Climate Data Forecasting -

Atmospheric CO_2 Concentration / Temperature / Precipitation

Wolfgang Vollmer

2026-01-08

Contents

1 Forecasting of Mannheim - 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

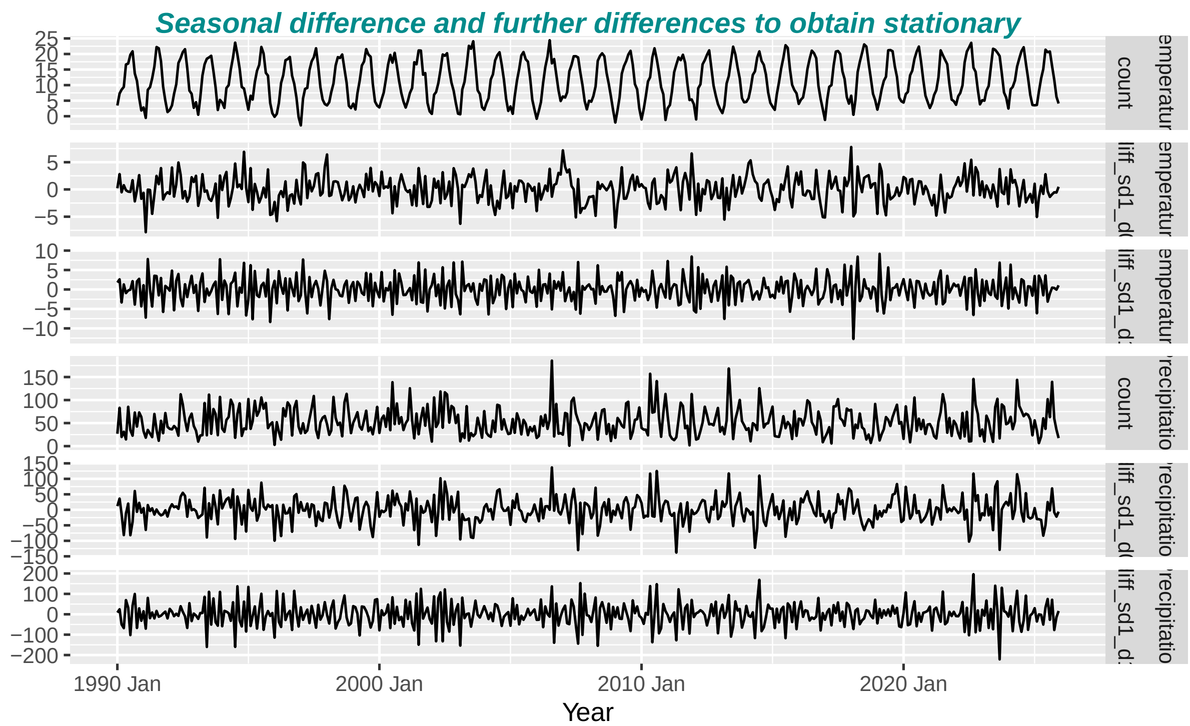
The model fit residuals have to be stationary. For good forecasting this has to be verified with residual diagnostics.

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        1.04        0.01       1      1
#> 2 Precipitation      2.98        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     0.00439         0.1
#> 2 Precipitation   0.00459         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

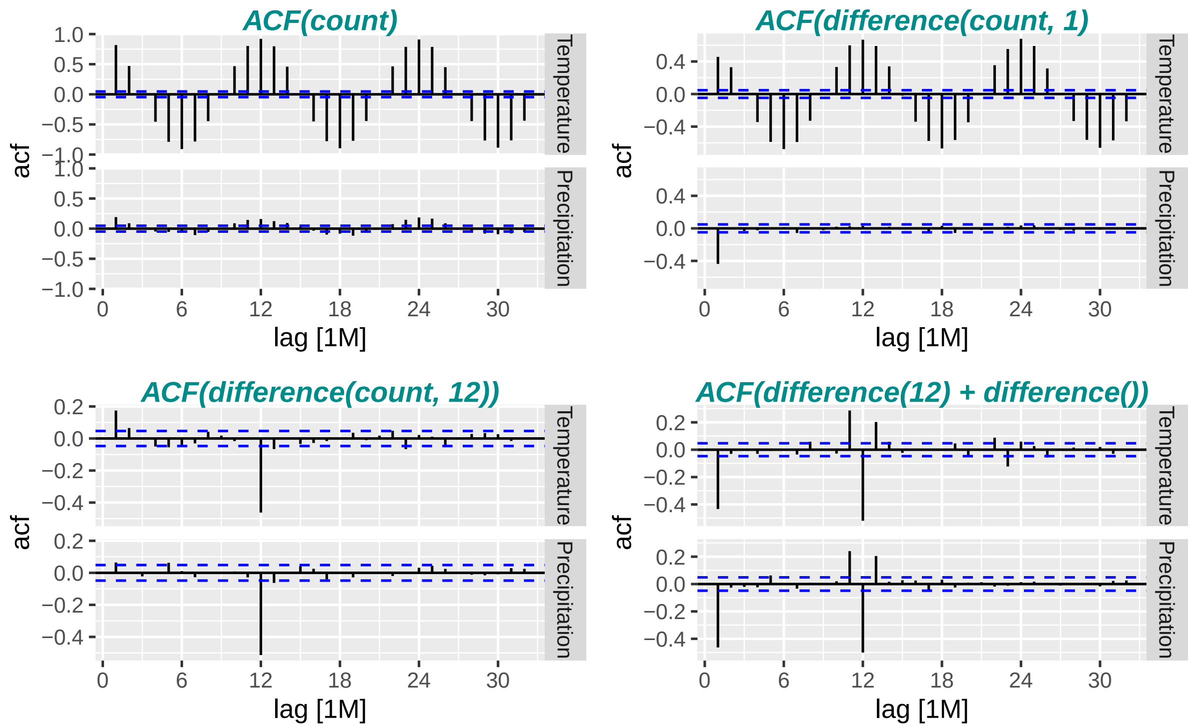
- augment(fit) |> features(.innov, ljung_box, lag=x, dof=y) (Ch. 5.4 Residaul diagnostics)
  - portmanteau test suggesting that the residuals are white noise
  - *Null Hypothesis* H_0: independence/white noise for residuals, i.e. each autocorrelation in the time series residuals of the model for lag 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     6247.         0
#> 2 Precipitation    125.         0
#> Ljung-Box test on (difference(count, 12))
```

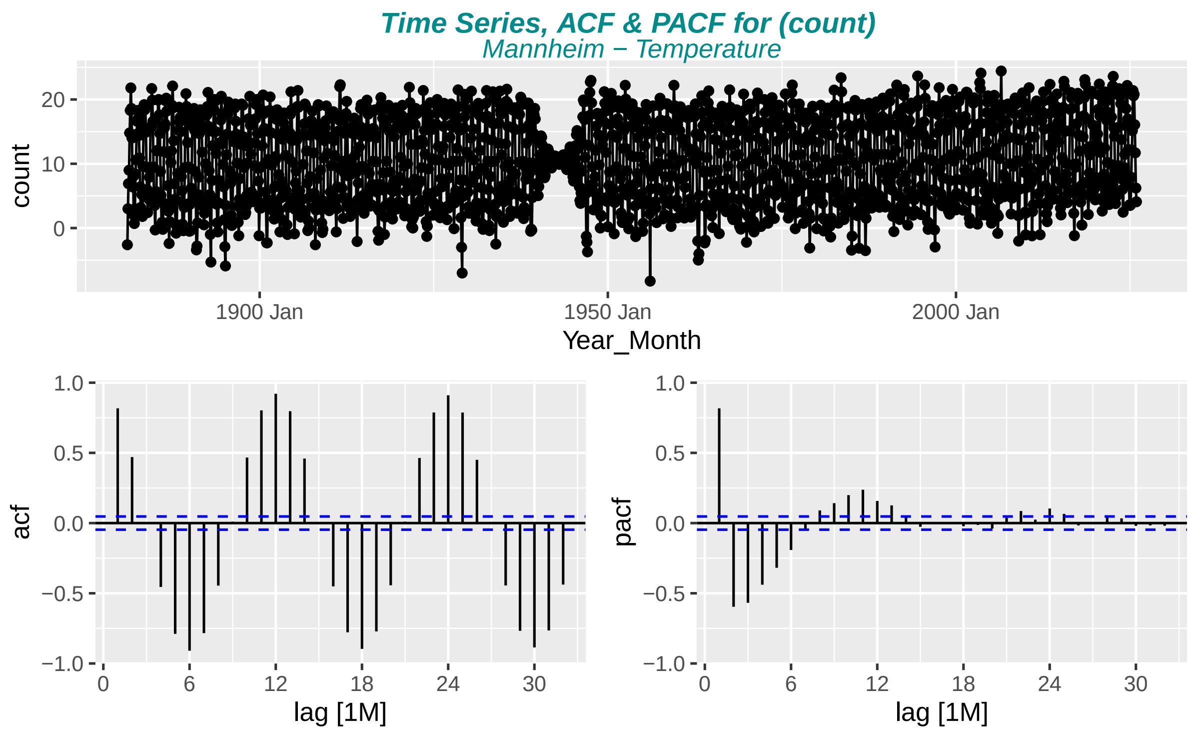
```
#> # A tibble: 2 x 3
#>   Measure       lb_stat lb_pvalue
#>   <fct>           <dbl>     <dbl>
#> 1 Temperature      79.2  7.26e-13
#> 2 Precipitation    15.6  1.12e- 1
#> Ljung-Box test on (difference(count, 12) + difference())
#> # A tibble: 2 x 3
#>   Measure       lb_stat lb_pvalue
#>   <fct>           <dbl>     <dbl>
#> 1 Temperature      338.         0
#> 2 Precipitation    360.         0
```

1.1.3 ACF (Autocorrelation Function) Plots of Differences

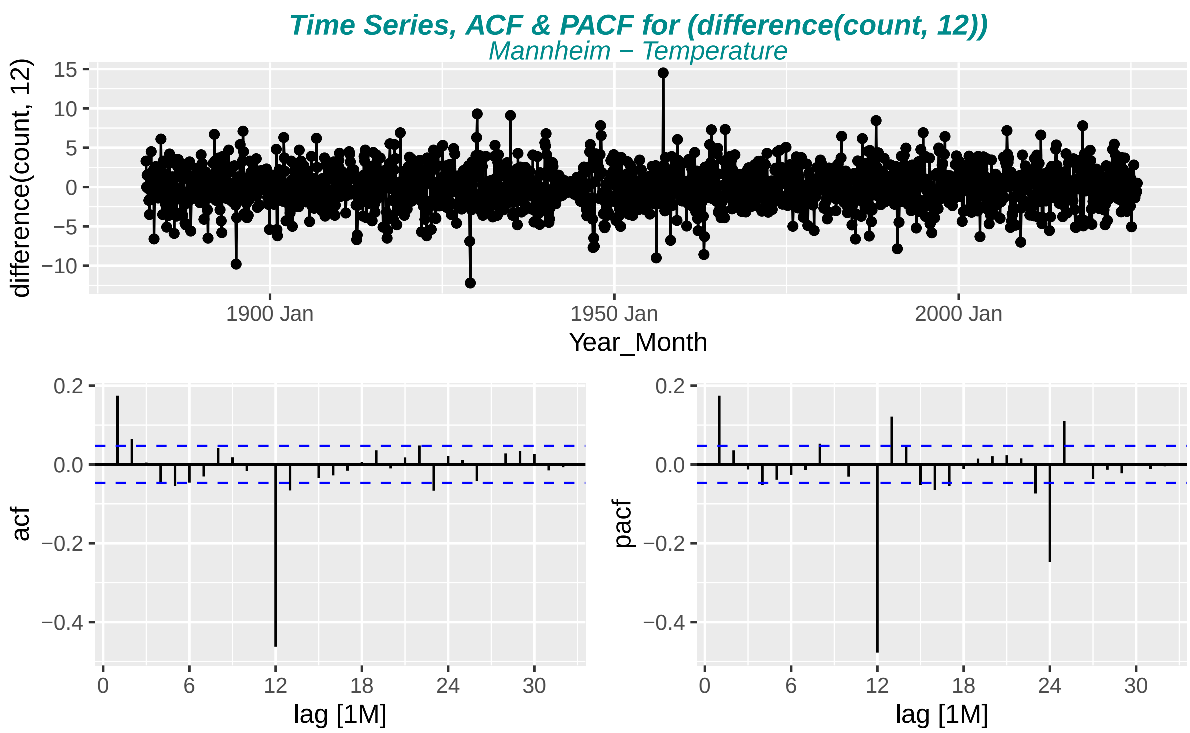


1.1.4 Time Series, ACF and PACF (Partial) Plots of Differences - for ARIMA p, q check

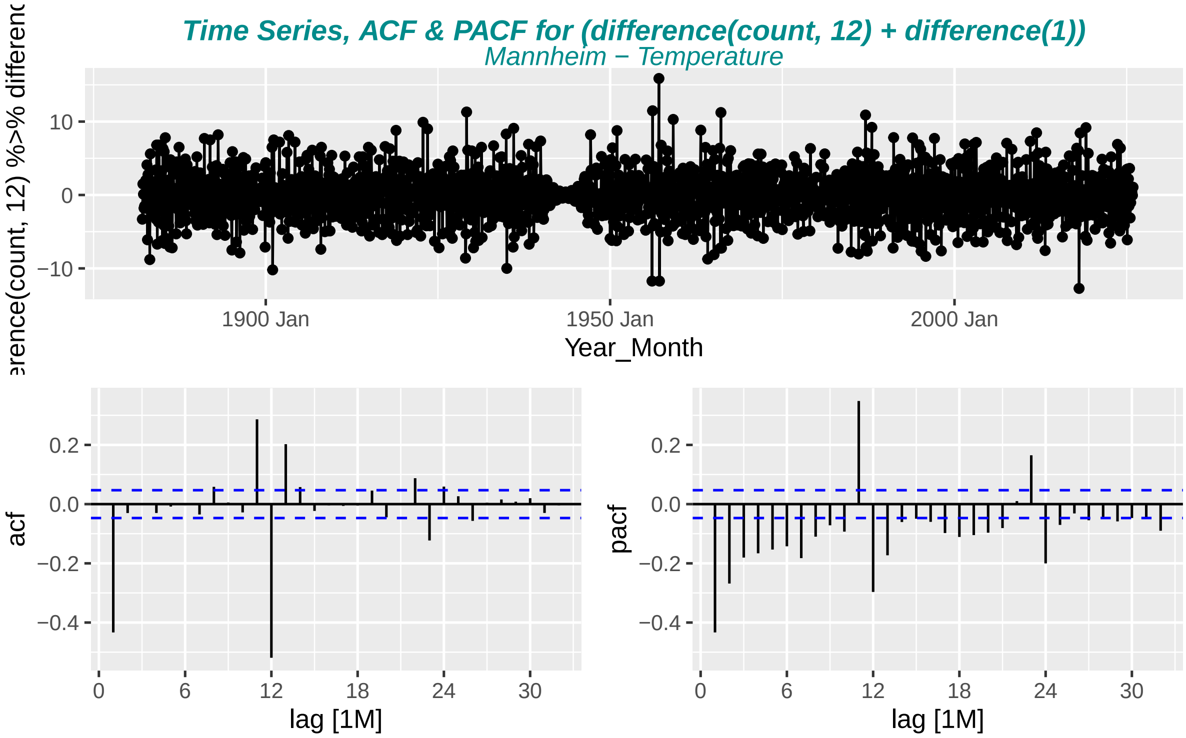
```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City     Measure        Sum  Mean
#>   <chr>    <fct>        <dbl> <dbl>
#> 1 Mannheim Temperature 18256.  10.5
```



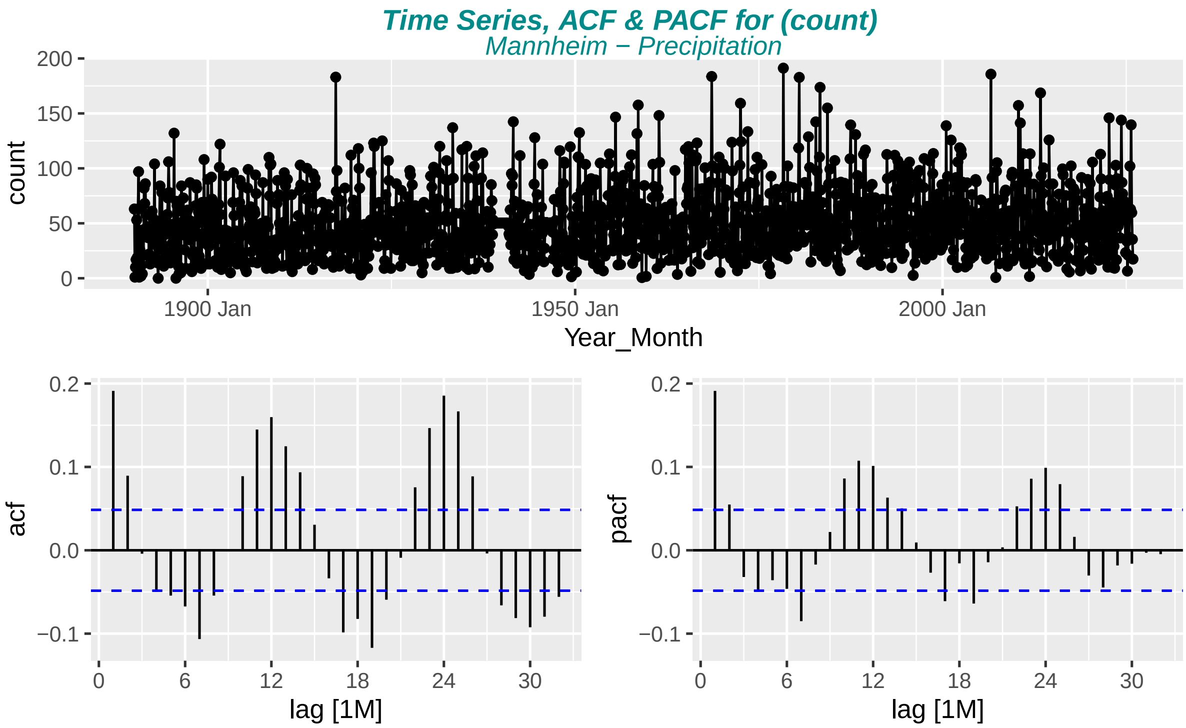

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City     Measure       Sum   Mean
#>   <chr>    <fct>       <dbl>  <dbl>
#> 1 Mannheim Temperature  23.0 0.0133
```



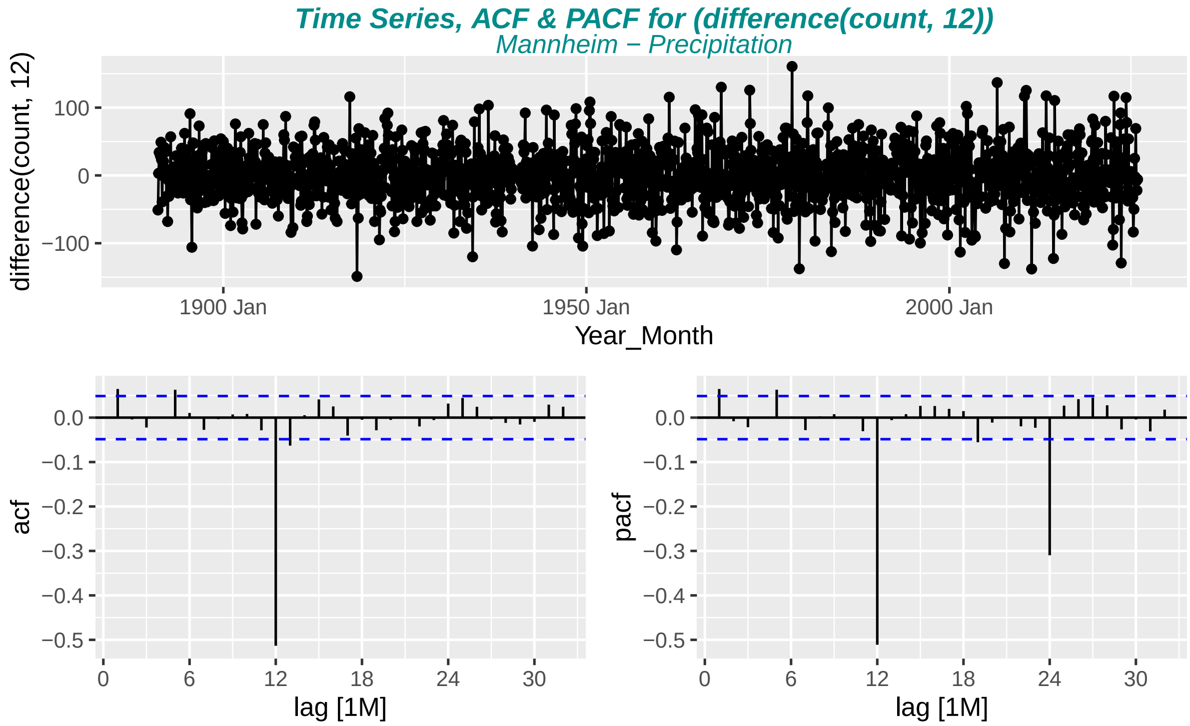

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City     Measure       Sum     Mean
#>   <chr>    <fct>       <dbl>    <dbl>
#> 1 Mannheim Temperature -2.81 -0.00163
```



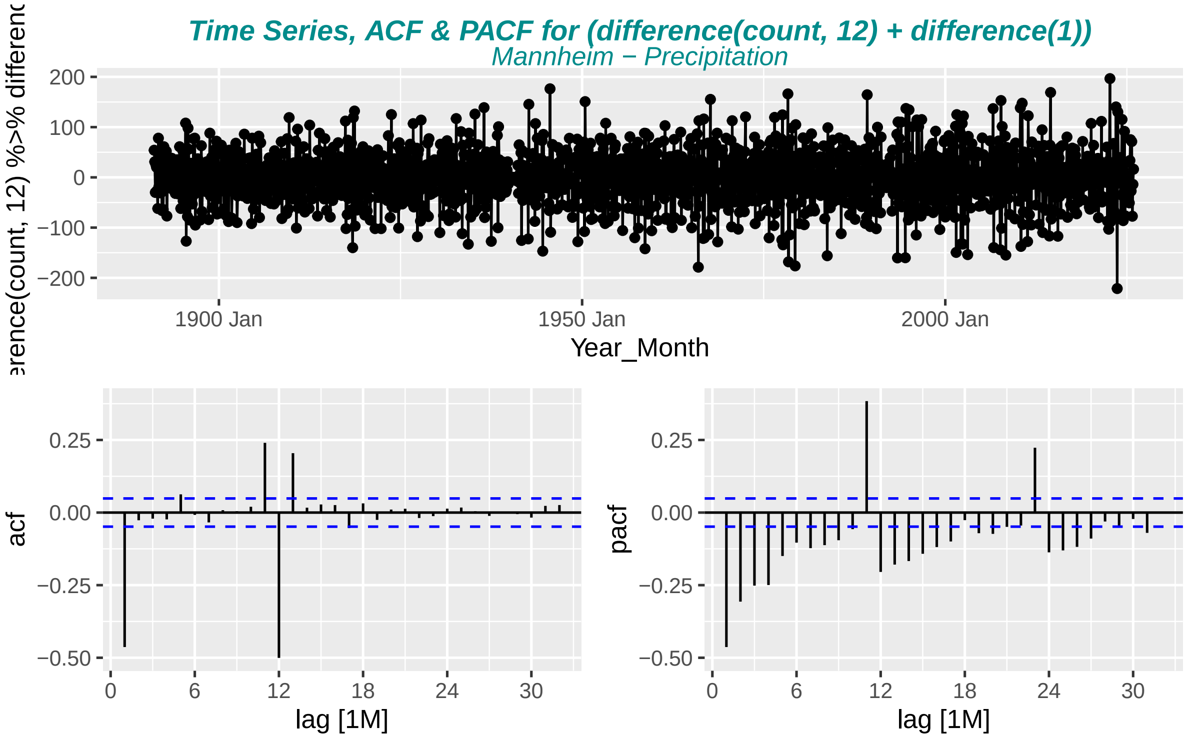

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City     Measure          Sum  Mean
#>   <chr>    <fct>          <dbl> <dbl>
#> 1 Mannheim Precipitation 82020.  50.3
```




```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City     Measure         Sum  Mean
#>   <chr>    <fct>         <dbl> <dbl>
#> 1 Mannheim Precipitation  241. 0.149
```




```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City     Measure         Sum   Mean
#>   <chr>    <fct>         <dbl>  <dbl>
#> 1 Mannheim Precipitation  45.1 0.0279
```



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 Mannheim Temperature   <ETS(A,N,A)>
#> 2 Mannheim Precipitation <ETS(M,N,A)>
#> [1] "Mannheim Temperature"
#> Series: count
#> Model: ETS(A,N,A)
#>   Smoothing parameters:
#>     alpha = 0.04222888
#>     gamma = 0.0001000266
#>
#>   Initial states:
#>      l[0]      s[0]     s[-1]       s[-2]    s[-3]    s[-4]   s[-5]    s[-6]
#>  11.04648 -7.862992 -5.172378 -0.08741953 4.689572 8.832004 9.35363 7.419503
#>     s[-7]      s[-8]     s[-9]    s[-10]    s[-11]
#>  4.071809 -0.4590285 -4.195475 -7.707884 -8.881342
#>
#>   sigma^2:  2.8739
#>
#>      AIC     AICc      BIC
#> 5513.016 5513.698 5581.705
#> [1] "Mannheim Precipitation"
#> Series: count
#> Model: ETS(M,N,A)
#>   Smoothing parameters:
#>     alpha = 0.0001001404
#>     gamma = 0.0001000175
#>
#>   Initial states:
#>      l[0]      s[0]     s[-1]     s[-2]     s[-3]    s[-4]    s[-5]   s[-6]
#>  55.72239 -3.574711 -3.051401 -2.126575 -0.827569 12.85195 17.71642 10.1291
#>     s[-7]     s[-8]     s[-9]    s[-10]    s[-11]
#>  19.30158 -8.729449 -12.71324 -13.72795 -15.24815
#>
#>   sigma^2:  0.278
#>
#>      AIC     AICc      BIC
#> 9589.060 9589.741 9657.748
#> # A tibble: 2 x 8
#>   City     Measure       .model   AIC  AICc   BIC    MSE   MAE
#>   <chr>    <fct>         <chr>  <dbl> <dbl> <dbl>  <dbl> <dbl>
#> 1 Mannheim Temperature   ETS    5513. 5514. 5582.   2.82 1.34
#> 2 Mannheim Precipitation ETS    9589. 9590. 9658. 879.   0.418
```

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

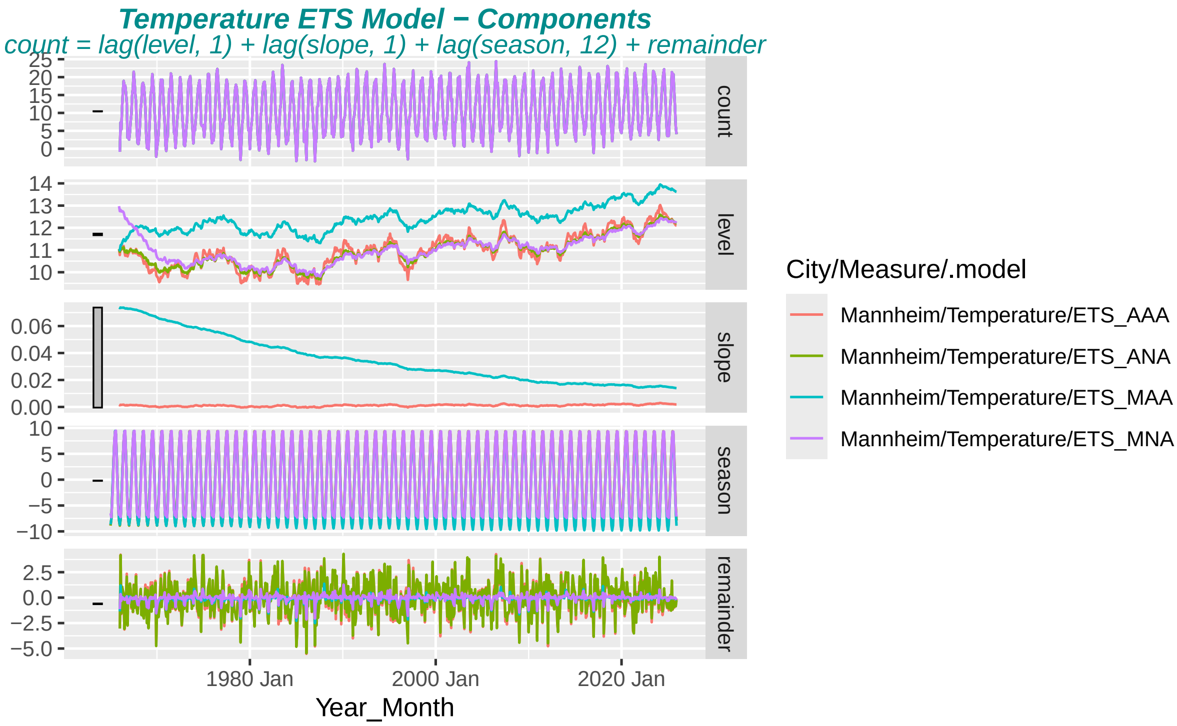
Model Selection by Information Criterion - lowest AIC, AICc, BIC

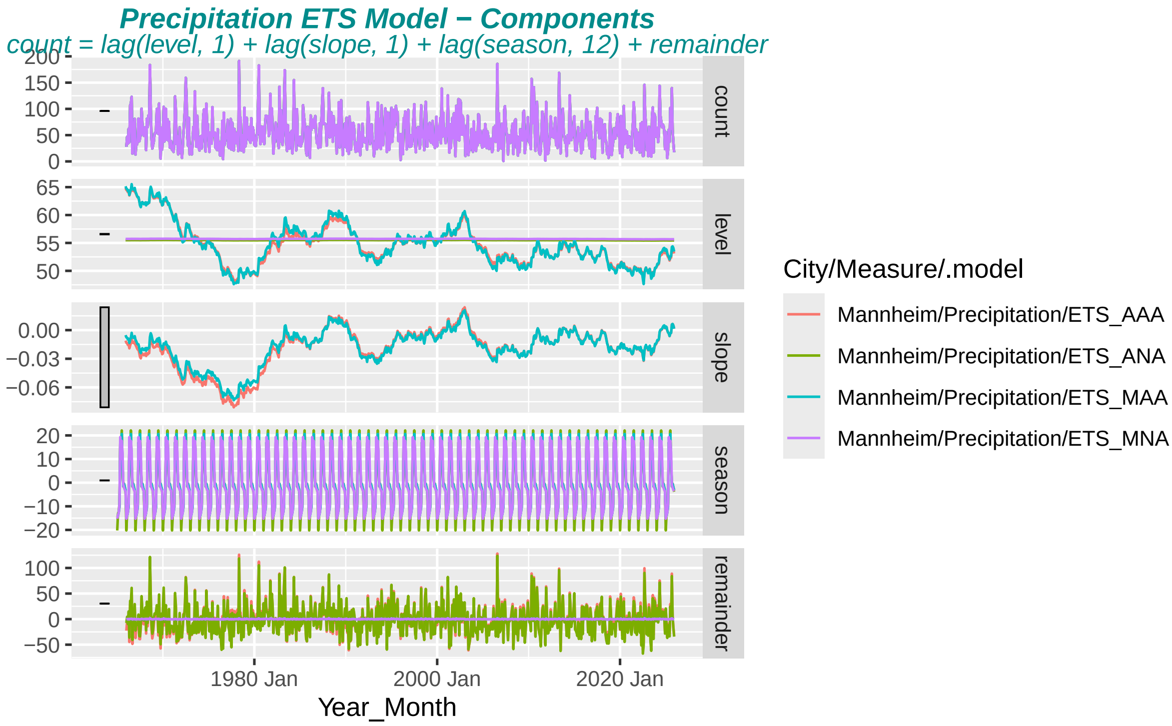
Best model fit with

- 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 Mannheim Temperature   ETS_ANA  5513. 5514. 5582.   2.82   2.82  1.34
#>  2 Mannheim Temperature   ETS_AAdA 5518. 5519. 5601.   2.81   2.82  1.34
#>  3 Mannheim Temperature   ETS_AAA  5521. 5522. 5599.   2.83   2.85  1.35
#>  4 Mannheim Temperature   ETS_AMA  5526. 5527. 5604.   2.85   2.88  1.35
#>  5 Mannheim Temperature   ETS_MAdA 6508. 6509. 6590.   4.25   4.22  0.231
#>  6 Mannheim Temperature   ETS_MNA  6527. 6527. 6595.   3.71   3.69  0.227
#>  7 Mannheim Temperature   ETS_MMA  6670. 6671. 6748.   3.80   3.87  0.224
#>  8 Mannheim Temperature   ETS_MAA  6670. 6671. 6748.   3.77   3.84  0.225
#>  9 Mannheim Precipitation ETS_MNA  9589. 9590. 9658. 879.   880.    0.418
#> 10 Mannheim Precipitation ETS_MAdA 9603. 9604. 9685. 878.   879.    0.416
#> 11 Mannheim Precipitation ETS_MMA  9613. 9614. 9691. 884.   886.    0.426
#> 12 Mannheim Precipitation ETS_MAA  9615. 9616. 9693. 885.   887.    0.425
#> 13 Mannheim Precipitation ETS_ANA  9645. 9646. 9714. 876.   877.   23.0
#> 14 Mannheim Precipitation ETS_AAdA 9649. 9650. 9731. 873.   874.   22.9
#> 15 Mannheim Precipitation ETS_AMA  9657. 9658. 9734. 885.   887.   23.0
#> 16 Mannheim Precipitation ETS_AAA  9657. 9658. 9735. 885.   887.   23.0
```





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)

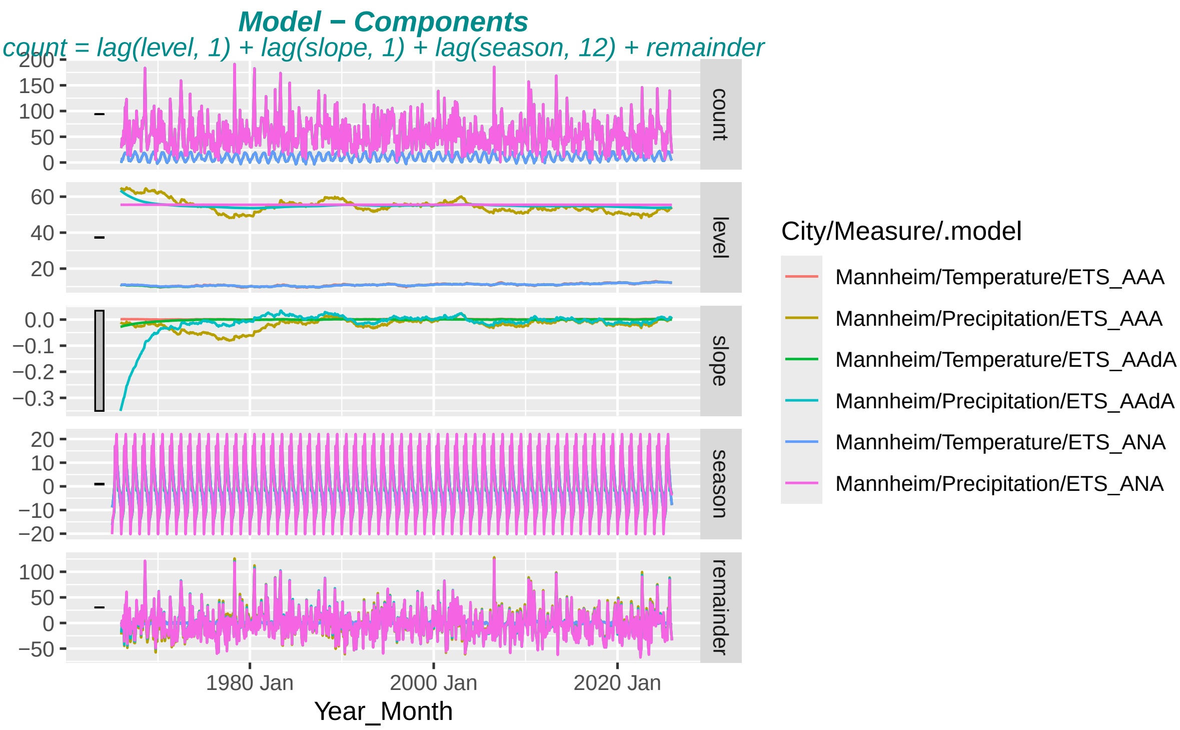
```
#> # A tibble: 16 x 7
#>    City     Measure       .model   .type          ME  RMSE   MAE
#>    <chr>    <fct>         <chr>    <chr>       <dbl> <dbl> <dbl>
#>  1 Mannheim Temperature   ETS_AAdA Training  0.0610   1.68  1.34
#>  2 Mannheim Temperature   ETS_ANA  Training  0.0379   1.68  1.34
#>  3 Mannheim Temperature   ETS_AAA  Training  0.00458  1.68  1.35
#>  4 Mannheim Temperature   ETS_AMA  Training -0.0724   1.69  1.35
#>  5 Mannheim Temperature   ETS_MNA  Training -0.0337   1.93  1.50
#>  6 Mannheim Temperature   ETS_MAA  Training -0.825    1.94  1.55
#>  7 Mannheim Temperature   ETS_MMA  Training -0.788    1.95  1.55
#>  8 Mannheim Temperature   ETS_MAdA Training  0.114    2.06  1.61
#>  9 Mannheim Precipitation ETS_AAdA Training -0.138   29.5  22.9
#> 10 Mannheim Precipitation ETS_ANA  Training -0.656   29.6  23.0
#> 11 Mannheim Precipitation ETS_MAdA Training -1.21    29.6  23.2
#> 12 Mannheim Precipitation ETS_MNA  Training -0.870   29.6  23.2
#> 13 Mannheim Precipitation ETS_MMA  Training  0.151   29.7  23.1
#> 14 Mannheim Precipitation ETS_AMA  Training  0.413   29.7  23.0
#> 15 Mannheim Precipitation ETS_AAA  Training  0.175   29.7  23.0
#> 16 Mannheim Precipitation ETS_MAA  Training  0.0928  29.8  23.1
```

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 Mannheim Precipitation ETS_AMA     15.1  8.19e- 1
```

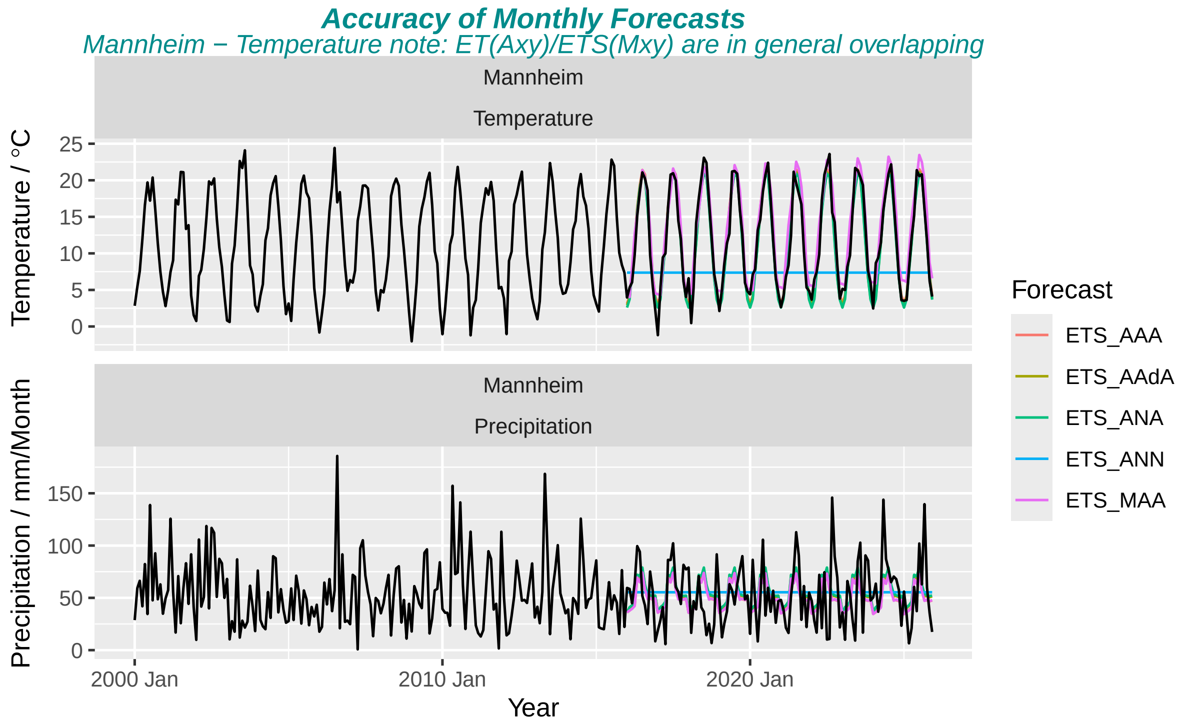
```
#>  2 Mannheim Precipitation ETS_MMA     15.2  8.14e- 1
#>  3 Mannheim Precipitation ETS_MAA     15.2  8.12e- 1
#>  4 Mannheim Precipitation ETS_ANA     15.3  8.06e- 1
#>  5 Mannheim Precipitation ETS_AAA     15.5  7.95e- 1
#>  6 Mannheim Precipitation ETS_AAdA    15.6  7.92e- 1
#>  7 Mannheim Precipitation ETS_MAdA    15.6  7.91e- 1
#>  8 Mannheim Precipitation ETS_MNA     16.1  7.62e- 1
#>  9 Mannheim Temperature   ETS_AMA     19.9  5.29e- 1
#> 10 Mannheim Temperature   ETS_AAA     19.9  5.26e- 1
#> 11 Mannheim Temperature   ETS_AAdA    21.8  4.14e- 1
#> 12 Mannheim Temperature   ETS_ANA     21.8  4.10e- 1
#> 13 Mannheim Temperature   ETS_MAA     97.6  7.54e-12
#> 14 Mannheim Temperature   ETS_MAdA   224.   0
#> 15 Mannheim Temperature   ETS_MMA    207.   0
#> 16 Mannheim Temperature   ETS_MNA    198.   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

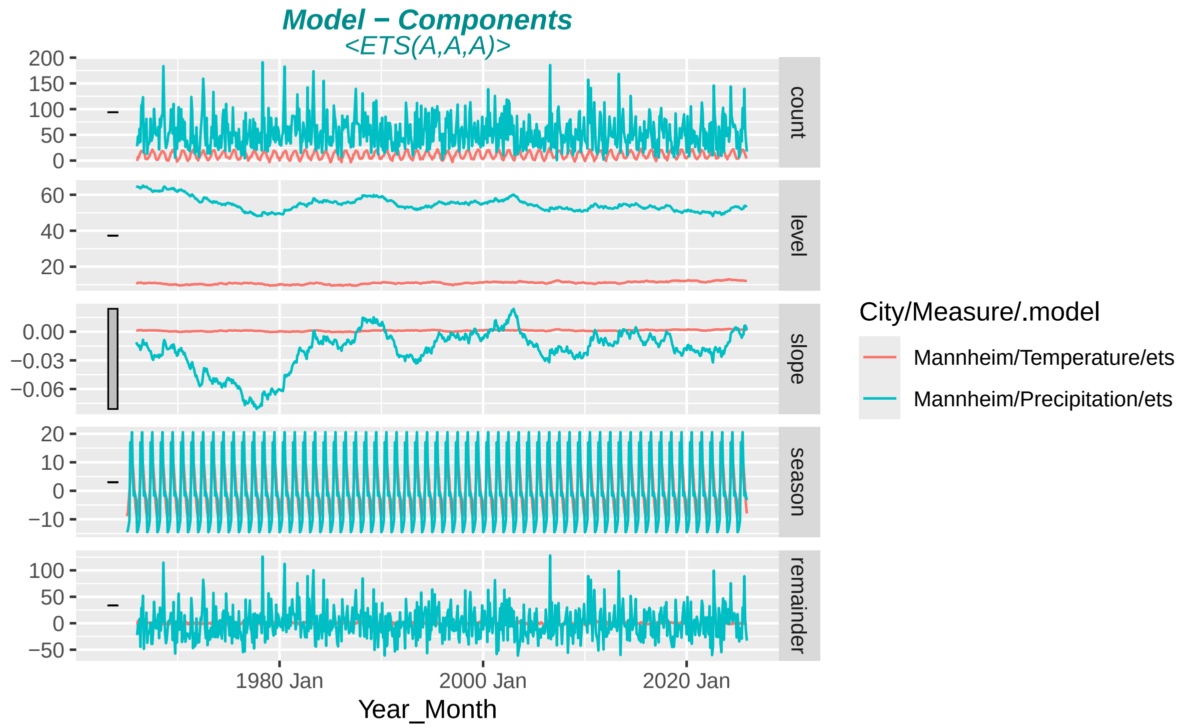
```
#> # A tibble: 10 x 7
#>    .model   City     Measure       .type       ME  RMSE   MAE
#>    <chr>    <chr>    <fct>         <chr>    <dbl> <dbl> <dbl>
#>  1 ETS_AAA  Mannheim Temperature   Test   -0.0253  1.55  1.20
#>  2 ETS_AAdA Mannheim Temperature   Test    0.243   1.56  1.20
#>  3 ETS_ANA  Mannheim Temperature   Test    0.458   1.62  1.25
#>  4 ETS_MAA  Mannheim Temperature   Test   -1.22    2.18  1.80
#>  5 ETS_ANN  Mannheim Temperature   Test    4.70    8.11  6.49
#>  6 ETS_MAA  Mannheim Precipitation Test    2.05   28.7  22.8
#>  7 ETS_AAdA Mannheim Precipitation Test   -1.29   28.8  22.7
#>  8 ETS_AAA  Mannheim Precipitation Test    1.95   28.8  22.8
#>  9 ETS_ANA  Mannheim Precipitation Test   -2.27   28.8  22.8
#> 10 ETS_ANN  Mannheim Precipitation Test   -2.52   30.3  24.2
```

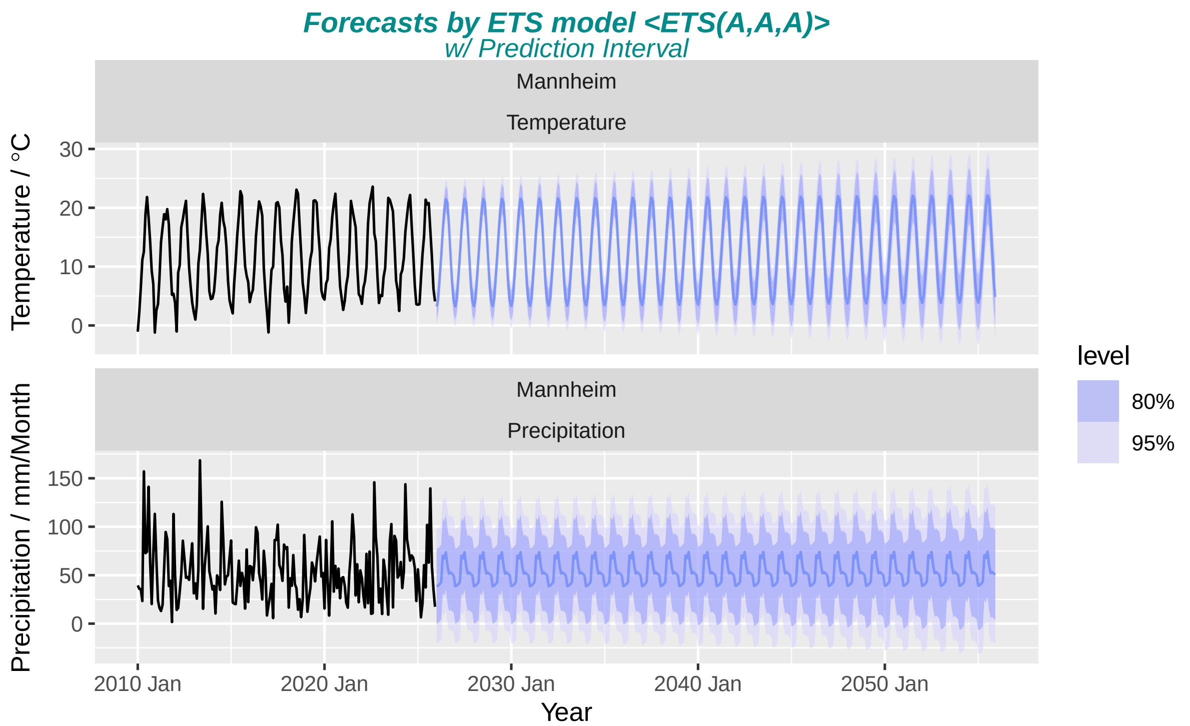

Accuracy of Monthly Forecasts
Mannheim − Temperature note: ET(Axy)/ETS(Mxy) are in general overlapping
Mannheim
Temperature
Mannheim
Precipitation
Temperature / °C
Precipitation / mm/Month
25
20
15
10
5
0
150
100
50
0
2000 Jan
2010 Jan
2020 Jan
Year
Forecast
ETS_AAA
ETS_AAdA
ETS_ANA
ETS_ANN
ETS_MAA

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 Mannheim Temperat~ ets      2.90  -2743. 5521. 5522. 5599.  2.83  2.85  1.35
#> 2 Mannheim Precipit~ ets    905.    -4811. 9657. 9658. 9735. 885.   887.   23.0
```

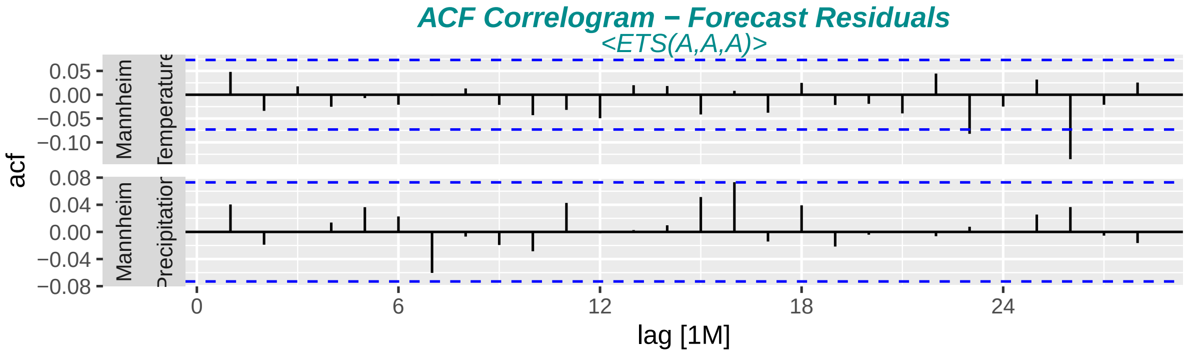


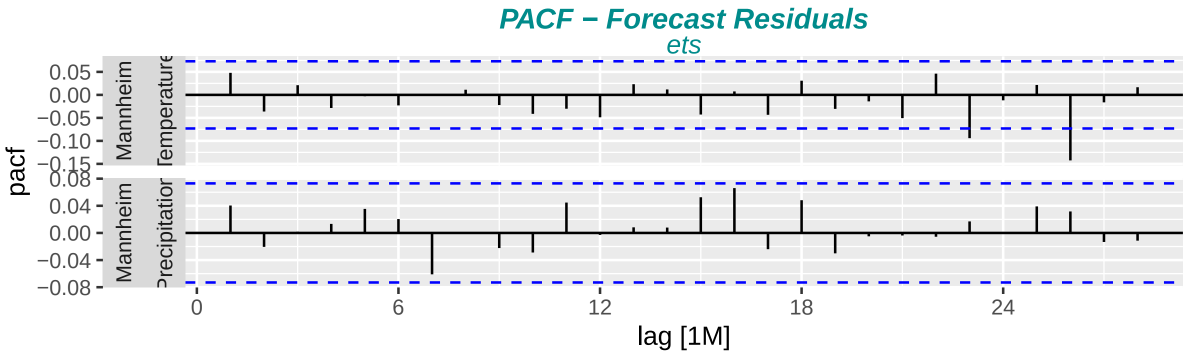


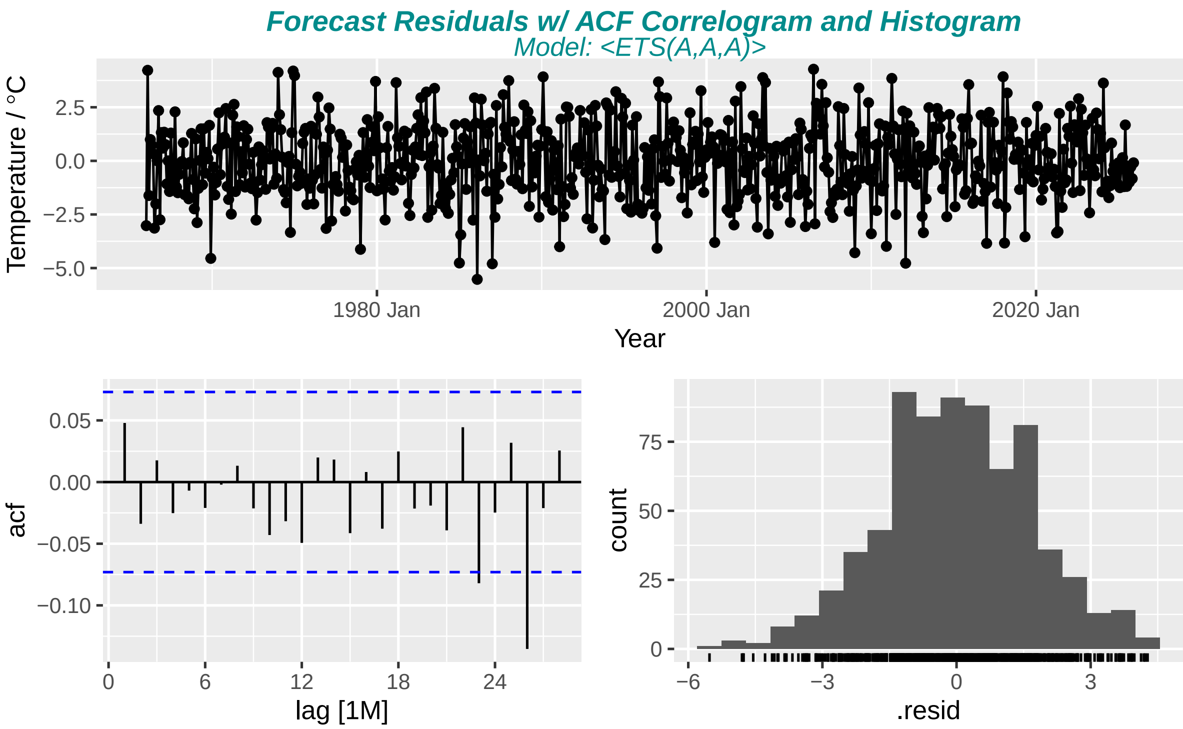
2.2.2 Residual Stationarity

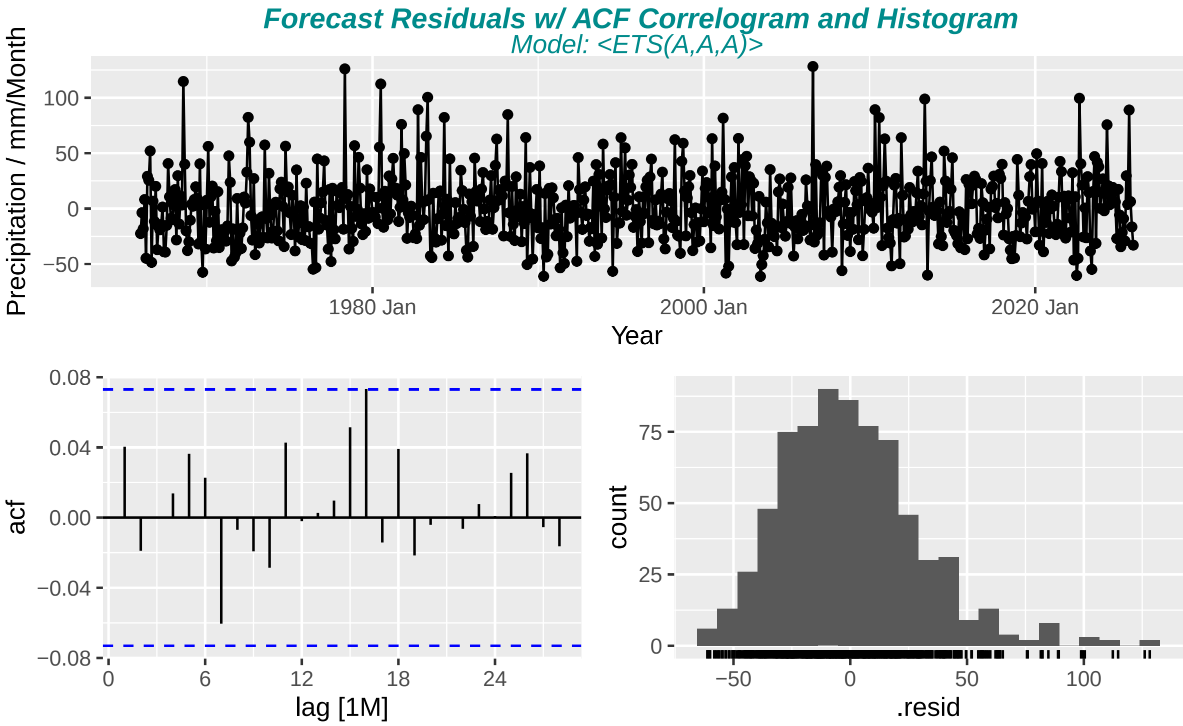
Required checks to be ready for forecasting:

- 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



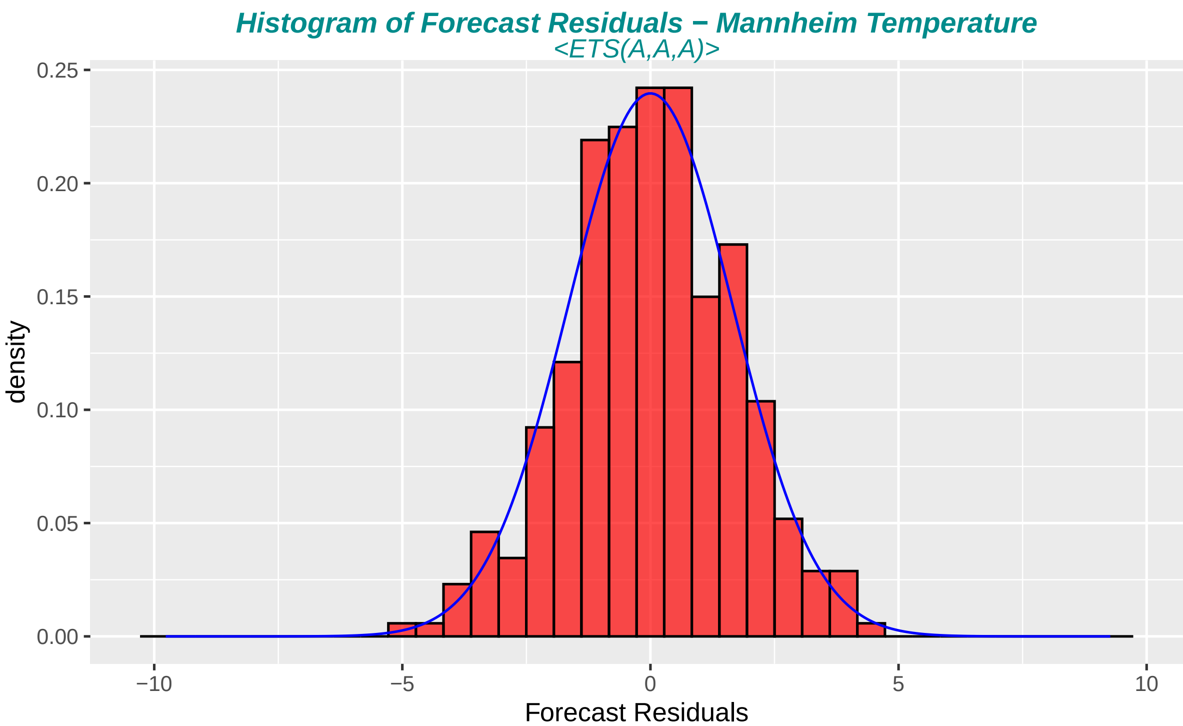


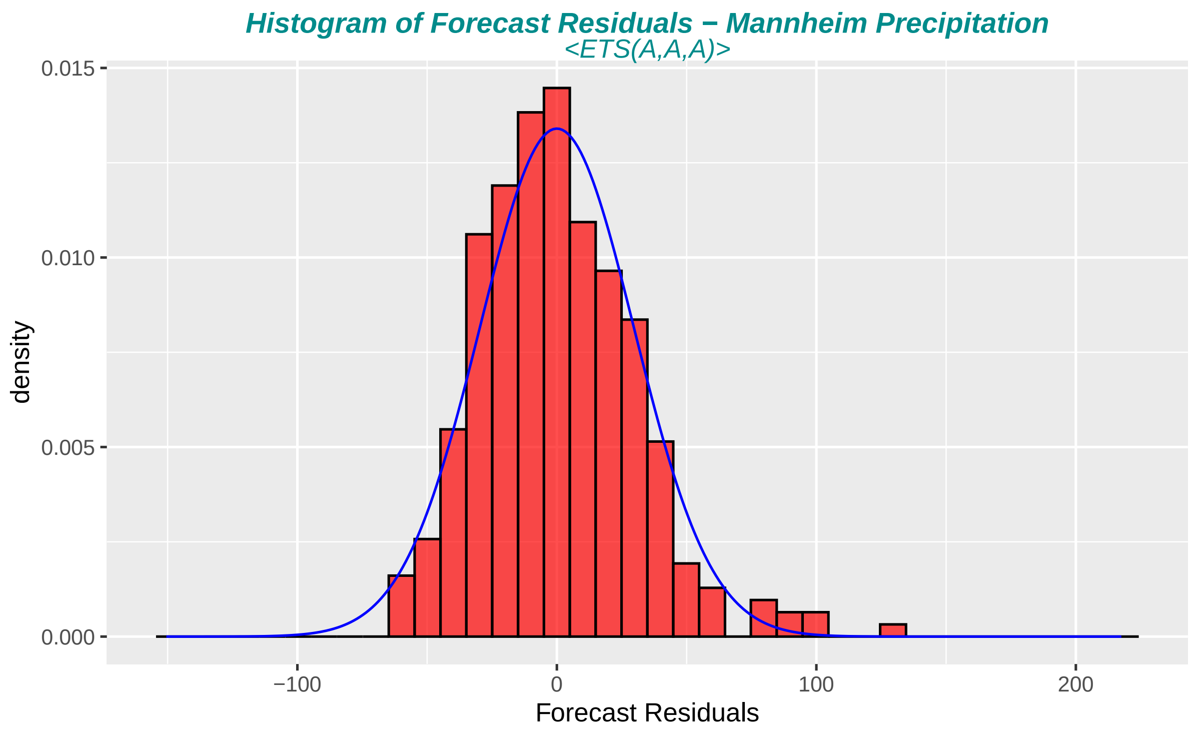




2.2.3 Ljung-Box Test and Histogram of forecast residuals with overlaid normal curve

```
#> 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 Mannheim Temperature   ets       27.3     0.290
#> 2 Mannheim Precipitation ets       20.0     0.696
```




Histogram of Forecast Residuals − Mannheim Precipitation
<ETS(A,A,A)>
0.015
0.010
0.005
0.000
density
−100
0
100
200
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 Mannheim Temperature arima_111_011     2.84  -1398. 2805. 2805. 2823.
#> 2 Mannheim Temperature arima_012_011     2.84  -1398. 2805. 2805. 2823.
#> 3 Mannheim Temperature arima_111_012     2.85  -1398. 2807. 2807. 2829.
#> 4 Mannheim Temperature arima_211_011     2.85  -1398. 2807. 2807. 2830.
#> 5 Mannheim Temperature arima_012_112     2.85  -1398. 2809. 2809. 2836.
#> 6 Mannheim Temperature arima_100_210     3.83  -1481. 2970. 2970. 2989.
#> 7 Mannheim Temperature arima_200_011     4.23  -1515. 3038. 3038. 3057.
#> 8 Mannheim Temperature arima_100_110_c   4.23  -1515. 3040. 3040. 3063.
#> # A tibble: 8 x 8
#>   City     Measure       .model        sigma2 log_lik   AIC  AICc   BIC
#>   <chr>    <fct>         <chr>          <dbl>   <dbl> <dbl> <dbl> <dbl>
#> 1 Mannheim Precipitation arima_012_011   888.  -3432. 6872. 6872. 6891.
#> 2 Mannheim Precipitation arima_111_011   888.  -3432. 6872. 6872. 6891.
#> 3 Mannheim Precipitation arima_211_011   889.  -3432. 6874. 6874. 6897.
#> 4 Mannheim Precipitation arima_111_012   889.  -3432. 6874. 6874. 6897.
#> 5 Mannheim Precipitation arima_012_112   890.  -3432. 6876. 6876. 6904.
#> 6 Mannheim Precipitation arima_001_002   979.  -3499. 7008. 7008. 7031.
#> 7 Mannheim Precipitation arima_100_210  1157.  -3503. 7015. 7015. 7038.
#> 8 Mannheim Precipitation arima_200_011  1313.  -3547. 7101. 7101. 7120.
```

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 Mannheim Temperature arima_111_012   Training 0.0797    1.67  1.31
#> 2 Mannheim Temperature arima_211_011   Training 0.0826    1.67  1.31
#> 3 Mannheim Temperature arima_012_112   Training 0.0781    1.67  1.31
#> 4 Mannheim Temperature arima_111_011   Training 0.0812    1.67  1.31
#> 5 Mannheim Temperature arima_012_011   Training 0.0793    1.67  1.31
#> 6 Mannheim Temperature arima_100_210   Training 0.0437    1.94  1.53
#> 7 Mannheim Temperature arima_100_110_c Training 0.000759  2.03  1.59
#> 8 Mannheim Temperature arima_200_110_c Training 0.000759  2.03  1.59
#> # A tibble: 8 x 7
#>   City     Measure       .model          .type          ME  RMSE   MAE
#>   <chr>    <fct>         <chr>           <chr>       <dbl> <dbl> <dbl>
#> 1 Mannheim Precipitation arima_211_011   Training  0.274    29.5  22.5
#> 2 Mannheim Precipitation arima_012_112   Training  0.273    29.5  22.5
#> 3 Mannheim Precipitation arima_012_011   Training  0.266    29.5  22.5
#> 4 Mannheim Precipitation arima_111_011   Training  0.266    29.5  22.5
#> 5 Mannheim Precipitation arima_111_012   Training  0.272    29.5  22.5
#> 6 Mannheim Precipitation arima_001_002   Training -0.00392  31.2  24.4
#> 7 Mannheim Precipitation arima_100_210   Training  0.0412   33.6  25.9
#> 8 Mannheim Precipitation arima_100_110_c Training -0.0787   35.9  27.6
```

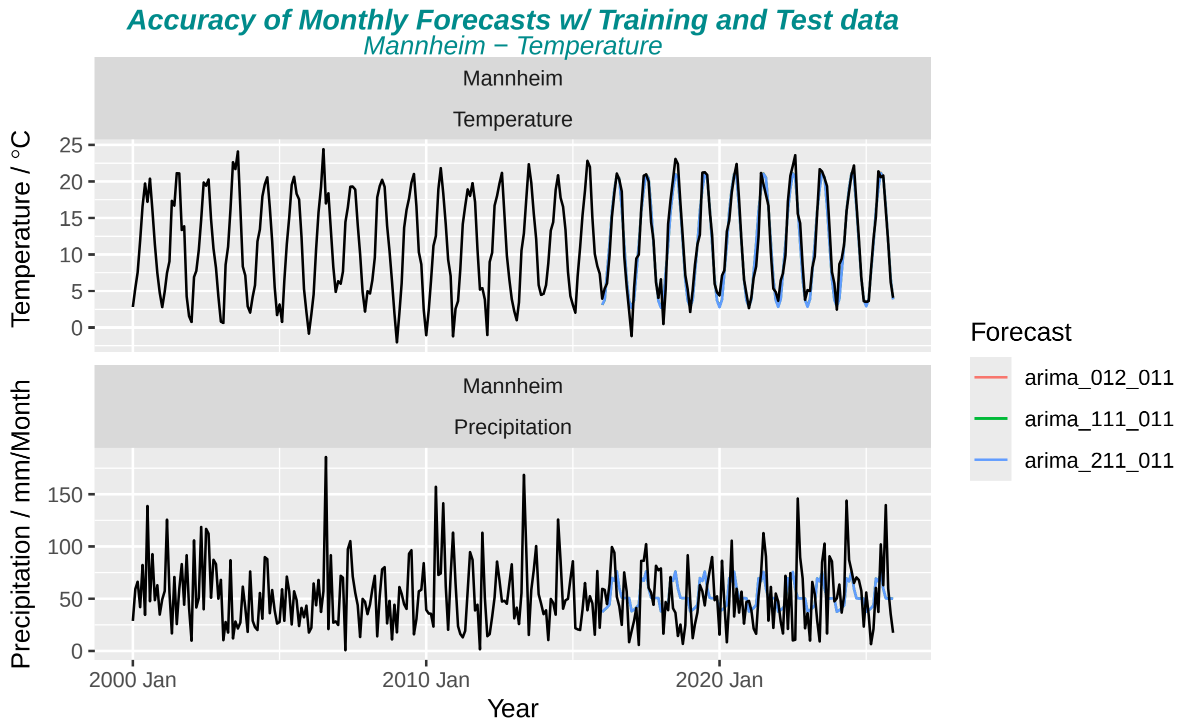
3.1.2 Ljung-Box Test - independence/white noise of the forecasts residuals

```
#> 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 Mannheim Temperature arima_111_012      20.9  4.67e- 1
#> 2 Mannheim Temperature arima_012_112      21.1  4.51e- 1
#> 3 Mannheim Temperature arima_211_011      21.2  4.47e- 1
#> 4 Mannheim Temperature arima_111_011      21.3  4.39e- 1
#> 5 Mannheim Temperature arima_012_011      21.4  4.32e- 1
#> 6 Mannheim Temperature arima_100_210      46.3  1.17e- 3
#> 7 Mannheim Temperature arima_200_011     111.   3.35e-14
#> 8 Mannheim Temperature arima_100_110_c   111.   3.24e-14
#> # A tibble: 8 x 5
#>   City     Measure       .model          lb_stat lb_pvalue
#>   <chr>    <fct>         <chr>             <dbl>     <dbl>
#> 1 Mannheim Precipitation arima_111_012      15.3  8.05e- 1
#> 2 Mannheim Precipitation arima_211_011      15.4  8.03e- 1
#> 3 Mannheim Precipitation arima_012_112      15.4  8.03e- 1
#> 4 Mannheim Precipitation arima_012_011      15.5  7.96e- 1
#> 5 Mannheim Precipitation arima_111_011      15.6  7.94e- 1
#> 6 Mannheim Precipitation arima_001_002      31.6  6.42e- 2
#> 7 Mannheim Precipitation arima_100_210      34.8  2.93e- 2
#> 8 Mannheim Precipitation arima_100_110_c   115.   5.77e-15
```

3.1.3 Forecast Accuracy with Training/Test Data

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

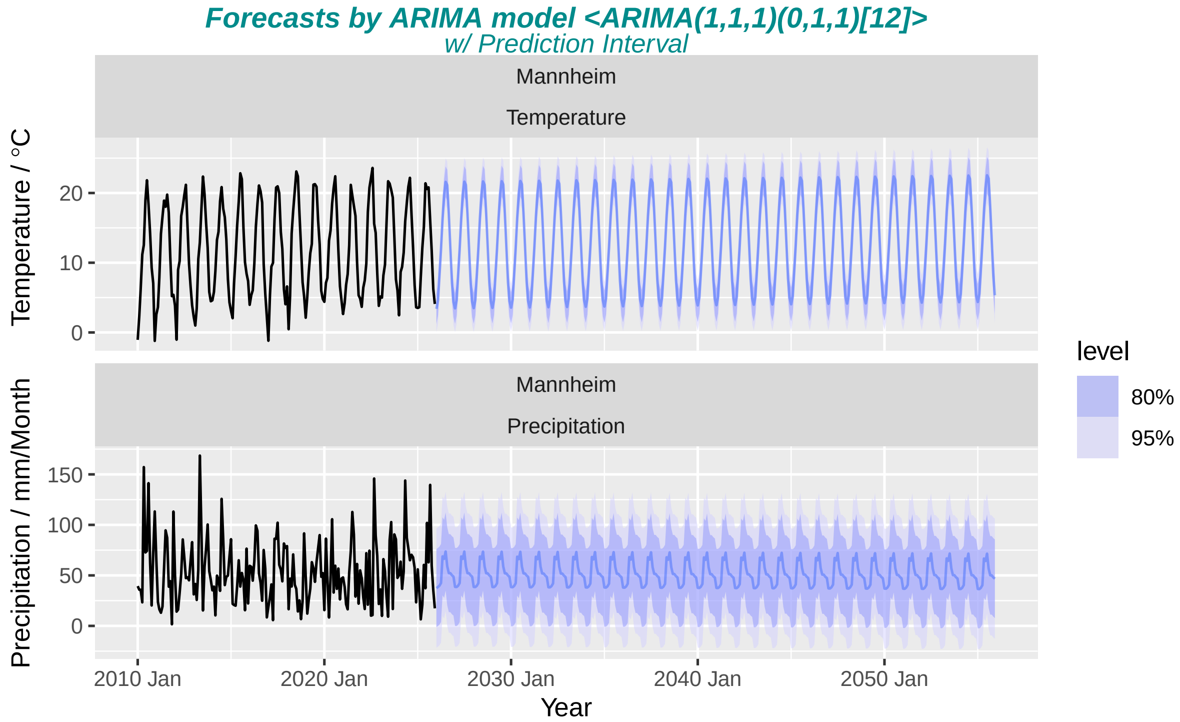
```
#>    .model         City     Measure       .type      ME  RMSE   MAE
#>    <chr>          <chr>    <fct>         <chr>   <dbl> <dbl> <dbl>
#> 1 arima_012_011 Mannheim Temperature   Test    0.325  1.57  1.21
#> 2 arima_111_011 Mannheim Temperature   Test    0.359  1.58  1.22
#> 3 arima_211_011 Mannheim Temperature   Test    0.386  1.58  1.22
#> 4 arima_111_011 Mannheim Precipitation Test   -0.216 28.6  22.7
#> 5 arima_012_011 Mannheim Precipitation Test   -0.216 28.6  22.7
#> 6 arima_211_011 Mannheim Precipitation Test   -0.216 28.6  22.7
```



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

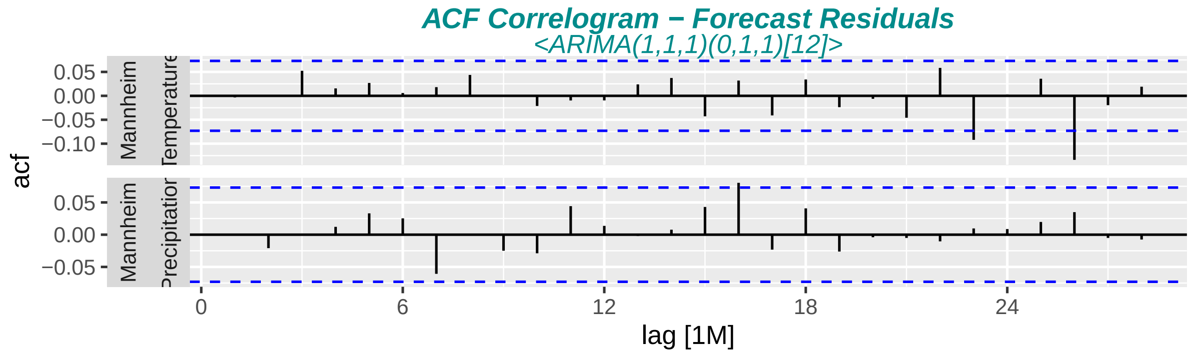
```
#> Provide model coefficients by report(fit_model)
#> # A tibble: 2 x 10
#>   City     Measure     .model sigma2 log_lik   AIC  AICc   BIC ar_roots ma_roots
#>   <chr>    <fct>       <chr>   <dbl>   <dbl> <dbl> <dbl> <dbl> <list>   <list>
#> 1 Mannheim Temperature arima    2.84  -1398. 2805. 2805. 2823. <cpl>    <cpl>
#> 2 Mannheim Precipitat~ arima  888.    -3432. 6872. 6872. 6891. <cpl>    <cpl>
```



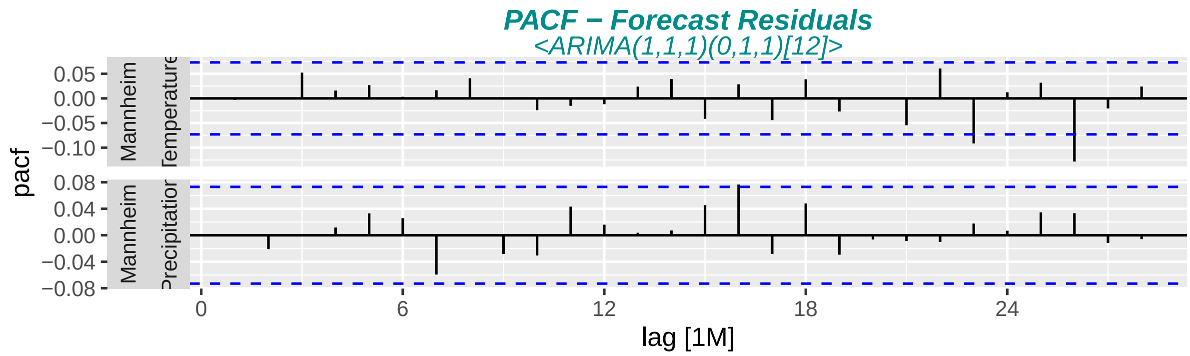
3.2.2 Residual Stationarity

Required checks to be ready for forecasting:

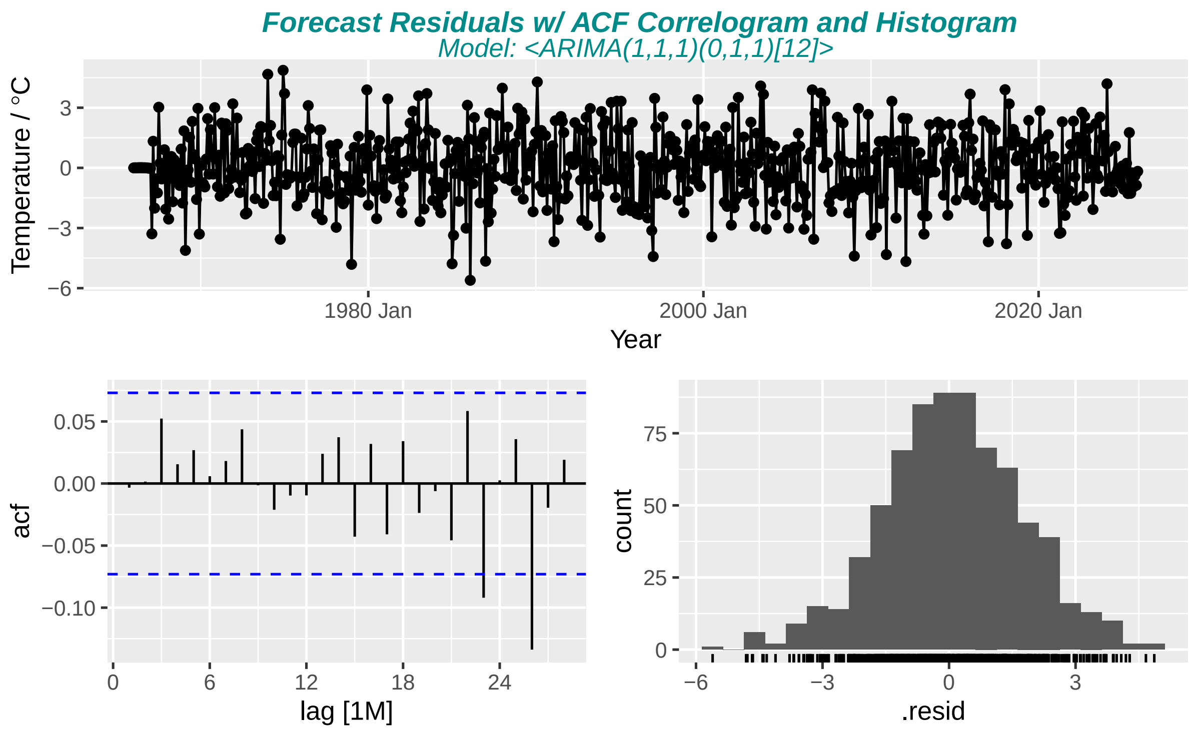
- ACF Forecast Residual: all spikes are within the significance limits, so the residuals appear to be white noise
- The Ljung-Box test also shows that the residuals have no remaining autocorrelations
- Forecast Residuals are more or less normally distributed with roughly centred on zero



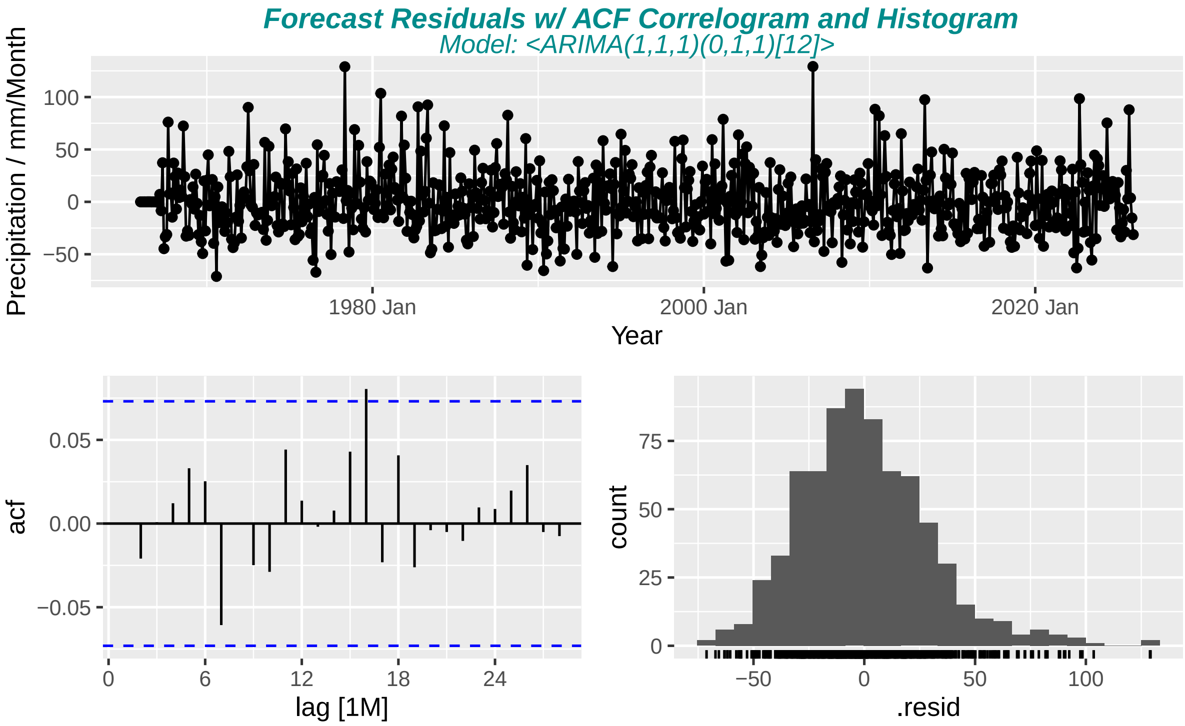
ACF Correlogram − Forecast Residuals
<ARIMA(1,1,1)(0,1,1)[12]>
acf
Mannheim
Temperature
Precipitation
lag [1M]



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

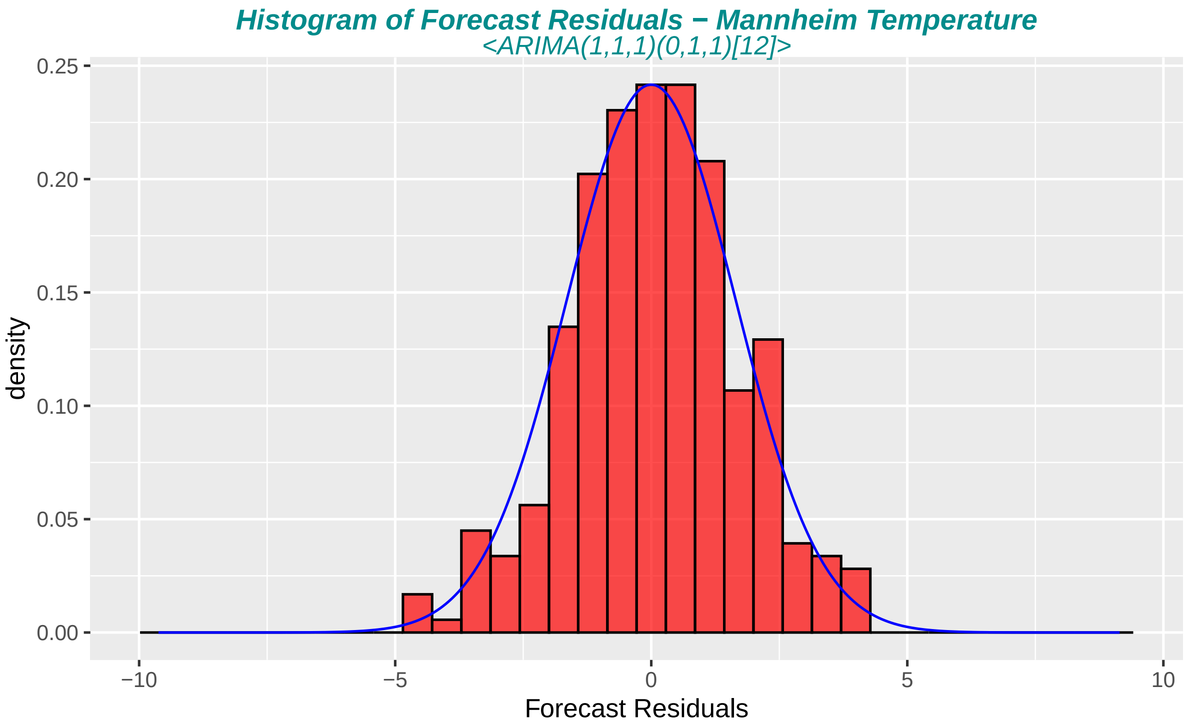


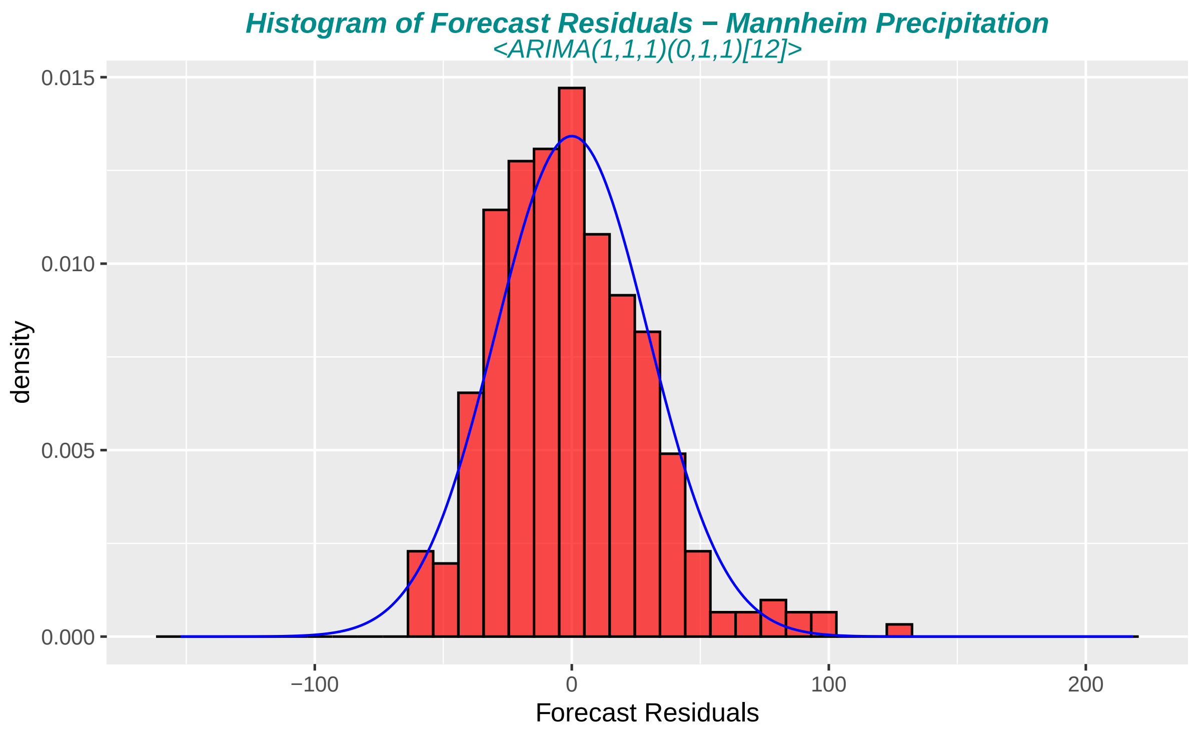
Forecast Residuals w/ ACF Correlogram and Histogram
Model: <ARIMA(1,1,1)(0,1,1)[12]>
Temperature / °C
1980 Jan
2000 Jan
2020 Jan
Year
acf
lag [1M]
count
.resid



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 Mannheim Temperature   arima     23.8     0.303
#> 2 Mannheim Precipitation arima     17.4     0.689
```




Histogram of Forecast Residuals − Mannheim Precipitation
<ARIMA(1,1,1)(0,1,1)[12]>
0.015
0.010
0.005
0.000
density
−100
0
100
200
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.

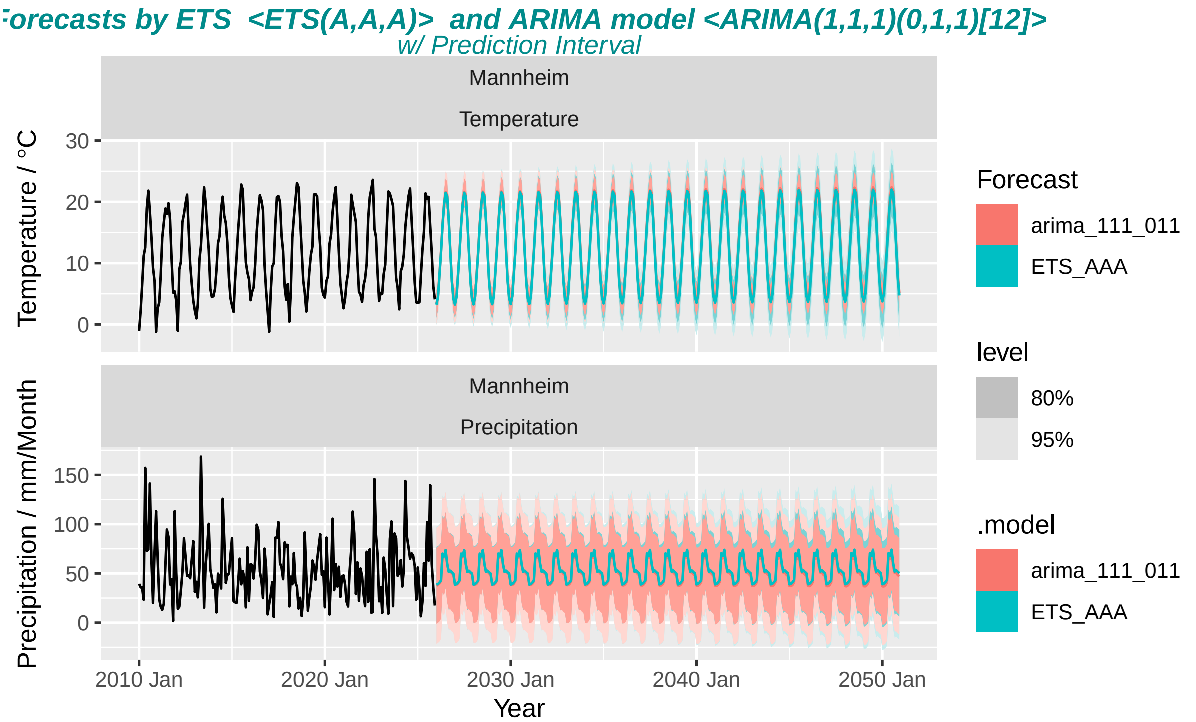
- Residual Accuracy with one-step-ahead fitted residuals
- Forecast Accuracy with Training/Test Data

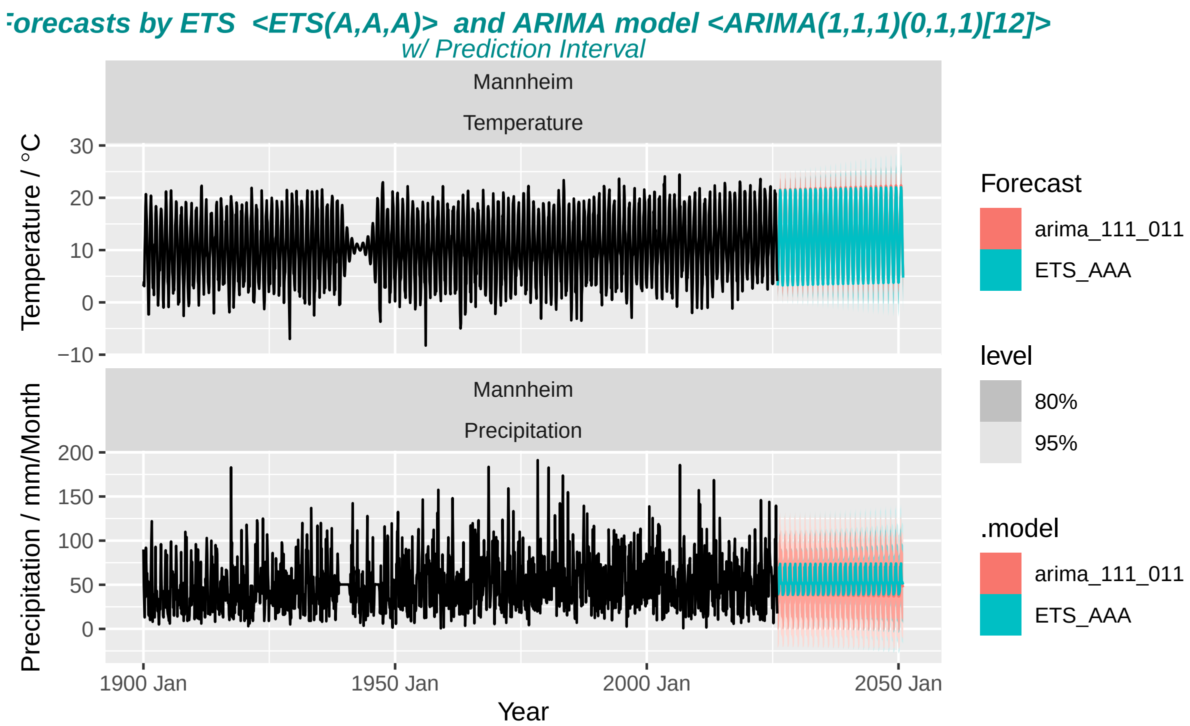
Note: a good fit to training data is never an indication that the model will forecast well. Therefore the values of the Forecast Accuracy are the more relevant one.

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 Mannheim Temperature   ETS_AAA       Test      1.55  1.20  22.6 0.637 0.651
#> 2 Mannheim Temperature   arima_111_011 Test      1.58  1.22  21.6 0.646 0.665
#> 3 Mannheim Temperature   arima         Training  1.67  1.31  42.5 0.695 0.700
#> 4 Mannheim Temperature   ets           Training  1.68  1.35  44.6 0.715 0.706
#> 5 Mannheim Precipitation arima_111_011 Test     28.6  22.7   87.2 0.724 0.689
#> 6 Mannheim Precipitation ETS_AAA       Test     28.8  22.8   83.8 0.729 0.692
#> 7 Mannheim Precipitation arima         Training 29.5  22.5   81.7 0.712 0.706
#> 8 Mannheim Precipitation ets           Training 29.7  23.0   83.2 0.729 0.712
```

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
Mannheim
Temperature
Temperature / °C
30
20
10
0
−10
Mannheim
Precipitation
Precipitation / mm/Month
200
150
100
50
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
1871-1900	10.0	42.2
1901-1930	10.1	44.3
1931-1960	10.4	49.9
1961-1990	10.3	55.6
1991-2020	11.3	53.4
2021-2025	12.2	54.8

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
Mannheim	Temperature	2026	12.12	12.29
Mannheim	Temperature	2030	12.21	12.42
Mannheim	Temperature	2035	12.32	12.58
Mannheim	Temperature	2040	12.43	12.75
Mannheim	Temperature	2045	12.54	12.91
Mannheim	Temperature	2050	12.65	13.07
Mannheim	Precipitation	2026	53.12	52.46
Mannheim	Precipitation	2030	53.19	52.28
Mannheim	Precipitation	2035	53.29	51.91
Mannheim	Precipitation	2040	53.38	51.53
Mannheim	Precipitation	2045	53.48	51.16
Mannheim	Precipitation	2050	53.57	50.79

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	12.12	12.29	12.65	13.07	0.54	0.79
Precipitation	2026	2050	53.12	52.46	53.57	50.79	0.46	-1.67

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	3.23	3.39	3.76	4.20	0.54	0.81
Temperature	Feb	2026	2050	4.50	4.59	5.04	5.38	0.54	0.79
Temperature	Mrz	2026	2050	7.94	8.09	8.47	8.87	0.54	0.78
Temperature	Apr	2026	2050	11.76	11.88	12.29	12.67	0.54	0.78
Temperature	Mai	2026	2050	16.09	16.33	16.63	17.11	0.54	0.78
Temperature	Jun	2026	2050	19.57	19.74	20.11	20.53	0.54	0.78
Temperature	Jul	2026	2050	21.47	21.60	22.01	22.38	0.54	0.78
Temperature	Aug	2026	2050	20.86	21.13	21.40	21.91	0.54	0.78
Temperature	Sep	2026	2050	16.90	17.02	17.43	17.80	0.54	0.78
Temperature	Okt	2026	2050	11.90	12.18	12.44	12.97	0.54	0.78
Temperature	Nov	2026	2050	6.99	7.13	7.53	7.92	0.54	0.78
Temperature	Dez	2026	2050	4.20	4.38	4.73	5.16	0.54	0.78

Measure	Month	Year.x	Year.y	ETS.x	ARIMA.x	ETS.y	ARIMA.y	Delta_ETS	Delta_ARIMA
Precipitation	Jan	2026	2050	38.56	37.40	39.02	36.91	0.46	-0.49
Precipitation	Feb	2026	2050	39.21	38.55	39.67	36.81	0.46	-1.73
Precipitation	Mrz	2026	2050	40.66	39.78	41.12	37.99	0.46	-1.78
Precipitation	Apr	2026	2050	42.84	42.47	43.30	40.69	0.46	-1.79
Precipitation	Mai	2026	2050	70.11	68.85	70.56	67.06	0.46	-1.79
Precipitation	Jun	2026	2050	67.29	66.39	67.74	64.60	0.46	-1.79
Precipitation	Jul	2026	2050	73.69	73.68	74.15	71.89	0.46	-1.79
Precipitation	Aug	2026	2050	59.88	58.98	60.33	57.20	0.46	-1.79
Precipitation	Sep	2026	2050	51.41	52.25	51.87	50.46	0.46	-1.79
Precipitation	Okt	2026	2050	52.77	52.19	53.22	50.40	0.46	-1.79
Precipitation	Nov	2026	2050	51.13	50.16	51.59	48.38	0.46	-1.79
Precipitation	Dez	2026	2050	49.86	48.85	50.31	47.07	0.46	-1.79

5 Yearly Data Forecasts with ARIMA and ETS

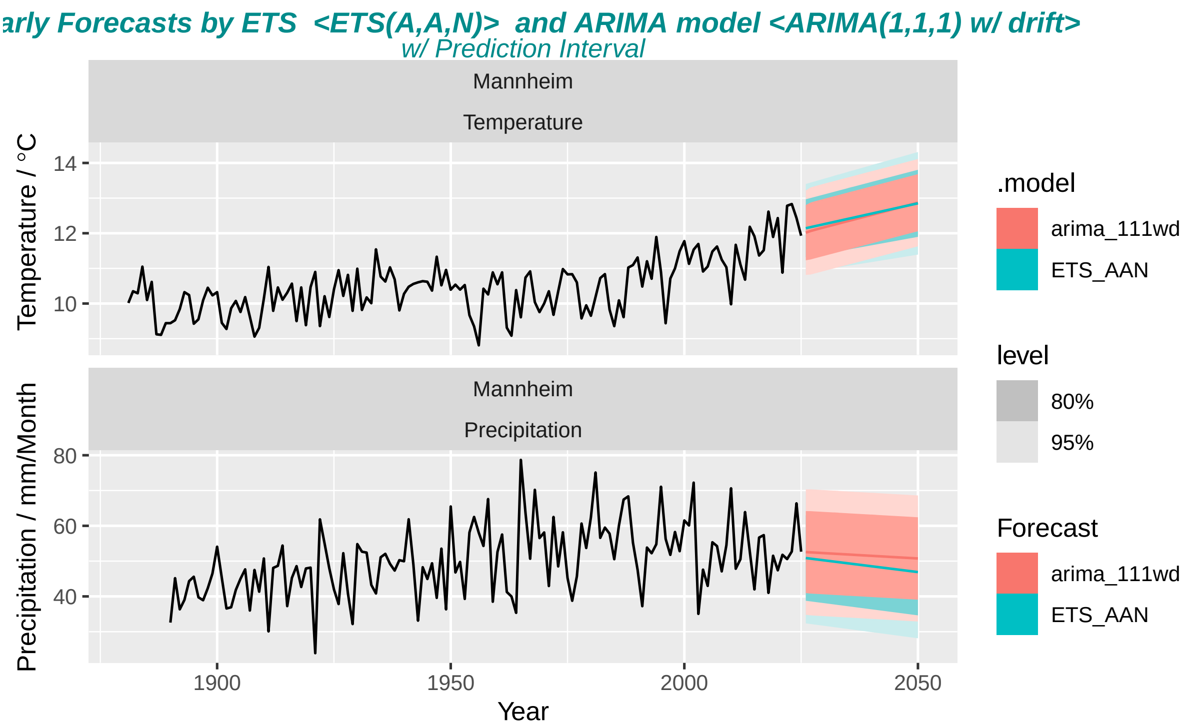
For yearly data the seasonal monthly data are replaced by the yearly average data. Therefore the seasonal component of the ETS and ARIMA model are to be taken out.

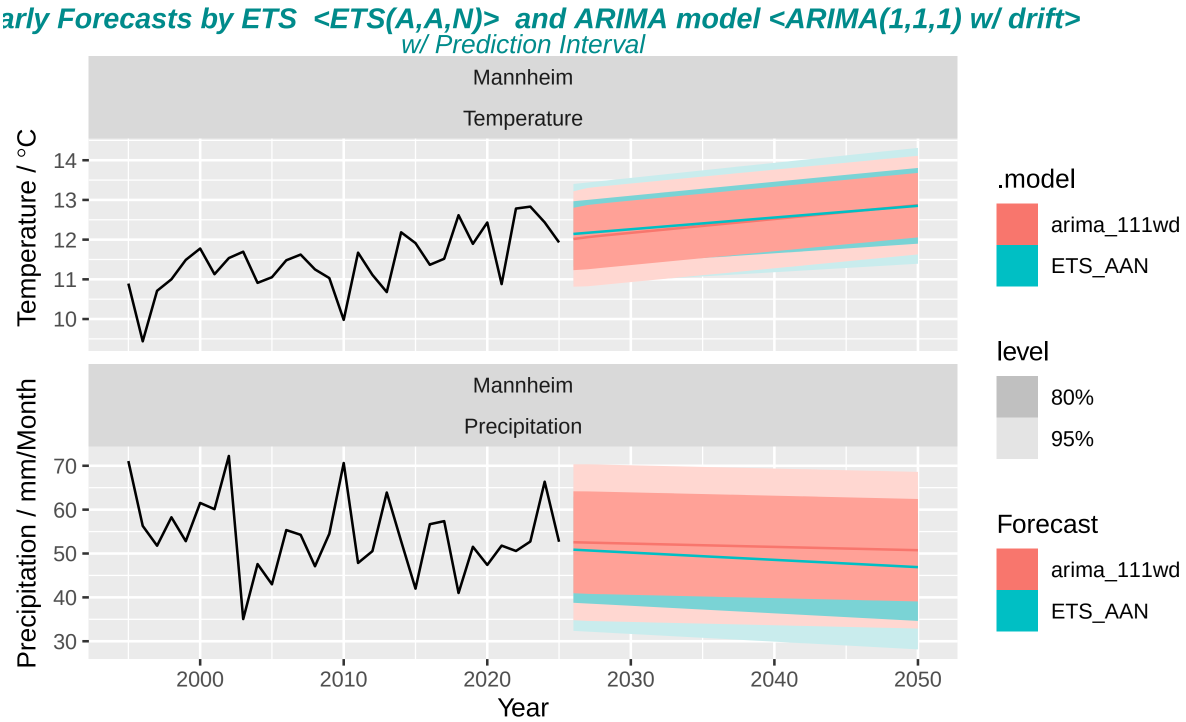
The ETS model $< ETS(A, A, N) >$ with seasonal term change “A” -> “N” is chosen. For ARIMA models the seasonal term (P,D,Q)m has to be taken out and an optimal ARIMA(p,1,q) with one differencing (d=1) is selected. However, for Mauna Loa two times differncing had to be selected $CO_2 <ARIMA(0,2,1) w/ 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 Mannheim Tempera~ arima~  0.370   -54.6  117.  118.  126. NA     NA     NA
#> 2 Mannheim Tempera~ ETS_A~  0.412   -94.2  198.  199.  209.  0.385  0.403  0.505
#> 3 Mannheim Precipi~ arima~ 80.8    -214.   435.  436.  444. NA     NA     NA
#> 4 Mannheim Precipi~ ETS_A~ 89.2    -255.   521.  522.  531. 83.2   84.5    6.84
#> # 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 Mannheim Temperature   ETS_AAN    22.2     0.568
#> 2 Mannheim Precipitation ETS_AAN    26.4     0.332
#> # A tibble: 2 x 5
#>   City     Measure       .model      lb_stat lb_pvalue
#>   <chr>    <fct>         <chr>         <dbl>     <dbl>
#> 1 Mannheim Temperature   arima_111wd    21.1     0.451
#> 2 Mannheim Precipitation arima_111wd    25.7     0.217
```

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
Mannheim	Temperature	2026	12.14	12.02
Mannheim	Temperature	2030	12.26	12.17
Mannheim	Temperature	2035	12.41	12.35
Mannheim	Temperature	2040	12.56	12.52
Mannheim	Temperature	2045	12.70	12.69
Mannheim	Temperature	2050	12.85	12.87
Mannheim	Precipitation	2026	50.87	52.56
Mannheim	Precipitation	2030	50.21	52.26
Mannheim	Precipitation	2035	49.38	51.88
Mannheim	Precipitation	2040	48.55	51.50
Mannheim	Precipitation	2045	47.72	51.13
Mannheim	Precipitation	2050	46.89	50.75

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	12.14	12.02	12.85	12.87	0.71	0.85
Precipitation	2026	2050	50.87	52.56	46.89	50.75	-3.97	-1.81