Climate Data Forecasting -

Atmospheric CO_2 Concentration / Temperature / Precipitation

Wolfgang Vollmer

2026-01-08

Contents

1 Forecasting of Giessen - 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

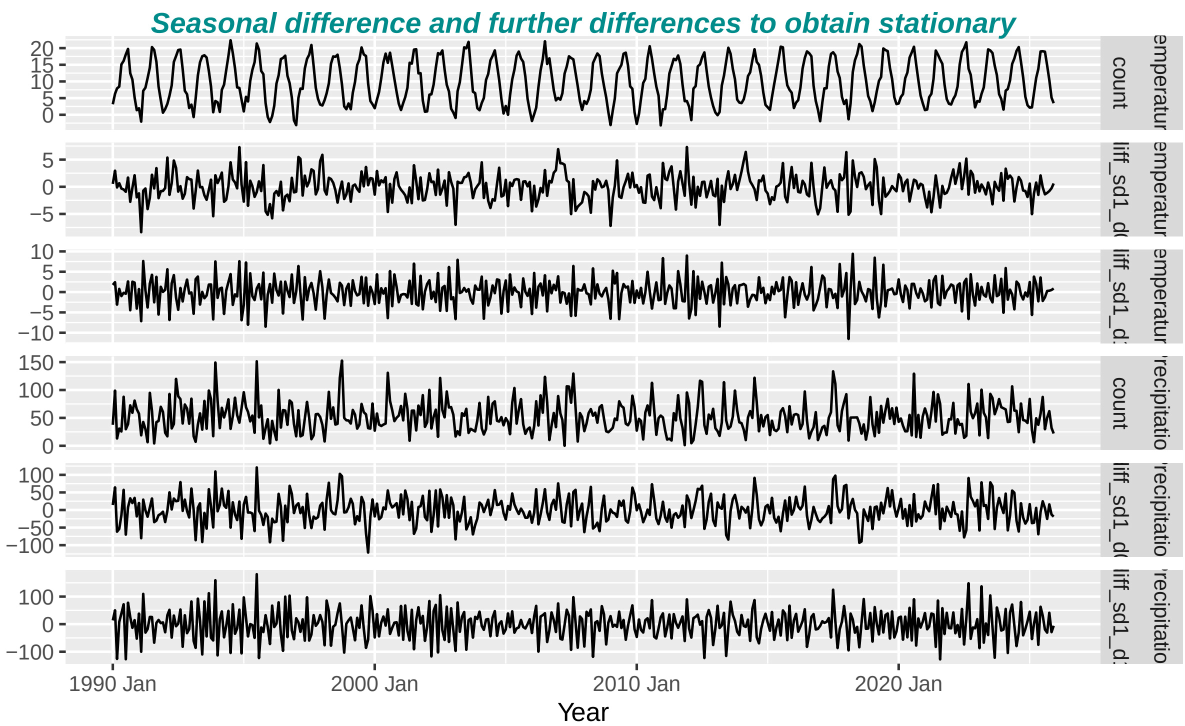
The model fit residuals have to be stationary. For good forecasting this has to be verified with residual diagnostics.

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       0.462      0.0503       1      0
#> 2 Precipitation     0.291      0.1          0      0
#> unitroot_kpss test on (difference(count, 12) + difference())
#> # A tibble: 2 x 3
#>   Measure       kpss_stat kpss_pvalue
#>   <fct>             <dbl>       <dbl>
#> 1 Temperature     0.0117          0.1
#> 2 Precipitation   0.00621         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

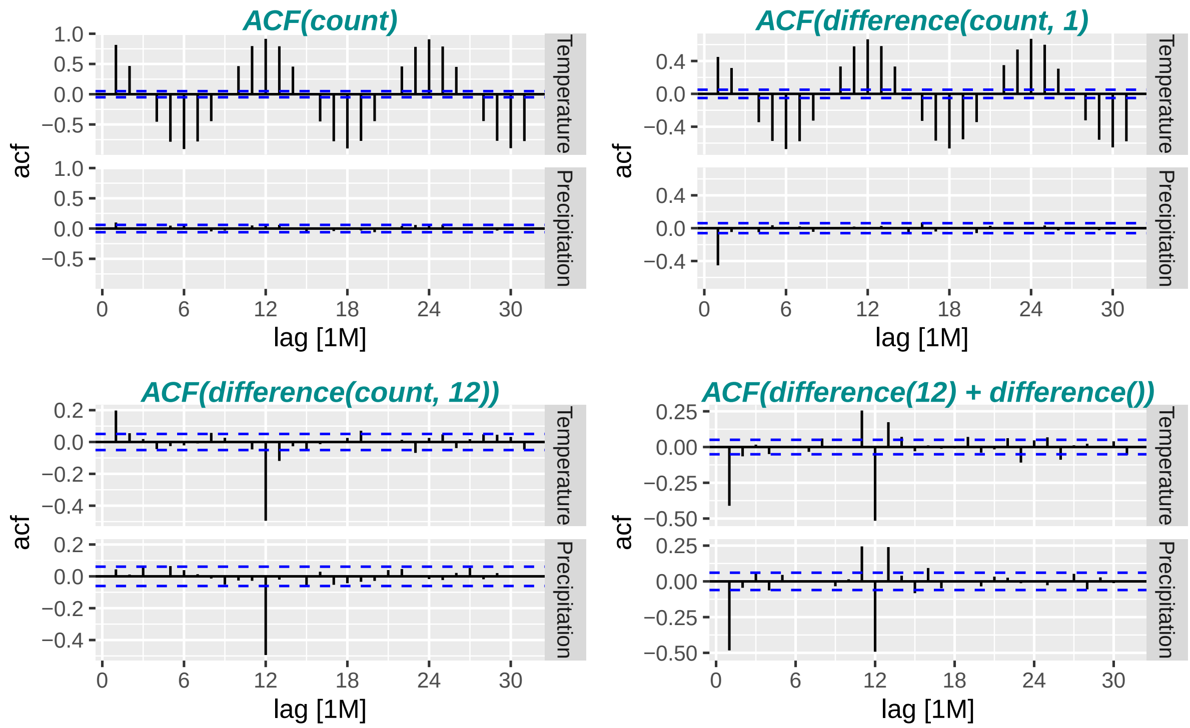
- augment(fit) |> features(.innov, ljung_box, lag=x, dof=y) (Ch. 5.4 Residaul diagnostics)
  - portmanteau test suggesting that the residuals are white noise
  - *Null Hypothesis* H_0: independence/white noise for residuals, i.e. each autocorrelation in the time series residuals of the model for lag 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    5352.     0
#> 2 Precipitation    19.6    0.0331
#> Ljung-Box test on (difference(count, 12))
```

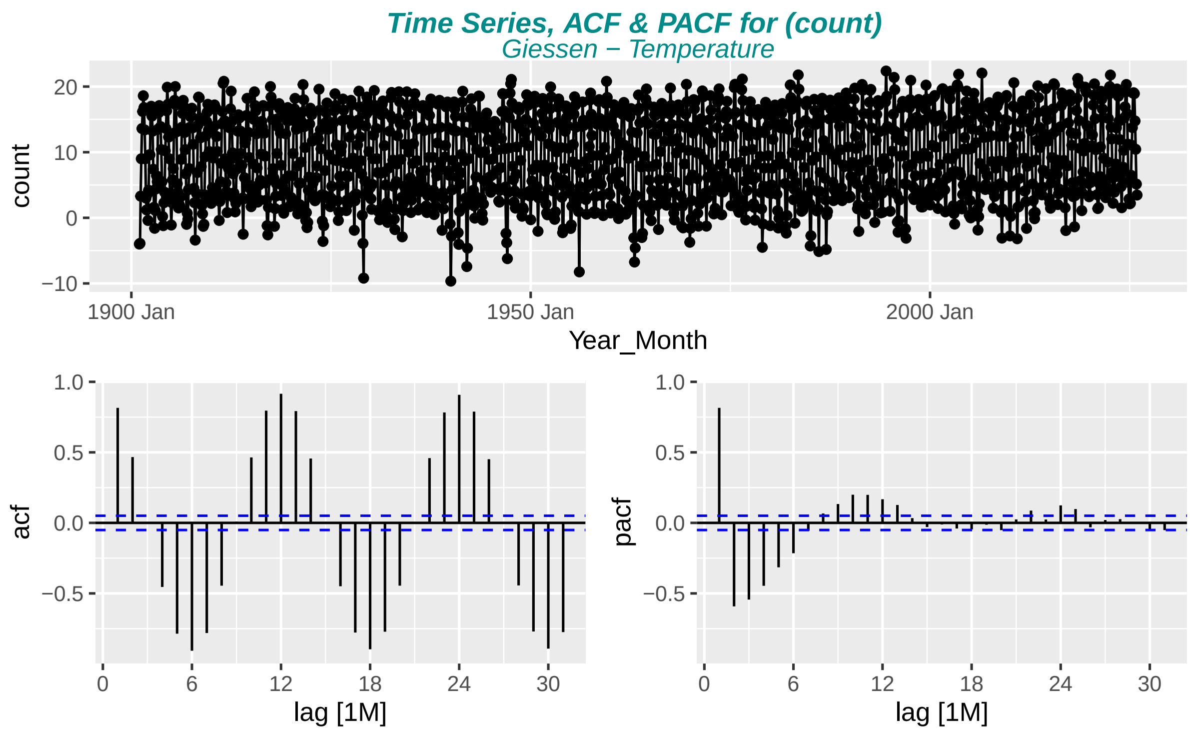
```
#> # A tibble: 2 x 3
#>   Measure       lb_stat lb_pvalue
#>   <fct>           <dbl>     <dbl>
#> 1 Temperature      74.1  7.09e-12
#> 2 Precipitation    16.3  9.24e- 2
#> Ljung-Box test on (difference(count, 12) + difference())
#> # A tibble: 2 x 3
#>   Measure       lb_stat lb_pvalue
#>   <fct>           <dbl>     <dbl>
#> 1 Temperature      269.         0
#> 2 Precipitation    254.         0
```

1.1.3 ACF (Autocorrelation Function) Plots of Differences

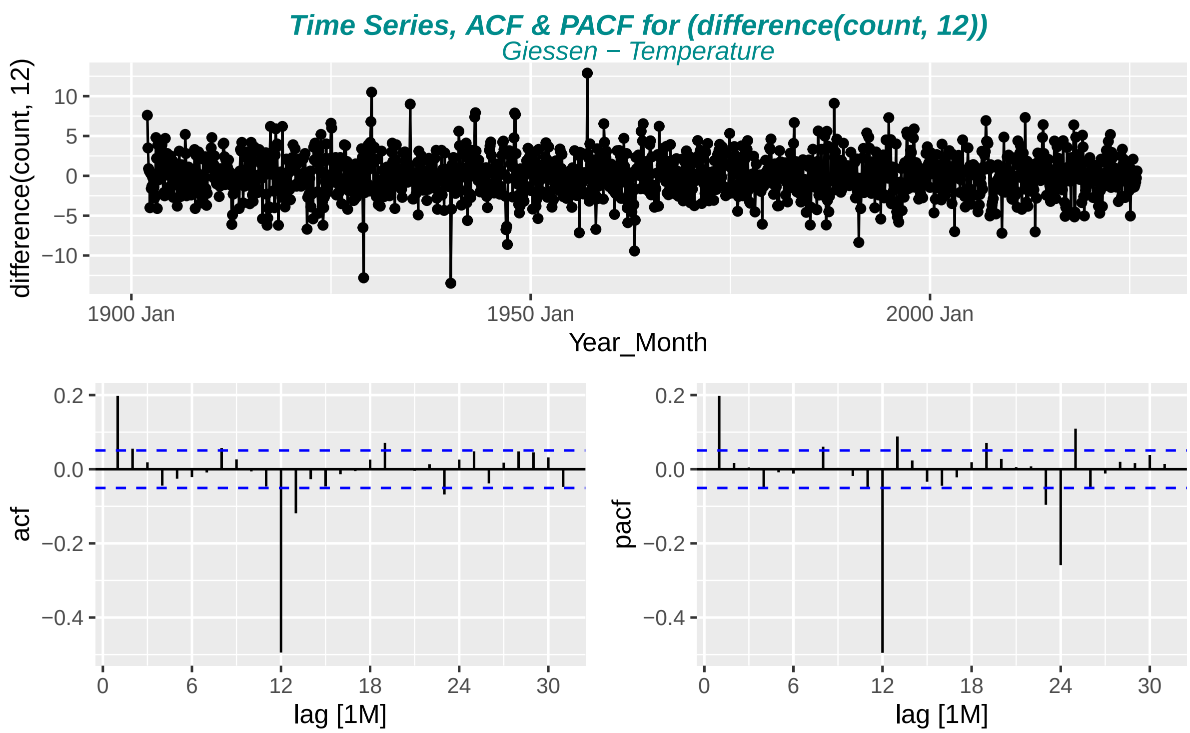


1.1.4 Time Series, ACF and PACF (Partial) Plots of Differences - for ARIMA p, q check

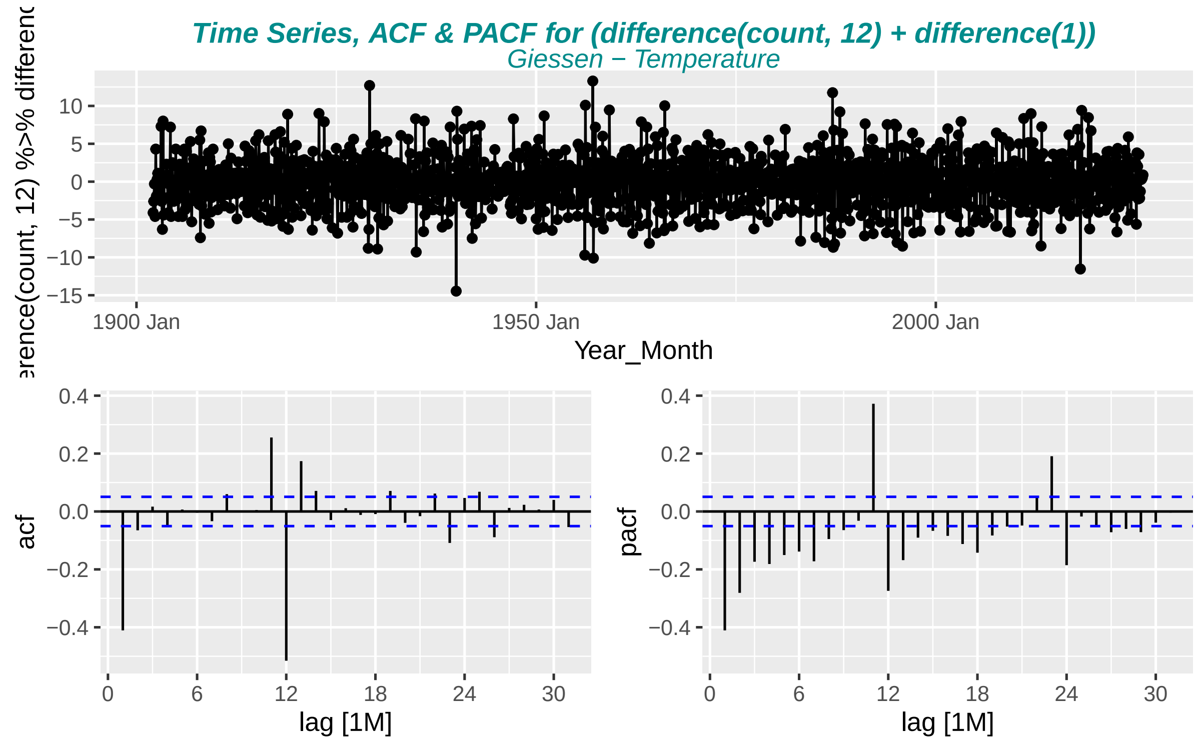
```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City    Measure        Sum  Mean
#>   <chr>   <fct>        <dbl> <dbl>
#> 1 Giessen Temperature 13933.  9.29
```



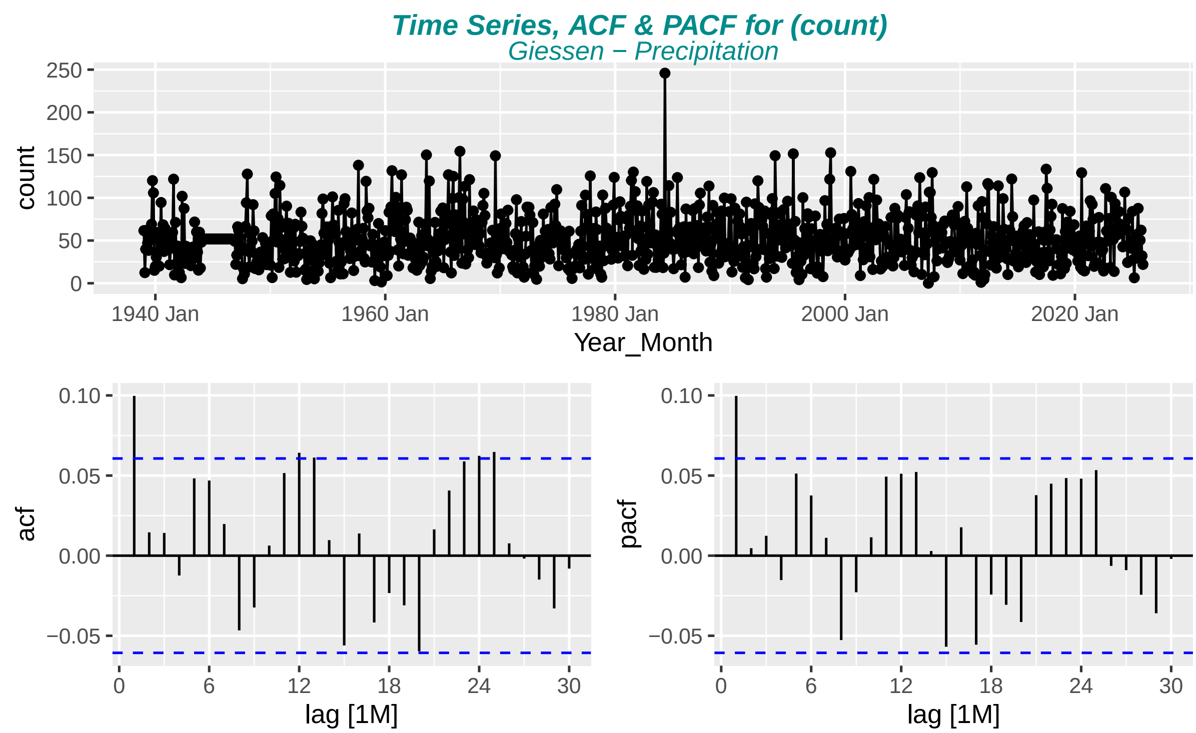

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City    Measure       Sum   Mean
#>   <chr>   <fct>       <dbl>  <dbl>
#> 1 Giessen Temperature  30.2 0.0203
```



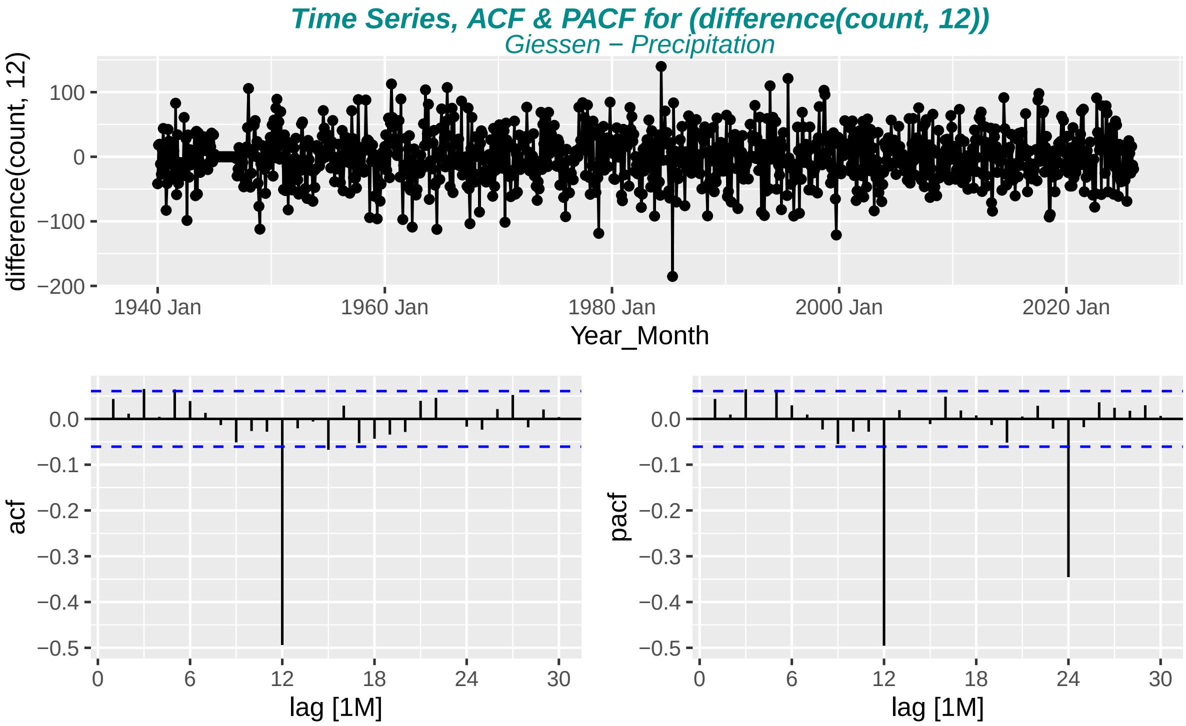

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City    Measure       Sum     Mean
#>   <chr>   <fct>       <dbl>    <dbl>
#> 1 Giessen Temperature -6.99 -0.00470
```



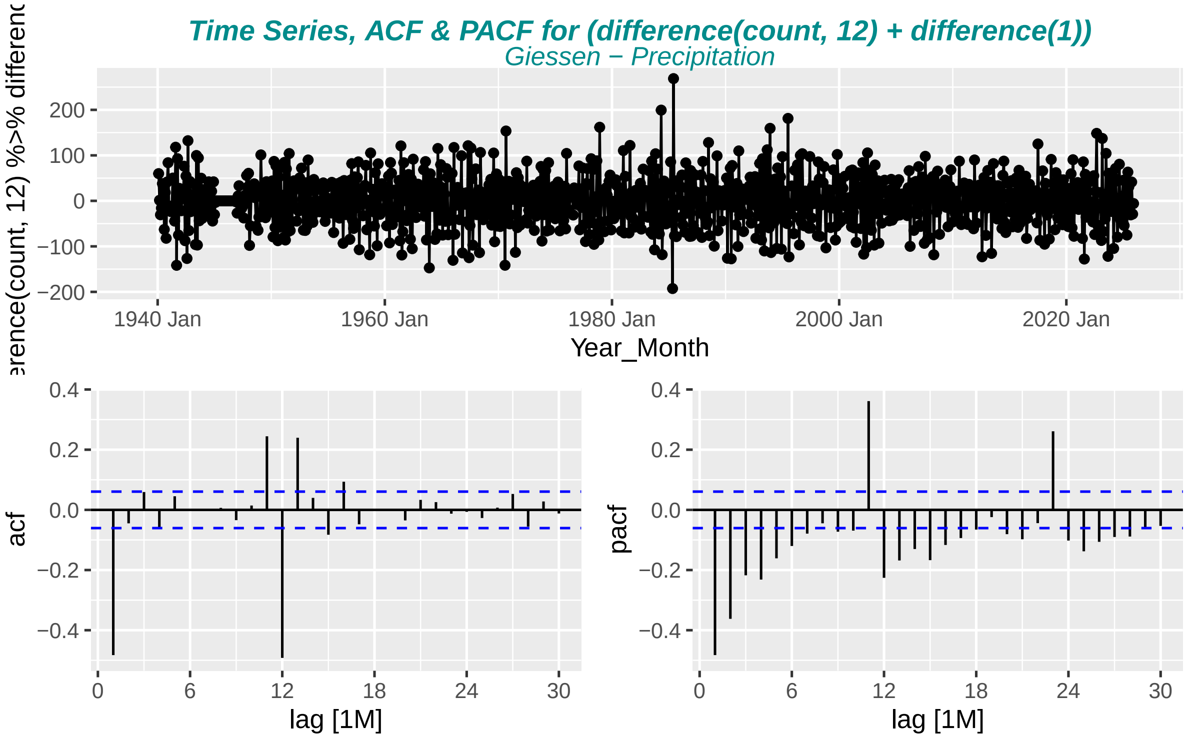

```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City    Measure           Sum  Mean
#>   <chr>   <fct>           <dbl> <dbl>
#> 1 Giessen Precipitation 54126.   51.8
```




```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City    Measure          Sum   Mean
#>   <chr>   <fct>          <dbl>  <dbl>
#> 1 Giessen Precipitation -134. -0.130
```




```
#> # A tibble: 1 x 4
#> # Groups:   City [1]
#>   City    Measure         Sum   Mean
#>   <chr>   <fct>         <dbl>  <dbl>
#> 1 Giessen Precipitation  22.9 0.0222
```



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 Giessen Temperature    <ETS(A,N,A)>
#> 2 Giessen Precipitation <ETS(A,Ad,A)>
#> [1] "Giessen Temperature"
#> Series: count
#> Model: ETS(A,N,A)
#>   Smoothing parameters:
#>     alpha = 0.06347289
#>     gamma = 0.000101769
#>
#>   Initial states:
#>      l[0]      s[0]     s[-1]       s[-2]    s[-3]    s[-4]    s[-5]    s[-6]
#>  10.04836 -7.524003 -4.762938 0.002065576 4.574083 8.545342 9.048699 7.208974
#>     s[-7]      s[-8]     s[-9]    s[-10]    s[-11]
#>  3.907522 -0.5490237 -4.337241 -7.606435 -8.507044
#>
#>   sigma^2:  2.9692
#>
#>      AIC     AICc      BIC
#> 5536.489 5537.170 5605.177
#> [1] "Giessen Precipitation"
#> Series: count
#> Model: ETS(A,Ad,A)
#>   Smoothing parameters:
#>     alpha = 0.0001255647
#>     beta  = 0.0001083295
#>     gamma = 0.0006636497
#>     phi   = 0.9527014
#>
#>   Initial states:
#>      l[0]       b[0]     s[0]     s[-1]     s[-2]     s[-3]    s[-4]    s[-5]
#>  71.32223 -0.9262571 4.331651 -1.135833 -2.358115 -1.241468 5.450903 15.14357
#>     s[-6]    s[-7]    s[-8]     s[-9]    s[-10]    s[-11]
#>  10.40632 9.472255 -10.3146 -8.481856 -11.34869 -9.924131
#>
#>   sigma^2:  800.1774
#>
#>      AIC     AICc      BIC
#> 9568.937 9569.913 9651.364
#> # A tibble: 2 x 8
#>   City    Measure       .model   AIC  AICc   BIC    MSE   MAE
#>   <chr>   <fct>         <chr>  <dbl> <dbl> <dbl>  <dbl> <dbl>
#> 1 Giessen Temperature   ETS    5536. 5537. 5605.   2.91  1.35
#> 2 Giessen Precipitation ETS    9569. 9570. 9651. 781.   21.8
```

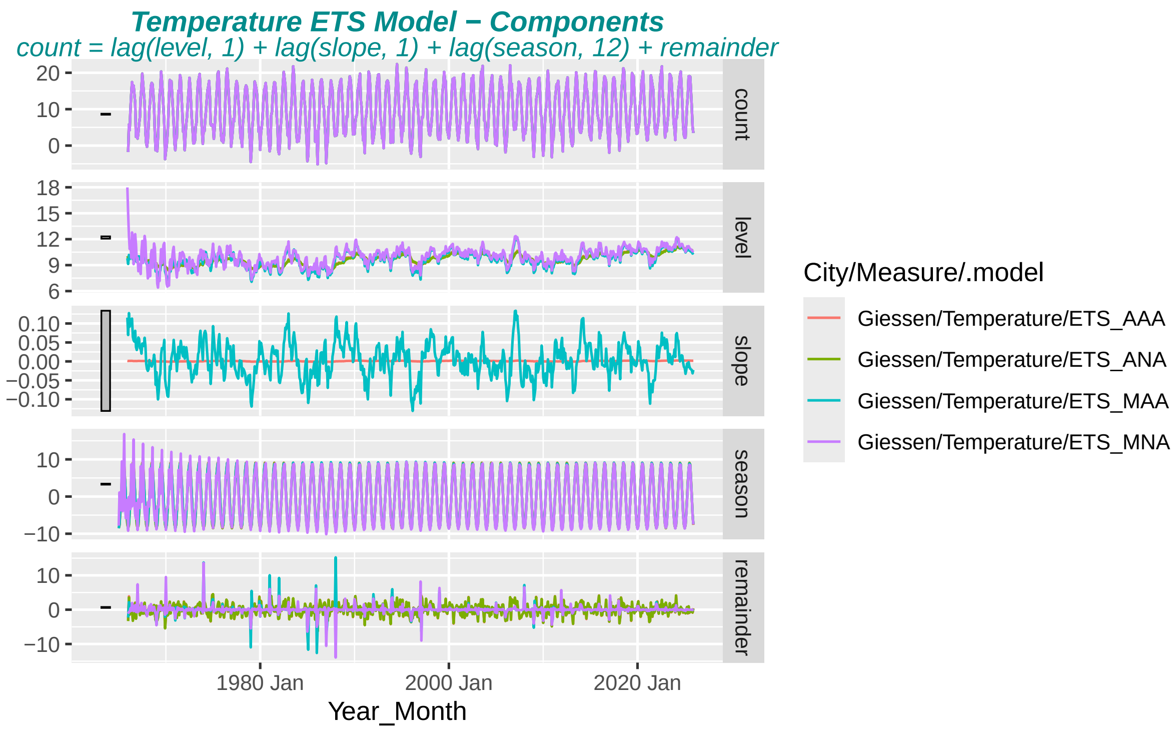
Fit of different pre-defined $ETS(A|M,N|A|M,N|A|M)$ models

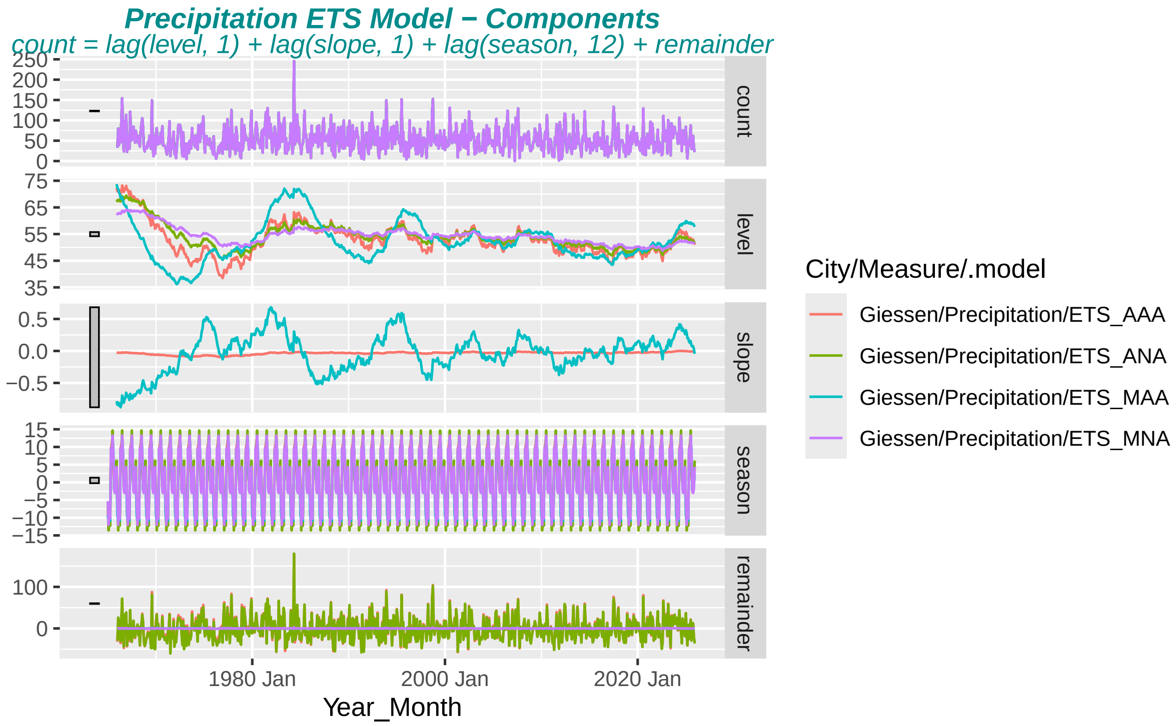
Model Selection by Information Criterion - lowest AIC, AICc, BIC

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 Giessen Temperature   ETS_ANA  5536. 5537. 5605.   2.91   2.93  1.35
#>  2 Giessen Temperature   ETS_AAA  5541. 5542. 5619.   2.91   2.94  1.36
#>  3 Giessen Temperature   ETS_AMA  5541. 5542. 5619.   2.91   2.94  1.36
#>  4 Giessen Temperature   ETS_AAdA 5542. 5543. 5624.   2.91   2.93  1.35
#>  5 Giessen Temperature   ETS_MNA  8093. 8093. 8161.   7.11   7.01  0.565
#>  6 Giessen Temperature   ETS_MAdA 8152. 8153. 8235.   3.57   3.80  0.536
#>  7 Giessen Temperature   ETS_MMA  8172. 8173. 8250.   3.48   3.62  0.588
#>  8 Giessen Temperature   ETS_MAA  8188. 8188. 8265.   3.45   3.62  0.558
#>  9 Giessen Precipitation ETS_MNA  9543. 9544. 9612. 790.   790.    0.407
#> 10 Giessen Precipitation ETS_MAdA 9549. 9550. 9631. 787.   787.    0.405
#> 11 Giessen Precipitation ETS_AAdA 9569. 9570. 9651. 781.   781.   21.8
#> 12 Giessen Precipitation ETS_MMA  9570. 9570. 9647. 812.   814.    0.427
#> 13 Giessen Precipitation ETS_ANA  9571. 9572. 9640. 790.   791.   22.0
#> 14 Giessen Precipitation ETS_MAA  9571. 9572. 9649. 810.   811.    0.430
#> 15 Giessen Precipitation ETS_AMA  9578. 9578. 9655. 793.   793.   21.9
#> 16 Giessen Precipitation ETS_AAA  9579. 9580. 9657. 794.   794.   21.9
```





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)

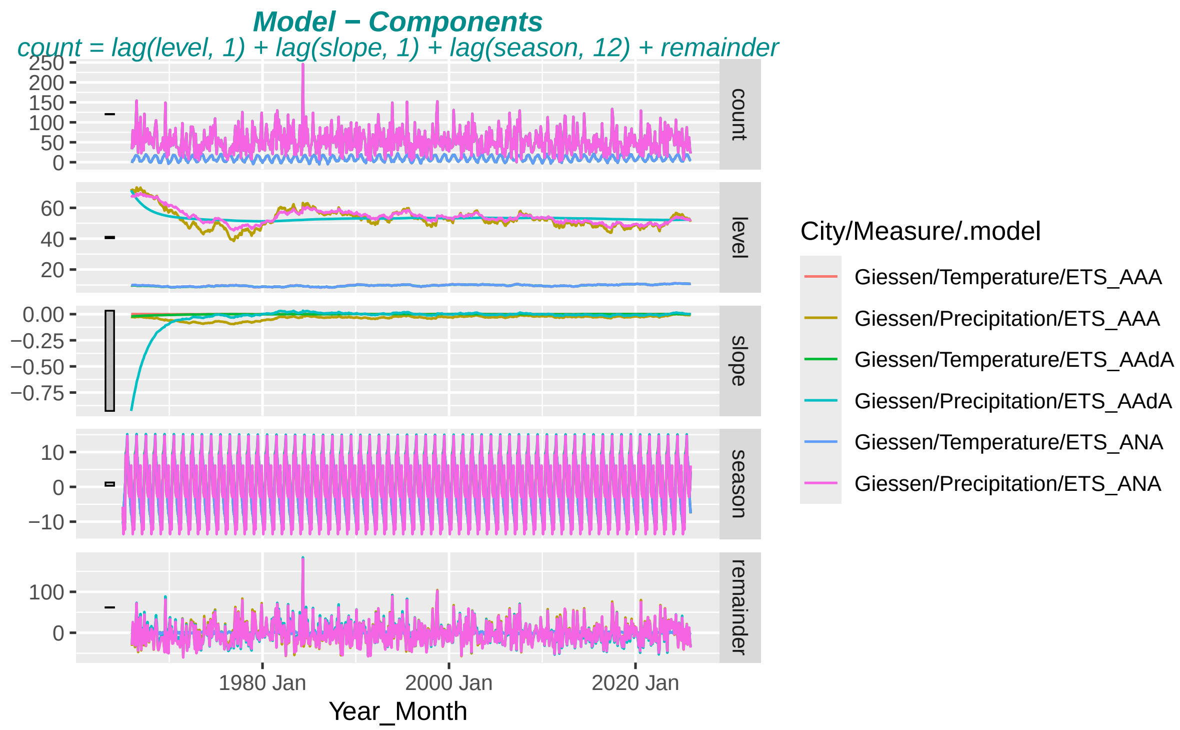
```
#> # A tibble: 16 x 7
#>    City    Measure       .model   .type          ME  RMSE   MAE
#>    <chr>   <fct>         <chr>    <chr>       <dbl> <dbl> <dbl>
#>  1 Giessen Temperature   ETS_AAdA Training  0.0470   1.71  1.35
#>  2 Giessen Temperature   ETS_ANA  Training  0.0143   1.71  1.35
#>  3 Giessen Temperature   ETS_AAA  Training  0.00702  1.71  1.36
#>  4 Giessen Temperature   ETS_AMA  Training  0.0132   1.71  1.36
#>  5 Giessen Temperature   ETS_MAA  Training -0.0132   1.86  1.47
#>  6 Giessen Temperature   ETS_MMA  Training -0.0432   1.87  1.49
#>  7 Giessen Temperature   ETS_MAdA Training -0.188    1.89  1.50
#>  8 Giessen Temperature   ETS_MNA  Training -0.0398   2.67  1.91
#>  9 Giessen Precipitation ETS_AAdA Training -0.250   28.0  21.8
#> 10 Giessen Precipitation ETS_MAdA Training -1.80    28.0  22.1
#> 11 Giessen Precipitation ETS_MNA  Training -1.41    28.1  22.1
#> 12 Giessen Precipitation ETS_ANA  Training -1.12    28.1  22.0
#> 13 Giessen Precipitation ETS_AMA  Training  0.542   28.2  21.9
#> 14 Giessen Precipitation ETS_AAA  Training  0.163   28.2  21.9
#> 15 Giessen Precipitation ETS_MAA  Training  0.609   28.4  22.2
#> 16 Giessen Precipitation ETS_MMA  Training -0.0138  28.5  22.4
```

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 Giessen Temperature   ETS_ANA     16.8     0.720
```

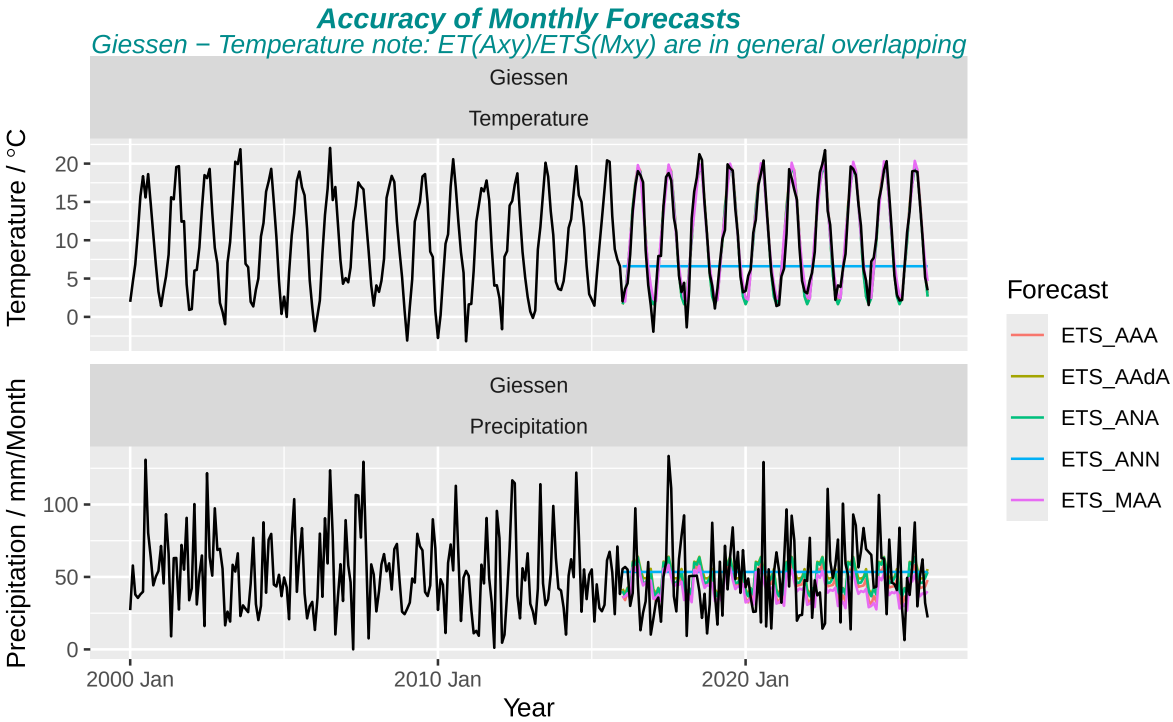
```
#>  2 Giessen Temperature   ETS_AAA     17.1   0.703
#>  3 Giessen Temperature   ETS_AMA     17.4   0.685
#>  4 Giessen Temperature   ETS_AAdA    18.1   0.645
#>  5 Giessen Precipitation ETS_AAA     20.8   0.471
#>  6 Giessen Precipitation ETS_ANA     20.9   0.464
#>  7 Giessen Precipitation ETS_AMA     21.0   0.460
#>  8 Giessen Precipitation ETS_AAdA    21.1   0.454
#>  9 Giessen Precipitation ETS_MAdA    21.5   0.431
#> 10 Giessen Precipitation ETS_MAA     21.9   0.408
#> 11 Giessen Precipitation ETS_MMA     22.1   0.395
#> 12 Giessen Precipitation ETS_MNA     22.4   0.379
#> 13 Giessen Temperature   ETS_MAA     30.2   0.0873
#> 14 Giessen Temperature   ETS_MMA     30.8   0.0775
#> 15 Giessen Temperature   ETS_MAdA    42.8   0.00334
#> 16 Giessen Temperature   ETS_MNA    355.    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  Giessen Temperature   Test  -0.0953  1.47  1.12
#>  2 ETS_AAdA Giessen Temperature   Test   0.196   1.49  1.14
#>  3 ETS_ANA  Giessen Temperature   Test   0.206   1.50  1.14
#>  4 ETS_MAA  Giessen Temperature   Test  -0.0963  1.63  1.30
#>  5 ETS_ANN  Giessen Temperature   Test   3.89    7.47  6.07
#>  6 ETS_AAdA Giessen Precipitation Test   0.454  26.0  21.0
#>  7 ETS_ANA  Giessen Precipitation Test   1.88   26.1  21.0
#>  8 ETS_ANN  Giessen Precipitation Test  -2.70   26.6  21.2
#>  9 ETS_AAA  Giessen Precipitation Test   5.29   26.8  21.5
#> 10 ETS_MAA  Giessen Precipitation Test   8.75   27.8  22.1
```

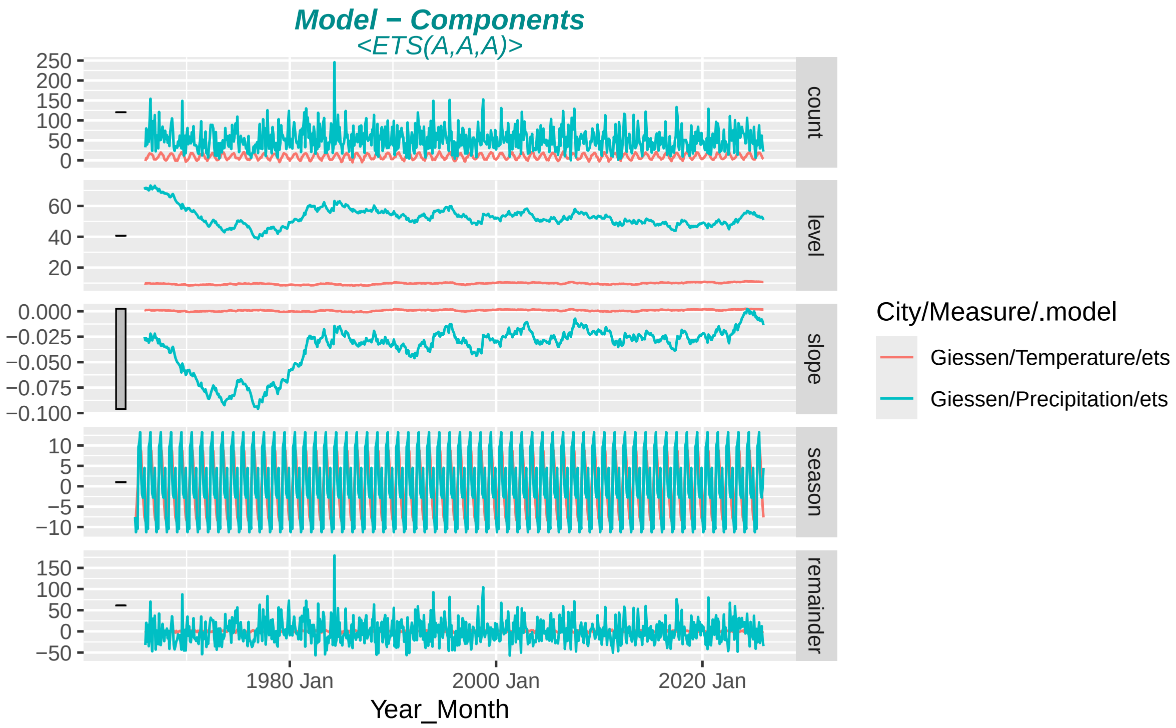


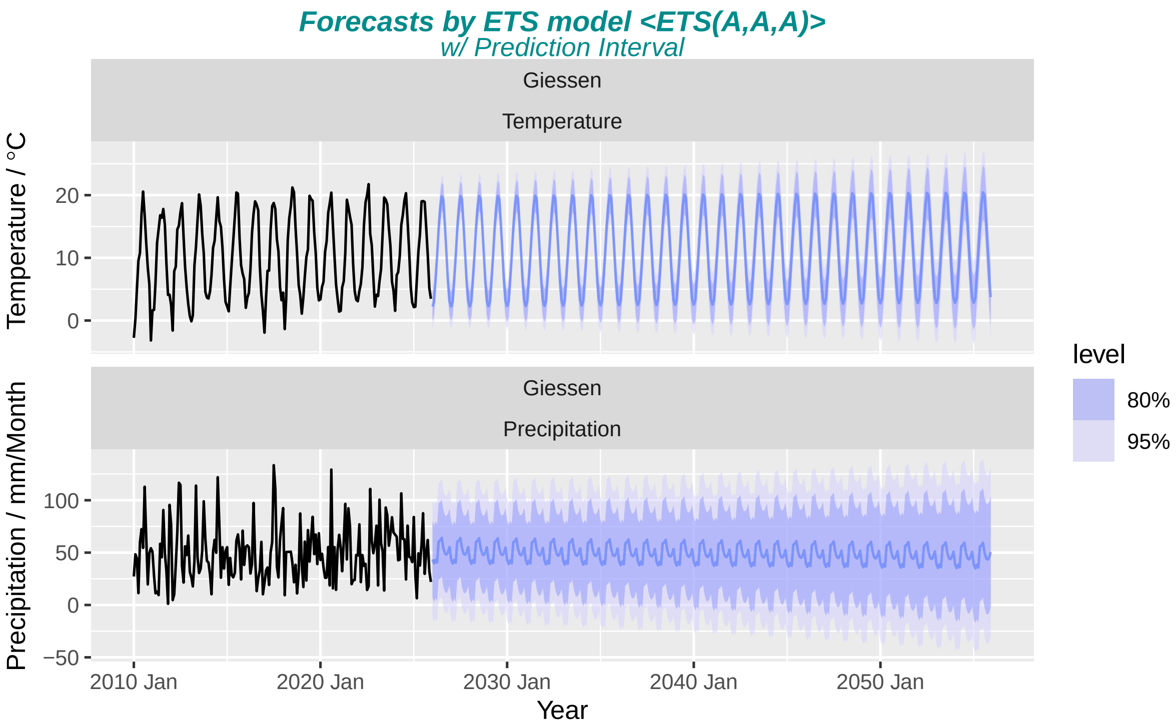
Accuracy of Monthly Forecasts
Giessen − Temperature note: ET(Axy)/ETS(Mxy) are in general overlapping
Giessen
Temperature
Temperature / °C
20
15
10
5
0
Giessen
Precipitation
Precipitation / mm/Month
100
50
0
2000 Jan
2010 Jan
2020 Jan
Year
Forecast
ETS_AAA
ETS_AAdA
ETS_ANA
ETS_ANN
ETS_MAA

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 Giessen Temperatu~ ets       2.98  -2754. 5541. 5542. 5619.  2.91  2.94  1.36
#> 2 Giessen Precipita~ ets     812.    -4772. 9579. 9580. 9657. 794.  794.   21.9
```

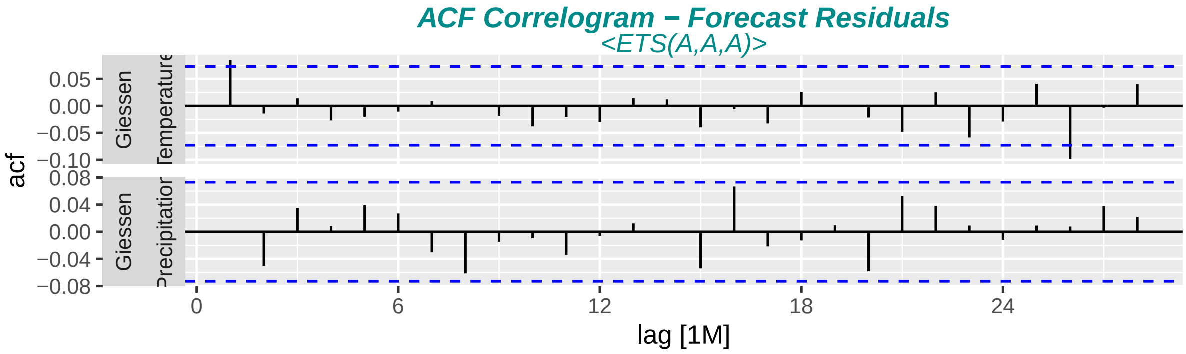


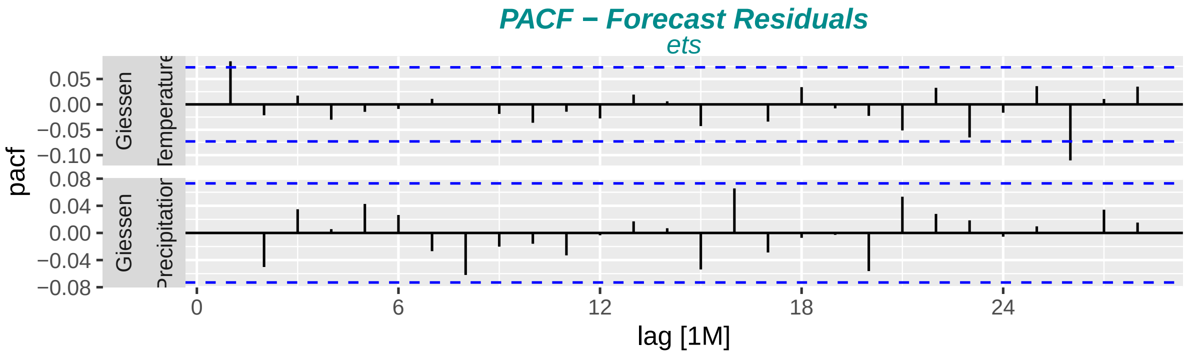


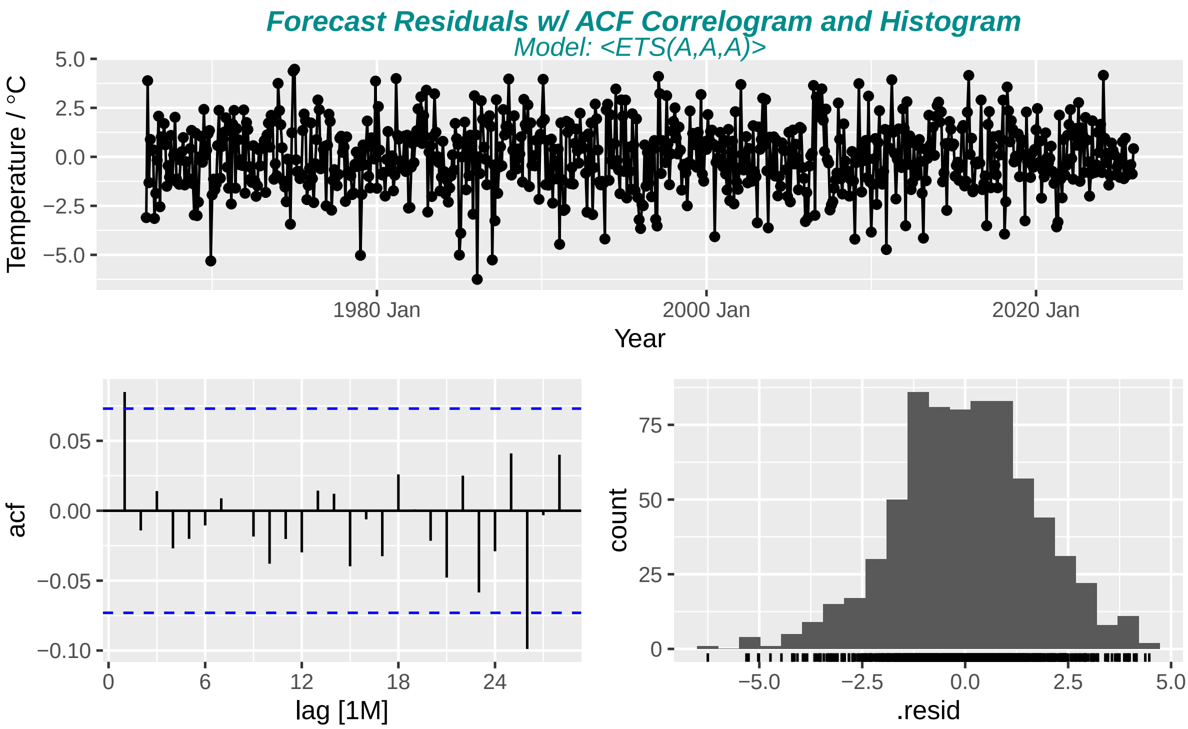
2.2.2 Residual Stationarity

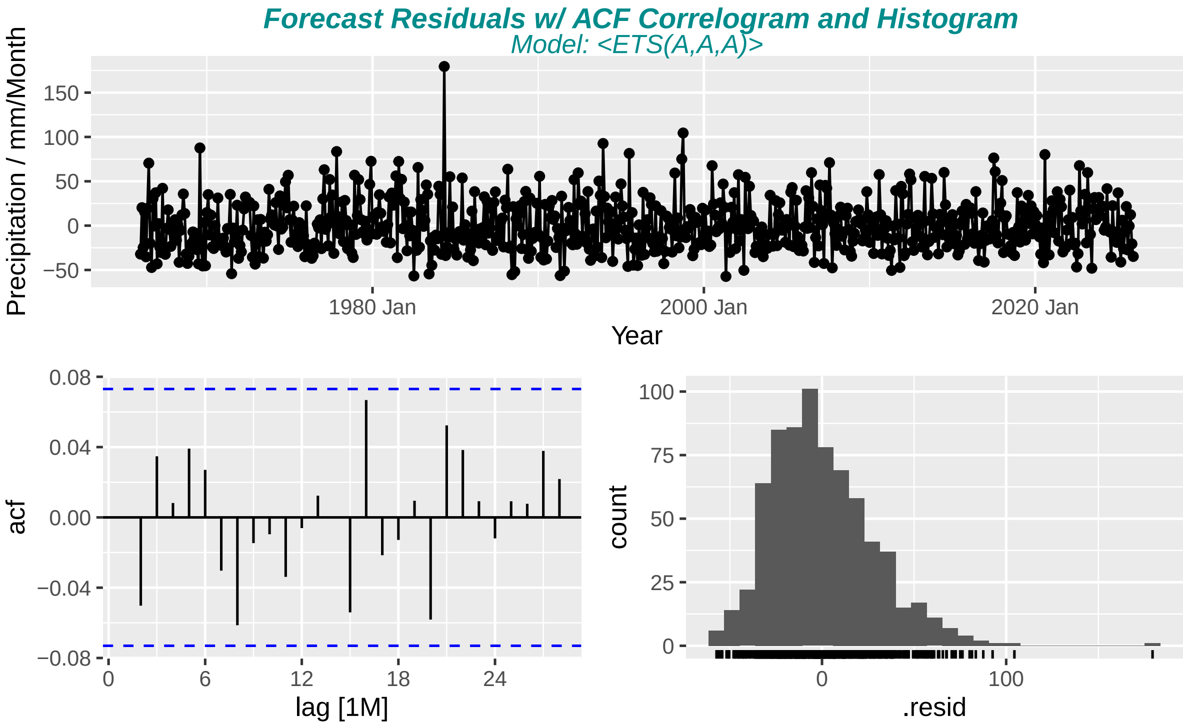
Required checks to be ready for forecasting:

- 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



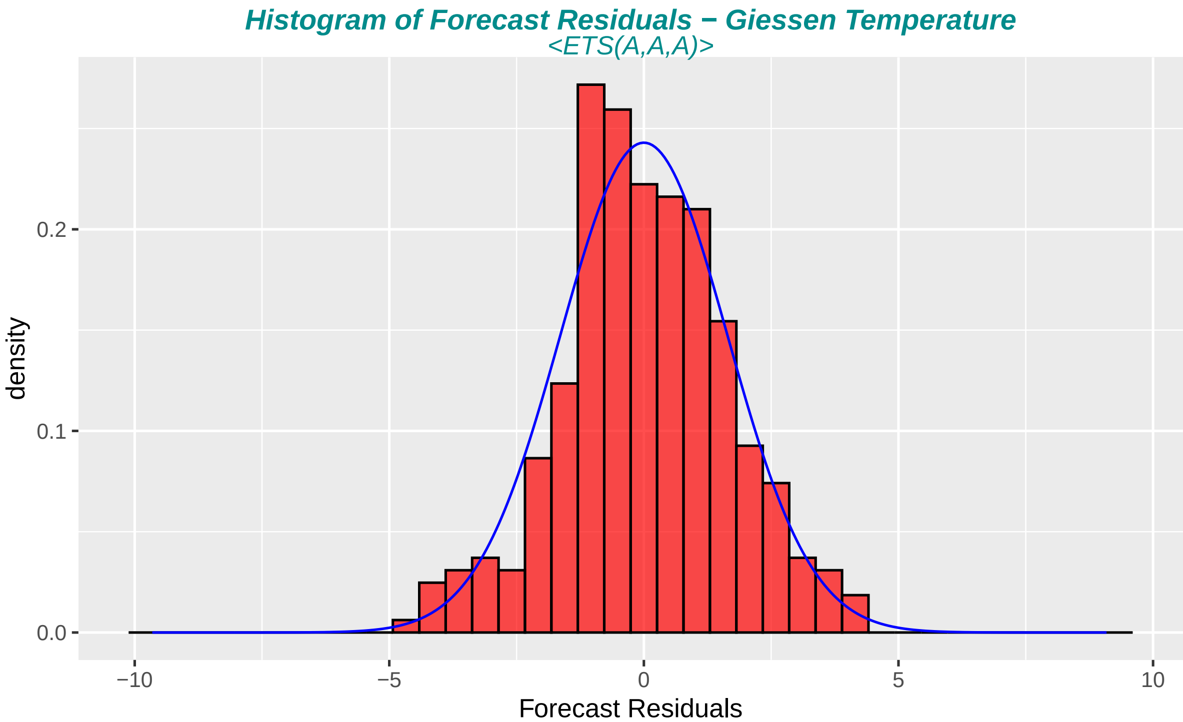


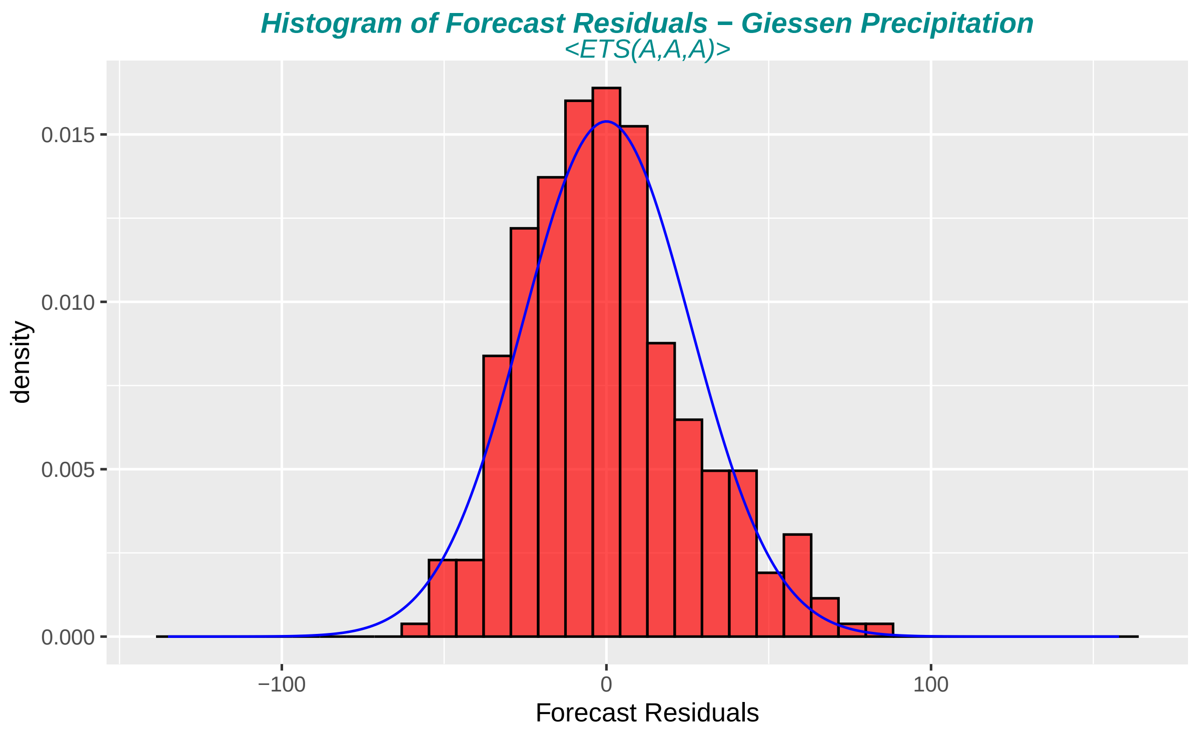




2.2.3 Ljung-Box Test and Histogram of forecast residuals with overlaid normal curve

```
#> 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 Giessen Temperature   ets       23.6     0.482
#> 2 Giessen Precipitation ets       18.5     0.776
```




Histogram of Forecast Residuals − Giessen Precipitation
<ETS(A,A,A)>
density
0.015
0.010
0.005
0.000
−100
0
100
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 Giessen Temperature arima_111_011     2.93  -1408. 2825. 2825. 2843.
#> 2 Giessen Temperature arima_012_011     2.93  -1408. 2825. 2825. 2843.
#> 3 Giessen Temperature arima_211_011     2.93  -1408. 2827. 2827. 2849.
#> 4 Giessen Temperature arima_111_012     2.93  -1408. 2827. 2827. 2849.
#> 5 Giessen Temperature arima_012_112     2.94  -1408. 2829. 2829. 2856.
#> 6 Giessen Temperature arima_100_210     3.96  -1493. 2994. 2994. 3013.
#> 7 Giessen Temperature arima_200_011     4.34  -1524. 3056. 3056. 3074.
#> 8 Giessen Temperature arima_100_110_c   4.34  -1524. 3058. 3058. 3081.
#> # A tibble: 8 x 8
#>   City    Measure       .model        sigma2 log_lik   AIC  AICc   BIC
#>   <chr>   <fct>         <chr>          <dbl>   <dbl> <dbl> <dbl> <dbl>
#> 1 Giessen Precipitation arima_012_011   803.   -3393. 6795. 6795. 6813.
#> 2 Giessen Precipitation arima_111_011   803.   -3393. 6795. 6795. 6813.
#> 3 Giessen Precipitation arima_211_011   803.   -3393. 6796. 6796. 6818.
#> 4 Giessen Precipitation arima_111_012   804.   -3393. 6797. 6797. 6820.
#> 5 Giessen Precipitation arima_012_112   805.   -3393. 6799. 6799. 6826.
#> 6 Giessen Precipitation arima_001_002   839.   -3443. 6897. 6897. 6920.
#> 7 Giessen Precipitation arima_100_210  1042.   -3465. 6939. 6939. 6957.
#> 8 Giessen Precipitation arima_200_011  1177.   -3508. 7023. 7023. 7042.
```

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 Giessen Temperature arima_211_011   Training 0.0505    1.69  1.32
#> 2 Giessen Temperature arima_111_011   Training 0.0499    1.69  1.32
#> 3 Giessen Temperature arima_111_012   Training 0.0499    1.69  1.32
#> 4 Giessen Temperature arima_012_011   Training 0.0489    1.69  1.32
#> 5 Giessen Temperature arima_012_112   Training 0.0489    1.69  1.32
#> 6 Giessen Temperature arima_100_210   Training 0.0384    1.97  1.54
#> 7 Giessen Temperature arima_100_110_c Training 0.000972  2.06  1.60
#> 8 Giessen Temperature arima_200_110_c Training 0.000972  2.06  1.60
#> # A tibble: 8 x 7
#>   City    Measure       .model          .type         ME  RMSE   MAE
#>   <chr>   <fct>         <chr>           <chr>      <dbl> <dbl> <dbl>
#> 1 Giessen Precipitation arima_211_011   Training  1.57    28.0  21.5
#> 2 Giessen Precipitation arima_012_112   Training  1.54    28.0  21.5
#> 3 Giessen Precipitation arima_012_011   Training  1.53    28.0  21.5
#> 4 Giessen Precipitation arima_111_011   Training  1.54    28.0  21.5
#> 5 Giessen Precipitation arima_111_012   Training  1.53    28.0  21.5
#> 6 Giessen Precipitation arima_001_002   Training -0.0274  28.9  22.4
#> 7 Giessen Precipitation arima_100_210   Training -0.628   31.9  25.1
#> 8 Giessen Precipitation arima_100_110_c Training -0.0882  33.9  26.5
```

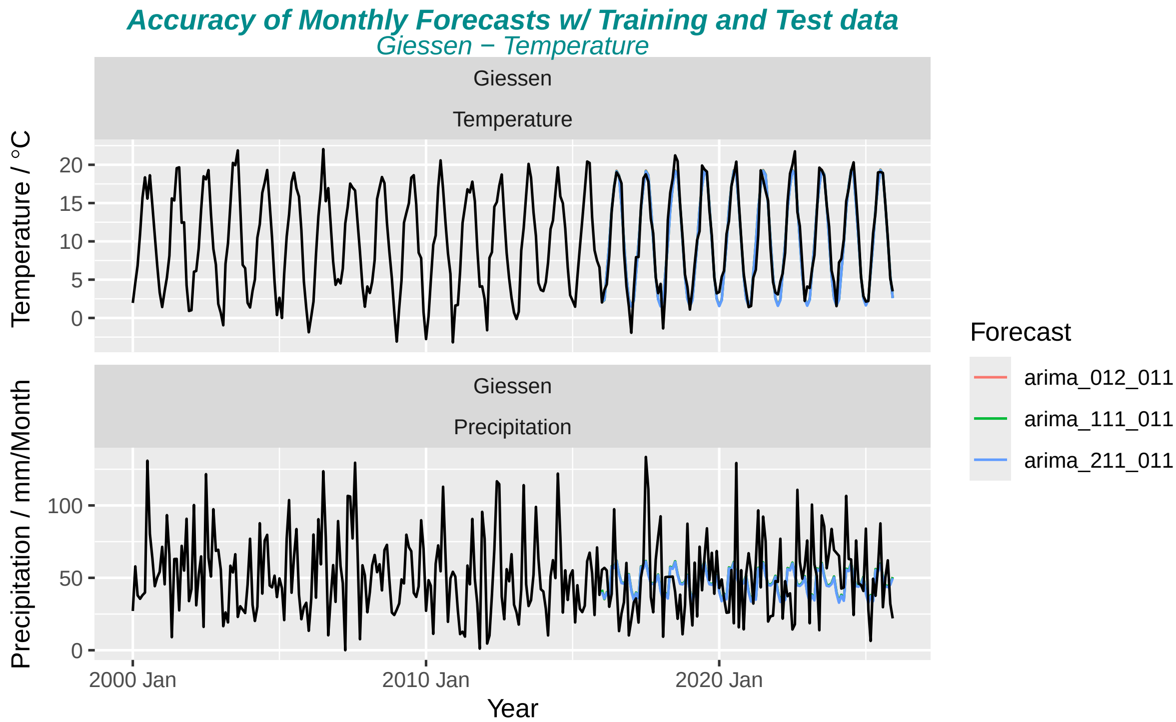
3.1.2 Ljung-Box Test - independence/white noise of the forecasts residuals

```
#> 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 Giessen Temperature arima_211_011      12.5  9.25e- 1
#> 2 Giessen Temperature arima_111_011      12.5  9.24e- 1
#> 3 Giessen Temperature arima_111_012      12.5  9.24e- 1
#> 4 Giessen Temperature arima_012_112      12.6  9.22e- 1
#> 5 Giessen Temperature arima_012_011      12.6  9.22e- 1
#> 6 Giessen Temperature arima_100_210      36.5  1.93e- 2
#> 7 Giessen Temperature arima_200_011      97.2  8.93e-12
#> 8 Giessen Temperature arima_100_110_c    97.3  8.76e-12
#> # A tibble: 8 x 5
#>   City    Measure       .model          lb_stat lb_pvalue
#>   <chr>   <fct>         <chr>             <dbl>     <dbl>
#> 1 Giessen Precipitation arima_211_011      19.8  5.31e- 1
#> 2 Giessen Precipitation arima_111_012      21.0  4.59e- 1
#> 3 Giessen Precipitation arima_012_112      21.2  4.50e- 1
#> 4 Giessen Precipitation arima_111_011      21.2  4.48e- 1
#> 5 Giessen Precipitation arima_012_011      21.2  4.47e- 1
#> 6 Giessen Precipitation arima_001_002      22.5  3.72e- 1
#> 7 Giessen Precipitation arima_100_210      47.0  9.42e- 4
#> 8 Giessen Precipitation arima_100_110_c   116.   4.55e-15
```

3.1.3 Forecast Accuracy with Training/Test Data

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

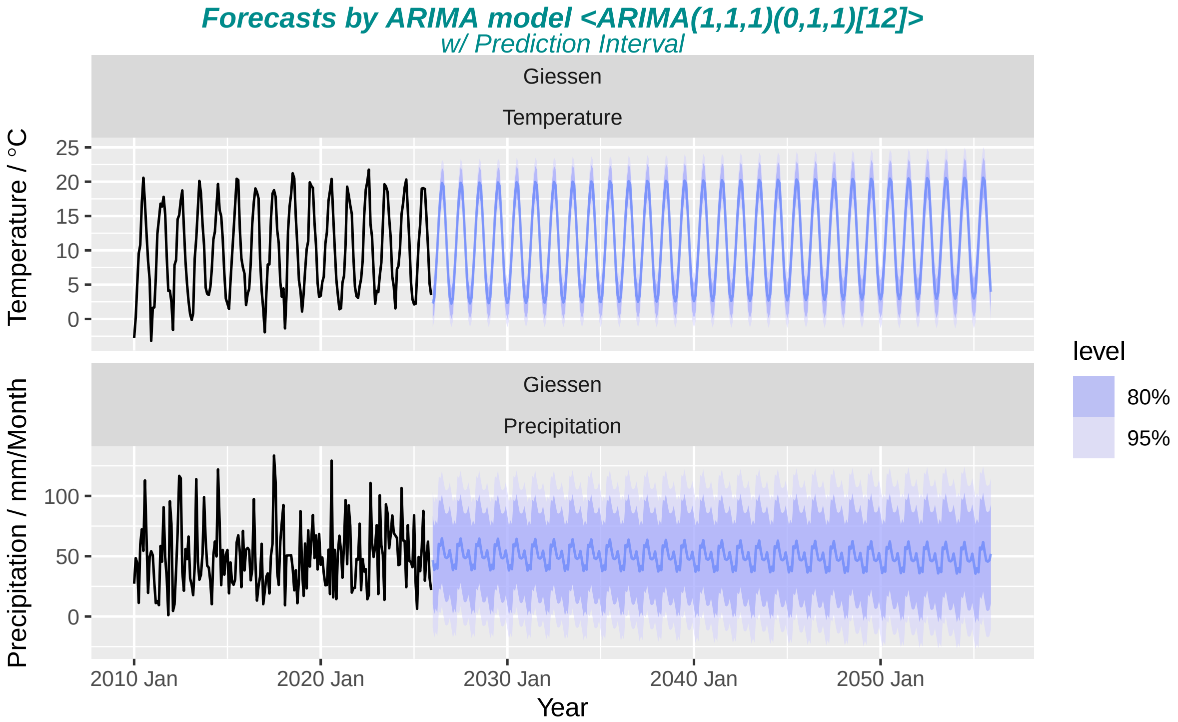
```
#>   .model         City    Measure       .type    ME  RMSE   MAE
#>   <chr>          <chr>   <fct>         <chr> <dbl> <dbl> <dbl>
#> 1 arima_012_011 Giessen Temperature   Test  0.289  1.50  1.14
#> 2 arima_111_011 Giessen Temperature   Test  0.337  1.51  1.15
#> 3 arima_211_011 Giessen Temperature   Test  0.378  1.52  1.15
#> 4 arima_012_011 Giessen Precipitation Test  3.66  26.2  21.0
#> 5 arima_111_011 Giessen Precipitation Test  3.66  26.2  21.0
#> 6 arima_211_011 Giessen Precipitation Test  4.24  26.3  21.1
```



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

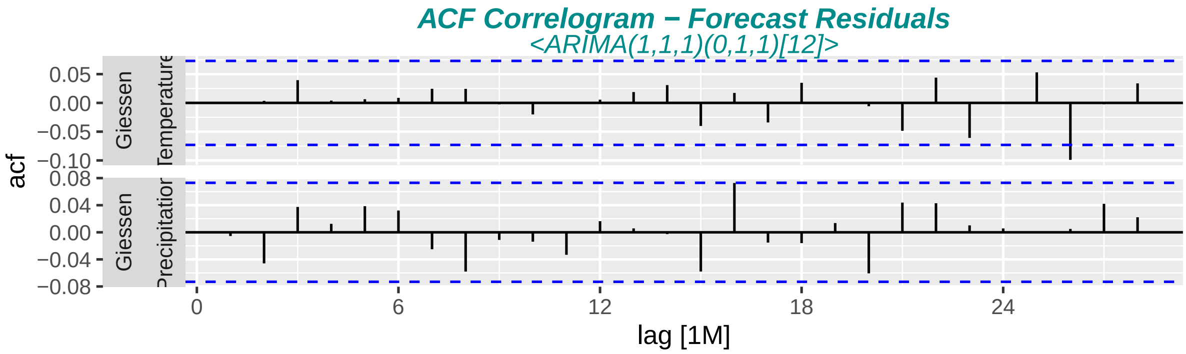
```
#> Provide model coefficients by report(fit_model)
#> # A tibble: 2 x 10
#>   City    Measure       .model sigma2 log_lik   AIC  AICc   BIC ar_roots ma_roots
#>   <chr>   <fct>         <chr>   <dbl>   <dbl> <dbl> <dbl> <dbl> <list>   <list>
#> 1 Giessen Temperature  arima    2.93  -1408. 2825. 2825. 2843. <cpl>    <cpl>
#> 2 Giessen Precipitati~ arima  803.    -3393. 6795. 6795. 6813. <cpl>    <cpl>
```

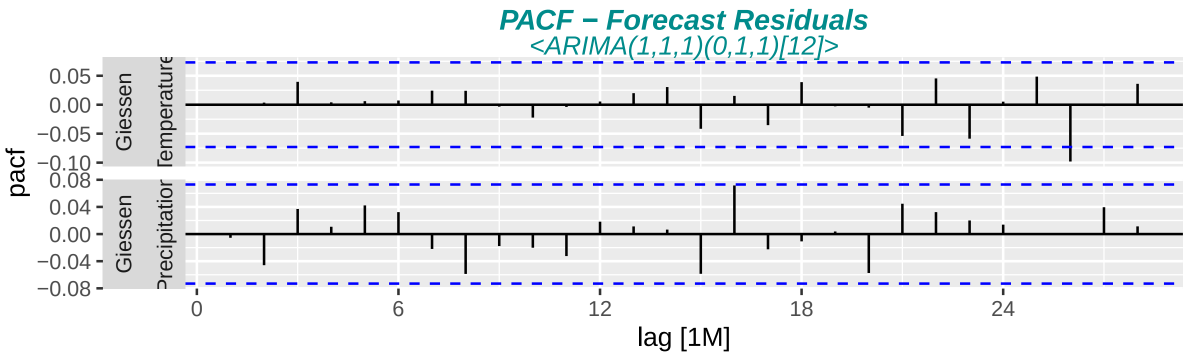


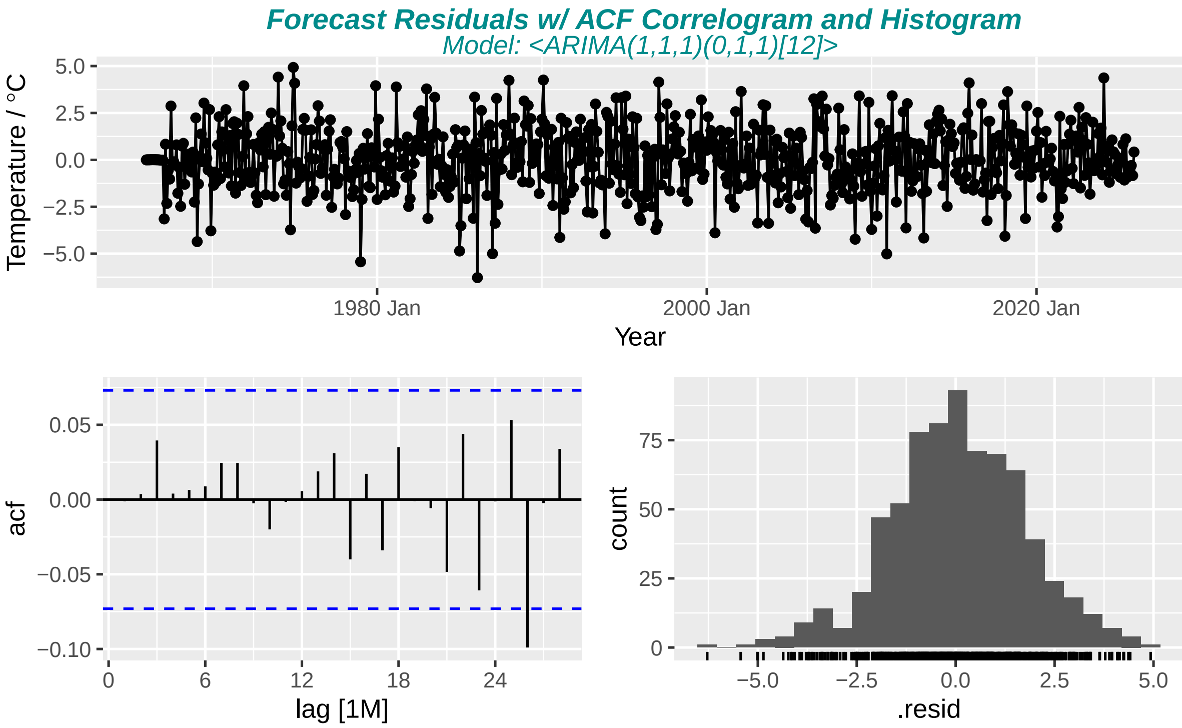
3.2.2 Residual Stationarity

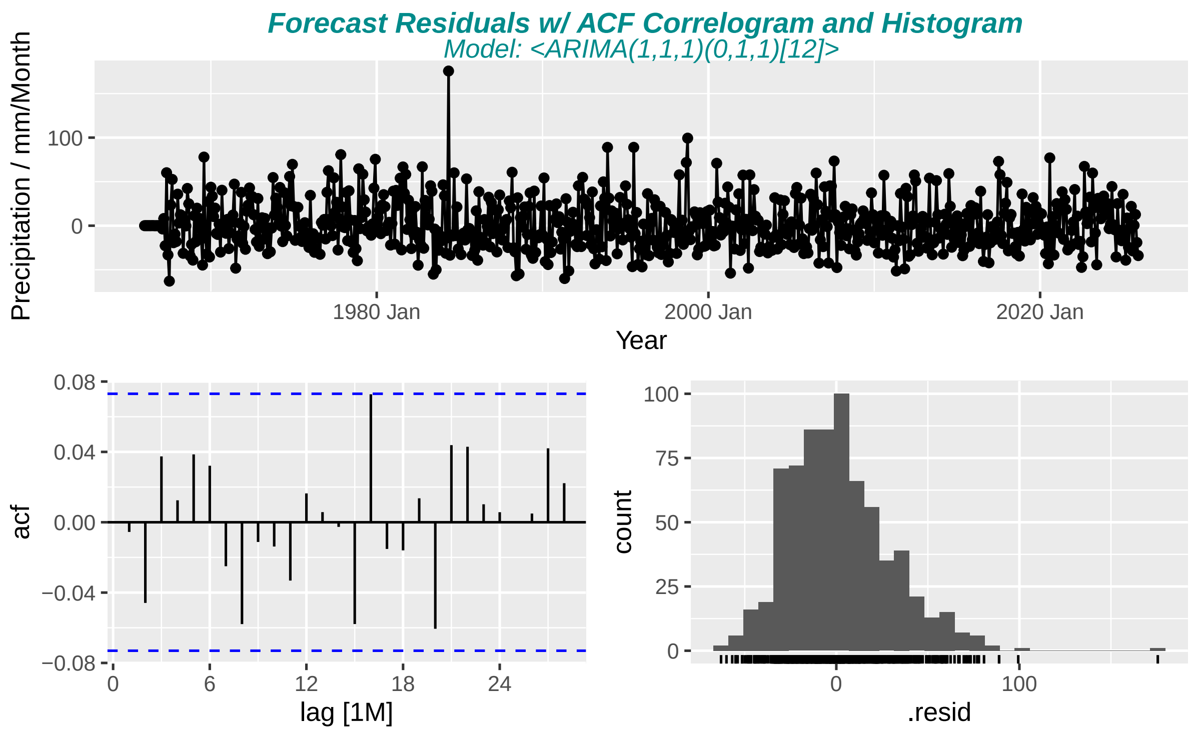
Required checks to be ready for forecasting:

- 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
Giessen
Giessen
lag [1M]

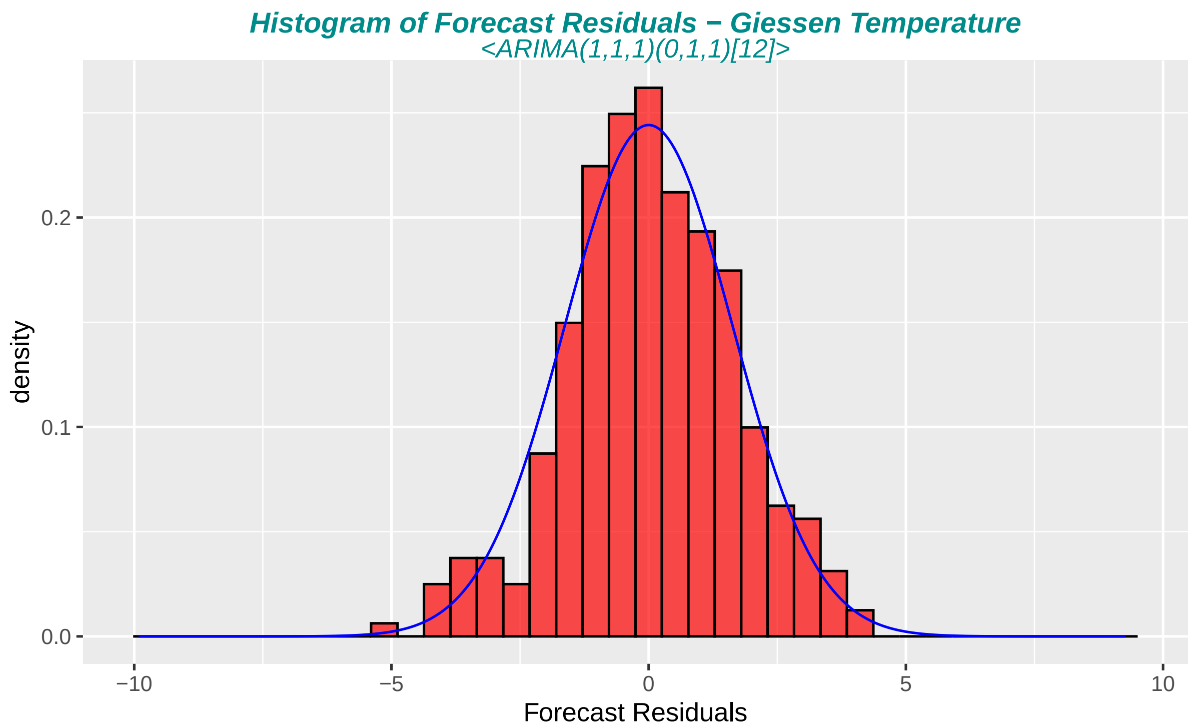

PACF − Forecast Residuals
<ARIMA(1,1,1)(0,1,1)[12]>
pacf
Giessen
Giessen
lag [1M]

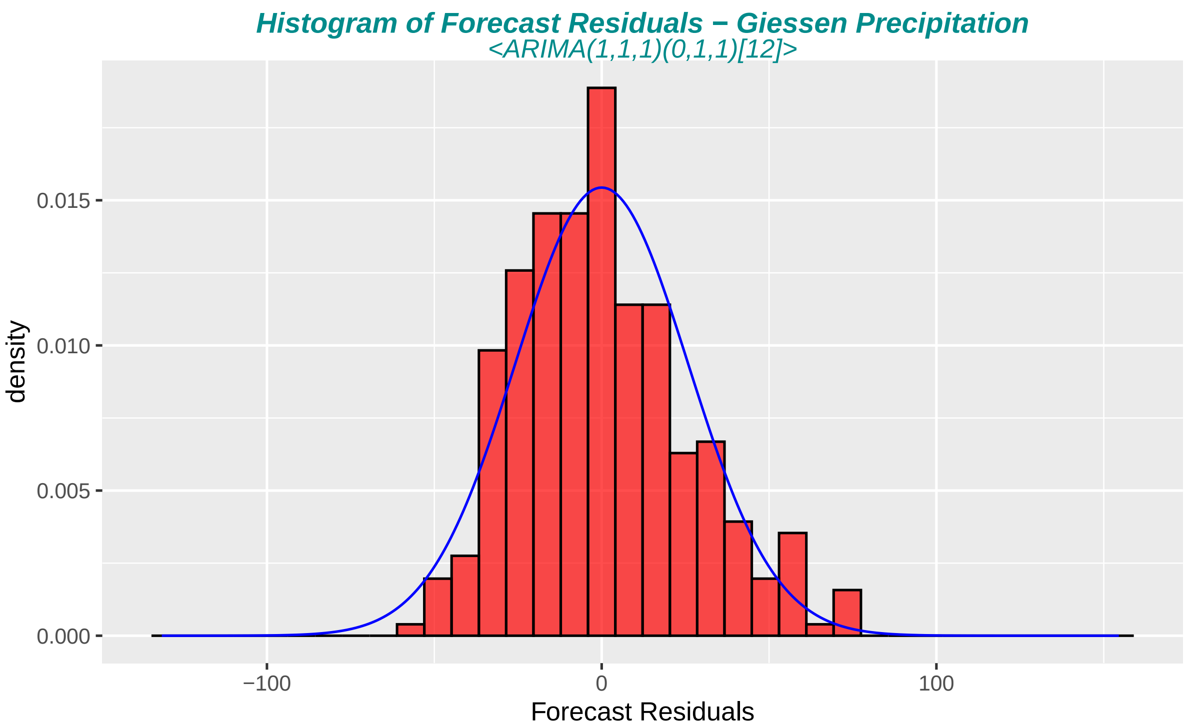

Forecast Residuals w/ ACF Correlogram and Histogram
Model: <ARIMA(1,1,1)(0,1,1)[12]>
Temperature / °C
1980 Jan
2000 Jan
2020 Jan
Year
acf
lag [1M]
count
.resid



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 Giessen Temperature   arima     19.1     0.582
#> 2 Giessen Precipitation arima     17.9     0.653
```




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

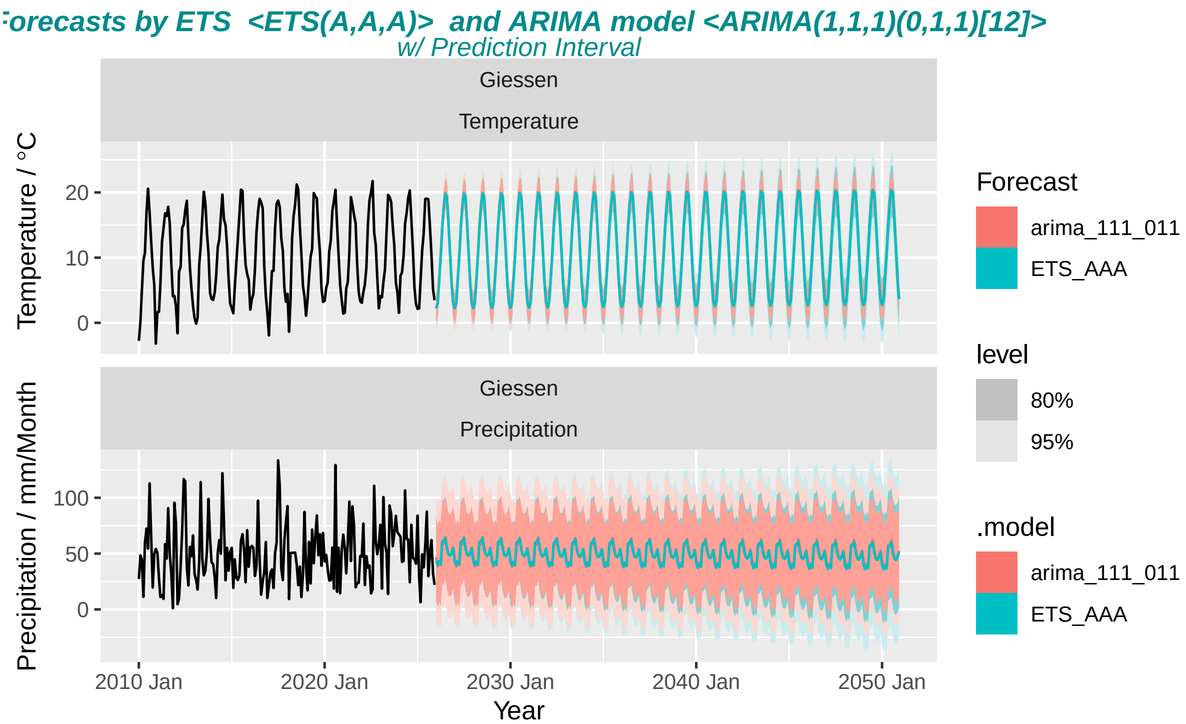
- Residual Accuracy with one-step-ahead fitted residuals
- Forecast Accuracy with Training/Test Data

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 Giessen Temperature   ETS_AAA       Test      1.47  1.12  20.2 0.597 0.606
#> 2 Giessen Temperature   arima_111_011 Test      1.51  1.15  19.5 0.610 0.622
#> 3 Giessen Temperature   arima         Training  1.69  1.32  78.0 0.708 0.704
#> 4 Giessen Temperature   ets           Training  1.71  1.36  74.3 0.728 0.710
#> 5 Giessen Precipitation arima_111_011 Test     26.2  21.0   62.3 0.679 0.665
#> 6 Giessen Precipitation ETS_AAA       Test     26.8  21.5   61.4 0.694 0.679
#> 7 Giessen Precipitation arima         Training 28.0  21.5  Inf   0.698 0.712
#> 8 Giessen Precipitation ets           Training 28.2  21.9  Inf   0.712 0.716
```

4.0.2 Forecast Plot of selected ETS and ARIMA model



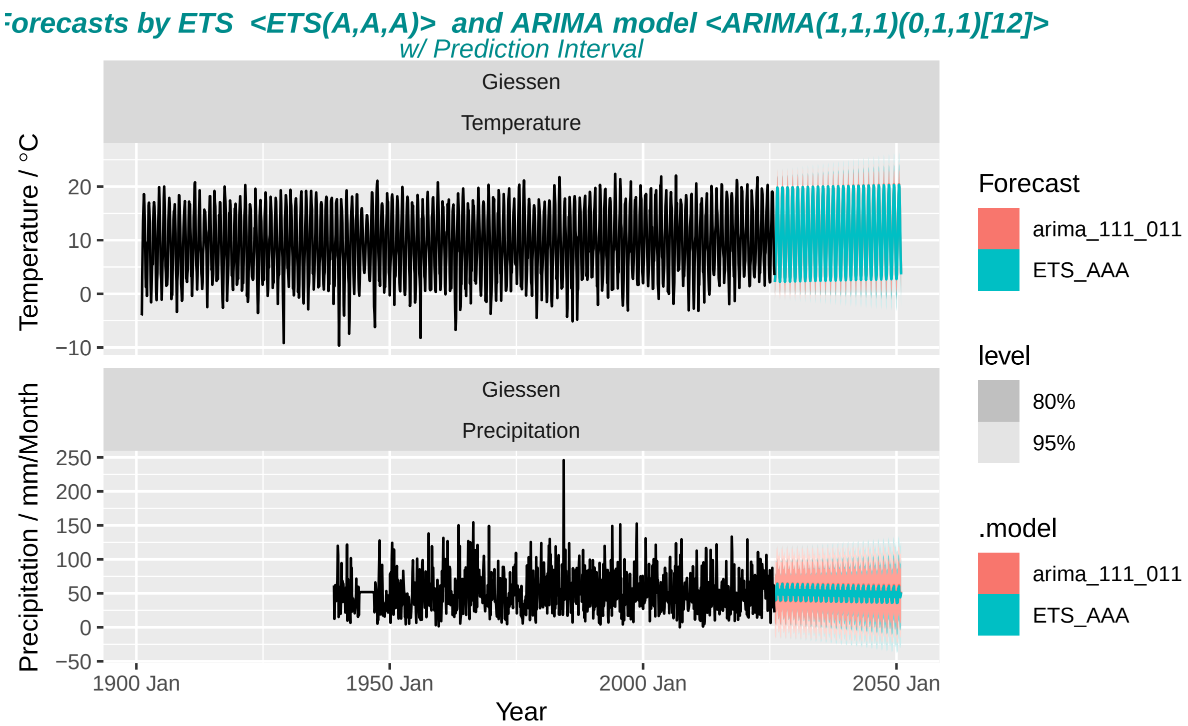

Forecasts by ETS <ETS(A,A,A)> and ARIMA model <ARIMA(1,1,1)(0,1,1)[12]>
w/ Prediction Interval
Giessen
Temperature
Temperature / °C
20
10
0
−10
Giessen
Precipitation
Precipitation / mm/Month
250
200
150
100
50
0
−50
1900 Jan
1950 Jan
2000 Jan
2050 Jan
Year
Forecast
arima_111_011
ETS_AAA
level
80%
95%
.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
1901-1930	8.9	NA
1931-1960	9.0	48.0
1961-1990	9.1	54.4
1991-2020	9.9	51.8
2021-2025	10.6	53.7

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
Giessen	Temperature	2026	10.73	10.76
Giessen	Temperature	2030	10.81	10.87
Giessen	Temperature	2035	10.92	11.00
Giessen	Temperature	2040	11.02	11.13
Giessen	Temperature	2045	11.13	11.27
Giessen	Temperature	2050	11.23	11.40
Giessen	Precipitation	2026	50.88	50.92
Giessen	Precipitation	2030	50.23	50.53
Giessen	Precipitation	2035	49.42	50.00
Giessen	Precipitation	2040	48.60	49.47
Giessen	Precipitation	2045	47.79	48.94
Giessen	Precipitation	2050	46.98	48.41

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	10.73	10.76	11.23	11.40	0.51	0.64
Precipitation	2026	2050	50.88	50.92	46.98	48.41	-3.90	-2.51

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	2.22	2.26	2.73	2.86	0.51	0.61
Temperature	Feb	2026	2050	3.03	3.10	3.54	3.74	0.51	0.64
Temperature	Mrz	2026	2050	6.41	6.47	6.92	7.12	0.51	0.64
Temperature	Apr	2026	2050	10.15	10.22	10.65	10.86	0.51	0.64
Temperature	Mai	2026	2050	14.66	14.66	15.17	15.31	0.51	0.64
Temperature	Jun	2026	2050	17.93	17.98	18.44	18.62	0.51	0.64
Temperature	Jul	2026	2050	19.82	19.81	20.32	20.45	0.51	0.64
Temperature	Aug	2026	2050	19.37	19.33	19.87	19.97	0.51	0.64
Temperature	Sep	2026	2050	15.26	15.29	15.76	15.93	0.51	0.64
Temperature	Okt	2026	2050	10.80	10.80	11.30	11.45	0.51	0.64
Temperature	Nov	2026	2050	6.00	6.02	6.51	6.67	0.51	0.64
Temperature	Dez	2026	2050	3.10	3.20	3.61	3.84	0.51	0.64
Precipitation	Jan	2026	2050	43.46	46.76	39.56	44.56	-3.90	-2.20

Measure	Month	Year.x	Year.y	ETS.x	ARIMA.x	ETS.y	ARIMA.y	Delta_ETS	Delta_ARIMA
Precipitation	Feb	2026	2050	39.66	38.62	35.76	36.08	-3.90	-2.54
Precipitation	Mrz	2026	2050	43.14	43.16	39.24	40.61	-3.90	-2.54
Precipitation	Apr	2026	2050	40.47	39.74	36.57	37.19	-3.90	-2.54
Precipitation	Mai	2026	2050	60.59	60.25	56.69	57.71	-3.90	-2.54
Precipitation	Jun	2026	2050	61.65	58.70	57.75	56.15	-3.90	-2.54
Precipitation	Jul	2026	2050	64.14	64.83	60.24	62.29	-3.90	-2.54
Precipitation	Aug	2026	2050	54.40	56.36	50.50	53.81	-3.90	-2.54
Precipitation	Sep	2026	2050	49.15	49.33	45.25	46.79	-3.90	-2.54
Precipitation	Okt	2026	2050	48.06	48.58	44.15	46.03	-3.90	-2.54
Precipitation	Nov	2026	2050	50.51	49.69	46.61	47.15	-3.90	-2.54
Precipitation	Dez	2026	2050	55.32	55.07	51.42	52.53	-3.90	-2.54

5 Yearly Data Forecasts with ARIMA and ETS

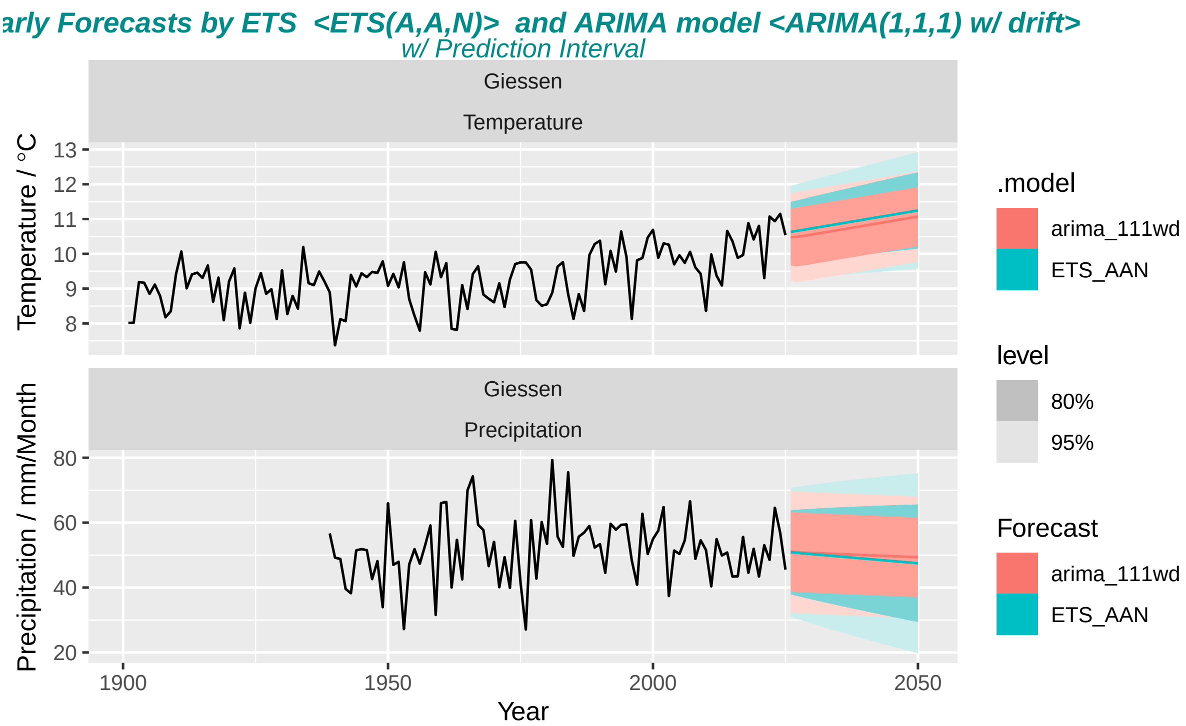
For yearly data the seasonal monthly data are replaced by the yearly average data. Therefore the seasonal component of the ETS and ARIMA model are to be taken out.

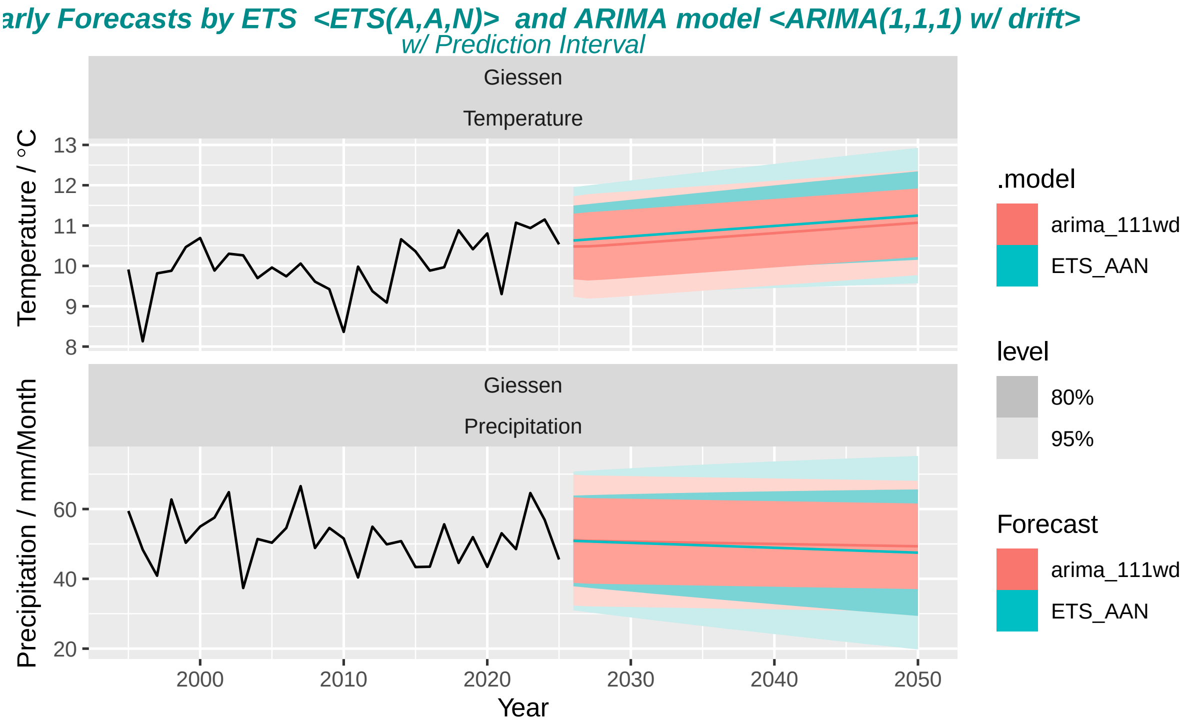
The ETS model $< ETS(A, A, N) >$ with seasonal term change "A" -> "N" is chosen. For ARIMA models the seasonal term (P,D,Q)m has to be taken out and an optimal ARIMA(p,1,q) with one differencing (d=1) is selected. However, for Mauna Loa two times differncing had to be selected $CO_2 <ARIMA(0,2,1) w/ 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 Giessen Tempera~ arima~   0.397   -56.7  121.  122.  130. NA     NA     NA
#> 2 Giessen Tempera~ ETS_A~   0.456   -97.2  204.  205.  215.  0.425  0.461  0.539
#> 3 Giessen Precipi~ arima~  90.1    -217.   442.  443.  450. NA     NA     NA
#> 4 Giessen Precipi~ ETS_A~ 103.     -260.   530.  531.  540. 96.1   99.8    7.60
#> # 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 Giessen Temperature   ETS_AAN    21.5     0.608
#> 2 Giessen Precipitation ETS_AAN    19.1     0.745
#> # A tibble: 2 x 5
#>   City    Measure       .model      lb_stat lb_pvalue
#>   <chr>   <fct>         <chr>         <dbl>     <dbl>
#> 1 Giessen Temperature   arima_111wd    17.9     0.658
#> 2 Giessen Precipitation arima_111wd    21.4     0.437
```

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
Giessen	Temperature	2026	10.63	10.48
Giessen	Temperature	2030	10.73	10.55
Giessen	Temperature	2035	10.86	10.68
Giessen	Temperature	2040	10.99	10.81
Giessen	Temperature	2045	11.12	10.94
Giessen	Temperature	2050	11.25	11.07
Giessen	Precipitation	2026	50.87	51.11
Giessen	Precipitation	2030	50.31	50.68
Giessen	Precipitation	2035	49.60	50.34
Giessen	Precipitation	2040	48.90	50.00
Giessen	Precipitation	2045	48.20	49.66
Giessen	Precipitation	2050	47.49	49.33

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	10.63	10.48	11.25	11.07	0.62	0.59
Precipitation	2026	2050	50.87	51.11	47.49	49.33	-3.37	-1.78