Coronavirus SARS-CoV-2 / Covid-19

Wolfgang Vollmer

2020-04-01

Contents

1 World Map

World Map with Confirmed and Death Cases

1.0.1 Total Confirmed for each Country

1.0.2 Total Deaths for each Country

2 Bar Chart

Bar Charts with descending order

2.1 Row

2.1.1 Bar Chart Confirmed

2.1.2 Bar Chart Table Deaths

2.2 Row

2.2.1 Data Table for Confirmed and Deaths - Cases and Daily Differences

3 Trend / Daily Change

Trend / Daily Changer

3.1 Row

3.1.1 World

World Cumulated Cases / Daily Changes over Time

3.2 Row

Selected Countries - Cumulated Cases / Daily Changes over Time

3.2.1 China

3.2.2 China

3.2.3 China

3.2.4 China

3.2.5 China

3.2.6 China

3.2.7 China

3.2.8 China

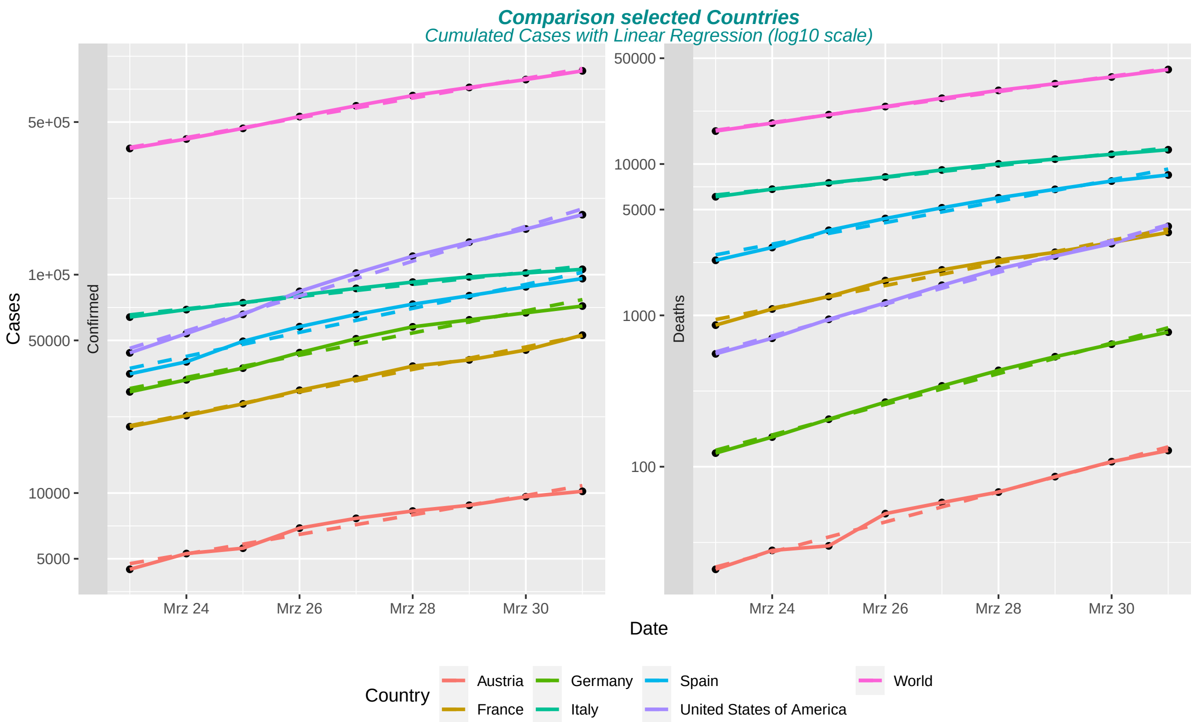
3.2.9 Germany - Confirmed

3.2.10 Germany - Deaths

4 Comparison Exponential Growth

4.1 Column

4.1.1 Comparison Exponential Growth of selected Countries



4.1.2 Estimation speed of spread of the Coronavirus with Linear Regression

Exponential Growth and Doubling Time T

Exponential growth over time can be fitted by linear regression if the logarithms of the case numbers is taken. Generally, exponential growth corresponds to linearly growth over time for the log (to any base) data.

The semi-logorithmic plot with base-10 log scale for the Y axis shows functions following an exponential law $y(t) = y_0 * a^{t/\tau}$ as straight lines. The time constant τ describes the time required for y to increase by one factor of a.

If e.g. the confirmed or death cases are growing in $t - days$ by a factor of 10 the doubling time $T \widehat{=} \tau$ can be calculated with $a \widehat{=} 2$ by

$T[days] = \frac{t[days]*log_{10}(2)}{log_{10}(y(t)) - log_{10}(y_0)}$

with

$$log_{10}(y(t)) - log_{10}(y_0) == log_{10}(y(t))/y_0) = log_{10}(10 * y_0/y_0) = 1$$

and doubling time

$$T[days] = t[days] * log_{10}(2) \approx t[days] * 0.30.$$

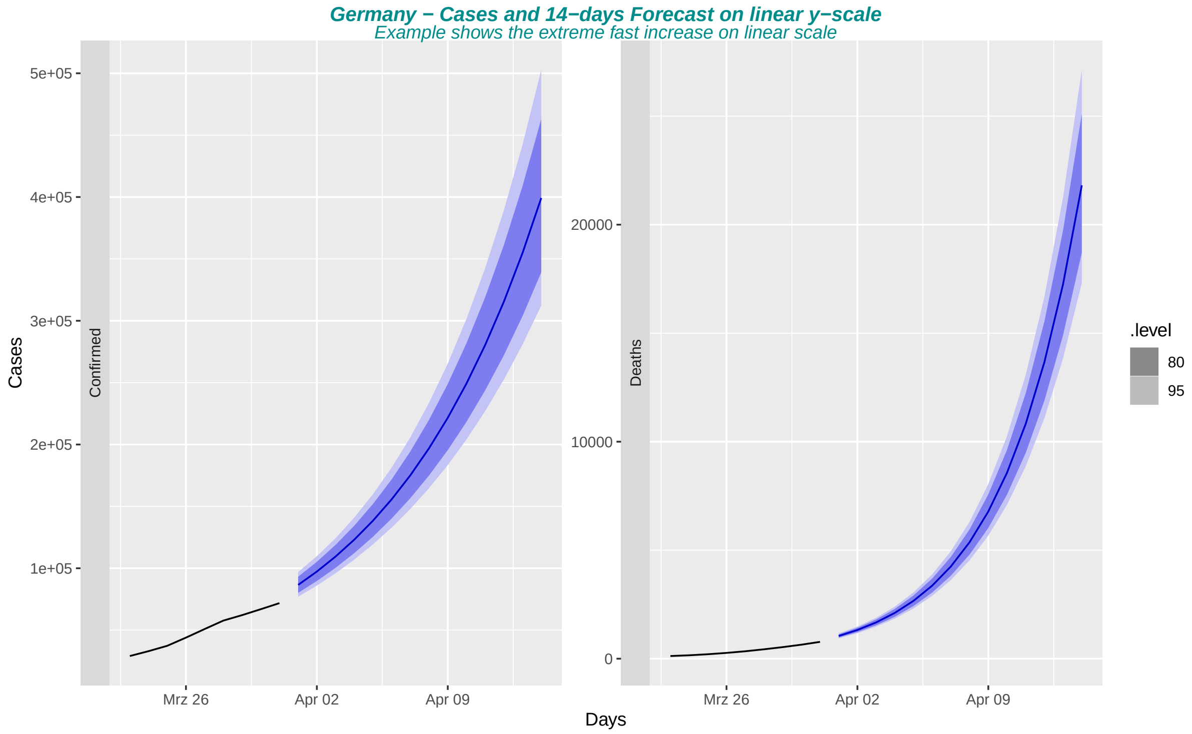
For Spain, Italy, Germany we have had a doubling time up to $T \approx 9 - 12days * 0.3 \approx 2.7 - 4days!!$.

The *doubling time T* and the *Forecast* is calculated for following selected countries: **Austria, France, Germany, Italy, Spain, United States of America**

4.2 Column

4.2.1 Germany - Example plot with ~linear slope on a log10 scale

4.2.2 Germany - Trend with Forecast on a linear scale



4.2.3 Forecast Plot - next 14 days

The plot shows the extreme forecast increase in case of unchecked exponentiell growth. The dark shaded regions show 80% rsp. 95% prediction intervals. These prediction intervals are displaying the uncertainty in forecasts based on the linear regression over the past 9 days.

5 Doubling Time / Forecast

5.1 Column

Doubling Time and Forecast

The forecasted cases for the next 14 days are calculated 'only' from the linear regression of the logarithmic data and are not considering any effects of measures in place. In addition data inaccuracies are not taken into account, especially relevant for the confirmed cases.

Therefore the 14 days forecast is only an indication for the direction of an unchecked exponentiell growth.

Table 1: Forecast (FC) with linear regression: Doubling Time (days), Forecasted cases tomorrow and Forecasted cases in 14 days

Country	Case_Type	T_doubling	last_day	FC_next_day	FC_14days
Austria	Confirmed	6.7	10'180	11'982	45'629
France	Confirmed	5.8	52'827	59'146	276'454
Germany	Confirmed	5.9	71'808	86'396	396'298
Italy	Confirmed	10.8	105'792	116'705	269'038
Spain	Confirmed	5.5	95'923	116'213	601'664
United States of America	Confirmed	3.8	188'172	240'480	2'622'639
World	Confirmed	6.7	857'487	975'302	3'768'654
Austria	Deaths	3.0	128	170	3'319
France	Deaths	4.0	3'532	4'420	41'456
Germany	Deaths	3.0	775	1'047	21'670
Italy	Deaths	7.7	12'428	13'986	44'833
Spain	Deaths	4.2	8'464	10'901	90'869
United States of America	Deaths	2.9	3'873	5'062	117'586
World	Deaths	5.9	42'107	48'076	221'001

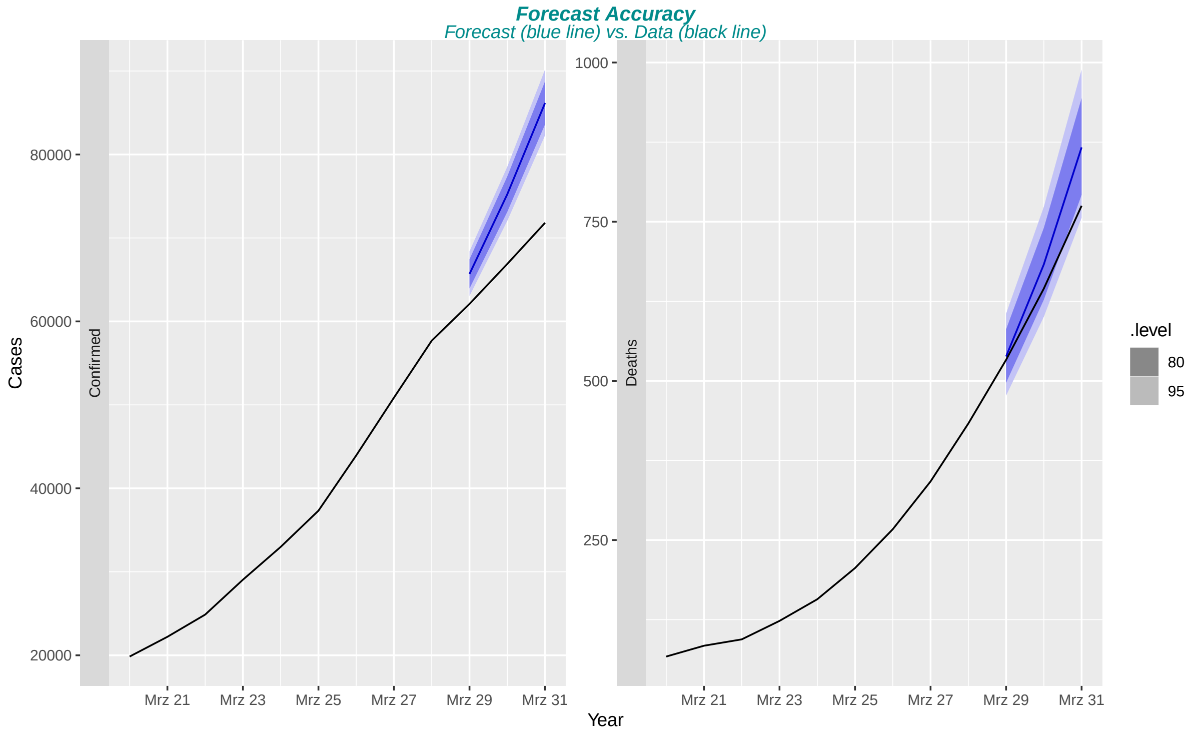
5.2 Column

5.2.1 Check of Forecast Accuracy

The forecast accuracy is checked by using the forecast method for the nine days before the past three days (*training data*). Subsequent forecasting of the past three days enables comparison with the real data of these days (*test data*).

The comparison is also an early indicator if the *exponential growth* is declining. However, possible changes in underreporting (in particular the proportion confirmed / actually infected) requires careful interpretation.

5.2.2 Germany - Forecast Accuracy for past three days



6 Forecast

6.1 Column

6.1.1 Forecasting with lagged Predictors

The number of confirmed cases can be used as a time delayed predictor of the number of deaths. This will allow comclusions on the time period confirmed to death. More inportant the country specific case fatality rate (CFR, proportion of deaths from confirmed cases) indicates the country specific testing.

Overall a rough conclusion on the country specific proportion of infected to confirmed cases is feasible if the infection fatality rate (IFR, confiremd cases plus all asymptomatic and undiagnosed infections) is assumed to be country independent and the IFR is known (assumption by RKI ~ 0.56%, bottom of existing estimates, see https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Modellierung_Deutschland.pdf?__blob=publicationFile).

In the model paper RKI assumes for the

- Incubation period ~ 5-6 days - Day of infection day until symptoms are upcoming)
- Hospitalisation +4 days - Admission to the hospital (if needed) after Incubation Period)
- Average period to death + 11 - if the patient dies, it takes an average of 11 days after admission to the hospital

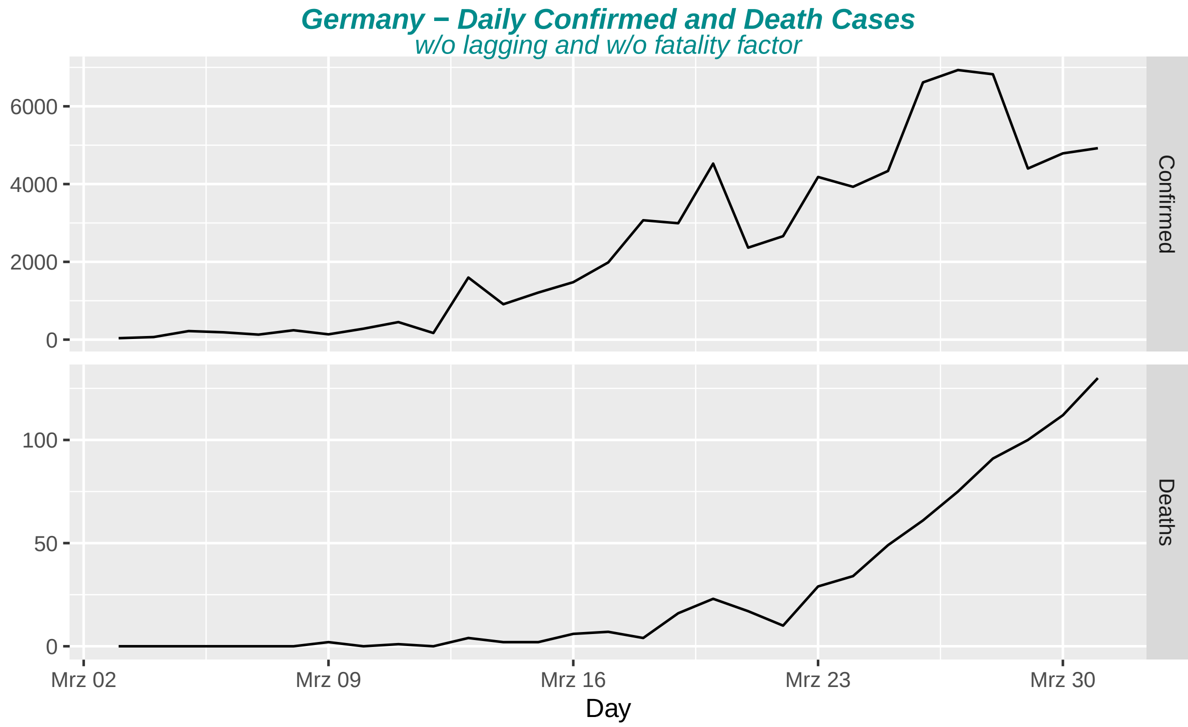
Depending on the country-specific test frequency (late or early tests), the

*lag_days - time from receipt of the confirmed test result to death, Confirmed to Death, is about 11-13 days.

Note: these methods are also used for example for advertising campaigns. The campaign impact on sales will be some time beyond the end of the campaign, and sales in one month will depend on the advertising expenditure in each of the past few months (see https://otexts.com/fpp3/lagged-predictors.html).

6.2 Column

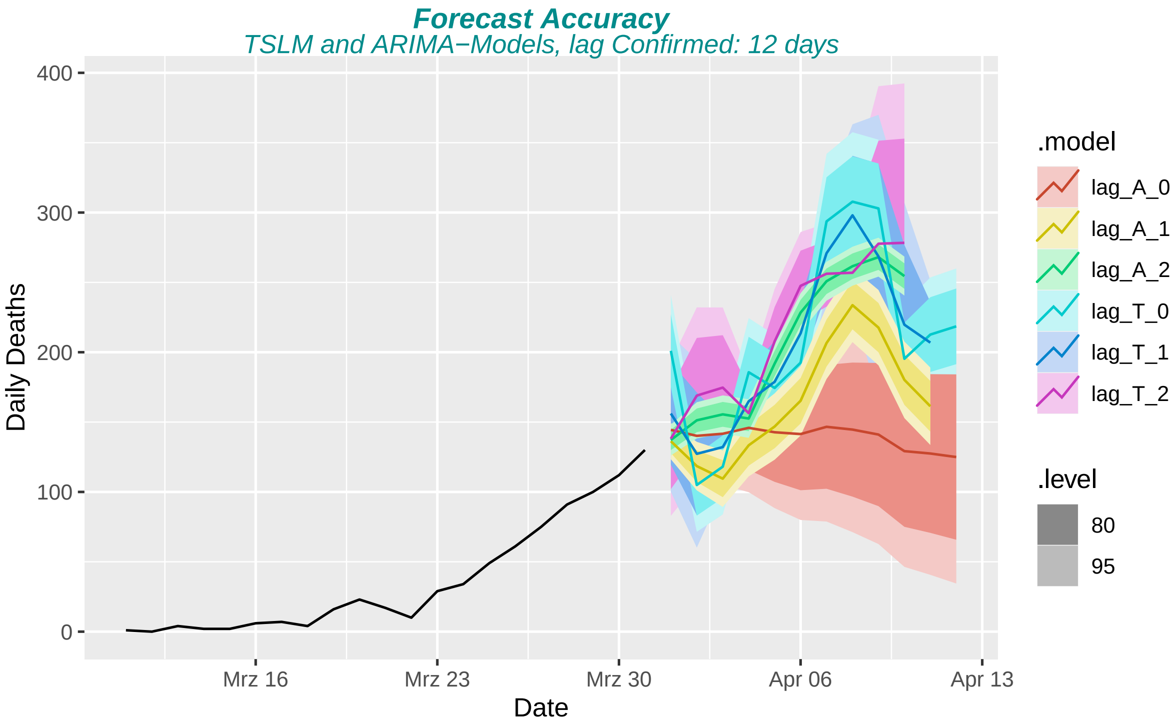
6.2.1 Forecasting Daily Deaths with lagged Daily Confirmed Cases

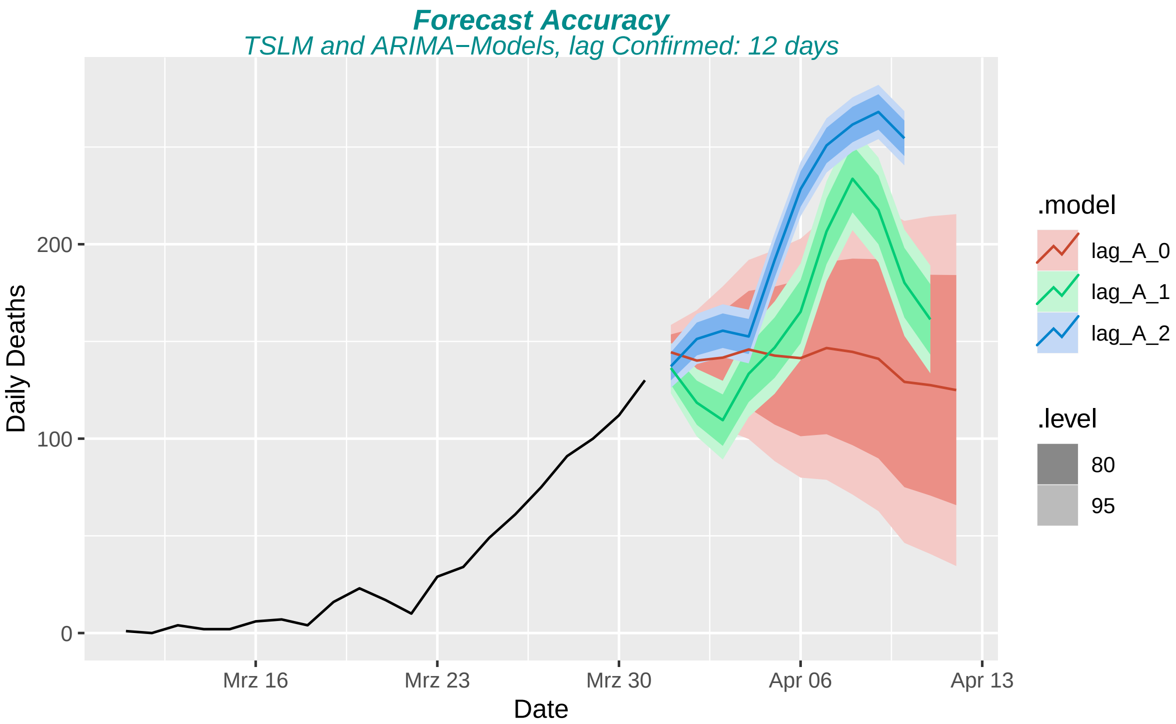


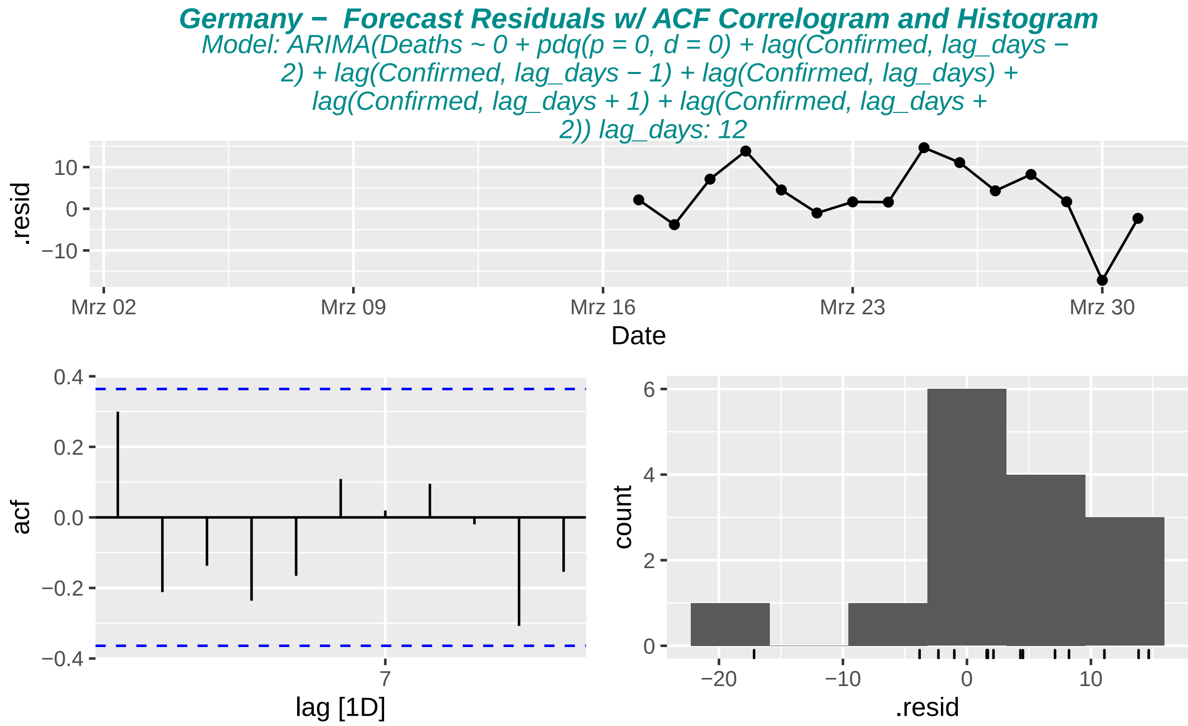

```
## # A tibble: 6 x 18
##   Country .model sigma2 log_lik   AIC  AICc   BIC ar_roots ma_roots r_squared
##   <chr>   <chr>   <dbl>   <dbl> <dbl> <dbl> <dbl> <list>   <list>       <dbl>
## 1 Germany lag_T~  102.    -52.9  75.2  85.7  79.5 <NULL>   <NULL>          NA
## 2 Germany lag_T~  177.    -62.4  87.5  91.1  90.6 <NULL>   <NULL>          NA
## 3 Germany lag_T~  246.    -70.4  96.6  97.4  98.3 <NULL>   <NULL>          NA
## 4 Germany lag_A~   31.7   -50.7 115.  121.  125.  <cpl [1~ <cpl [0~        NA
## 5 Germany lag_A~   43.7   -57.3 125.  127.  131.  <cpl [1~ <cpl [0~        NA
## 6 Germany lag_A~   51.4   -63.7 135.  137.  141.  <cpl [2~ <cpl [0~        NA
## # ... with 8 more variables: adj_r_squared <dbl>, statistic <dbl>,
## #   p_value <dbl>, df <int>, CV <dbl>, deviance <dbl>, df.residual <int>,
## #   rank <int>
```

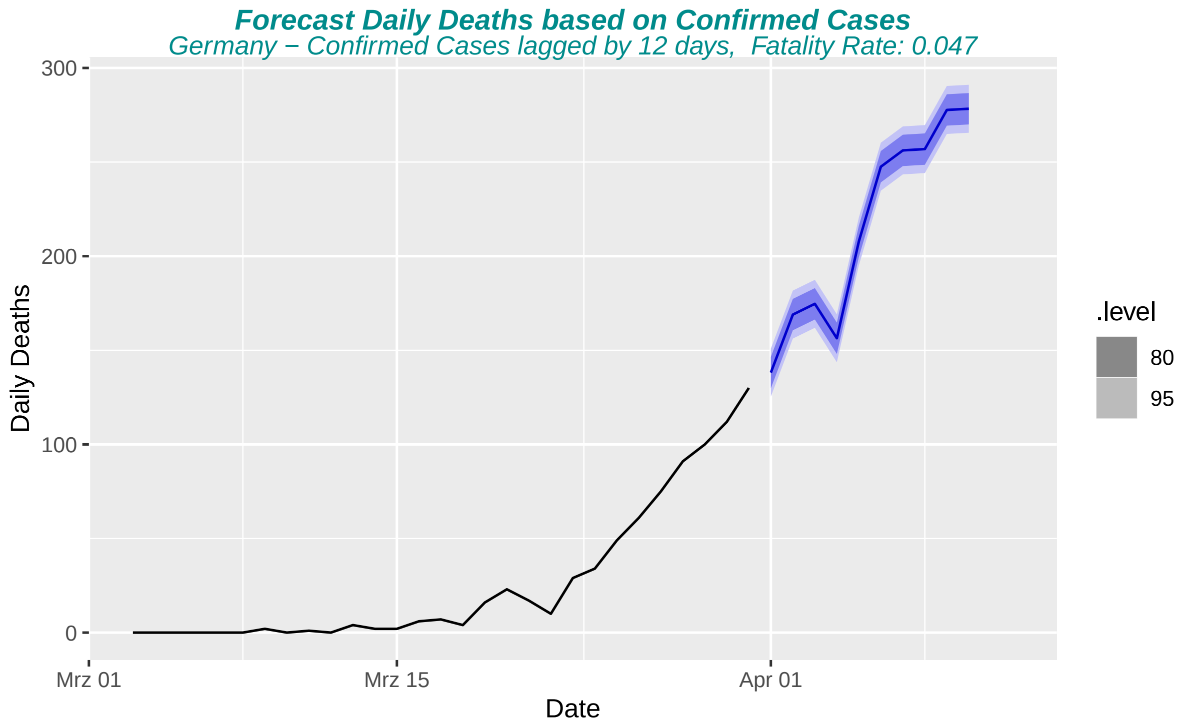
```
## # A tibble: 22 x 7
##    Country .model  term                    estimate std.error statistic  p.value
##    <chr>   <chr>   <chr>                      <dbl>     <dbl>     <dbl>    <dbl>
##  1 Germany lag_A_0 ar1                      1.55      0.218       7.13  5.11e- 6
##  2 Germany lag_A_0 ar2                     -0.563     0.220      -2.56  2.27e- 2
##  3 Germany lag_A_0 lag(Confirmed, lag_day~  0.00361   0.00363     0.995 3.37e- 1
##  4 Germany lag_T_0 lag(Confirmed, lag_day~  0.0444    0.00287    15.4   4.91e-11
##  5 Germany lag_A_1 ar1                      0.895     0.175       5.12  2.52e- 4
##  6 Germany lag_A_1 lag(Confirmed, lag_day~  0.00695   0.00449     1.55  1.48e- 1
##  7 Germany lag_A_1 lag(Confirmed, lag_day~  0.0158    0.00556     2.83  1.51e- 2
##  8 Germany lag_A_1 lag(Confirmed, lag_day~  0.0104    0.00513     2.04  6.45e- 2
##  9 Germany lag_T_1 lag(Confirmed, lag_day~  0.0124    0.00822     1.51  1.55e- 1
## 10 Germany lag_T_1 lag(Confirmed, lag_day~  0.0217    0.00784     2.77  1.59e- 2
```

```
## # ... with 12 more rows

## # A tibble: 22 x 7
##    Country .model  term                     estimate std.error statistic  p.value
##    <chr>   <chr>   <chr>                       <dbl>     <dbl>     <dbl>    <dbl>
##  1 Germany lag_A_0 ar1                       1.55      0.218       7.13  5.11e- 6
##  2 Germany lag_A_0 ar2                      -0.563     0.220      -2.56  2.27e- 2
##  3 Germany lag_A_0 lag(Confirmed, lag_day~  0.00361   0.00363     0.995 3.37e- 1
##  4 Germany lag_T_0 lag(Confirmed, lag_day~  0.0444    0.00287    15.4   4.91e-11
##  5 Germany lag_A_1 ar1                       0.895     0.175       5.12  2.52e- 4
##  6 Germany lag_A_1 lag(Confirmed, lag_day~  0.00695   0.00449     1.55  1.48e- 1
##  7 Germany lag_A_1 lag(Confirmed, lag_day~  0.0158    0.00556     2.83  1.51e- 2
##  8 Germany lag_A_1 lag(Confirmed, lag_day~  0.0104    0.00513     2.04  6.45e- 2
##  9 Germany lag_T_1 lag(Confirmed, lag_day~  0.0124    0.00822     1.51  1.55e- 1
## 10 Germany lag_T_1 lag(Confirmed, lag_day~  0.0217    0.00784     2.77  1.59e- 2
## # ... with 12 more rows

## # A tibble: 6 x 10
##   Country .model  .type       ME  RMSE   MAE     MPE  MAPE  MASE    ACF1
##   <chr>   <chr>   <chr>    <dbl> <dbl> <dbl>   <dbl> <dbl> <dbl>   <dbl>
## 1 Germany lag_A_2 Training  1.96  6.97  5.88 -0.0215  25.8 0.182  0.0468
## 2 Germany lag_T_2 Training  3.10  8.23  6.34  7.60    24.0 0.196  0.300
## 3 Germany lag_A_1 Training  2.70  8.26  6.46  0.682   29.2 0.200  0.0714
## 4 Germany lag_A_0 Training  4.00  8.86  6.63  4.84    29.3 0.205 -0.198
## 5 Germany lag_T_1 Training  4.81 12.0   9.62 11.2     31.2 0.297  0.410
## 6 Germany lag_T_0 Training  5.47 15.2  12.4  13.0     41.4 0.382 -0.0320

## Null Hypothesis of independence/white noise for residuals - for p < 0.05: reject H_0

## # A tibble: 6 x 4
##   Country .model  lb_stat lb_pvalue
##   <chr>   <chr>     <dbl>     <dbl>
## 1 Germany lag_T_0    5.84     0.828
## 2 Germany lag_A_0    7.67     0.661
## 3 Germany lag_T_2   10.4      0.404
## 4 Germany lag_T_1   11.1      0.348
## 5 Germany lag_A_2   11.6      0.312
## 6 Germany lag_A_1   13.2      0.213
```


Forecast Accuracy
TSLM and ARIMA−Models, lag Confirmed: 12 days
Daily Deaths
400
300
200
100
0
Mrz 16
Mrz 23
Mrz 30
Apr 06
Apr 13
Date
.model
lag_A_0
lag_A_1
lag_A_2
lag_T_0
lag_T_1
lag_T_2
.level
80
95


Forecast Accuracy
TSLM and ARIMA−Models, lag Confirmed: 12 days
Daily Deaths
200
100
0
Mrz 16
Mrz 23
Mrz 30
Apr 06
Apr 13
Date
.model
lag_A_0
lag_A_1
lag_A_2
.level
80
95





7 References

7.0.2 Data Source

Data Source

Data files are provided by **Johns Hopkins University** on GitHub
https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series

- Data files:
  time_series_covid19_confirmed_global.csv,
  time_series_covid19_deaths_global time_series_covid19_recovered_global.csv

Note: as of 2020-03-27 recovered cases are provided again

The data are visualized on their excellent Dashboard
Johns Hopkins University Dashboard
https://coronavirus.jhu.edu/map.html

7.0.3 Links

Further links

WHO Dashboard
https://experience.arcgis.com/experience/685d0ace521648f8a5beeeee1b9125cd

Robert Koch Institut, Germany
https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html?nn=13490888

Wikipedia - Exponential Growth
https://en.wikipedia.org/wiki/Exponential_growth

Code Source

Code is based on ideas from https://rpubs.com/TimoBoll/583802