Instructions for using the tool.

Objective

The main purpose of the tool is to forecast the ICU and hospital bed occupancy by COVID-19 patients over a period of time in the future. Here “hospital bed occupancy” comprises both ICU and non-ICU beds. Also, the tool can be validated using past periods in order to assess its predictive power for the hospital or region under study.

Hardware/software request

The program has been coded in Java, for its execution you only need a jre (java runtime environment) 1.8 or higher. All the required libraries are included in the distributed executable version of the program.

Data files

The tool requires a file with data of Covid-19 patients in the region in the previous months. This includes the relevant dates of patients and their classification in groups depending on their characteristics (sex, age…), both in Groupings I and II. See Section ?? in the Supplementary Material file for details on the definitions and usefulness of Groupings I and II. The file with patients’ information must be in .csv format. Each line contains the information of a patient:

Column1: id_number. An integer from 0 to 99999999. This number is not used by the tool.

Columns 2, 3. Group number of the individual in Groupings I and II. Group number must be an integer.

Columns 4-8. Dates of positive testing, admission to hospital, admission to ICU, discharge from ICU and discharge from hospital. The date format is “yyyy-mm-dd”; missing dates must be written as NA.

See the Suplementary material document for the exact definition of each column. Click here for a sample file with 5 patients.

Usage

Main menu

In the main menu window the user is asked to choose Forecasting or Validation. Check an option and click on *Proceed*.

Forecasting

> Select the data file with the patients’ information by clicking on Select and then click on Upload. The tool will upload the file and will show the range of dates included in the data file. Then select the period for model estimation (values of tI and tF in the Supplementary Material file) and the forecast period. The period for model estimation

must be within the range of dates included in the file and the forecast period must begin after the last day included in the file.

The tool needs a projection of the new positive cases during the forecast period for each of the groups defined in Grouping II. If you have your own projections, then check the box "*Use file with projected number of COVID-19 positive cases*". Then select a file containing the information of the projected cases. The file must be in .csv format. The number of lines in the file is the number of days in the forecast period and it can have either a single column or as many columns as the number of groups in Grouping II. In the latter case, the number in Row i and Column j indicates the number of projected new positive cases diagnosed on day i of the forecast period in group j. Click here for an example of a 7-day forecast period and 3 groups in Grouping II. In the example, it is projected that the first day of the period there will be 16 new positives cases in group 1, 32 in group 2 and 17 in group 3. If the file has a single column, the number in Row i is the total number of projected new positive cases diagnosed on day i of the forecast period; the tool internally splits this number for estimating the number in each group using the proportion of cases by group in the last weeks of the estimation period. In this example, we see that the number of total projected new cases for the first day is 65.

Click on Upload and choose a random seed to initialize simulations (e.g. 3231321) and the number of replications (e.g. 1000). Then click on Proceed.

If you do not have a projection of the new positive cases for the forecast period, then leave the box "*Use file with projected number of COVID-19 positive cases*" unchecked. The tool will project the number of positive cases using time-series methodology. Choose a random seed to initialize simulations (e.g. 3231321) and the number of replications (e.g. 1000). Then click on Proceed. A new window will pop up, requesting for the time frequency; this is an integer specifying the seasonal parameter of the series of positive cases, which will be used for projections during the forecast period. The default is 7, corresponding to week seasonality; if your series has no seasonality, write 1.

Once simulations have been completed, you can access the results, both in text and graph formats.

Result: **Text**. For each day of the forecast period, the following information is given:

- Projected positive cases (from the user file or the time-series projections)
- Mean, percentiles 5, 95 of forecasted ICU bed occupancy
- Mean, percentiles 5, 95 of forecasted hospital bed occupancy

  Warning: The 5-95 percentile bands in forecasting take into account only the variability of the evolution of the system considering the number of positive cases in the period as fixed; in particular, they do not include the error in the time-series projections of the number of positive cases.

Result: **Text (Group)**. Detailed information for each group (in Grouping II), including mean, sd 5, 10, 50, 90, 95 percentiles of forecasted ICU and hospital bed occupancy.

The information can be exported as a txt file.

Result: Graph. An interactive graph is produced, where the x-axis represents time and the y-axis represents the number of cases or occupied beds. The lines on the graph are drawn by checking the corresponding names at the left panel:

- **Positive**. Either projected positive cases from the user file or time-series projections of positive cases.
- **Forecast Hospital.** Forecast of hospital bed occupancy together with 5-95 percentile bands.
- **Forecast ICU.** Forecast of ICU bed occupancy together with 5-95 percentile bands.

Drag the *Previous period* slider to modify the initial date in the x-axis and draw the observed data in the extended period.

Result: **Graph (Group)**. Same graph as above, separated by groups (in Grouping II). Check the desired information and select the groups in the left panel of the graph.

Validation

Select the data file with the patients' information by clicking on Select and click on Upload. The tool will upload the file and will show the range of dates included in the file. Then select the period for model estimation (values of tI and tF in the Supplementary Material file) and the period of validation Note that both the periods for model estimation and for validation must be within the range of dates included in the file, and that the period for validation must begin after the end of the period for model estimation. Choose a random seed to initialize simulations (e.g. 3231321) and the number of replications (e.g. 1000). Then click on Proceed. A window will pop up, requesting for the time frequency; this is an integer specifying the seasonal parameter of the series of positive cases, which will be used for projections during the validation period. The default is 7, corresponding to week seasonality; if your series has no seasonality, write 1.

Once simulations have been completed, you can access the results, both in text and graph formats.

Result: **Text**. For each day of the validation period, the following information is given:

- Observed and time series projection of the number of positive cases
- Observed ICU bed occupancy
- Mean, percentiles 5, 95 of forecasted ICU bed occupancy both when the observed number of positive cases is used (column OPC) and when the time-series projection of the number of positive cases is used (column PPC).
- Observed hospital bed occupancy
- Mean, percentiles 5, 95 of forecasted hospital bed occupancy both when the observed number of positive cases is used (column OPC) and when the

time-series projection of the number of positive cases is used (column PPC).

Also, the MAPE (mean absolute percentage errors) for ICU and hospital occupancy, both using the observed and the time-series projection of the number of positive cases, are shown.

Warning: The MAPE can be high when the number of positive cases is projected by the time series and the validation period is long, since the time series algorithm cannot predict a change of trend in the number of positive cases. If such a change occurs in the validation period, the projected positive cases will be very different from the observed cases and this will result in bad forecasts of hospital and ICU bed occupancy. Also, note that the 5-95 percentile bands in forecasting take into account only the variability of the evolution of the system considering the number of positive cases in the period as fixed; in particular, they do not include the error in the time-series projections of the number of positives.

Result: **Text (Group)**. Detailed information for each group (in Grouping II), including mean, sd 5, 10, 50, 90, 95 percentiles of forecasted ICU and hospital bed occupancy.

Result: Graph. An interactive graph is produced, where the x-axis represents time and the y-axis represents the number of cases or occupied beds. The lines on the graph are drawn by checking the corresponding names at the left panel:

- **Forecast (P). Positive**. Time-series projection of the number of positive cases in the validation period.
- **Forecast (P). Hospital.** Forecast of hospital bed occupancy using the time-series projections of the number of positive cases in the validation period, together with 5-95 percentile bands.
- **Forecast (P). ICU.** Forecast of ICU bed occupancy using the time-series projections of the number of positive cases in the validation period, together with 5-95 percentile bands.
- **Forecast (O).** Hospital. Forecast of hospital bed occupancy using the observed number of positive cases in the validation period, together with 5-95 percentile bands.
- **Forecast (O).** ICU. Forecast of ICU bed occupancy using the observed number of positive cases in the validation period, together with 5-95 percentile bands.
- **Observed data. Positive.** Number of observed positive cases.
- **Observed data. Hospital.** Observed hospital bed occupancy.
- **Observed data. ICU.** Observed ICU bed occupancy.

Drag the *Previous period* and *Future period* sliders to modify the initial and final dates in the x-axis and draw the observed data in the extended periods.