

Forecasting interrupted time series

Rob J Hyndman

Department of Econometrics & Business Statistics

Monash University

Clayton VIC 3800

Australia

Email: Rob.Hyndman@monash.edu

Corresponding author

Bahman Rostami-Tabar

Cardiff Business School

Wales CF10 3EU

United Kingdom

Email: rostami-tabarb@cardiff.ac.uk

8 January 2024

JEL classification: C10,C14,C22

Forecasting interrupted time series

Abstract

A brief summary of our ideas

Keywords: blah, blah

1 Introduction

Time series are sometimes interrupted by unusual events; for example, a natural disaster may occur, or there may be a temporary policy change. Many forecasters have faced this issue recently with the COVID-19 pandemic, where historical patterns have been severely disrupted due to lockdowns and other restrictions. In this paper, we consider the problem of forecasting during such an event, and after it has occurred.

This is a different problem from change point detection. In the situations we consider, we know that a change has occurred, and we want to forecast the future after the change. Change point detection is about identifying *when* a change has occurred.

Time series models usually assume that the data evolve in the future in a similar way to how they have evolved in the past. But a big event (such as COVID-19) can result in a future that is different from the past, at least in the short term.

The disruption may result in relatively simple changes in the series; for example a level shift at the start of the disruption, and another at the end of the disruption. Or they may be more complex, with changes to the seasonal patterns, and changes to the level, which evolve over time. In this paper, we consider a range of models that can be used to handle such changes, and compare their performance on some real data sets.

We introduce several approaches to handling interruptions when forecasting in Section 2. We then apply these approaches to two real data sets in [?@sec-sec-examples](#). Finally, in [?@sec-conclusions](#), we discuss some of the advantages and disadvantages of each approach, and provide some recommendations for practitioners.

2 Handling interruptions when forecasting

We consider several possible ways to handle interruptions when forecasting.

2.1 Use a highly adaptive model

Highly adaptive models can adjust to the interruption as it is happening, and will therefore be able to approximate the data generating process relatively well. For example, an ETS model with large smoothing parameters will be able to adjust to the interruption relatively quickly. This has the advantage of being a very simple solution that is easy to implement and fast to compute. There is no need to explicitly model the interruption, and so the model can be used for forecasting even if the timing or effect of the interruption is unknown.

However, the prediction intervals will be large, because the model will have heavily discounted past data. In fact, the model will largely forget the past data other than the most recent observations, so there is no memory of the seasonal patterns and other dynamics that were present before the interruption. Consequently, the approach works best if there is no assumption that the post-interruption period will be similar to the pre-interruption period.

2.2 Use an intervention model

A dynamic regression model with intervention covariates can be used to model the interruption explicitly. For example, if the intervention involves a simple level shift, with a reverse level shift at the end of the intervention, we can use a dummy variable to indicate the interruption period, and allow the model to adjust to the interruption. More complicated interventions can be handled by using more covariates.

This has the advantage of retaining the memory of the past, and so the seasonal patterns and other dynamics will be retained. This allows the change period to be effectively modelled provided the intervention variables are chosen well. However, the model will assume that the post-interruption period will be similar to the pre-interruption period, and so the prediction intervals may be too narrow, especially if there is a lasting effect beyond the end of the interruption.

2.3 Set to missing

Many time series models will handle periods of missing values. So the problematic observations that occur during the period of disruption can be set to missing, and the model should continue to produce forecasts as if the interruption had not occurred. Of course, the forecasts will not be accurate for the period of disruption, but they can be interpreted as “what might have been”. This solution requires a judgement to be made about when the disruption has begun, and when normality resumes.

Because no information is retained during the disruption, the prediction intervals will become large during the disruption, and after the disruption they will remain large until the model has enough data to estimate the forecast distribution more accurately.

2.4 Estimate what might have been

A fourth solution is to estimate what might have been during the period of disruption, and then use the adjusted data to fit a model. The estimates could be made using any convenient method. One approach would be to set the estimates to equal the forecasts made using only pre-interruption data. This then becomes almost equal to the previous solution except that the prediction intervals will be narrower because the estimation uncertainty has not been taken into account. On the other hand, it is more flexible than the previous approach because models that do not handle missing values can then be used.

An alternative would be to take the model estimated under the previous solution (setting the observations during the disruption to missing), and use it to estimate what might have been during the disruption. Then the model can be re-estimated using the adjusted data. Note that forecasts obtained in this way during the disruption will not be true forecasts, because they will have used data from the future in computing the adjusted data during the disruption. But post-disruption forecasts will be true forecasts, and should be almost the same as those obtained using the previous solution.

2.5 Downweight the interruption period

Intuitively, we want the model to be more influenced by the patterns observed before the disruption period than those during the disruption, but we still want to take some account of the observations during the disruption period.

Using this weighting approach, we can leverage the insights from the pre-disruptive period while still accounting for the disruptive period's influence. It allows the model to better capture the underlying dynamics of the time series in non-disrupted periods, which may be essential for accurate forecasting in the post-disruption era.

If the weights during the disruption are set to zero, and the weights at other times are set to 1, this becomes equivalent to the missing value solution above. Rather than ignoring the data during the disruption, the weighting approach allows the model to not be completely blind to the disruption's effects, and adapt to changes or shifts in the time series caused by the event.

While conceptually simple to implement, in practice, this approach may require more work than the other methods discussed here as most forecasting software does not allow for the explicit use of weights.

2.6 Ensemble models

As with many forecasting problems, taking an ensemble approach often leads to more accurate forecasts (Wang et al. 2023). Here, we could combine some or all four of the approaches discussed above to obtain an ensemble forecast. In fact, if we were unable to implement the weighted approach, by using an ensemble of the other approaches we are effectively downweighting the observations during the disruption period as they are only explicitly used in the first two approaches (using a highly adaptive model, or using an intervention model).

However, a disadvantage of this approach is that it is averaging forecasts that aim to achieve different objectives. For example, the highly adaptive model and the intervention model are trying to forecast what happened during the disruption, while the missing value and estimation approaches are trying to forecast what might have happened. So the ensemble approach will lie somewhere between these two objectives.

3 Examples

Let's consider some examples to data observed during the last few years, where the effect of the COVID-19 pandemic has been particularly evident.

3.1 Australian tourism

The Australian tourism data set is a monthly time series showing the number of short-term overseas visitors to Australia. The data (Australian Bureau of Statistics 2024) are available from January 2000 to May 2023, and are shown in Figure 1. As the borders closed in March 2020, the number of visitors to Australia dropped to near zero, and remained there until towards the end of 2021. The borders officially reopened on 21 February 2022, although it seems visitors began to arrive earlier than that.

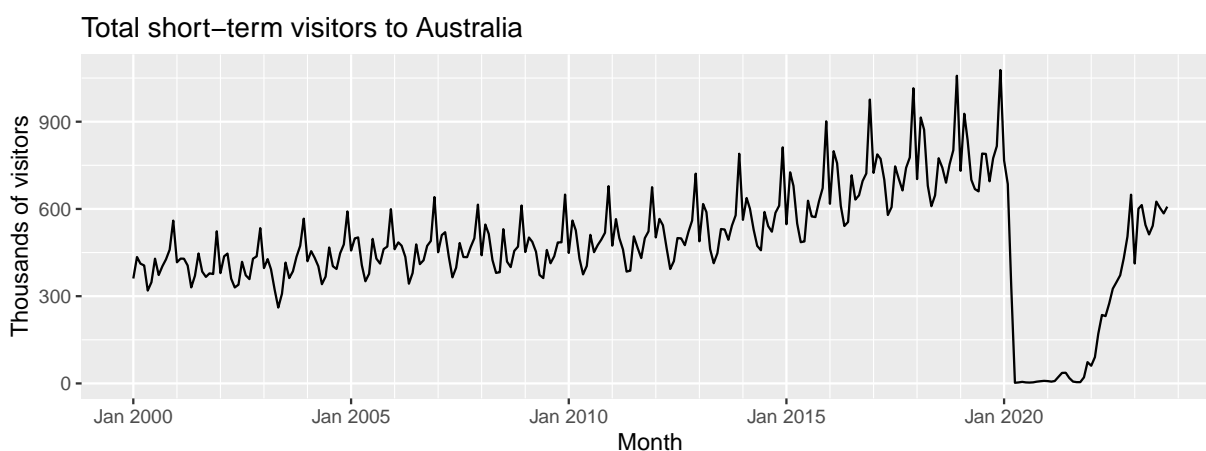


Figure 1: Short-term visitor arrivals to Australia (monthly): Jan 2000 – May 2023.

We apply the first four solutions discussed in the previous section to these data, making 12 month forecasts at the end of each year from 2019 to 2022. We fit ETS, ARIMA and dynamic ARIMA models to the data (Hyndman & Athanasopoulos 2021), first applying a log transformation to ensure the resulting forecasts are positive. This is possible, because the observations never reach exactly zero, with the smallest number of visitors per month equal to 2250 in April 2020. For all forecasts, we also show 90% prediction intervals.

In Figure 2, we show forecasts by applying ETS and ARIMA models to the data. ETS, in particular, is well-known to be relatively adaptive to changes in the series, and this is evident in these forecasts. The forecasts for 2020 were made using data before COVID-19 had any affect, and so they show

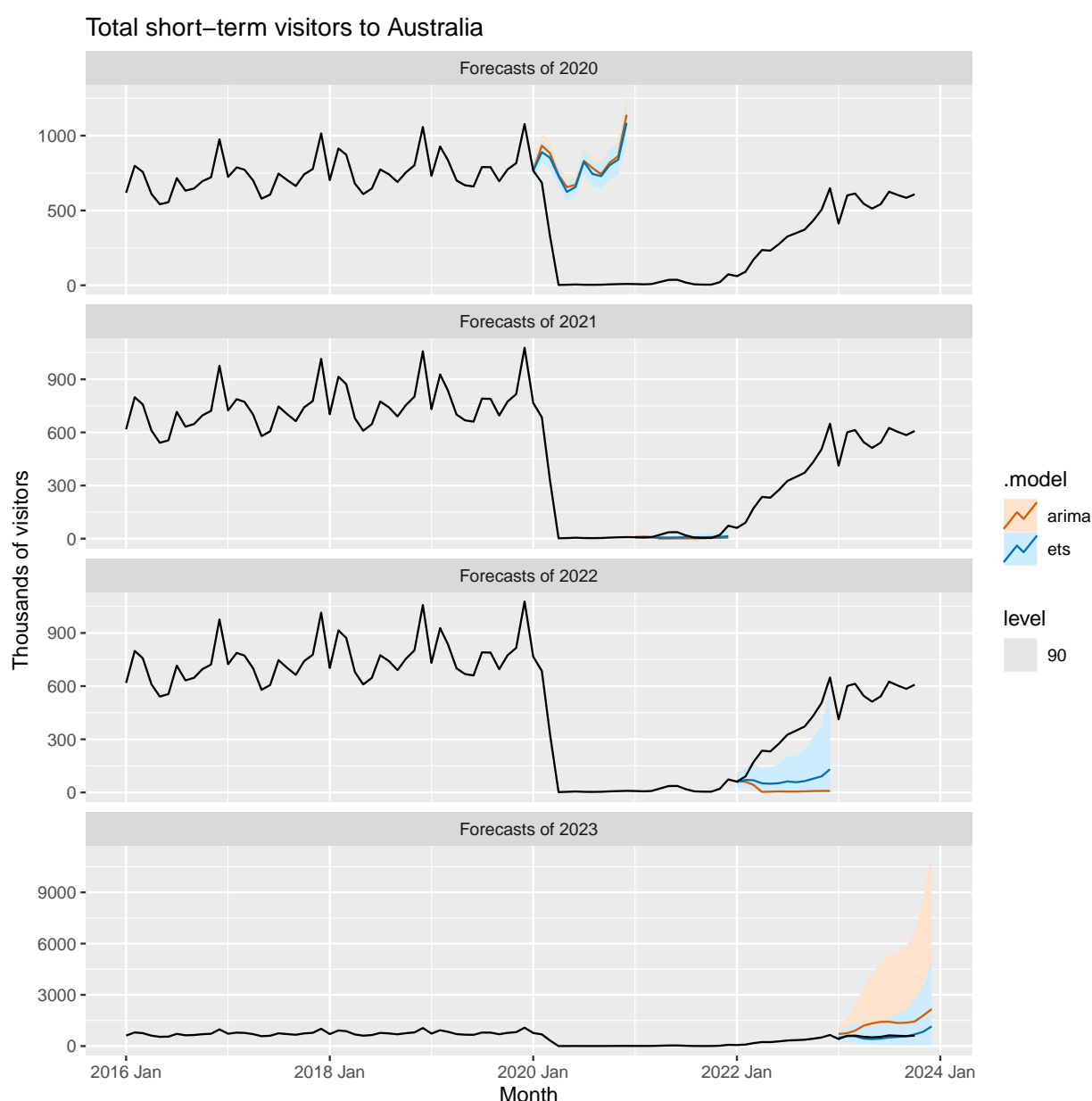


Figure 2: Highly adaptive models. Forecasts from ETS and ARIMA models. Neither works particularly well for disruptions of this magnitude.

similar patterns to the past. The forecasts for 2021 were made after 9 months of very low levels of arrivals, and both models show forecasts consistent with the recent history. The use of logarithms is particularly important here as the variance is much smaller during 2020 than previously, but after taking logarithms, the variance is more stable over time. The forecasts made at the start of 2022 have struggled to detect the small increase in traffic at the end of 2021, and both models have forecast relatively flat trajectories as a result, although the prediction intervals are wide indicating model uncertainty. Finally, the forecasts made at the end of 2022 have captured the increasing trend, but ETS is much closer to the reality, adapting more quickly to the changing patterns. Again, the wide prediction intervals indicate a high level of uncertainty. It is possible to make ETS more adaptive to changes in the data by increasing the value of the smoothing parameters. For example, a high value of β (the smoothing parameter for the slope), will result in changes of trend being incorporated into the forecasts more quickly, at the risk of over-reacting to noise in the data, and increasing the size of the prediction intervals even more.

Figure 3 shows forecasts obtained using a dynamic regression model (i.e., a regression with ARIMA errors) using two intervention variables: a level shift from March 2022 to November 2022, and a ramp from October 2021 to November 2022. It is evident that the level shift variable has worked well, giving relatively good forecasts for 2021. However, the forecasts for 2022 are particularly poor, because the ramp slope has been greatly overestimated, as it was based on only three observations (Oct – Dec 2021). The forecasts for 2023 are much better, and the relatively large prediction intervals are appropriate given the uncertainty in the industry at the end of 2022.

The third solution (Figure 4) involved setting the observations during the disruption period to missing, and then fitting an ARIMA model to the series. Consequently, the first three years of forecasts show what might have been without the COVID-19 pandemic, based on the history to the end of 2019. The final forecasts for 2023 use the data to the end of 2019, and the two observations in November and December 2022. These are much better, but the prediction intervals are too narrow, as there hasn't been sufficient recent data.

We show in Figure 5 the forecasts obtained using solution 4. Here we have replaced observations between March 2020 and October 2022 with estimates based on the average of the same month in the three years prior to March 2020. The resulting forecasts are similar to those from solution 3, but with narrower prediction intervals because the model is (falsely) assuming that the “observations” during the disruption are real.

Finally, in Figure 6, we show the ensemble forecasts obtained by averaging the forecasts from the four solutions discussed above. The first and fourth solutions carry double weight, because both the

ETS and ARIMA forecasts were included in the ensemble. This time, no prediction intervals have been computed. The first set of forecasts have not taken any account of the disruption, and so are almost identical to those obtained in the previous plots. The last set of forecasts are reasonably good, showing what is obtained by combining the forecasting methods used to handle the disruption period. However, the forecasts for 2021 and 2022 are particularly poor, because they are averaging forecasts that aim to achieve different objectives, and are affected by the poor forecasts obtained using the dynamic regression model.

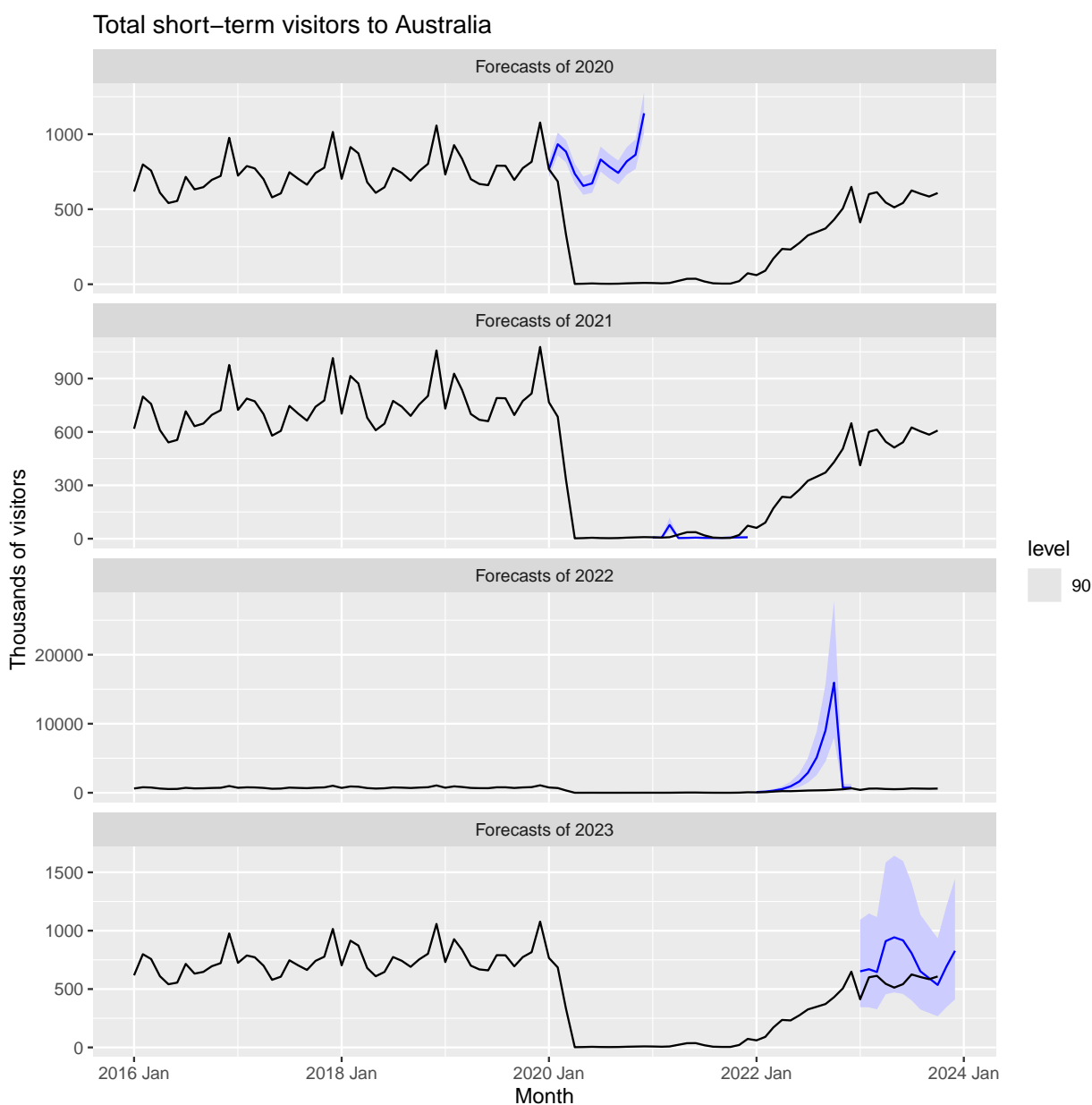


Figure 3: Intervention model. We use a level shift from March 2020 to October 2022, and a ramp from October 2021 to October 2022. After October 2022, all intervention variables are set to zero.

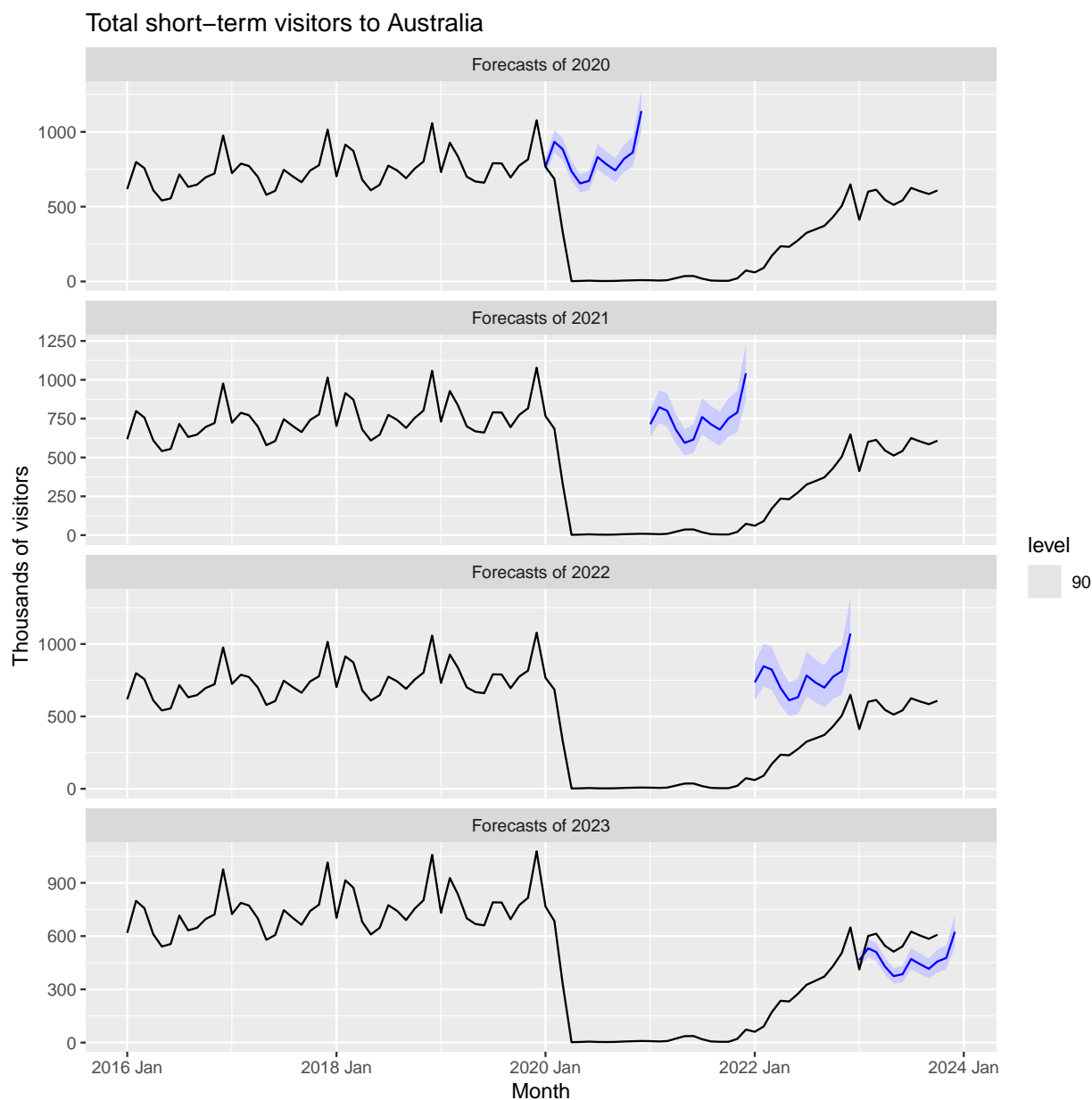


Figure 4: Set to missing. The period from March 2020 to October 2022 is set to missing. So the first three years of forecasts show what might have been. The fourth set of forecasts uses observations from November and December 2022, and so the forecasts have been adjusted downwards.

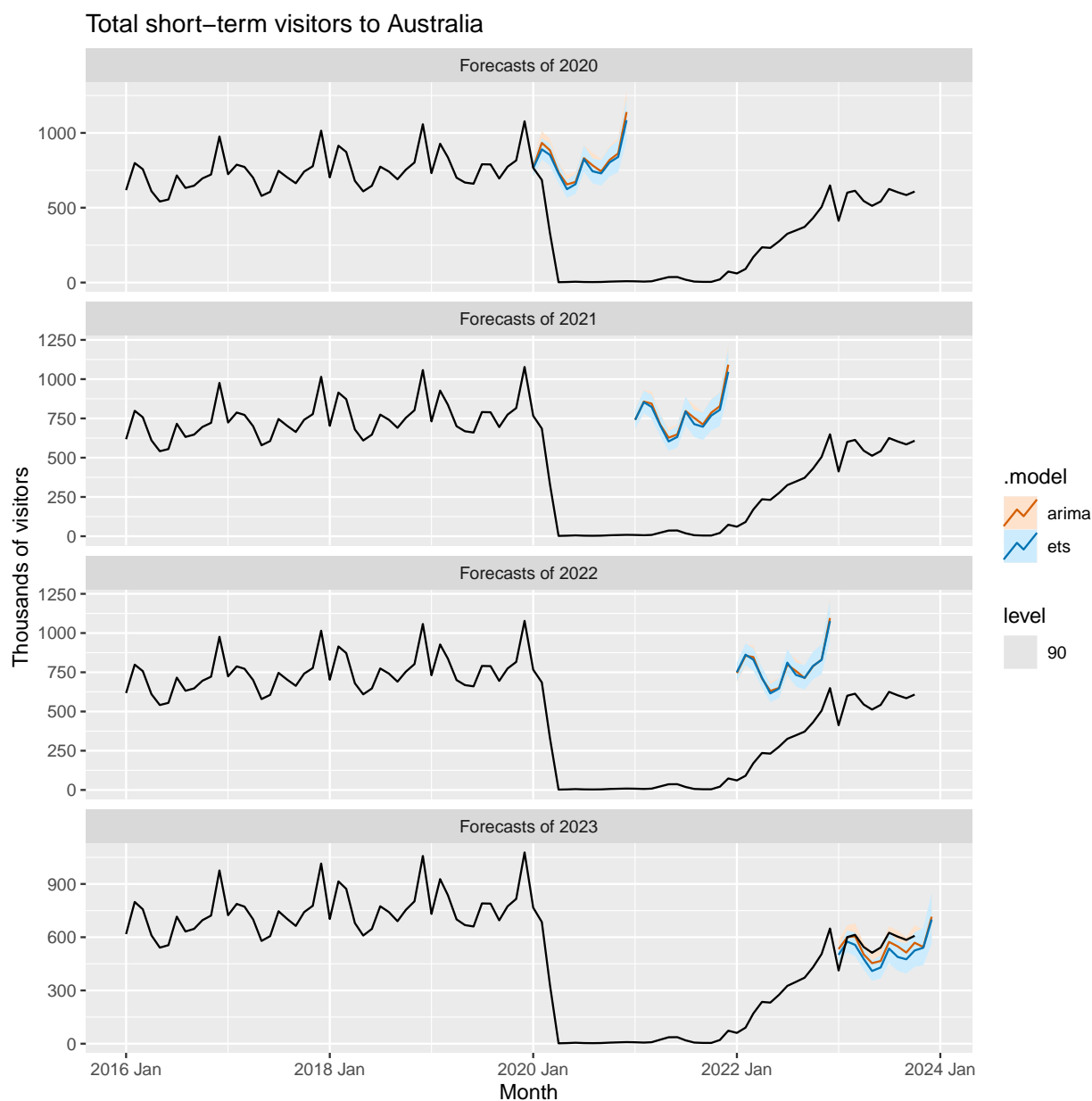


Figure 5: Estimate what might have been. We replace the observations from March 2020 to October 2022 with the average of the same month in the three years prior to March 2020.

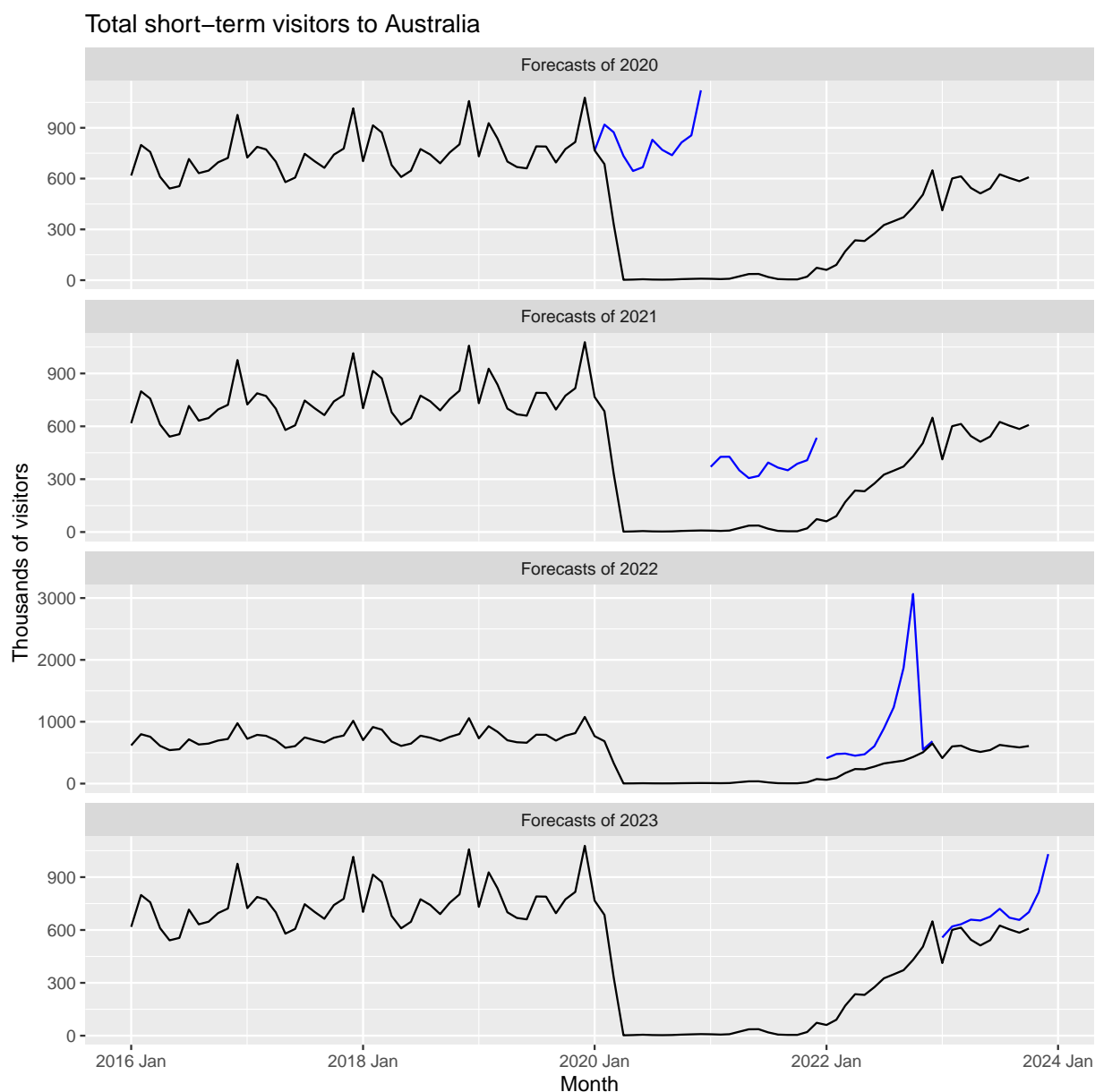


Figure 6: *Ensemble approach*, combining the previous four approaches shown in Figures 2–5.

3.2 Example: Pedestrians

The pedestrian data set is a daily time series showing the number of pedestrians per day in Melbourne, Australia. The data were obtained from the Pedestrian Counting System maintained by the City of Melbourne (City of Melbourne 2024), and are based on automated hourly counts from sensors at 66 locations over the period from 2019-01-01 to 2021-12-31. All sensors had some missing data, and we chose to use sensors with no more than 720 missing hours, equivalent to 30 days of missing data. This resulted in 25 sensors being used in the analysis. The data were aggregated to daily counts for each sensor, and then the results were averaged across sensors to obtain a measure of pedestrian traffic per day. The resulting time series is shown in Figure 7.

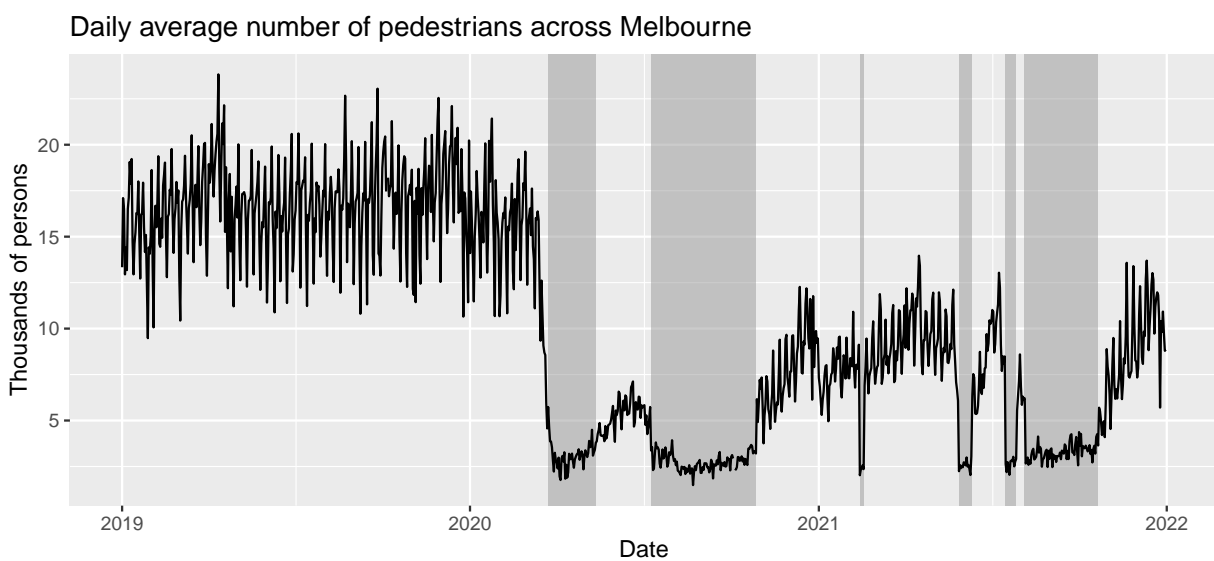


Figure 7: Pedestrian traffic in Melbourne, Australia, from 2018 to 2023. The gray shaded regions indicate periods of lockdown due to COVID-19.

We have also shown the periods of lockdown due to COVID-19. In Melbourne, there were six separate lockdowns ranging from 5 days to 111 days in length. The first official lockdown period was from 31 March 2020 to 12 May 2020, and the last lockdown was from 05 Aug 2021 to 21 Oct 2021. However, the first period was preceded by over a week when most people elected to stay and work at home, so we have chosen to start the first period on 23 Mar 2020, to better reflect human behaviour. Otherwise, we have used the official lockdown dates (Wikipedia 2024).

We evaluate the solutions discussed in Section 2 by using time series cross-validation, with the initial training set comprising the whole of 2019, and subsequent training sets growing by 1 week at a time, to the end of 2021. The test sets are always 1 week long. Thus, we have evaluated the forecasts during 2020 and 2021, covering all lockdown periods, and the start of the recovery period. No transformations of the data have been used. The results are shown in Figures 8–11. Prediction intervals have not been shown to avoid cluttering the plots.

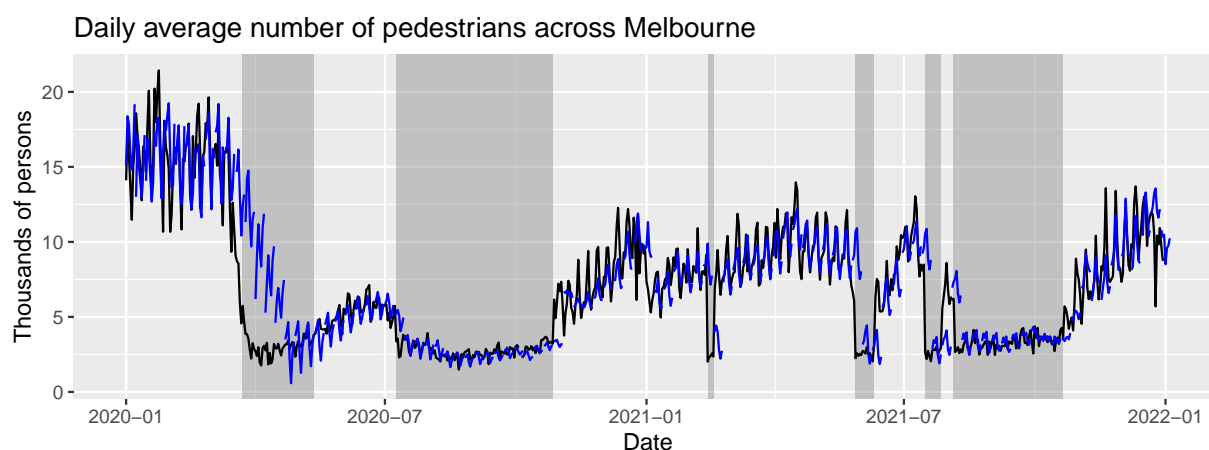


Figure 8: *Highly adaptive model:* the forecasts do not take into account the lockdown periods, but the ARIMA model is largely able to adjust to the changing level of the series.

We use ARIMA models in Figure 8, which have mostly been able to adapt to the changing level of the series, apart from during the first few weeks of the first lockdown, and the first day or two of subsequent lockdowns.

In Figure 9, we have used an intervention model, with a dummy variable indicating the lockdown periods. This has worked a little better than the ARIMA model used in Figure 8, apart from around the first lockdown period. Many people were in self-imposed lockdown before the start of the first lockdown period, and were reluctant to return to the city after it ended, resulting in poor forecasts on either side of the first lockdown.

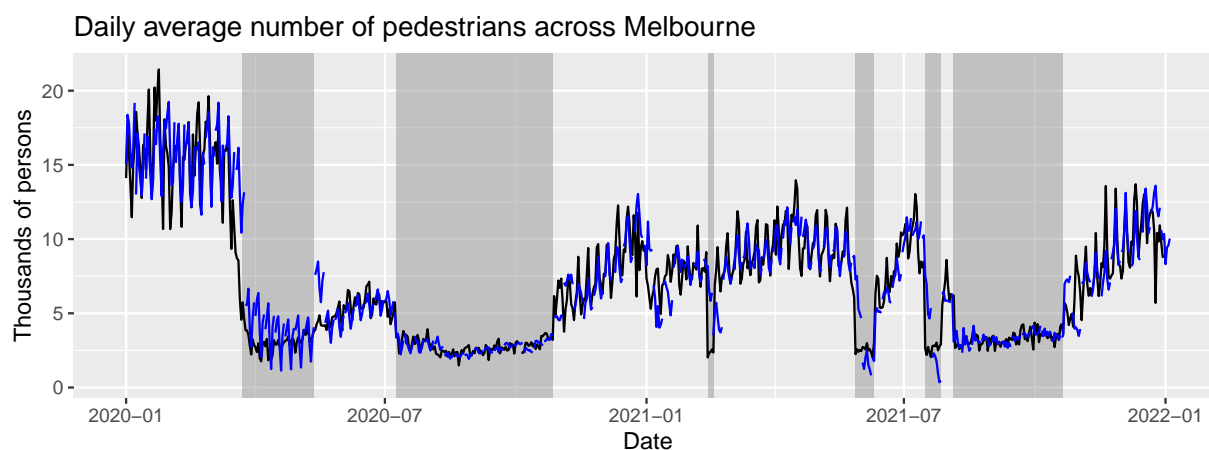


Figure 9: *Intervention model:* the forecasts are based on an ARIMA model with a dummy variable indicating when the lockdown periods occurred. Many people were in self-imposed lockdown before the start of the first lockdown period, and were reluctant to return to the city after it ended, resulting in poor forecasts on either side of the first lockdown. The model was able to adjust to the changing level of the series during subsequent lockdowns.

Figure 10 shows the results obtained by setting the observations during the lockdown periods to missing, and then fitting an ARIMA model to the remaining data. The forecasts during lockdowns show

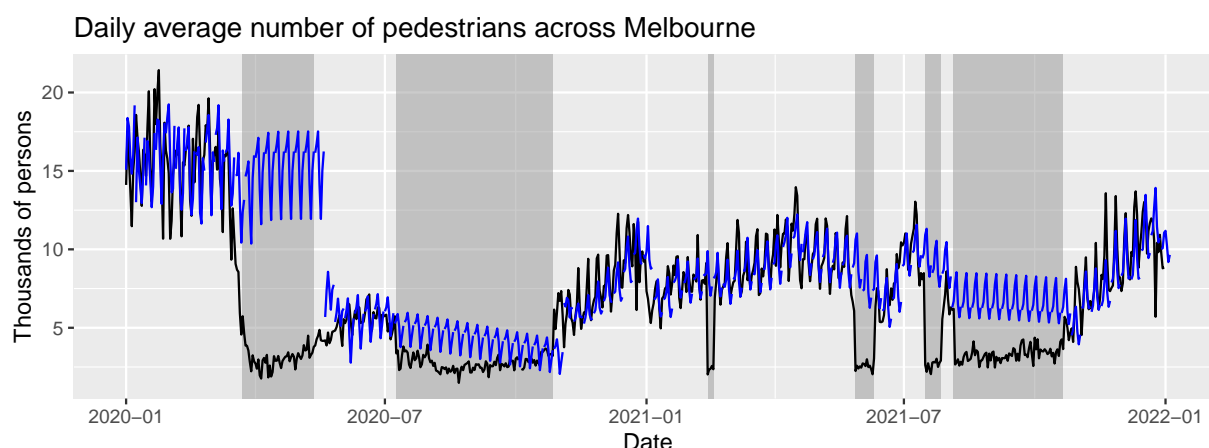


Figure 10: *Set to missing*: the lockdown periods are set to missing, and the ARIMA model uses the remaining data to produce forecasts. Thus, the forecasts are based on only previous non-lockdown data. Naturally, the forecasts during the lockdown periods do not reflect reality, but they can be interpreted as what might have been.

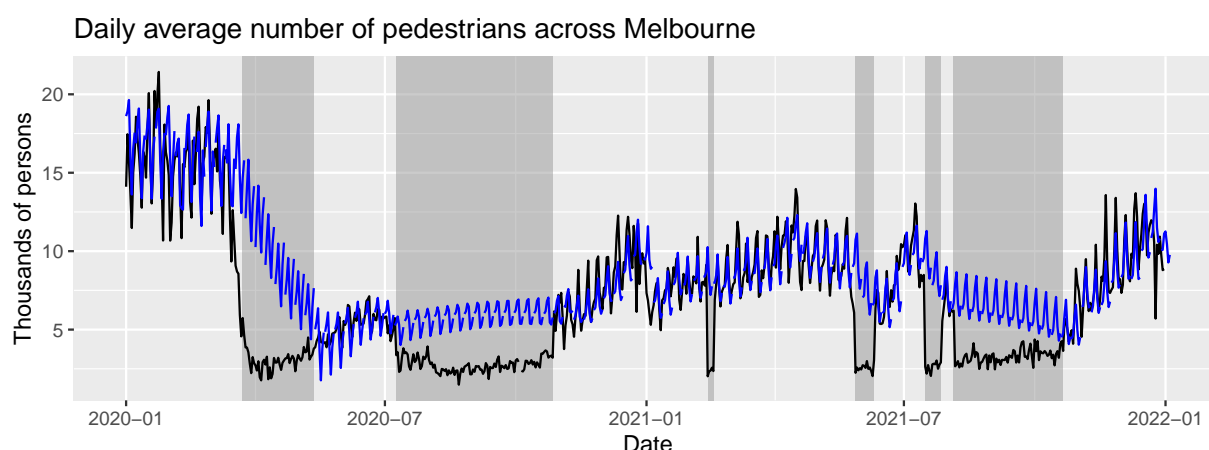


Figure 11: *Estimate what might have been*: the lockdown periods are replaced with estimated counts based on an ARIMA model applied to the remaining data. Then a new ARIMA model is fitted to the whole data set, and used to produce forecasts. As a result, the forecasts during the lockdown periods are not true forecasts, as they use non-lockdown data from the future.

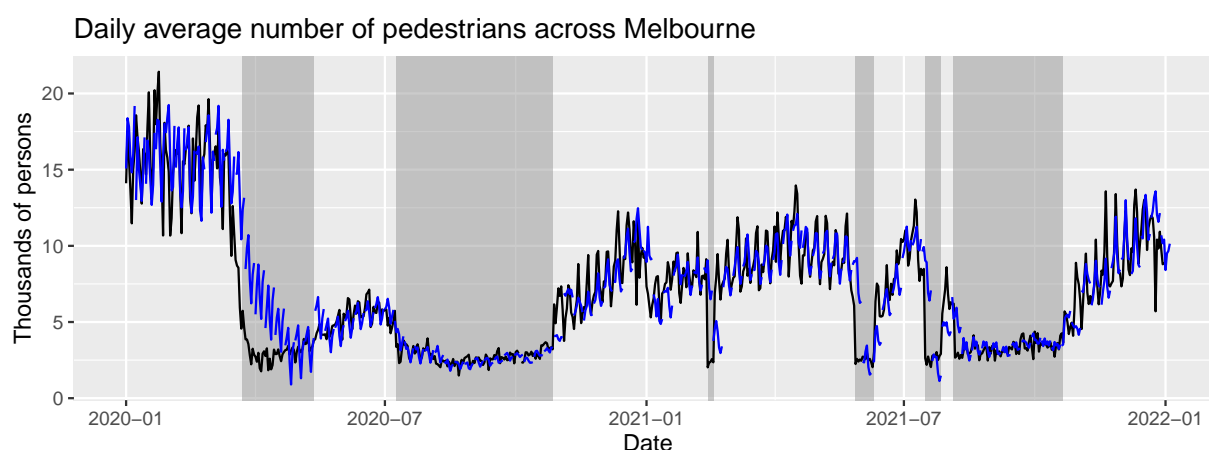


Figure 12: *Ensemble forecasts* based on the first two solutions shown in Figures 8 and 9. Other than having some difficulty around the period of the first lockdown, the results are relatively good.

what might have been if that particular lockdown hadn't occurred, based on previous non-lockdown data.

The forecasts shown in Figure 11 are obtained by replacing the observations during the lockdown periods with estimates based on an ARIMA model applied to the remaining data. Then a new ARIMA model is fitted to the whole data set, and used to produce forecasts. As a result, the "forecasts" appear to interpolate across the lockdown periods, reflecting neither the true lockdown pattern, or the pattern that might have occurred without the lockdowns.

Finally, in Figure 12, we show the ensemble forecasts obtained by averaging the forecasts from the first two solutions discussed above. The results are relatively good, other than having some difficulty around the period of the first lockdown.

4 Discussion

We have presented several approaches to forecasting in the presence of an interruption. Each has its advantages and disadvantages, and the choice of which to use will depend on the situation. In fact, we have used almost all of these approaches in our own consulting work. An intervention model is often a good solution, provided the intervention can be modelled relatively simply. However, if the intervention is complex, then a highly adaptive method is often better. The highly adaptive model is also useful when the timing of the interruption is unknown, or when the post-interruption period is expected to be very different from the pre-interruption period. The missing value approach is particularly useful when only post-interruption forecasts are required, and forecasts during the interruption period are not needed.

We have not previously used the weighted approach in practice, and software to implement it is not currently available. However, it seems to be a good compromise between the highly adaptive model and the intervention model, and we expect it to be useful in many situations, once software is available to implement it easily.

Finally, the ensemble approach is useful when there is uncertainty about which approach to use, and will often result in greater accuracy due the power of averaging.

We have not discussed probabilistic forecasting in any detail in this paper, other than producing some prediction intervals for some of the examples. One would expect forecast uncertainty to increase during, and to some extent after, a major disruption. The highly adaptive model and innovation model should produce relatively good prediction intervals, although the variance is often underestimated. Because the missing value approach is producing counterfactual ("what might have

been”) forecasts, the prediction intervals will often not include the actual observations during the disruption period. Prediction intervals for ensemble forecasts can be obtained by averaging the end points of the component prediction intervals (Lichtendahl Jr, Grushka-Cockayne & Winkler 2013), for those models where the prediction intervals seem reasonable.

References

- Australian Bureau of Statistics (2024). *Overseas Arrivals and Departures, Australia*. Accessed 8 January 2024. <https://www.abs.gov.au/statistics/industry/tourism-and-transport/overseas-arrivals-and-departures-australia/oct-2023>.
- City of Melbourne (2024). *Pedestrian counting system*. Accessed 5 January 2024. Melbourne, Australia. <https://www.pedestrian.melbourne.vic.gov.au/>.
- Hyndman, RJ & G Athanasopoulos (2021). *Forecasting: principles and practice*. 3rd ed. Melbourne, Australia: OTexts. <http://OTexts.org/fpp3>.
- Lichtendahl Jr, KC, Y Grushka-Cockayne & RL Winkler (2013). Is it better to average probabilities or quantiles? *Management Science* **59**(7), 1594–1611.
- Wang, X, RJ Hyndman, F Li & Y Kang (2023). Forecast combinations: an over 50-year review. *International J Forecasting* **39**(4), 1518–1547. <http://robjhyndman.com/publications/combinations/>.
- Wikipedia (2024). *COVID-19 pandemic in Victoria*. Accessed 8 January 2024. https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Victoria#Lockdown_statistics.