberg Precipitation arima_111_~  Test  47.5  37.2   57.3 0.768 0.750
#> 8 Hohenpeissenberg Precipitation ETS_AAA      Test  49.2  38.0   52.8 0.786 0.777
```

4.0.2 Forecast Plot of selected ETS and ARIMA model





Forecasts by ETS <ETS(A,A,A)> and ARIMA model <ARIMA(1,1,1)(0,1,1)[12]>
w/ Prediction Interval
Hohenpeissenberg
Temperature
Temperature / °C
20
10
0
−10
Hohenpeissenberg
Precipitation
Precipitation / mm/Month
300
200
100
0
1900 Jan
1950 Jan
2000 Jan
2050 Jan
Year
Forecast
arima_111_011
ETS_AAA
level
80%
95%
.model
arima_111_011
ETS_AAA

Table 1: Mean values for the given time periods; Units: Temperature (degree C), Precipitation (mm/Month), CO2 (ppm)

Period_Time	Temperature	Precipitation
1781-1810	6.1	54.7
1811-1840	5.6	50.8
1841-1870	5.8	49.2
1871-1900	5.9	72.7
1901-1930	6.1	90.4
1931-1960	6.4	95.6
1961-1990	6.5	100.8
1991-2020	7.7	97.2
2021-2025	8.9	98.5

Table 2: Mean Yearly ARIMA and ETS Forecast values (next 25 years); Units: Temperature (degree C), Precipitation (mm/Month), CO2 (ppm)

City	Measure	Year	ETS_AAA	arima_111_011
Hohenpeissenberg	Temperature	2026	9.10	8.90
Hohenpeissenberg	Temperature	2030	9.30	9.09
Hohenpeissenberg	Temperature	2035	9.56	9.33
Hohenpeissenberg	Temperature	2040	9.82	9.57
Hohenpeissenberg	Temperature	2045	10.08	9.81
Hohenpeissenberg	Temperature	2050	10.33	10.06
Hohenpeissenberg	Precipitation	2026	95.73	97.68
Hohenpeissenberg	Precipitation	2030	94.48	97.40
Hohenpeissenberg	Precipitation	2035	92.93	97.17
Hohenpeissenberg	Precipitation	2040	91.37	96.94
Hohenpeissenberg	Precipitation	2045	89.81	96.71
Hohenpeissenberg	Precipitation	2050	88.25	96.47

Table 3: Forecast increase/decrease over the next 25 years; Units: Temperature (degree C), Precipitation (mm/Month), CO2 (ppm)

Measure	Year.x	Year.y	ETS.x	ARIMA.x	ETS.y	ARIMA.y	Delta_ETS	Delta_ARIMA
Temperature	2026	2050	9.10	8.90	10.33	10.06	1.23	1.16
Precipitation	2026	2050	95.73	97.68	88.25	96.47	-7.48	-1.21

Table 4: Forecast increase/decrease over the next 25 years; Units: Temperature (degree C), Precipitation (mm/Month), CO2 (ppm)

Measure	Month	Year.x	Year.y	ETS.x	ARIMA.x	ETS.y	ARIMA.y	Delta_ETS	Delta_ARIMA
Temperature	Jan	2026	2050	0.95	0.80	2.18	1.92	1.23	1.12
Temperature	Feb	2026	2050	1.57	1.39	2.80	2.55	1.23	1.16
Temperature	Mrz	2026	2050	4.41	4.22	5.64	5.39	1.23	1.16
Temperature	Apr	2026	2050	7.83	7.70	9.06	8.86	1.23	1.16
Temperature	Mai	2026	2050	12.33	12.01	13.56	13.17	1.23	1.16
Temperature	Jun	2026	2050	15.65	15.64	16.88	16.80	1.23	1.16
Temperature	Jul	2026	2050	17.60	17.43	18.83	18.59	1.23	1.16
Temperature	Aug	2026	2050	17.53	17.29	18.76	18.46	1.23	1.16
Temperature	Sep	2026	2050	13.96	13.63	15.20	14.79	1.23	1.16

Measure	Month	Year.x	Year.y	ETS.x	ARIMA.x	ETS.y	ARIMA.y	Delta_ETS	Delta_ARIMA
Temperature	Okt	2026	2050	10.25	9.94	11.49	11.11	1.23	1.16
Temperature	Nov	2026	2050	4.99	4.77	6.23	5.93	1.23	1.16
Temperature	Dez	2026	2050	2.12	1.95	3.35	3.11	1.23	1.16
Precipitation	Jan	2026	2050	54.71	61.91	47.24	59.56	-7.48	-2.35
Precipitation	Feb	2026	2050	50.76	52.69	43.29	51.62	-7.48	-1.07
Precipitation	Mrz	2026	2050	60.13	62.74	52.66	61.62	-7.48	-1.11
Precipitation	Apr	2026	2050	80.63	76.02	73.15	74.91	-7.48	-1.11
Precipitation	Mai	2026	2050	126.11	135.27	118.63	134.16	-7.48	-1.11
Precipitation	Jun	2026	2050	156.81	154.98	149.34	153.87	-7.48	-1.11
Precipitation	Jul	2026	2050	156.96	156.93	149.48	155.82	-7.48	-1.11
Precipitation	Aug	2026	2050	155.86	159.98	148.38	158.87	-7.48	-1.11
Precipitation	Sep	2026	2050	102.41	104.94	94.93	103.83	-7.48	-1.11
Precipitation	Okt	2026	2050	73.50	75.32	66.03	74.21	-7.48	-1.11
Precipitation	Nov	2026	2050	69.53	69.36	62.05	68.25	-7.48	-1.11
Precipitation	Dez	2026	2050	61.33	62.08	53.85	60.97	-7.48	-1.11

5 Yearly Data Forecasts with ARIMA and ETS

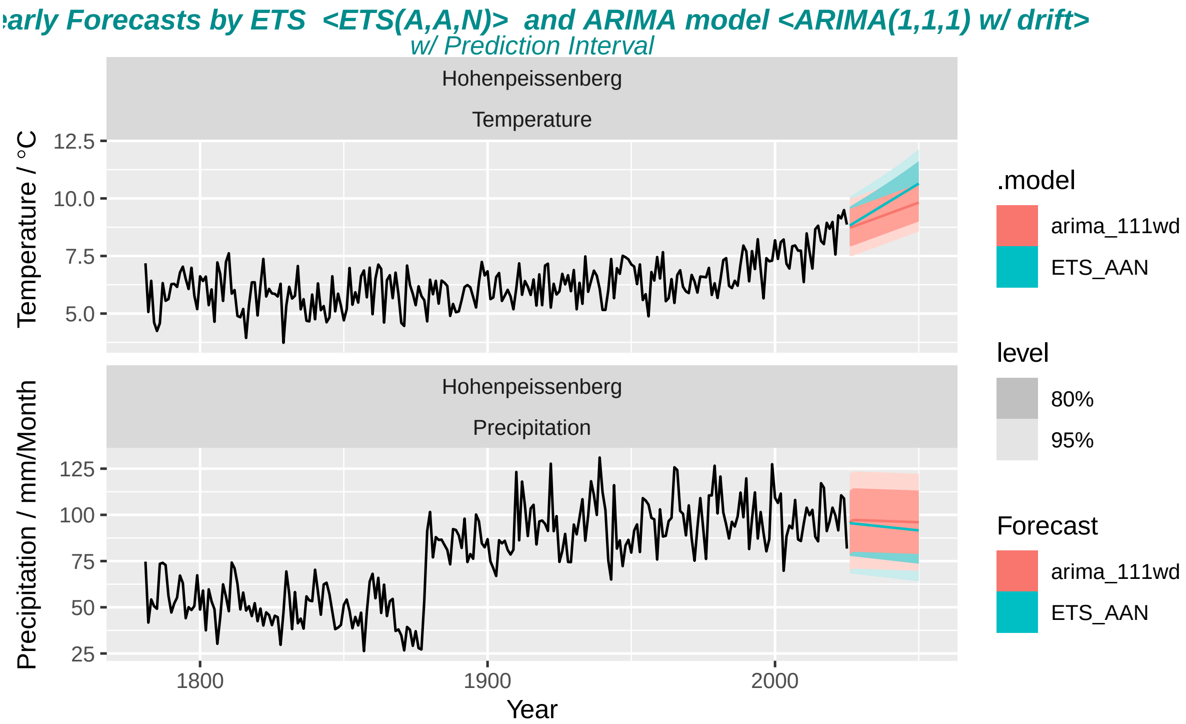
For yearly data the seasonal monthly data are replaced by the yearly average data. Therefore the seasonal component of the ETS and ARIMA model are to be taken out.

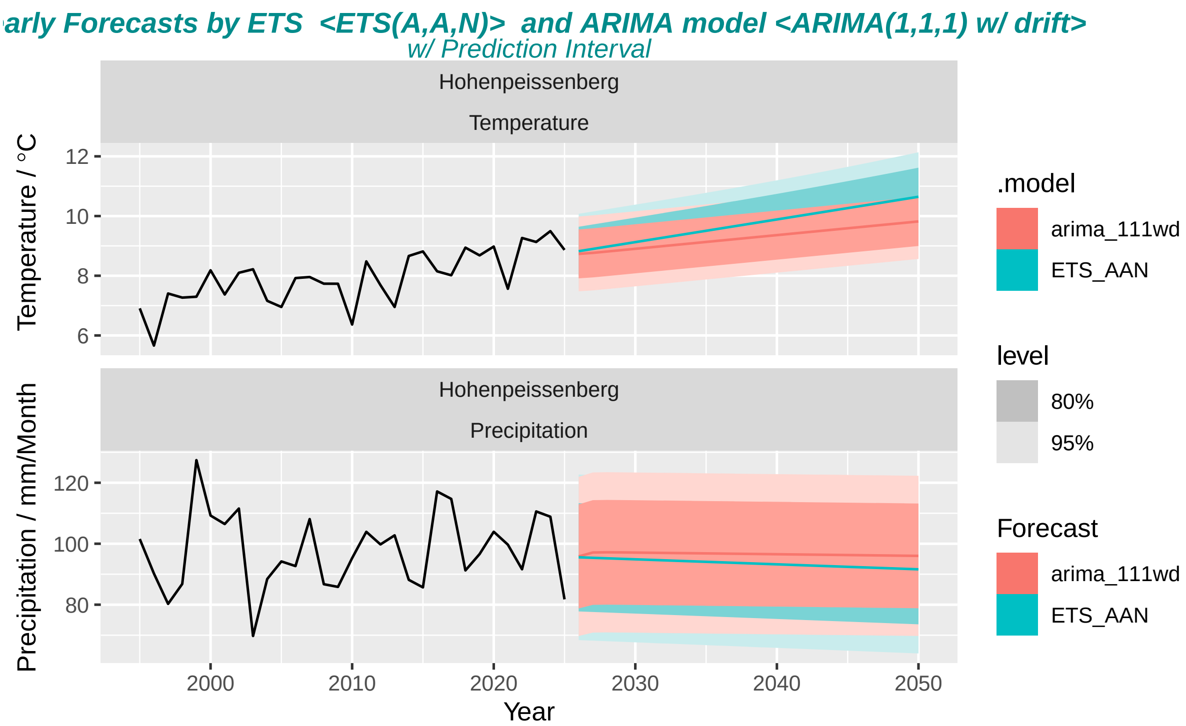
The ETS model $< ETS(A, A, N) >$ with seasonal term change “A” -> “N” is chosen. For ARIMA models the seasonal term (P,D,Q)m has to be taken out and an optimal ARIMA(p,1,q) with one differencing (d=1) is selected. However, for Mauna Loa two times differncing had to be selected $CO_2 <ARIMA(0,2,1) w/ poly>. For Temperature and Precipitation the same model as for monthly data can be taken by leaving out the seasonal term $< ARIMA(0, 1, 2)w/drift >$.

5.0.1 Comparing Residual and Forecast Accuracy of selected ETS and ARIMA model

5.0.2 Forecast Plot of selected ETS and ARIMA model

```
#> selected: ETS_AAN and arima_111 w/ drift
```






```
#> # A tibble: 4 x 13
#>   City   Measure .model  sigma2 log_lik   AIC  AICc   BIC     MSE    AMSE    MAE
#>   <chr>  <fct>   <chr>    <dbl>   <dbl> <dbl> <dbl> <dbl>   <dbl>   <dbl>  <dbl>
#> 1 Hohen~ Temper~ arima~   0.402   -57.3  123.  123.  131.  NA      NA      NA
#> 2 Hohen~ Temper~ ETS_A~   0.409   -93.9  198.  199.  208.   0.382   0.381  0.502
#> 3 Hohen~ Precip~ arima~ 175.     -236.   481.  482.  489.  NA      NA      NA
#> 4 Hohen~ Precip~ ETS_A~ 192.     -278.   567.  568.  577. 179.    180.     11.1
#> # i 2 more variables: ar_roots <list>, ma_roots <list>
```

Ljung-Box Test - independence/white noise of the forecasts residuals

```
#> # A tibble: 2 x 5
#>   City             Measure       .model  lb_stat lb_pvalue
#>   <chr>            <fct>         <chr>     <dbl>     <dbl>
#> 1 Hohenpeissenberg Temperature   ETS_AAN    19.7     0.714
#> 2 Hohenpeissenberg Precipitation ETS_AAN    14.1     0.945
#> # A tibble: 2 x 5
#>   City             Measure       .model      lb_stat lb_pvalue
#>   <chr>            <fct>         <chr>         <dbl>     <dbl>
#> 1 Hohenpeissenberg Temperature   arima_111wd    18.5     0.617
#> 2 Hohenpeissenberg Precipitation arima_111wd    12.7     0.919
```

Table 5: Mean Yearly ARIMA and ETS Forecast values of mean yearly data (next 25 years); Units: Temperature (degree C), Precipitation (mm/Month), CO2 (ppm)

City	Measure	Year	ETS_AAN	arima_111wd
Hohenpeissenberg	Temperature	2026	8.82	8.73
Hohenpeissenberg	Temperature	2030	9.13	8.90
Hohenpeissenberg	Temperature	2035	9.51	9.13
Hohenpeissenberg	Temperature	2040	9.89	9.36
Hohenpeissenberg	Temperature	2045	10.27	9.59
Hohenpeissenberg	Temperature	2050	10.65	9.82
Hohenpeissenberg	Precipitation	2026	95.54	95.84
Hohenpeissenberg	Precipitation	2030	94.88	97.09
Hohenpeissenberg	Precipitation	2035	94.06	96.82
Hohenpeissenberg	Precipitation	2040	93.24	96.55
Hohenpeissenberg	Precipitation	2045	92.42	96.28
Hohenpeissenberg	Precipitation	2050	91.60	96.01

Table 6: Forecast increase/decrease over the next 25 years; Units: Temperature (degree C), Precipitation (mm/Month), CO2 (ppm)

Measure	Year.x	Year.y	ETS.x	ARIMA.x	ETS.y	ARIMA.y	Delta_ETS	Delta_ARIMA
Temperature	2026	2050	8.82	8.73	10.65	9.82	1.82	1.09
Precipitation	2026	2050	95.54	95.84	91.60	96.01	-3.94	0.18