Journal of the Operational Research Society

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjor20

Forecasting interrupted time series

Rob J. Hyndman & Bahman Rostami-Tabar

To cite this article: Rob J. Hyndman & Bahman Rostami-Tabar (04 Sep 2024): Forecasting interrupted time series, Journal of the Operational Research Society, DOI: 10.1080/01605682.2024.2395315

To link to this article: https://doi.org/10.1080/01605682.2024.2395315

Published online: 04 Sep 2024.

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
https://doi.org/10.1080/01605682.2024.2395315

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Forecasting interrupted time series

Rob J. Hyndman[a] and Bahman Rostami-Tabar[b]

[a]Monash University, Clayton, Australia; ARC Training Centre in Optimisation Technologies, Integrated Methodologies, and Applications (OPTIMA); [b]Data Lab for Social Good Group, Cardiff Business School, Cardiff University, Cardiff, UK

ABSTRACT

Forecasting interrupted time series data is a major challenge for forecasting teams, especially in light of events such as the COVID-19 pandemic. This paper investigates several strategies for dealing with interruptions in time series forecasting, including highly adaptable models, intervention models, marking interrupted periods as missing, forecasting what may have been, downweighting the interruption period, and ensemble models. Each approach offers specific advantages and disadvantages, such as adaptability, memory retention, data integrity, flexibility, and accuracy. We evaluate the effectiveness of these strategies using two actual datasets that were interrupted by COVID-19, and we provide recommendations for how to handle these interruptions. This work contributes to the literature on time series forecasting, offering insights for academics and practitioners dealing with interrupted data in numerous domains.

ARTICLE HISTORY
Received 20 February 2024
Accepted 11 August 2024

KEYWORDS
Counterfactual forecasting; COVID-19; disruptive events; intervention analysis; structural breaks

1. Introduction

Time series forecasting models use historical data to estimate future values (Fildes et al., 2008) based on consistent patterns observed in the past. However, time series are sometimes disrupted by unusual events that jeopardise the regularity of the time series pattern. This can be defined as an *interrupted time series*. Time series data may be disrupted by a variety of reasons, ranging from temporary policy changes to natural disasters. Outbreaks, epidemics, and pandemics, for example, may have a considerable impact on population behaviour, causing discrepancies and interruptions of data in a variety of industries such as healthcare, pharmaceuticals, transportation, retail, tourism, and traffic management. Similarly, transportation-related data may be interrupted due to equipment failures, such as malfunctions or breakdowns in vehicles, aircraft, or rail systems, resulting in missing or inaccurate data points. Furthermore, cybersecurity breaches data quality by possibly skewing readings in datasets relating to financial transactions, network traffic, and user activity, complicating time series analysis. Supply chain disruptions induced by factors such as strikes or shortages cause irregularities in sales, inventory levels, production, and delivery, often resulting in data gaps. For example, the war in Ukraine caused disruptions in sunflower oil and wheat supply chains, which were most likely driven by social media and people's recent memories of the COVID-19 toilet paper shortage. A power outage in Eastern Germany halted production at a Tesla facility and cut off electricity to a large retailer distribution centre, disrupting supplies to around 500 stores in and around Berlin for several weeks. Recent strikes in various forms of transport, including air and rail in Germany, France, and the United Kingdom, have had an impact on passenger flow and retail sales at airports and train stations. These issues highlight the significance of robust forecasting systems in addressing such disruptions effectively.

The disruptions 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.

The presence of such disrupted events poses a significant challenge to time series forecasting approaches, restricting their ability to reliably capture systematic information and forecast future values. The implications of these temporal interruptions are far-reaching: time series forecasting techniques, which are a key component of the forecaster's toolbox, may no longer be viable options since time series models usually assume that the data will evolve in the future in a similar way to how it has 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

CONTACT Bahman Rostami-Tabar rostami-tabarb@cardiff.ac.uk 3 Colum Drive, CF10 3EU, Cardiff

short term. Therefore, practitioners face a critical question: how to effectively forecast time series data that are impacted by such disruptive events? Many forecasters have faced this issue recently with the COVID-19 pandemic, where historical patterns have been severely disrupted due to policy interventions, lockdowns, and other restrictions. The motivation for this study originates from numerous conversations with academics and practitioners, who have expressed widespread concern about dealing with data impacted by events like the COVID-19 pandemic. In this paper, we consider a range of models that can be used to forecast time series influenced by disrupted events and compare their performance on some real data sets.

Although several approaches to the problem exist, the literature lacks recommendations and comparisons of alternative strategies for forecasting time series data influenced by interruptions such as the COVID-19 pandemic. This paper aims to close this gap by proposing several methodologies for forecasting time series influenced by such disruptions.

We consider the problem of forecasting interrupted time series both during and after the event. This is a separate problem from change point detection (Aue & Horváth, 2013; Truong et al., 2020) and anomaly detection (Talagala et al., 2020). In the situations we investigate, we know that a change has occurred, which has caused some unusual observations, and we want to forecast what will happen next. Change point detection refers to determining *when* a change occurred (Blázquez-García et al., 2021), and anomaly detection aims to identify unusual observations with the intention of either minimising their value or paying particular attention to them and doing a root-cause investigation. Our study focus also differs from *intervention modelling*, sometimes referred to as *interrupted time series analysis and modelling* (McDowall et al., 2019), which is used when we have data about an outcome over time and the aim is to assess the impact of an intervention, policy, or program implemented at a specific time point. We are not trying to measure the impact of the interruption; rather, we are trying to produce sensible forecasts during and after the interruption.

To our knowledge, this is the first study to describe and compare general strategies for forecasting interrupted time series, such as COVID-19-affected data. The study proposes and evaluates multiple forecasting strategies designed to handle interruptions in time series forecasting caused by disruptive events. These approaches include a range of models capable of capturing different types of changes, from simple level shifts to more complex alterations in seasonal patterns and trends. By comparing the performance of these models on two real data sets affected by the COVID-19 pandemic, the paper offers valuable insights into the effectiveness and practical implications of each approach. By discussing the advantages and disadvantages of each strategy, and providing recommendations based on the findings, the paper equips practitioners with actionable insights for addressing the challenges associated with forecasting during and after disruptive events. Further, by adhering to reproducibility principles and providing both data and R code for the forecasting models, the research contributes to transparency and promotes the generalisability of the suggested strategies to a variety of domains. This openness allows for the replication of results, ensuring the reliability of findings, and enhancing the accessibility and utility of study outputs for the larger research community.

We introduce several strategies to handling interruptions when forecasting in Section 2. We then apply these approaches to two real data sets in Section 3. Finally, in Section 4, we discuss some of the advantages and disadvantages of each approach and provide some recommendations for practitioners.

2. Handling interruptions when forecasting

In this section, we describe several possible strategies to handle interruptions when forecasting time series data. The handling of time series interruptions is a complex and multifaceted issue, and the choice of the most appropriate strategy depends on the specific characteristics of the data and the nature of the interruption. Therefore, in this paper, we propose multiple strategies rather than a single approach to effectively address different situations and cases.

2.1. Use a highly adaptive model

Highly adaptive models offer flexibility in adjusting to changes in time series characteristics over time. These models can adjust to the interruption as it is happening and will therefore be able to approximate the data generating process relatively well.

Exponential Smoothing State Space models (ETS) represent an adaptive method that has demonstrated competitiveness in time series forecasting (Gardner, 2006). These models allow for the adaptation of model parameters over time to accommodate shifts in time series characteristics (Hyndman et al., 2002). 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.

Similarly, time-varying parameter models exhibit high adaptability (Harvey, 2006), enabling forecasting models to better accommodate interruptions in data over time. Time-varying parameters are frequently utilised in dynamic forecasting models, such as state-space models or Bayesian structural time series models (Talih & Hengartner, 2005). Some models employ approaches to reduce the influence of affected observations during model training by assigning different weights. The utilisation of weights extends beyond merely including or excluding observations and can also balance the degree of influence each observation has on the forecasting model (Khoshgoftaar et al., 2007).

The prediction intervals from highly adaptive models tend to be wide 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.

The disadvantage of smoothing and other highly adaptive techniques is that they need time to adapt after a disturbance and can eventually end up chasing noise. The latter is an intriguing issue since individuals often prefer adaptability while failing to recognise that it does not necessarily improve forecasting.

2.2. Use an intervention model

One modelling approach that has gained popularity in handling disruptions is the intervention model, also known as interrupted time series (ITS) modelling, introduced by Box and Tiao (1975). This approach focusses on understanding how and whether outcomes change following the implementation of an intervention, policy, or programme at a specific time point. ITS models provide a robust tool for analysing policy and programme evaluation across diverse domains (Bernal et al., 2017; McDowall et al., 2019). These models can forecast future values by developing models that incorporate a range of components, such as pre-intervention levels, trends, seasonality, covariates related to the intervention time, and other external factors that might affect the outcome variable. A related approach is the multiple regimes model, which is particularly useful in forecasting due to its ability to capture structural changes and transitions observed in interrupted time series data (Koop & Potter, 2007).

A special case of the intervention models proposed by Box and Tiao (1975) is a dynamic regression (or regression with ARIMA errors) model. This model includes covariates describing the interruption explicitly, with an ARIMA process describing the remaining time series dynamics. 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. Intervention models are useful because they enable the forecasting of similar future disruptions, but distinguishing between a disruption and a general driving factor, such as a promotion or a seasonal storm, may become subjective. 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

Disruptive events may potentially alter time series by causing gaps. Alternatively, problematic observations that occur during the period of disruption can be set to missing, and a model fitted to the remaining observations 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.

While many time series forecasting models struggle with missing data, several approaches can internally handle missing data (Twala et al., 2008). Wu et al. (2015) suggested a forecasting approach based on a Least Squares Support Vector Machine (LSSVM) that was particularly designed to handle time series forecasting with missing data. A different study by Tang et al. (2020) focused on local and global temporal dynamics for multivariate time series forecasting with missing data. Their proposed framework uses a memory network to capture global temporal patterns with local data as key components.

In this approach, no information is retained during the disruption, so 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. 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.

These approaches are sometimes called "counterfactual" methods, as they examine what would have happened to the forecast variable of interest if the disruptive event had not occurred. Athanasopoulos et al. (2023) proposed a variant of this approach for forecasting tourism recovery from COVID-19, combining forecast reconciliation and forecast combinations applied to historical data to generate COVID-free counterfactual forecasts of what might have been if the pandemic never occurred. Then scenario-based judgemental probabilistic forecasts were compared with the counterfactual forecasts to better understand the future recovery of the tourism sector from the pandemic.

Inventory stockouts are a common cause of data interruption in the supply chain. To deal with these interruptions, a counterfactual approach is popular, whereby the demand corresponding to the stock-out period is estimated as if there were no stock-outs, and then forecasts are obtained. Bell (2000) presented an adjustment strategy for stock-out periods that involves smoothing demand volatility and correcting for stockouts using predicted variance conditioned on the observed stockout. Trapero et al. (2024) also used the Tobit Kalman filter (TKF) for models presented in a state space framework for forecasting purposes. This method can efficiently deal with trends, seasonality, and exogenous influences in censored data, as long as the forecasting model functions within a state space framework.

In the context of forecasting disruptions such as the COVID-19 pandemic, Bayesian approaches have also been proposed for detecting changing points in outbreaks and forecasting case numbers under counterfactual scenarios (Dehning et al., 2020).

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 based on 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

In this section, we implement and evaluate the strategies described in Section 2 using two daily and monthly datasets collected over the last several years, where the impact 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.

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. (A bias correction is also applied to ensure the back-transformed forecasts are the means of the forecast distributions.) A log transformation 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 effect, and so they show 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 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 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 in trend being incorporated into the forecasts more quickly, at the risk of overreacting 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

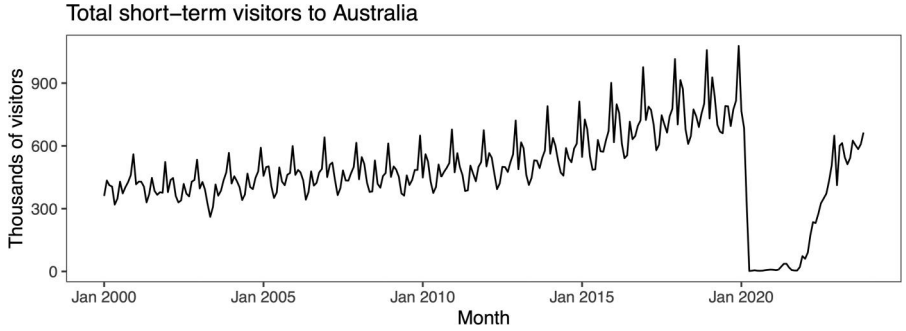

Figure 1. Short-term visitor arrivals to Australia (monthly): Jan 2000 – May 2023. Source: Australian Bureau of Statistics (2024). The Australian international borders closed in March 2020 and officially reopened in February 2022.

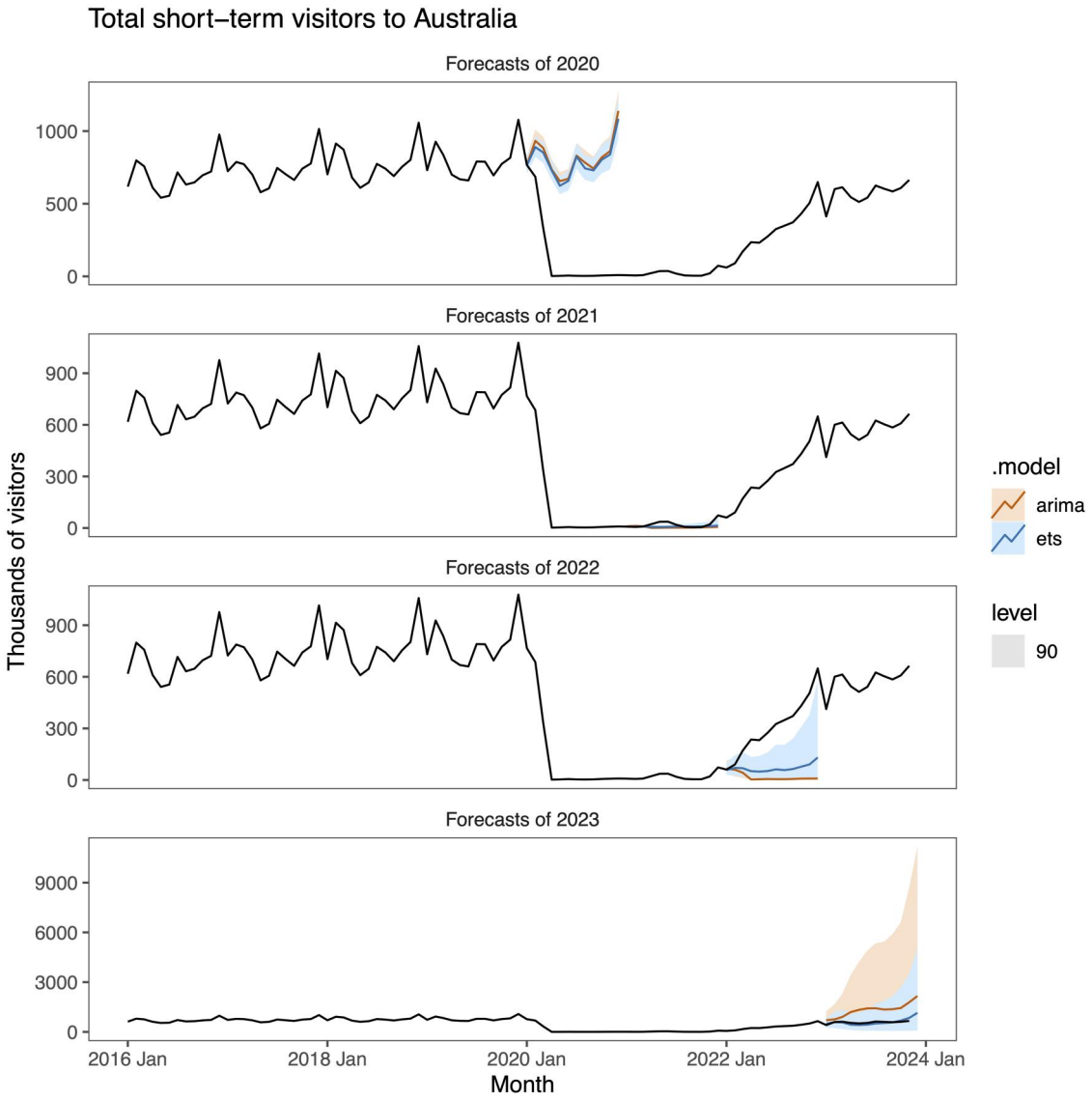

Figure 2. Highly adaptive models. Each panel shows one year of monthly forecasts from ETS and ARIMA models. Neither works particularly well for disruptions of this magnitude.

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. The prediction intervals are formed from the mixture distributions obtained by combining the four forecast distributions for each time horizon. The first set of forecasts has not taken any account of the disruption, and so they are almost identical to those obtained in the previous plots. The last set of forecasts is 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.

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

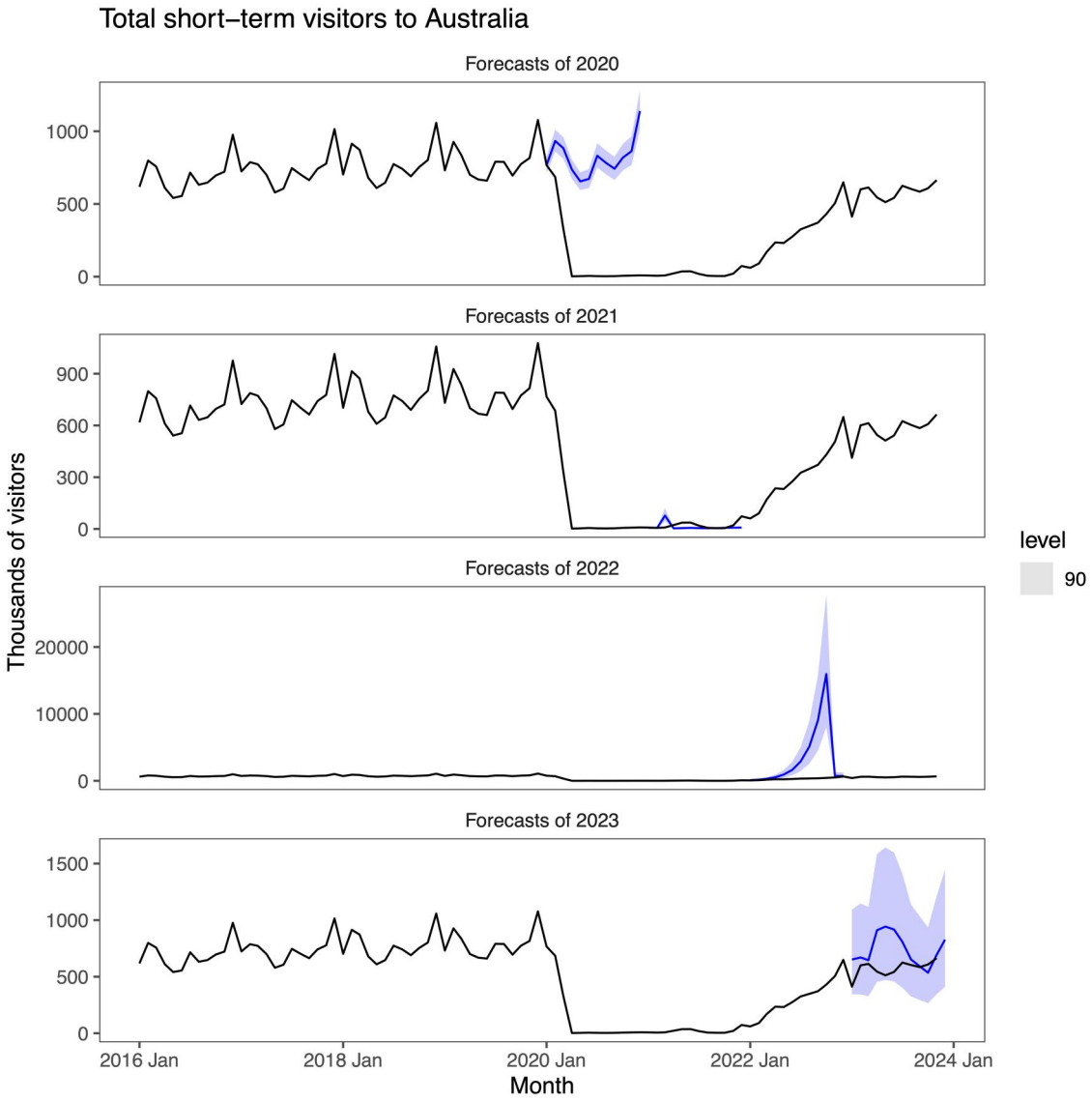

Figure 3. Intervention model. Each panel shows one year of monthly forecasts from an intervention model containing a level shift from March 2020 to October 2022, and a ramp from October 2021 to October 2022. After October 2022, all intervention variables are set to zero.

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.

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 March 31, 2020, to May 12, 2020, and the last lockdown was from August 5, 2021, to October 21, 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 March 23, 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 one week long. Thus, we have evaluated the forecasts for 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.

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.

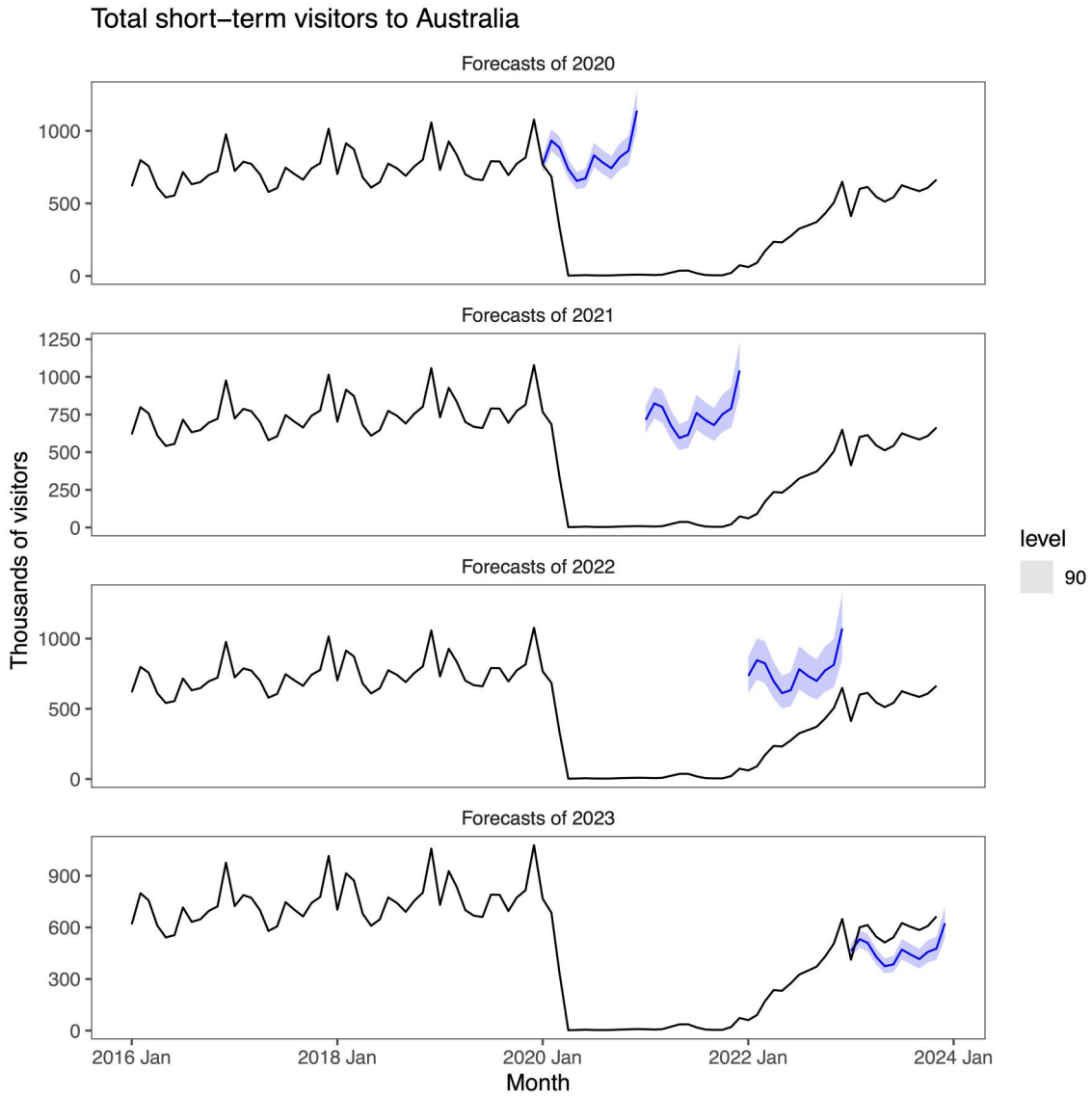

Figure 4. Set to missing. Each panel shows one year of monthly forecasts from an ARIMA model. The period from March 2020 to October 2022 is set to missing. So the first three years of forecasts show what might have been. The fourth set of forecasts uses observations from November and December 2022, and so the forecasts have been adjusted downward.

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 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 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 nor 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 and conclusions

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

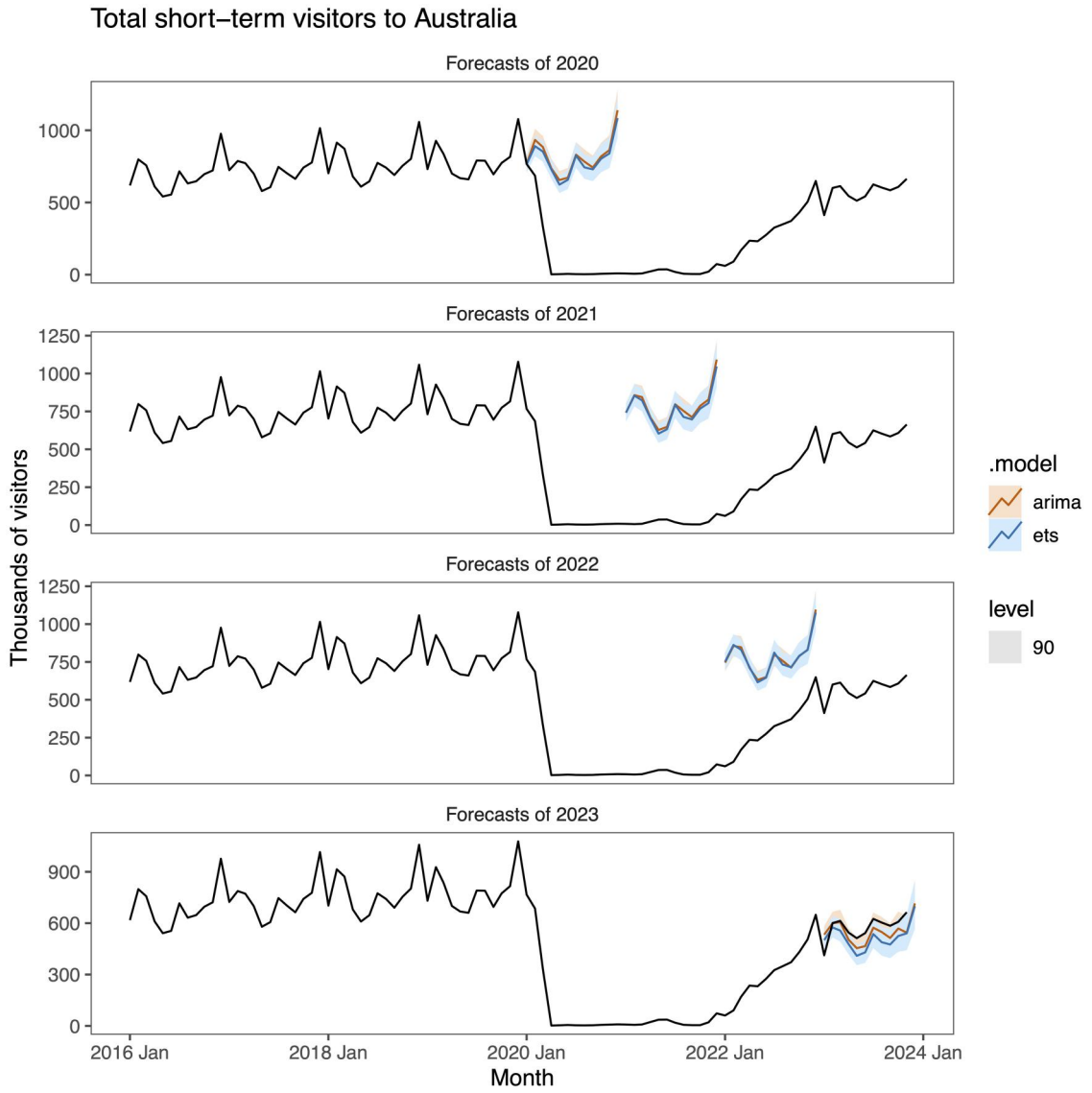

Figure 5. Estimate what might have been. Each panel shows one year of monthly forecasts from ETS and ARIMA models. We have replaced the observations from March 2020 to October 2022 with the average of the same month in the three years prior to March 2020.

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 open-source 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 to the power of averaging.

When forecasting interrupted time series data, exogenous variables associated with specific events should be identified and incorporated where possible, regardless of which approach is employed, since they may improve forecasting model performance.

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 et al., 2013), for those models where the prediction intervals seem reasonable.

It is not uncommon for modellers to experience interruptions in time series data. Both those in

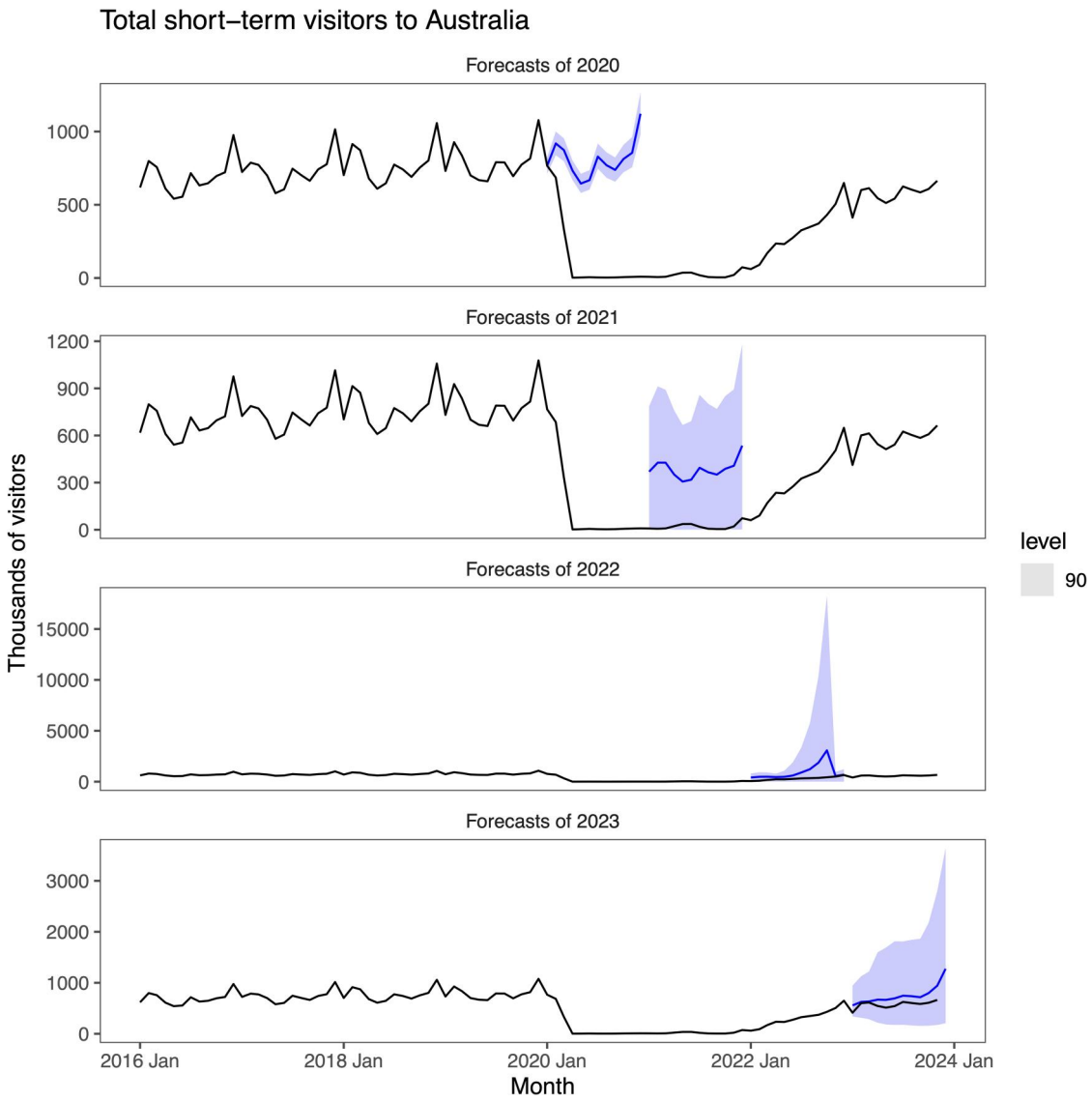

Figure 6. Ensemble approach, combining the previous four approaches shown in Figures 2–5. The point forecasts are the average of the four approaches, while the prediction intervals are formed by combining the four forecast distributions at each horizon to form a mixture distribution.

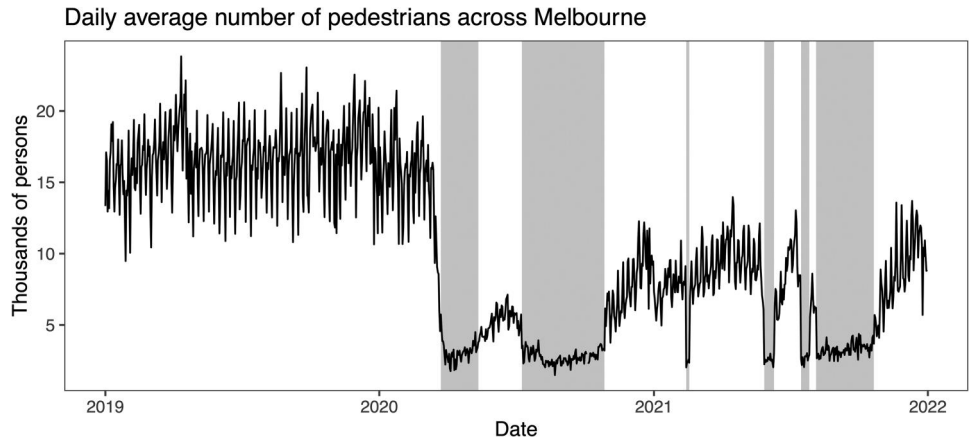

Figure 7. Average pedestrian traffic in Melbourne, Australia, from 2018 to 2023, based on automated hourly counts from sensors around the city. The counts are averaged over 25 sensors, each of which had no more than 30 days of missing data. The grey-shaded regions indicate periods of lockdown due to COVID-19. Source: City of Melbourne (2024).

charge of developing forecasting models and managers relying on forecasts to inform decisions and policies have recently faced the issue of effectively forecasting time series data influenced by the COVID-19 pandemic. While this can be a critical challenge for many forecasting teams, it is important

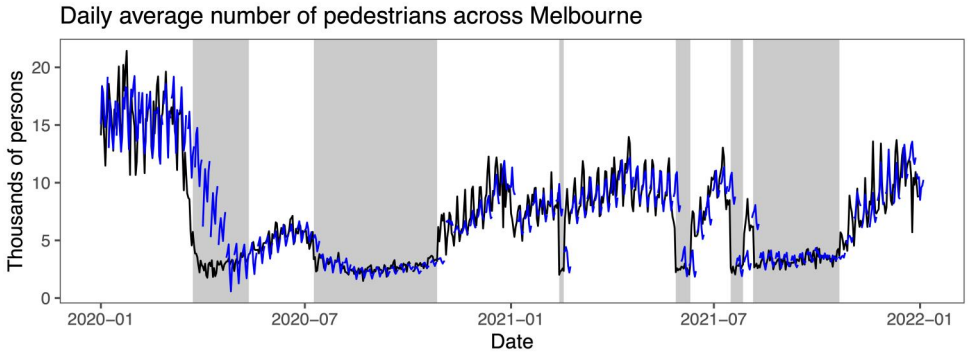

Figure 8. Highly adaptive model: Rolling 7-day forecasts of daily pedestrian data, made each week, using ARIMA models. The forecasts do not take into account the lockdown periods, but the ARIMA models are largely able to adjust to the changing level of the series.

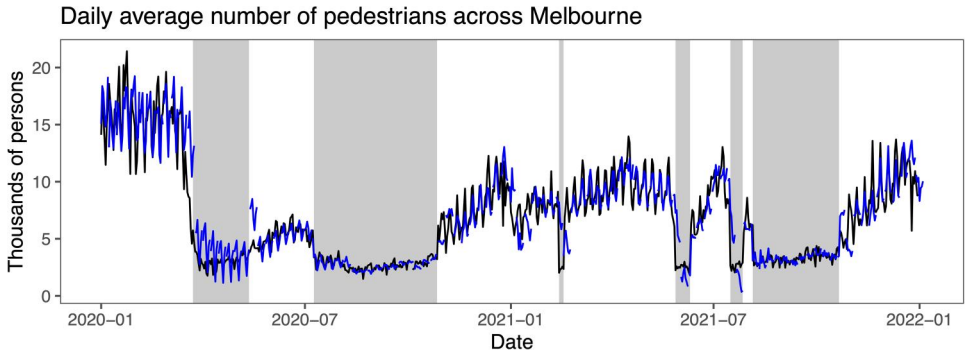

Figure 9. Intervention model: Rolling 7-day forecasts of daily pedestrian data, made each week, using ARIMA models with a dummy variable indicating when the lockdown periods occurred. Many people were in self-imposed lockdown before the start of the first lockdown period, and were reluctant to return 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.

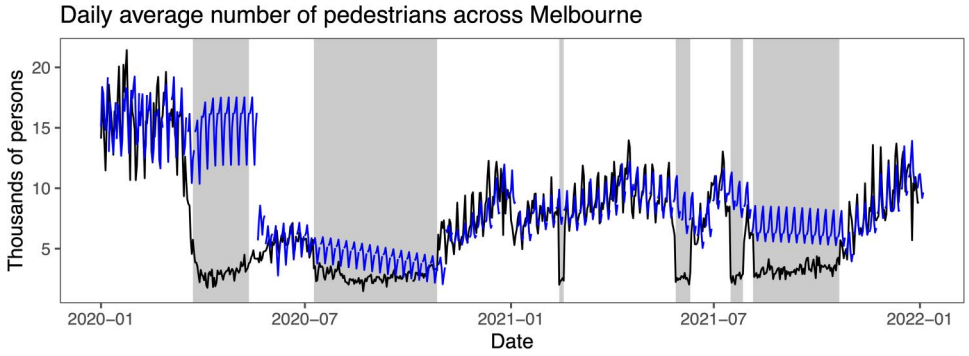

Figure 10. Set to missing: Rolling 7-day forecasts of daily pedestrian data, made each week, using ARIMA models, where the lockdown periods are set to missing. The ARIMA models use the remaining data to produce forecasts. Thus, the forecasts are based on only previous non-lockdown data. Naturally, the forecasts during the lockdown periods do not reflect reality, but they can be interpreted as what might have been.

to recognise that similar challenges could arise in circumstances other than COVID-19. Time series can be disrupted by various factors, such as natural disasters, policy changes, price fluctuations, definition changes, sensor failures in the IoT (Internet of Things), maintenance, and breakdowns in factories, just to mention a few.

The aim of this study is not to compile a comprehensive list of forecasting models applicable to interrupted time series data. Rather, we intend to propose overarching strategies to address such situations, with each strategy allowing for a variety of forecasting methodologies. As we provide publicly accessible R code, readers are encouraged to adapt the study by employing other forecasting methods tailored to their specific contexts.

The proposed strategies to forecast interrupted time series offer various advantages and disadvantages. Highly adaptive models offer swift adjustments to interruptions with ease of implementation and speedy computation. However, this simplicity comes at the expense of wider prediction intervals

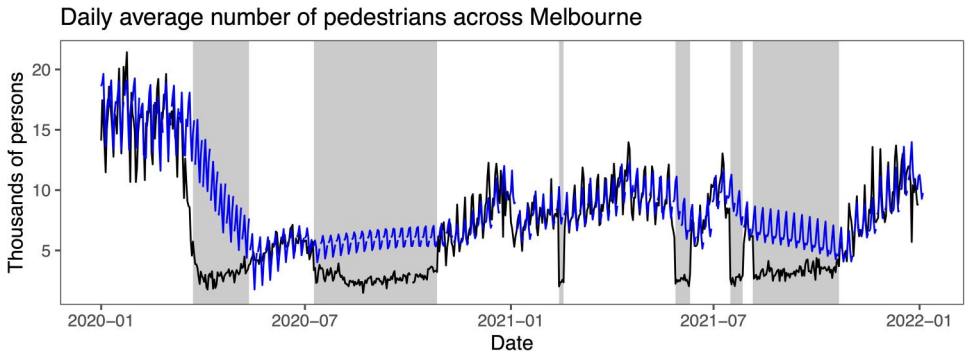

Figure 11. Estimate what might have been: Rolling 7-day forecasts of daily pedestrian data are made each week. The lockdown periods are replaced with estimated counts based on an ARIMA model applied to the remaining data. Then a new ARIMA model is fitted to each data set and used to produce forecasts. As a result, the forecasts during the lockdown periods are not true forecasts, as they use non-lockdown data from the future.

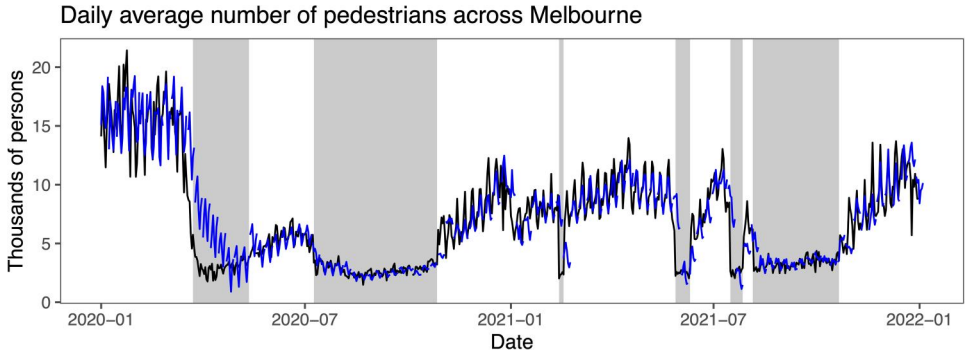

Figure 12. Ensemble forecasts combining the first two solutions shown in Figures 8 and 9. Other than having some difficulty around the period of the first lockdown, the results are relatively good.

and the loss of historical pattern memory, limiting their effectiveness in capturing pre-interruption dynamics. Intervention models explicitly account for disruptions, retaining past patterns, yet may yield overly narrow prediction intervals by assuming post-interruption similarity without proper calibration. Setting interrupted periods to missing values preserves data integrity but may result in inaccurate forecasts during interruptions. Estimating what might have been provides flexibility but may underestimate uncertainty. Downweighting the interruption period balances pre- and post-disruption data but may necessitate additional implementation complexity. Ensemble models combine various approaches for enhanced accuracy but may introduce challenges in interpretation.

When determining a suitable approach to forecasting interrupted time series, several factors should be considered. To begin, the complexity of the interruption itself is critical to approach selection. Intervention models are a suitable approach for interruptions with known features since they can represent the disruption explicitly. In contrast, highly adaptable models give a simpler option for fast adaptations to interruptions without explicit modelling, which is especially useful when the time or effect of the interruption is unknown. Balancing simplicity and flexibility is critical to ensuring that the selected technique is consistent with the features of the interruption and the required forecasting accuracy. Further, the width of prediction intervals should be carefully assessed to determine forecast uncertainty, particularly during and after interruptions. Finally, for greater robustness and accuracy, investigating ensemble models that integrate various forecasting methods might be beneficial.

Moreover, in determining which strategy to apply in each case, it is important to think through what is being assumed and what data are being used. This is a qualitative and subjective process and cannot simply be tackled with a quantitative comparison of forecast accuracy. The strategies we discuss have different ways of handling the disrupted period, and many of them do not seek to provide accurate forecasts during the disruption. For example, the "set to missing" strategy ignores the disrupted period, while the "estimate what might have been" strategy replaces the observations during the disruption with interpolations. In both cases, it does not make sense to evaluate the performance of the method during the disruption. Instead, the analyst needs to decide what approach makes the most sense for the application at hand.

This study proposes multiple viable strategies for forecasting time-series data interrupted by different types of events. Future research could focus on

exploring new strategies for handling such disrupted time series or enhancing probabilistic forecasting methods that are designed specifically for such data. Large-scale empirical research on further exploring which strategy may work best in different scenarios also provides useful insights. Comparative studies that assess the accuracy, robustness, and computing efficiency of different techniques across several datasets and disruption situations can help discover the most effective approaches for a range of practical applications.

5. Reproducibility

To enhance reproducibility and facilitate the adoption of proposed strategies, we provide the data and R code, as well as the entire paper written in R using Quarto and the targets package for R (Landau, 2021). All materials required to reproduce this paper is accessible *via* the GitHub repository.

Acknowledgment

We thank Dr. Stephan Kolassa and four anonymous referees for providing feedback and suggestions on the first version of our paper. These comments have helped us to refine and enhance our work.

Disclosure statement

No potential competing interest was reported by the authors.

Funding

No funding was received.

References

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., & O'Hara-Wild, M. (2023). Probabilistic forecasts using expert judgment: The road to recovery from covid-19. *Journal of Travel Research*, *62*(1), 233–258. https://doi.org/10.1177/00472875211059240

Aue, A., & Horváth, L. (2013). Structural breaks in time series. *Journal of Time Series Analysis*, *34*(1), 1–16. https://doi.org/10.1111/j.1467-9892.2012.00819.x

Australian Bureau of Statistics (2024). *Overseas arrivals and departures, Australia.*

Bell, P. C. (2000). Forecasting demand variation when there are stockouts. *Journal of the Operational Research Society*, *51*(3), 358–363. https://doi.org/10.1057/palgrave.jors.2600877

Bernal, J. L., Cummins, S., & Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: A tutorial. *International Journal of Epidemiology*, *46*(1), 348–355.

Blázquez-García, A., Conde, A., Mori, U., & Lozano, J. A. (2021). A review on outlier/anomaly detection in time series data. *ACM Computing Surveys*, *54*(3), 1–33. https://doi.org/10.1145/3444690

Box, G. E., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, *70*(349), 70–79. https://doi.org/10.1080/01621459.1975.10480264

City of Melbourne (2024). *Pedestrian counting system.*

Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M., & Priesemann, V. (2020). Inferring change points in the spread of covid-19 reveals the effectiveness of interventions. *Science*, *369*(6500), eabb9789. https://doi.org/10.1126/science.abb9789

Fildes, R., Nikolopoulos, K., Crone, S. F., & Syntetos, A. A. (2008). Forecasting and operational research: A review. *Journal of the Operational Research Society*, *59*(9), 1150–1172. https://doi.org/10.1057/palgrave.jors.2602597

Gardner, E. S. Jr, (2006). Exponential smoothing: The state of the art—part ii. *International Journal of Forecasting*, *22*(4), 637–666. https://doi.org/10.1016/j.ijforecast.2006.03.005

Harvey, A. (2006). Forecasting with unobserved components time series models. *Handbook of Economic Forecasting*, *1*, 327–412.

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice. 3rd ed. OTexts.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, *18*(3), 439–454. https://doi.org/10.1016/S0169-2070(01)00110-8

Khoshgoftaar, T. M., Golawala, M., & Van Hulse, J. (2007). *An empirical study of learning from imbalanced data using random forest* [Paper presentation]. 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007) (vol. 2, pp. 310–317). IEEE.

Koop, G., & Potter, S. M. (2007). Estimation and forecasting in models with multiple breaks. *Review of Economic Studies*, *74*(3), 763–789. https://doi.org/10.1111/j.1467-937X.2007.00436.x

Landau, W. M. (2021). The targets r package: A dynamic make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. *Journal of Open Source Software*, *6*(57), 2959. https://doi.org/10.21105/joss.02959

Lichtendahl, K. C., Jr, Grushka-Cockayne, Y., & Winkler, R. L. (2013). Is it better to average probabilities or quantiles? *Management Science*, *59*(7), 1594–1611. https://doi.org/10.1287/mnsc.1120.1667

McDowall, D., McCleary, R., & Bartos, B. J. (2019). *Interrupted time series analysis*. Oxford University Press.

Talagala, P. D., Hyndman, R. J., Smith-Miles, K., Kandanaarachchi, S., & Muñoz, M. A. (2020). Anomaly detection in streaming nonstationary temporal data. *Journal of Computational and Graphical Statistics*, *29*(1), 13–27. https://doi.org/10.1080/10618600.2019.1617160

Talih, M., & Hengartner, N. (2005). Structural learning with time-varying components: Tracking the cross-section of financial time series. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(3), 321–341. https://doi.org/10.1111/j.1467-9868.2005.00504.x

Tang, X., Yao, H., Sun, Y., Aggarwal, C., Mitra, P., & Wang, S. (2020). Joint modeling of local and global temporal dynamics for multivariate time series forecasting with missing values. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(04), 5956–5963. https://doi.org/10.1609/aaai.v34i04.6056

Trapero, J. R., de Frutos, E. H., & Pedregal, D. J. (2024). Demand forecasting under lost sales stock policies. *International Journal of Forecasting*.

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, *167*, 107299. https://doi.org/10.1016/j.sigpro.2019.107299

Twala, B. E., Jones, M., & Hand, D. J. (2008). Good methods for coping with missing data in decision trees. *Pattern Recognition Letters*, *29*(7), 950–956. https://doi.org/10.1016/j.patrec.2008.01.010

Wang, X., Hyndman, R. J., Li, F., & Kang, Y. (2023). Forecast combinations: An over 50-year review. *International Journal of Forecasting*, *39*(4), 1518–1547. https://doi.org/10.1016/j.ijforecast.2022.11.005

Wikipedia (2024). *COVID-19 pandemic in Victoria*.

Wu, S.-F., Chang, C.-Y., & Lee, S.-J. (2015). *Time series forecasting with missing values* [Paper presentation]. 2015 1st International Conference on Industrial Networks and Intelligent Systems (INISCom) (pp. 151–156). https://doi.org/10.4108/icst.iniscom.2015.258269