Consolidation Nowcasting Analysis

Tommaso Bertola

September 24, 2025

1 Nowcasting

Starting from the 2023-2024 season, ISS started updating the ILI data retrospectively. This means that the incidence for a given week is not final when first published, but it gets updated in the following weeks. Often times the incidence is first underestimated. We can leverage this information to our advantage, by trying to estimate the final value of the incidence for the current week, based on the partial data we have so far.

Only for the 2024-2025 season this data is available at the regional level, while for the previous seasons it is only available at the national level, meaning that for the previous seasons, we only had one value of incidence which will never be updated.

2 Analysis

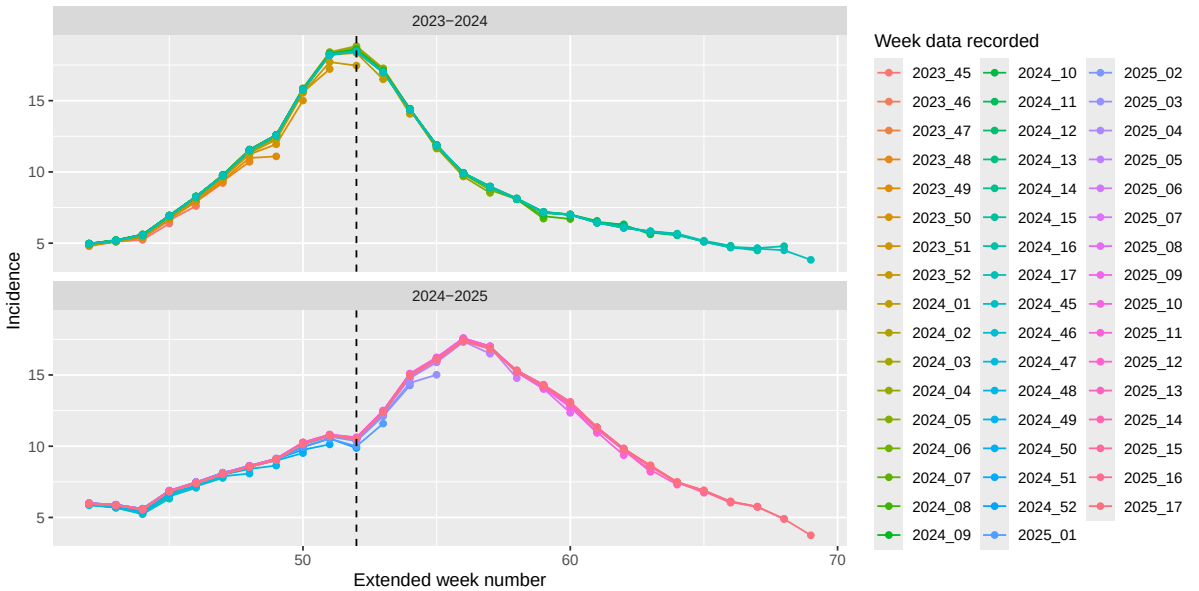

Figure 1: Raw data of the incidence for each week for Italy, colored by the week the data was recorded. The dashed vertical line indicates the end of the year.

As we can see from Figure 1, the incidence for a given week tends to increase in the following weeks, until it stabilizes. The question is: how long does it take for the incidence to stabilize? To answer this question, we can look at the data from the previous seasons,

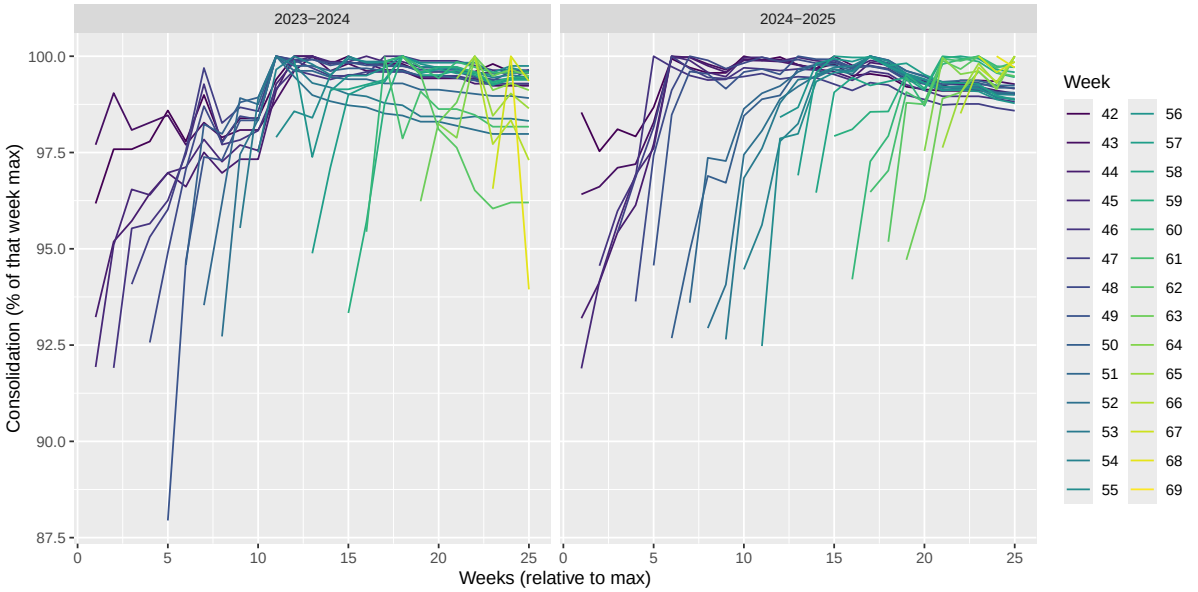

Figure 2: Ratio of the incidence for each week for Italy, colored by the week the data was recorded. It always happens that the first incidence recorded is the lowest.

where we have the final value of the incidence for each week. We can then compare the incidence for a given week, recorded at different times, to the final value of the incidence for that week.

2.1 National

We first started looking at the national data. To do so and inspired by the graph (maybe from some other works on COVID), we compute for each week in the season the ratio with the maximum incidence for that week in that season. This way, we can see how the incidence for a given week evolves over time, relative to its highest value.

Figure 2 shows the ratio of the incidence for each week, colored by the week the data was recorded. The first data point always shows that the incidence is the relative lowest. Only after 5-10 weeks it stabilizes around 100%. However it does not remain constant but sometimes it goes slightly down again. In general the first data point is around 92%.

In Figure 3 we can see the average consolidation of the incidence for Italy, as a function of the number of weeks since the incidence was recorded. On top of the single trajectories we plot the average consolidation as a function of weeks since the incidence was recorded. We would expect the average consolidation to be a monotonically increasing function, but sometimes it decreases. We further distinguish between seasons and trajectories with "many" and "few" data points, where the threshold is arbitrarily set to 15 weeks. This distinction does not provide much additional insights, besides showing that the consolidation is more or less consistent during the season.

We then grouped together all the trajectories for all weeks. The result is reported in Figure 4. Only after week 12 (about 3 months) the incidence stabilizes around 100% (99.5%), which is quite a long time. One could assume that at that point the consolidation is complete, and therefore fix the consolidation to 100% after week 12.

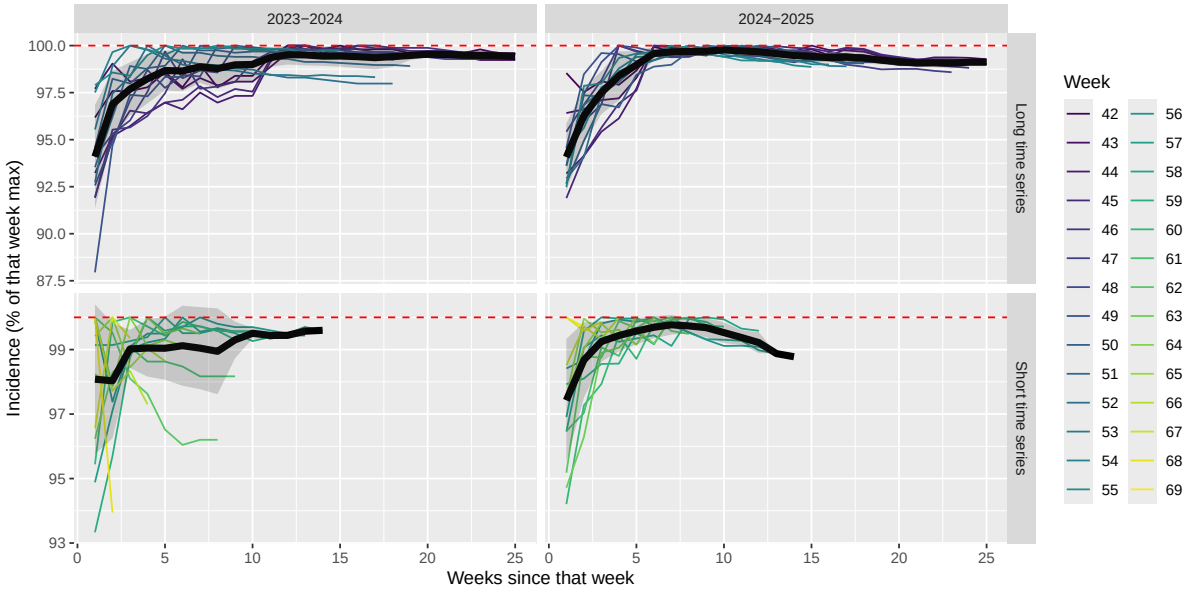

Figure 3: Average consolidation of the incidence for Italy, as a function of the number of weeks since the incidence was recorded.

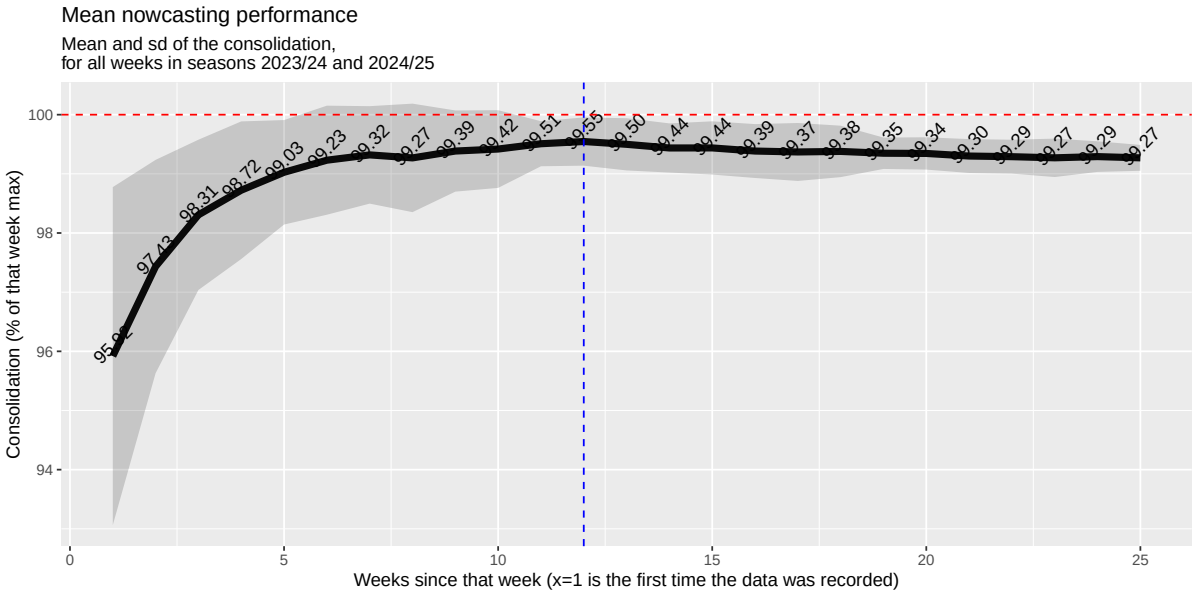

Figure 4: Smoothed average consolidation of the incidence for Italy, as a function of the number of weeks since the incidence was recorded.

2.2 Regional

Figure 5 shows the raw data of the incidence for each week for each region, colored by the week the data was recorded. This time, despite the fact that for the first season data was retrospectively updated, no clear trend is obviously visible. In some cases there is an increase over time of the incidence, corresponding to the consolidation function we observed in the national data, but in other cases a decrease is observed. This tends to happen more often for smaller regions, while larger regions tend to show a more consistent consolidation function.

Separating the "trajectories" for long and short series in Figure 6 is does not help

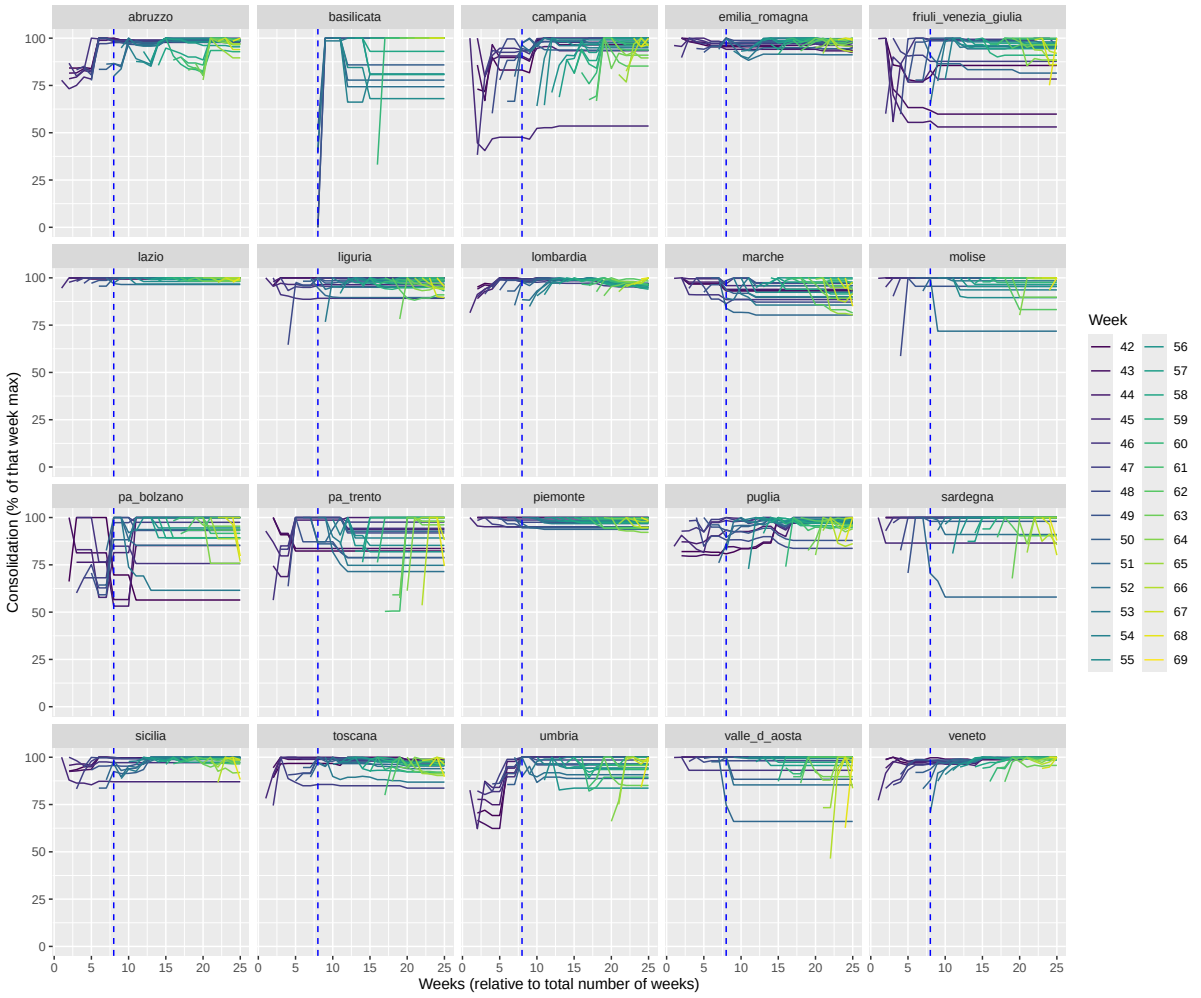

Figure 5: Raw data of the incidence for each week for each region, colored by the week the data was recorded. The dashed vertical line indicates the end of the year.

much, again. As in Figure 3 we aligned the trajectories to the left, such that we could see the overall consolidation function.

Finally in Figure 7 we can see the average consolidation for each region. Some regions appear to have flat consolidation functions, others are counterintuitive, while others display a decreasing trend, which is not what we would expect.

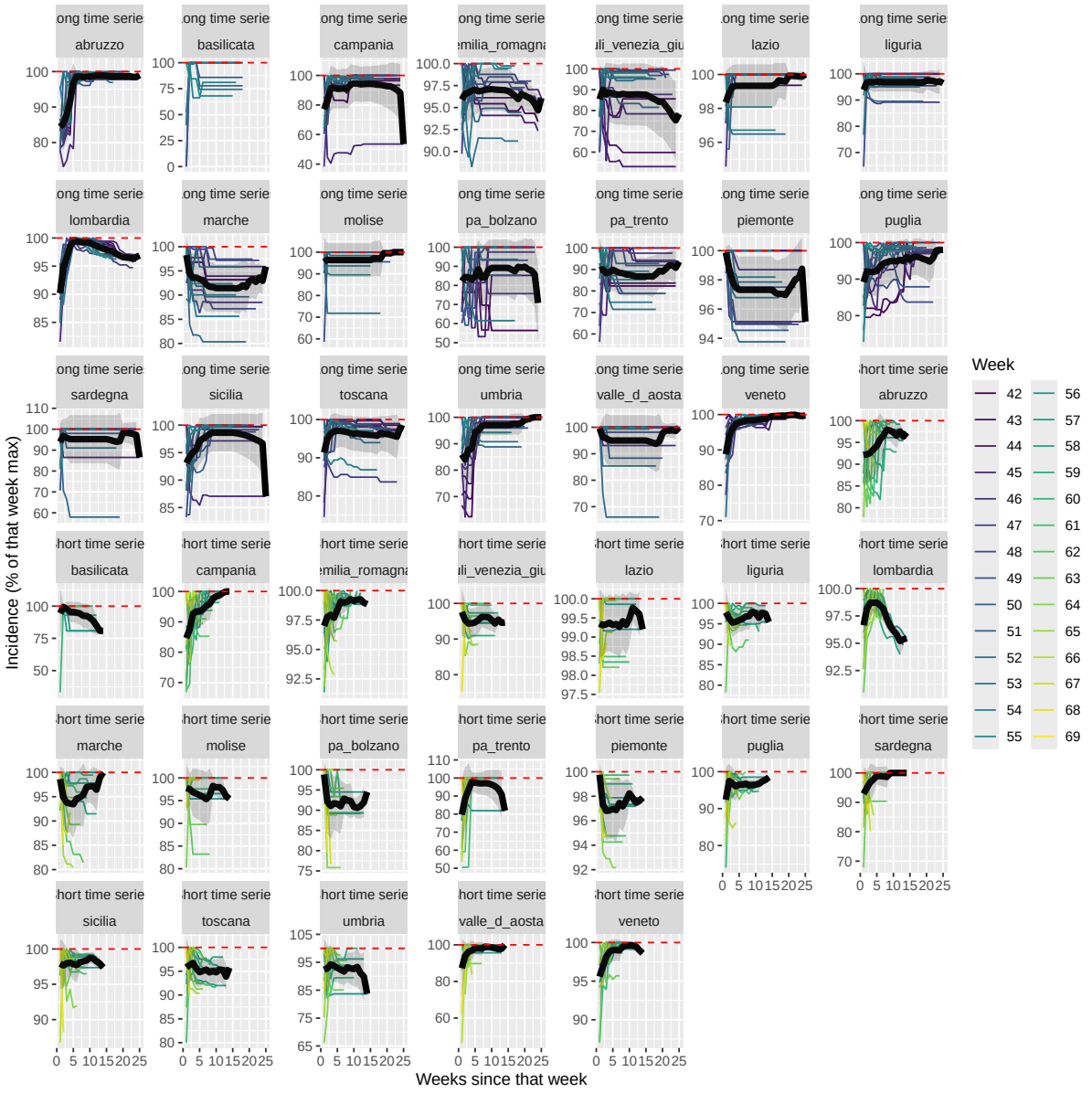

Figure 6: Average consolidation of the incidence for each region, as a function of the number of weeks since the incidence was recorded.

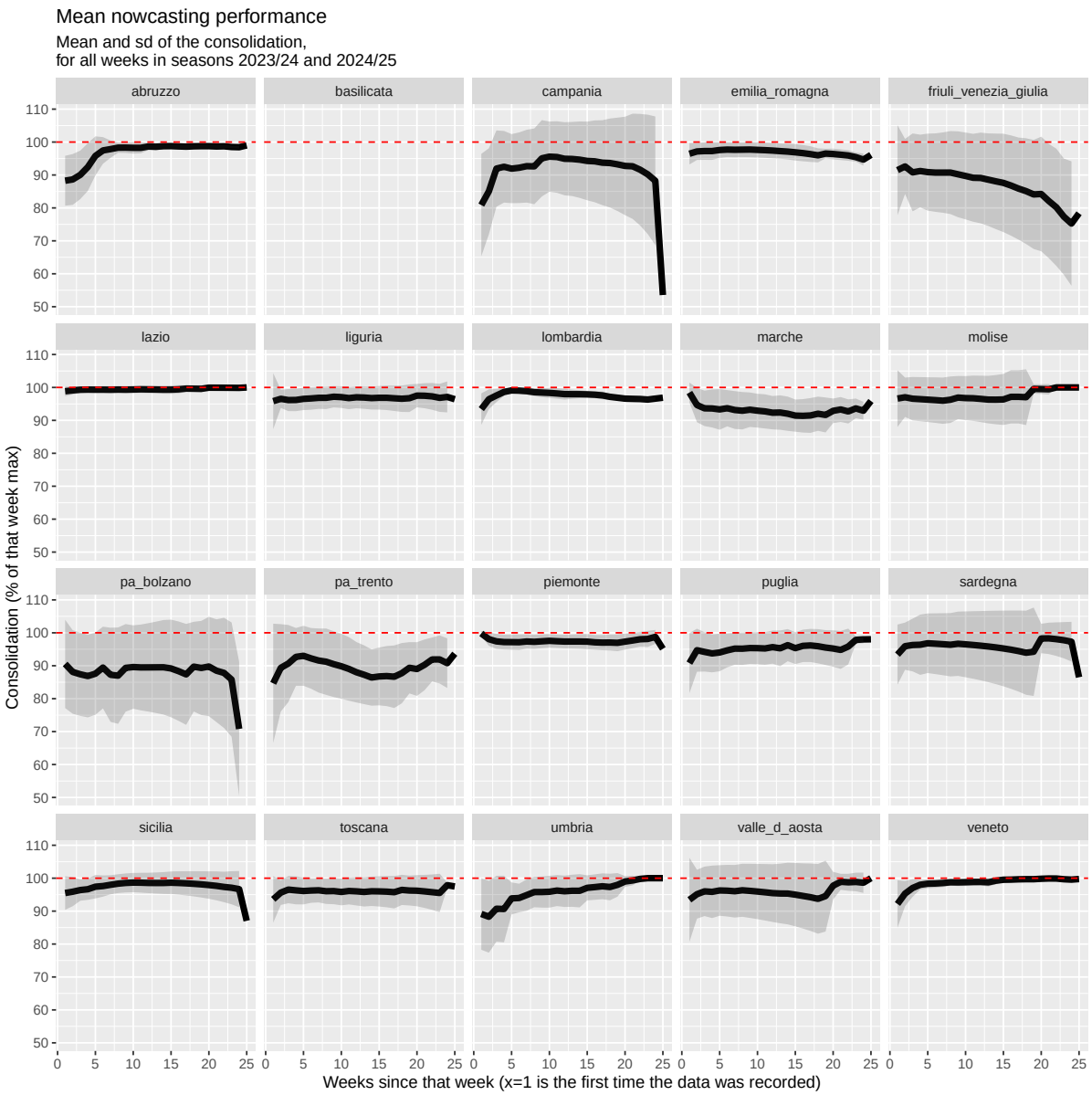

Figure 7: Smoothed average consolidation of the incidence for each region, as a function of the number of weeks since the incidence was recorded.

2.3 ILI

For the seasons 2023-2024 and 2024-2205 it is possible to check the consolidation of the ILI data, which is referred only at the national level. Figure 8 shows the raw data of the

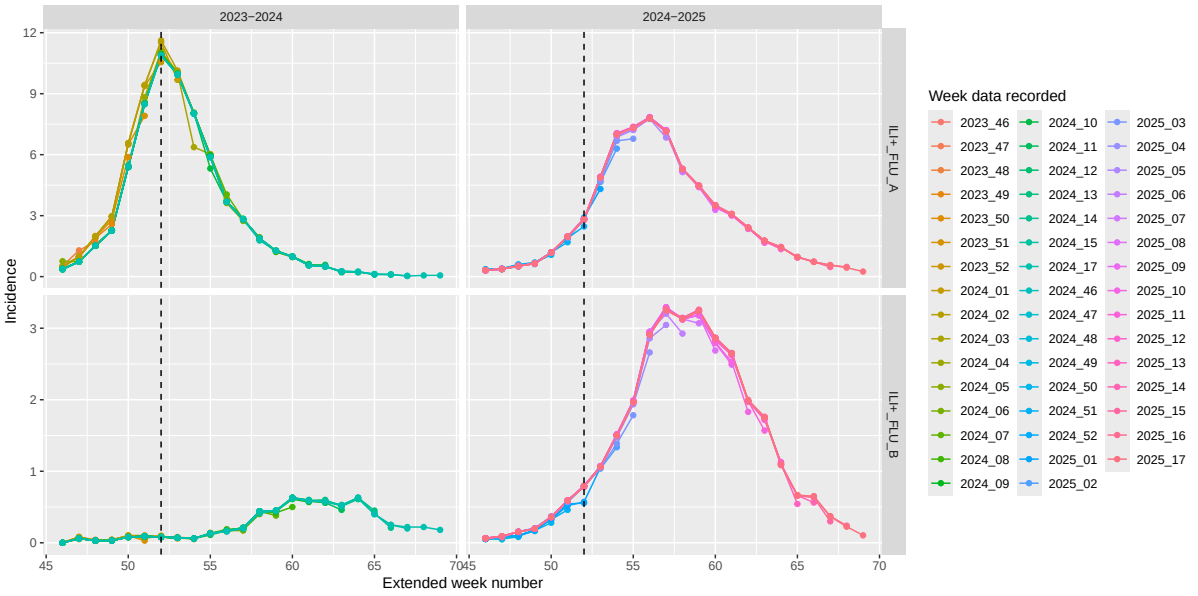

Figure 8: Raw data of the ILI for each week for Italy, colored by the week the data was recorded. The dashed vertical line indicates the end of the year.

ILI for each week for Italy, colored by the week the data was recorded. The trends are similar to what we expect and by rapid eye analysis seems overall consistent.

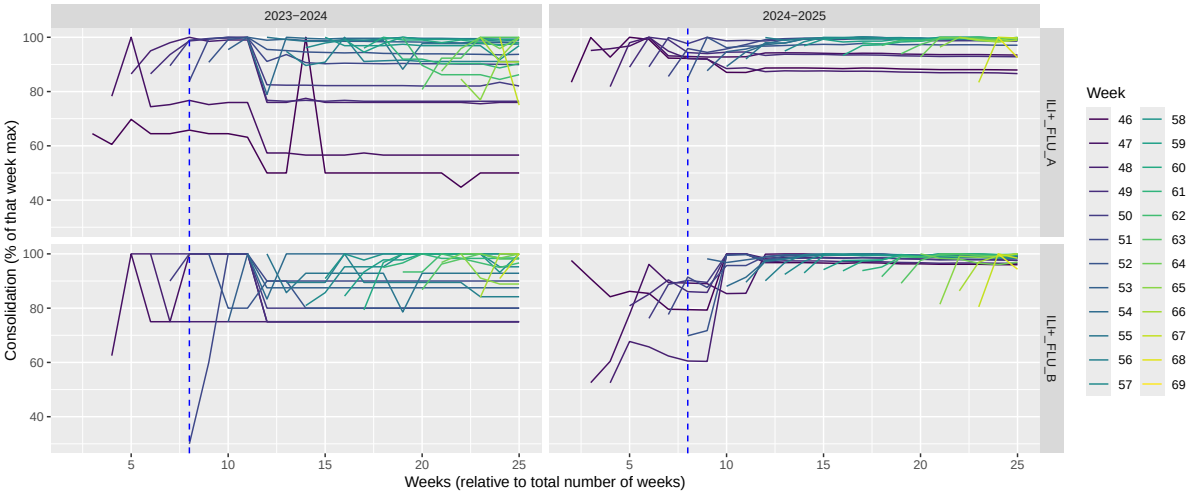

Figure 9: Ratio of the ILI for each week for Italy, colored by the week the data was recorded

From Figure 9 it is clear that the 2024-2025 season is the most "stable". As always the first week the incidence is the lowest. However, it happens that for ILI A and the 2024-2025 season the incidence tends to decrease again after some weeks.

Again in Figure 10 we can see the average consolidation of the ILI for Italy, as a function of the number of weeks since the incidence was recorded and separated by ILI A

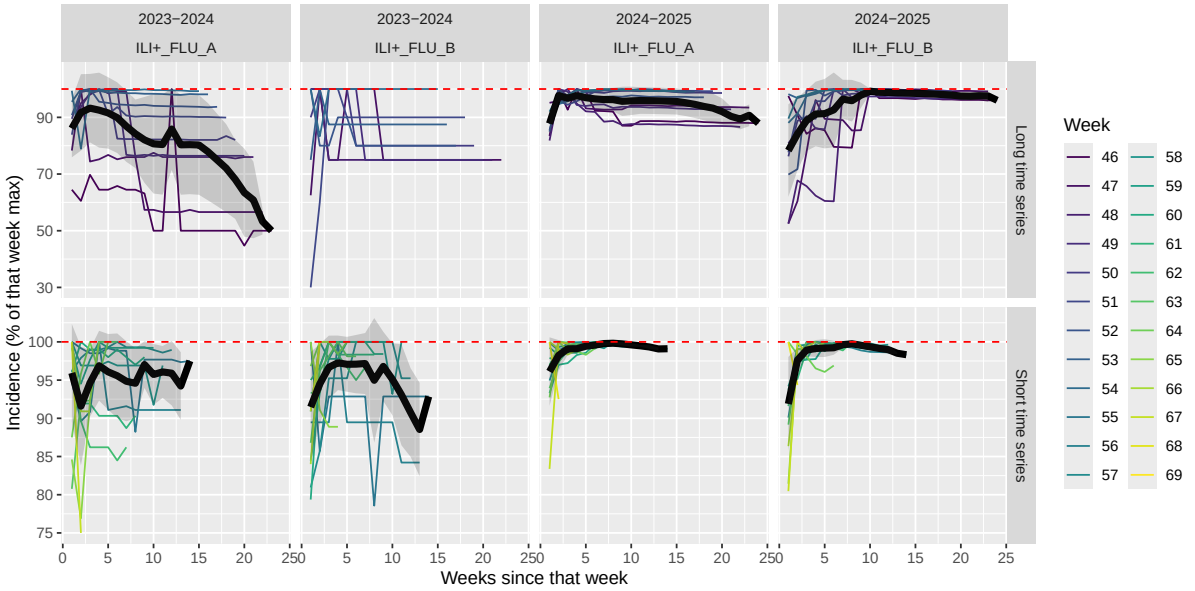

Figure 10: Average consolidation of the ILI for Italy, as a function of the number of weeks since the incidence was recorded.

and ILI B. For the 2023-2024 season the consolidation appears to be quite inconsistent, while for the 2024-2025 season it is more stable, as observed above.

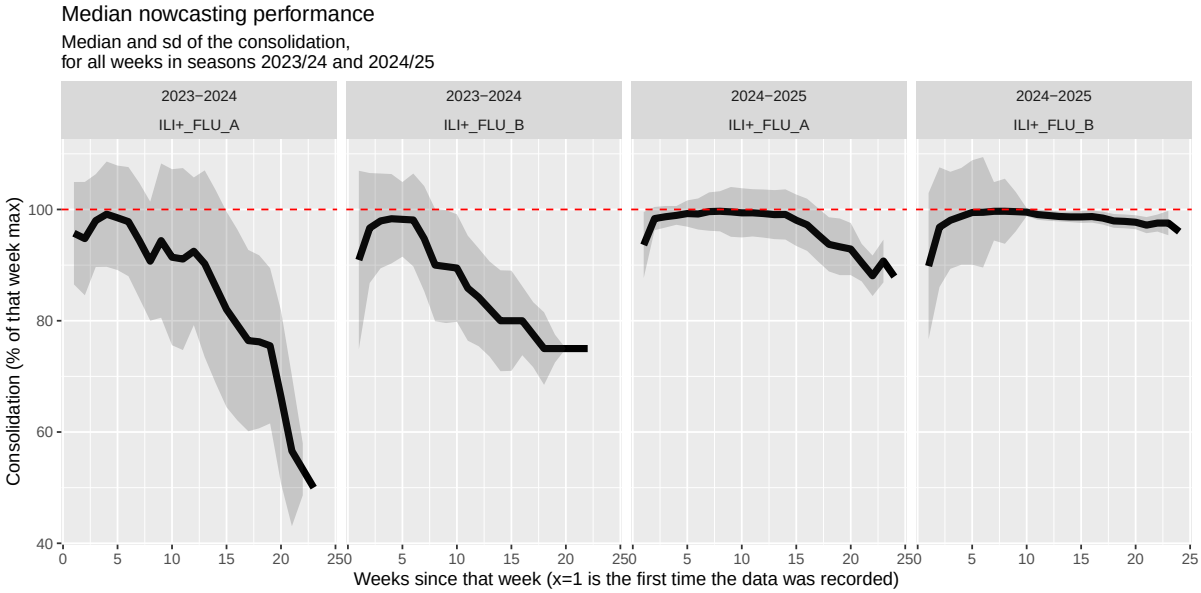

Figure 11: Smoothed average consolidation of the ILI for Italy, as a function of the number of weeks since the incidence was recorded.

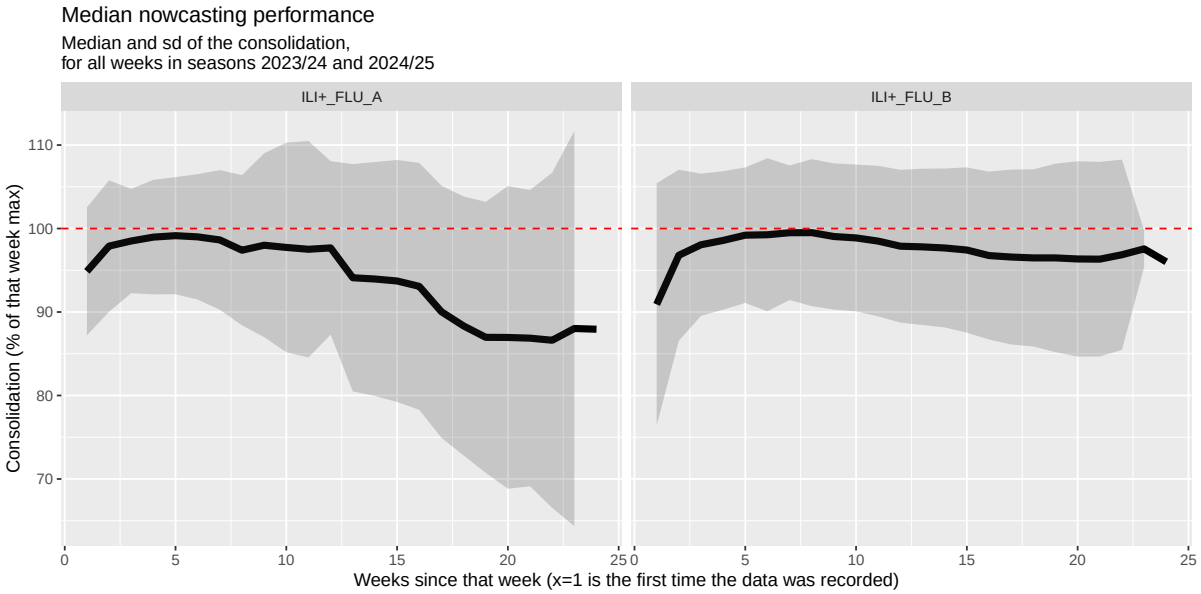

Figure 12: Smoothed average consolidation of the ILI for Italy, as a function of the number of weeks since the incidence was recorded.

3 How to implement it in the model

This is the easiest part: we just need to divide the incidence for the current week by the consolidation factor. For example, if the incidence for the current week is 100, and the consolidation factor is 95%, then the nowcasted incidence is 100 / 0.95 = 105.26. This means that we expect the final incidence for the current week to be around 105.26.

In principle we could also compute a confidence interval for the nowcasted incidence, by looking at the distribution of the consolidation factor for a given number of weeks since the incidence was recorded.

Will the nowcasting change much the predictions?