Suicide Mortality in Australia: Estimating and Projecting Monthly Variation and Trends From 2007 to 2018 and Beyond

Nicholas Biddle[1]; Lucy Ellen[1]; Rosemary Korda[2]; and Karuna Reddy[1]

1. Centre for Social Research and Methods, Australian National University
2. National Centre for Epidemiology and Population Health, Australian National University

Author Note

Nicholas Biddle https://orcid.org/0000-0002-4765-4445

Lucy Ellen https://orcid.org/0000-0002-6010-2901

Rosemary Korda https://orcid.org/0000-0002-9390-2171

Karuna Reddy https://orcid.org/0000-0002-9467-4668

Correspondence concerning this report should be addressed to Professor Nicholas Biddle.

Email: Nicholas.Biddle@anu.edu.au

Abstract

Suicide is a leading cause of death for a number of age groups and populations in Australia, and of important public health and policy interest. Accurately identifying trends in suicide for Australia and for smaller population sub-groups is vital for tracking our health as a nation, the effectiveness of our current policy frameworks, and the impact of a range of external shocks. Three somewhat exogenous factors that may have impacted on the level of suicide in Australia or impact on suicide into the future are the recent droughts in many parts of the country; the catastrophic 2019/20 Black Summer bushfires; and the ongoing SARS-CoV-2 or COVID-19 pandemic. Identifying whether those or similar external shocks have impacted on suicide in Australia requires an estimate of what suicide rates in Australia would have been in the absence of such shocks (the counterfactual), and then a measure of whether the observed suicide rates are in fact different to what we might otherwise have expected. This is made difficult by the apparently random variation in suicide on a daily, weekly and monthly basis (or at the very least variation for which the cause is unobservable) and other potential factors that may be impacting on suicide.

The present study investigates monthly deaths by suicide in Australia from 2007-2018, with the aim of identifying trends in suicide over the period and whether there are any consistent monthly patterns. These monthly deaths are analysed separately for males and females, and by State/Territory. The mortality data on suicide and intentional self-harm have been documented and provided by the Australian Bureau of Statistics (ABS) and the Australian Institute of Health and Welfare (AIHW). Apart from exploring and identifying the overall trends and seasonality in the death rates per 100,000 persons, several best-fitting time series models are constructed in order to forecast the number of deaths beyond the observed time period. After checking the forecast accuracies of a few plausible models, a final satisfactory model was chosen for forecasting purposes.

Keywords

Suicide; Australia; Projection; Forecasts

Acknowledgements

This work was commissioned and funded by the Australian Institute of Health and Welfare (AIHW). This work was undertaken in support of the Australian Government’s Suicide and Self-Harm Monitoring System initiative. Data used for this study were provided by the AIHW and we acknowledge colleagues at the AIHW for their support in fulfilling our data requests. We further acknowledge the support of colleagues at the Centre for Social Research and Methods, Australian National University for their peer review. We have no known conflicts to report.

1 The problem of suicide

According to the World Health Organisation (WHO, 2020), the global age-standardized suicide rate for 2016 was 10.5 per 100 000 persons. More than half (52.1%) of these deaths occurred before the deceased reached 45 years of age, and suicide was the second leading cause of death globally among young people aged 15 to 29 year. Furthermore, as a result of suicide deaths during 2016, an estimated 34.6 million years of life were lost (Naghavi, 2019; WHO, 2020).

Turning to Australia, 3,046 deaths were registered as caused by intentional self-harm during 2018; the most recent year for which data are currently available (Australian Bureau of Statistics [ABS], 2019). Death by intentional self-harm (or suicide) was the 14th most common cause of deaths registered in Australia during this period. Significantly, the median age for those dying by suicide (44.4 years of age), was 27.3 years younger than any of the other top 20 causes of death in Australia during 2018. For deaths registered in 2018, suicide was the leading cause among Australians aged 15 to 44 years, and the second leading cause of death for those 45 to 54-year-old (ABS, 2019). On average, Australians dying by suicide in 2018, lost 36.7 years of life. A total of 105,730 years of life were lost to suicide in Australia during the period (ABS, 2019).

In addition to the human cost, suicide places a very real monetary burden on the Australian economy. The Productivity Commission (2019) estimated that suicide and suicide attempts cost the Australian economy between 16 and 34 billion Australian dollars during the 2018-2019 financial year. It is clear that suicide poses significant public health and economic problems globally and for Australia.

1.1 Suicide and Australian Communities

While suicide presents challenges for the whole Australian community; it is also the case that some Australian communities and individuals are more vulnerable than others to the impacts of suicide. An empirical understanding of who in Australia is more or less vulnerable to death by intentional self-harm continues to be developed. Some of the factors, which have been investigated (to a greater or lesser extent) and may contribute to suicide vulnerability or resilience include: sex, Aboriginal and Torres Strait Islander status, geospatial location, socio-economic status, level of education, social isolation or loneliness, and health service use. It is not in the scope of the current study to measure these factors. However, this study is designed with these factors in mind, with a view to their incorporation within future iterations of our work. This study aims to develop a robust preliminary model, to which further explanatory variables can subsequently be added towards a more detailed understanding of suicide forecasting.

It is well established that males have a substantially higher rate of suicide compared to females in Australia and globally (ABS, 2019; WHO, 2020). In Australia, more than three-quarters of the deaths registered by suicide in 2018 (76%) were male deaths (Australian Institute of Health and Welfare [AIHW], 2020). Even so, females are more likely to be hospitalised due to intentional self-harm (AIHW, 2020). Males tend to make more lethal suicide attempts, and this likely explains the higher death and lower hospitalisation rates when compared to females (Choo et al., 2017). A significant disparity in suicide death rates between Indigenous and non-indigenous Australians has also been well established. Aboriginal and Torres Strait Islander Australians were almost twice as likely to die by suicide compared to non-indigenous Australians, between 2014 and 2018. These higher rates of

Aboriginal and Torres Strait Islander deaths were concentrated within the age groups 15-24, 25-34, and 35-44 years (ABS, 2019). This significant disparity in suicide rates between Indigenous and non-indigenous Australians extends well prior to the 2014 data reported here (for example see: De Leo et al. 2011).

There may be geospatial differences in suicide rates; particularly between metropolitan, rural, and remote areas within Australia. However, the current evidence in support of these differences is somewhat mixed. Qi et al. (2010) and Cheung et al. (2012) found stronger evidence for the clustering of suicide deaths in more remote parts of Australia. Whereas, Torok et al. (2017) examined both fatal and non-fatal intentional self-harm within New South Wales and found a tendency for suicide clusters to occur within metropolitan and coastal regions rather than rural areas. Too et al. (2017) also examined both fatal and non-fatal intentional self-harm data, though for Western Australia. Too and colleagues found evidence for two WA suicide clusters, both in remote locations; and that approximately 64% of suicide attempt clusters identified were in urban or regional areas (Too et al., 2017). Finally, Robinson et al. (2016) conducted a national study to assess spatial suicide clusters for young people (24yrs or younger) and adults (25yrs or older). Three of the five youth suicide clusters found occurred in remote areas, though most identified adult clusters occurred in major cities (Robinson et al., 2016). While mixed, findings suggest geographical variation is an important consideration when determining which Australian communities may be more or less vulnerable to suicidal behaviours. It is possible that differences in suicide rate found by geography are, at least in part, accounted for by associated variations in socioeconomic status.

In Australia, those living in lower socioeconomic status communities may be at higher risk of dying by suicide. In addition, residing within a higher socioeconomic status community might serve as a protective factor against suicide (AIHW: Henley & Harrison, 2019; Kõlves et al., 2015; Too et al., 2018). The observed effect of community-level socioeconomic status on risk of suicide is likely more pronounced for males (as compared to females) (AIHW: Henley & Harrison, 2019; Kõlves et al., 2015; Too et al., 2018). Furthermore, the community level socioeconomic inequalities of suicide risk appear to have widened (AIWH: Henley & Harrison, 2019; Too et al., 2018). Too et al. (2018) examined suicide data for deaths from 1979 till 2013 and found growing social-economic inequality in suicide was primarily associated with declines in suicide rates for high socioeconomic status areas. Though, they also found, increasing rates of suicide for low socioeconomic status areas was linked to growing socioeconomic inequality in suicide for older males. Somewhat incongruously, AIHW: Henley & Harrison (2019) found that the suicide deaths rate for our least socioeconomically advantaged communities had increased from 11.4 deaths per 100,000 in 2009–10 to 14.6 deaths per 100,000 in 2015–16. This rise in the modelled suicide death rate averaged 3.5% per year and was statistically significant. They found little change in death rate, over the same period, for our most socioeconomically advantaged communities (AIHW: Henley & Harrison, 2019).

Interrogations of American data have underscored the significance of a college-level education on risk of suicide (Case & Deaton, 2017; Phillips & Hempstead, 2020). One recent study concluded that in 2014 American men with a high school education were twice as likely to die by suicide compared with those with a college degree (Phillips & Hempstead, 2020). The authors are not aware of any similar research having been published within an Australian context, mainly due to a historic lack of mortality data with individual-level socioeconomic outcomes. There is, however, ongoing work being undertaken as part of an NHMRC Partnership Grant.[1] Initial published findings (Korda et al. 2020) have demonstrated a strong link at the individual level between education and mortality in Australia, with additional preliminary analysis showing that for men suicide had the highest number of excess deaths due to education-related inequalities amongst the 25 to 44 year old population, and the third highest number for those aged 45 to 64 years.

Social isolation, loneliness, and related constructs (e.g. social connectedness) are also potentially important. A recent international systematic narrative review found that both the condition of being alone and the subjective feeling of being alone were associated with suicide attempts and suicidal ideation (Calati et al., 2019). The authors are not aware of any population or community level research investigating any possible connection between social isolation or loneliness and risk of suicide death.

Economists Anne Case and Angus Deaton (2020) have coined the term 'deaths of despair', referring to deaths from suicide, drug overdose, and alcohol-related harm. These deaths, the authors argue, are not primarily due to absolute economic outcomes, which have improved or at least remained steady for the vast majority of people in the developed and developing world over recent decades (also documented by Angus Deaton in an earlier (2013) book – The Great Escape). Rather, these deaths of despair are said to be caused by an absence of hope in communities and for individuals, and a deepening sense that current economic and societal structures are not resulting in a fair distribution of resources and status. Case and Deaton (2020) focus their analysis mainly on the United States. Indeed, they argue that the economic system of the USA is one of the primary causes of deaths of despair, and that other developed countries have avoided such high and worsening mortality rates due to a better balance between the role of the market and the state.

Finally, understanding patterns of contact with health services prior to death by suicide may be important towards identifying Australian communities and individuals who are more vulnerable. A recent systematic literature review, undertaken by Stene-Larsen and Reneflot (2019), concluded that accessing primary health care prior to suicide is common even in the final month before death. This finding perhaps highlights an opportunity for suicide prevention within primary healthcare settings. However, a more nuanced investigation of health service use patterns, within an Australian context, is needed to provide more meaningful information. Towards this end, Chitty et al. (2020) have protocolled an Australian population-based case series study designed to describe: health service utilisation in the year prior to death by suicide, and prescribed medicines use in the year prior to death by suicide.

Hospitalised self-harm or psychiatric hospitalisation with a history of self-harming behaviour has been recognised as a risk factor for suicide death (Bickley et al., 2013; Hunt et al., 2009; Mitchell & Cameron, 2017; Singhal et al., 2014). The authors are aware of one study investigating service use and mortality of Australian hospitalised self-harm patients, with a matched population comparison group (Mitchell & Cameron, 2017). Mitchell and Cameron (2017) found that individuals hospitalised for intentional self-harm experienced higher hospital-based health service use in the 12 months pre and post the index admission and a higher suicide rate compared to their matched non-injured counterparts. Service use due to non-fatal intentional self-harm may be important to understanding vulnerability to suicide.

[1] https://researchdata.edu.au/whole-of-population-care-australia/1347953

1.2 COVID-19 and Bushfires: Recent challenges to the resilience of Australian Communities

At the time of writing, Australia had not yet recovered from a severe drought and the most devastating summer of bushfires (in terms of hectares burned in close proximity to urban centres) in recorded history. As far as the authors are aware, there is no data available on the effects of suicide due to these two external shocks. However, social survey data (reported in Biddle, Edwards, et al., 2020) showed that more than half (53.6 per cent) of adult Australians felt anxious and worried due to the bushfires.

Following on from the devastating bushfires, Australia is currently experiencing its worse public health crisis since at least 1918/19 and its worst economic downturn since the Great Depression in the 1930s. The spread of SARS-CoV-2 and the associated condition COVID-19 has led to significant social, economic and health impacts. At the time of writing (August 28th), there were 25,451 confirmed cases and 584 confirmed deaths from COVID-19. After a substantial decline in May and June, rates of both were higher than they had ever been in early to mid-August, particularly in Melbourne and parts of regional Victoria, but had begun to decline again in late August.

There has been speculation that the COVID-19 pandemic and associated economic and social dislocation will lead to higher rates of suicide in Australia (Deady et al., 2020). Having never experienced a crisis like this in modern times, it is hard to know for sure whether this will be the case as the effect of individual or even community economic and social decline is not necessarily a good indication of the effect of national or international economic, social and public health shocks. Survey data has, however, demonstrated an increase in some of the factors that predict suicide, with Biddle et al. (2020a) showing that the Kessler-6 measure of psychological distress had increased significantly and substantially between February 2017 and April/May 2020, including for a longitudinal sample of respondents. Furthermore, the increase in psychological distress was much greater for young Australians with older Australians experiencing steady, or even slightly improving mental health on this measure. More directly, Biddle et al. (2020b) showed that in April 2020 almost exactly two-thirds (66.6 per cent) of Australians report that they ‘felt anxious or worried for the safety of yourself, close family members or friends, due to COVID-19.’

There are, however, some measures that may be predictive of suicide that have declined for certain population groups, potentially mitigating or counteracting the more negative effects of COVID-19. Biddle et al. (2020b) showed that people were more confident in a range of institutions, more satisfied with the direction of the country, and more likely to trust those in the community. Biddle et al. (2020c) also showed that people were more likely to report that their consumption of alcohol and other illicit substances had declined during the initial months of the spread of COVID-19 than say that it had increased (though there is significant uncertainty around this change, as other estimates in the paper suggest a slight increase in consumption).

The only publicly available data at the moment suggests that suicide rates in Australia have been steady, at least in the early part of the COVID-19 period. The Coroners Court of Victoria (2020) reported 466 deaths by suicide between January 1st and August 26th 2020, roughly equivalent to the 468 deaths over the same period in 2019 and 461 deaths over the same period in 2018 (despite a slightly larger population).

Capturing the actual impact of these external shocks on suicide in Australia requires accurate and sufficiently disaggregated data. If increases (or decreases) in suicide for a given month, year, or population group are within the range of observed fluctuations in the absence of such shocks, then it is more difficult to claim that any changes are due to observed factors.

1.3 The importance of suicide monitoring and forecasting

Monitoring suicide deaths facilitates a more comprehensive understanding of Australia's community level health status and needs. This, in turn, has the potential to inform public health and broader service planning. Accurate suicide monitoring also enables the efficacy of suicide prevention and intervention initiatives to be assessed. Suicide monitoring further provides opportunities to measure the impact of large events such as the recent Black Summer bushfires and the COVID-19 pandemic.

Forecasting suicide deaths, at a meaningful aggregate level, provides additional potential benefit. Primarily, a comparison of observed data with forecast data allows for an assessment as to whether the number of deaths observed is more or less than what might be expected given historical data. If robustly and carefully implemented, improved capabilities in these areas have the potential to profoundly and positively impact Australian individuals and communities

1.4 The Australian Government National Suicide and Self-Harm Monitoring Project

The Australian Government held a National Suicide Prevention Summit in December 2018 and subsequently committed to prioritising suicide prevention as a whole-of-government issue and a Council of Australian Governments (COAG) priority (NMHC, 2019). This commitment included both strengthening the delivery of suicide prevention interventions and establishing a national system for the timely collection and communication of self-harm and suicide data. Subsequently, the National Mental Health Commission (NMHC), the Australian Institute of Health and Welfare (AIHW), and the Department of Health have collaborated to progress the National Suicide and Self-Harm Monitoring Project (NMHC, 2019). The current study is being undertaken, independently by the authors, towards supporting the AIHW in their work towards the National Suicide and Self-Harm Monitoring Project.

1.5 Aims and overview of the present study

The focus of this paper is explaining the observed patterns of suicide in Australia, as well as what the rates of suicide might be if these patterns continue. With the increased availability of (hopefully) more accurate data on suicide (as outlined below), as well as a real policy need to understand changes in as close to real-time as possible, it is important to better understand these patterns so any new data that does become available can be compared to existing trends and variation.

Our specific aims are to make the best use of currently available Australian suicide deaths data from 2007 – 2018 inclusive to:

1. Investigate possible trends and seasonality in suicide death rates;
2. Train and test various forecasting models to robustly predict suicide death rates; and
3. Use the most robust forecasting model to predict suicide death rates beyond currently available Australian suicide deaths data.

With these aims in mind, the remainder of the paper is structured as follows. In Section 2, we outline the data and methods used in this paper, including definitions and data sources. In

Section 3 we present our descriptive analysis, whereas in Section 4 we present our forecasting model (including an analysis of the performance of five different, commonly used models). Section 5 provides some concluding comments and discussion of future work.

2 Methods, definition, production, and reporting of Australian suicide statistics

2.1 Definition of suicide and production of Australian suicide statistics

Deaths by suicide in Australia are currently coded, by the Australian Bureau of Statistics, to International Classification of Disease 10th Revision (ICD-10, 2004) codes X60-X84 and Y87.0. These codes stipulate that a person intentionally harmed themself and that death resulted from this action. Accordingly, this includes individuals who intended to end their lives, as well as those who intended to harm themselves but did not necessarily intend to die. For the current study, authors rely on ABS coding to these ICD-10 codes. Subsequently, suicide will be defined as deaths determined to have resulted from intentional self-harm; regardless of whether the deceased intended to die.

The production of Australian suicide statistics involves the collaboration of different jurisdictions and organisations. The legal requirements governing whether a death is reportable to the coroner vary between Australian state and territory jurisdictions. Nonetheless, where a death may be by suicide, the process is such that it should generally be reported to the local coroner for investigation and determination (ABS, 2019; AIHW: Harrison et al., 2009).

Each jurisdiction, again, has different (though similar) legislation governing coronial inquiries and reporting. Once a case has been opened with the coroner, preliminary details will be recorded at the level of the state or territory jurisdiction using a local case management system. Each jurisdiction then later enters information to the National Coroners Information System (NCIS). Information may be uploaded, by the local jurisdiction, to the NCIS as the coronial enquiry progresses or only at the completion of the investigation (ABS, 2019; AIHW: Harrison et al., 2009).

Australian Bureau of Statistics (ABS) officers assess the NICS and code information for the cause of death according to the ICD-10. The current coding and data revisions processes applied by the ABS have been applied to deaths registered after 1 January 2007. Where a coronial enquiry remains open on the NCIS or where the coroner has not made a determination of intentional self-harm or assault and the mechanism of death indicates a possible suicide; ABS coders assess available data as to whether there is sufficient evidence to indicate the death was a suicide. For deaths registered prior to 1 January 2007, coding rules required a coroner determination around intent before a death could be coded as a suicide. However, for a range of reasons, in some instances, the coroner does not make a finding on intent. As such, these new coding guidelines were designed to improve data quality (ABS, 2012; AIHW: Harrison et al., 2009).

Currently, the ABS aims to release Preliminary data to the public 15 months after the end of the reference period. The ABS then reviews coronial cases for updated information on the NCIS; producing a Revised data 12 months after the initial processing, and Final data 24 months after initial processing (ABS, 2019; AIHW: Harrison et al., 2009). Reference dates and the revisions processes are undertaken according to the year in which the death was registered (not necessarily the year in which the death occurred). This revisions process was instigated for deaths registered from 1 January 2007. Prior to this, deaths were processed only once by the ABS (ABS, 2019; AIHW: Harrison et al., 2009).

After consideration of probable impacts to the data as a result of both improvements made to coding practices and the implementation of the revisions, this study will include data for deaths registered from 1 January 2007 onwards only.

The ABS's annual Causes of Death publications, with the exception of a single available download 'Causes of Death by Year of Occurrence (Australia)', present all deaths (including those resultant from intentional self-harm) by the year in which the death was registered; not necessarily the year in which the death occurred. Year of registration indicates the collection cycle in which a record was included in ABS counts. This includes all deaths registered in Australia and received by the ABS prior to the end of the March quarter the subsequent year; and deaths registered prior to the reference year, but not previously received by the ABS (ABS, 2019). The current study will utilise date of death information; not year of death registration (except as it is relevant to data quality considerations due to the data revisions processes). Date of death provides a more accurate information upon which to monitoring suicide and forecast future suicide deaths, though it does create challenges for the most recent years of data (which we address below).

2.2 Method of analysis

2.2.1 Sample

The sample included individuals deceased in Australia as a result of intentional self –harm, where the individuals' died and their death was registered between the years 2007 and 2018 inclusive.

All data were extracted from the National Mortality Database by the Australian Institute of Health and Welfare (AIHW) and provided to the authors. Suicide deaths were defined as those, by the ABS using information available from the NCIS, to ICD-10 codes X60-X84 and Y87.0. National monthly suicide deaths count data were provided for each month where the date of death and the year of registration fell between 2007 and 2018 inclusive. Data were disaggregated by state/territory of usual residence and by sex.

In order to adjust for delays in the registration of deaths, we estimated an adjustment factor based on historical data. Specifically, we estimated the proportion of deaths that occurred in a given month for each year from 2007 to 2016 that were registered in the subsequent year, or the subsequent two years. We then averaged these proportions across the sample, and took the reciprocal of these proportions as the adjustment factor. For 2017 data, we use an estimate of the proportion of deaths that are typically registered in two or more years after the given month of death as the adjustment factor. For 2018 data, we use an estimate of the proportion of deaths that were registered in one year or more after the given month of death as the adjustment factor. Table 1 gives the one and two-year adjustment factors for each month.

Table 1 Adjustment factors for delays in death registration – By month

Month	Two-year adjustment (applied to 2017)	One-year adjustment (applied to 2018)
January	1.00157	1.00755
February	1.00145	1.00984
March	1.00262	1.01195
April	1.00297	1.01060
May	1.00486	1.01488
June	1.00097	1.01904
July	1.00285	1.02343
August	1.00380	1.03610
September	1.00529	1.05114
October	1.00458	1.07768
November	1.00646	1.15935
December	1.00697	3.15317

2.3 Analytic Strategy

As outlined in Table 2, no preliminary data were provided to the authors for death registered in 2008 or 2009; no revised data were provided for deaths registered in 2007 or 2008. No further data were reported as missing. Nonetheless, death count data for December 2018 were excluded from the data set. Although we applied the above adjustment factors to this data point, the estimated value was still substantially lower than what would have been expected based on previous observations for December. This may have been due to actual variation in suicide, but we could not rule out particularly long delays in reporting. Data analyses were undertaken using "R" and the 'forecasting' package (Hyndman & Athanasopoulos, 2018) or STATA.

After adjusting for delays in registration, a time-series plot was generated for Australian monthly suicide death counts, including all available Preliminary, Revised, and Final suicide death count data versions separately. A visual analysis (provided below) was used to determine whether Preliminary and Revised death count data are sufficiently similar to the Final version, and therefore useable for further analysis.

2.4 Calculation of monthly suicide death rates per 100, 000 persons

The Australian population has grown over the time period under study and, all else being equal, this will lead to an increase in the number of deaths for a given year. There are two ways in which this can be taken into account in the analysis. The first option is to convert the counts of the number of deaths into rates, by dividing by a relevant population estimate. This creates a continuous variable that can be modelled using relevant techniques (described below).

The second alternative with a changing population is to keep the data as counts and model using either a Poisson or Negative Binomial (if there is over-dispersion) with an exposure variable. An exposure variable in these models indicates the number of times an event could have happened (i.e. the total population at risk) and is transformed into a natural-log, and then incorporated in the model with the coefficient restricted to be equal to 1. In this paper, we focus on the use of rates, as modelling rates is less reliant on distributional assumptions. However, we test the sensitivity of this choice by also estimating a Negative Binomial model. Regardless of the model chosen, predictions can be converted to and from rates and counts at the post-estimation stage.

The ABS publishes quarterly Australian resident population estimates; nationally for each state and territory. Population estimates are based on the results of the most recent Australian Census of Population and Housing, with adjustments to account for interim natural increase, overseas migration, and interstate migration. Population estimates for all quarters commencing January 2007 to January 2019 were retrieved from the ABS website, by state/territory and by sex.

To account for population growth and create equivalence across state/territory jurisdictions (of different-sized populations) monthly death rates per 100,000 persons were computed. Death rates were calculated as DR = (# deaths/population) x 100,000; where '# deaths' are monthly suicide death counts and 'population' is a monthly population estimate generated through linear interpolation of the relevant quarterly ABS population estimates (ABS, 2020). Linear Interpolation was chosen because quarterly population estimates appeared to have a linear growth within the periods. Monthly population estimates and death rates were calculated separately for each state/territory jurisdiction, nationally, and nationally by sex.

It should be noted that we do not adjust for the changing age structure for the Australian population when constructing rates through time. In our paper, this will be captured by the estimated trend component (either positive or negative, depending on whether more or fewer Australians are alive in the age brackets with the highest suicide rates). Changes in age structure can be captured using age-specific suicide rates, which is beyond the scope of this paper.

2.5 Investigate trend and seasonality within monthly suicide death rates data

Firstly, a visual investigation of possible seasonality or trends within monthly suicide death rates data was undertaken. To allow visual investigation of any trends in monthly suicide death rates per 100,000 persons for the period January 2007 to November 2018 inclusive; time series plots with overlayed locally weighted scatterplot smoothing ([Loess] Cleveland and Devlin 1988; Jacoby, 2000) curve were generated. Loess curves provide a graphical summary of the relationship between variables; without the requirement for prior specifications about the nature of the structure of any relationship (e.g. a linear relationship). Loess, therefore, allows for the exploration of relatively complex relationships between variables. Seasonal, polar seasonal, and seasonal subseries plots were drawn to visualize possible seasonality.

Secondly, Seasonal and Trend decomposition using Loess (STL) was used as a more formal means of identifying and isolating trend, seasonal, and remainder components of the monthly suicide death rates time series (for a detailed overview of the method see Cleveland et al. 1990). For the current study, the trend component refers to any significant pattern of increased or decreased rate of suicide deaths across the years 2007 to 2018. Seasonality refers to any significant pattern of increased or decreased rate of suicide deaths within the years 2007 to 2018 (i.e. are particular months associated with consistently higher or lower rates of death). The remainder is the component of the data not accounted for by neither trend nor seasonality. The STL decomposition model is robust to outliers and is designed for additive time series (where the time series data is equal to the sum of the trend, seasonal, and remainder components). Being able to handle additive and multiplicative time series, it is entirely automatic (i.e. data-driven) and tends to be highly robust to outliers and level shifts in the data. It would result in a stationary process that would be easier to understand, model, investigate and to predict, as the way they change becomes predictable. Once the decomposition had isolated the patterns in the series and extracted the three components of the time series, forecasting models were built and displayed.

Thirdly, two time series regression models were implemented as a further means of quantifying possible seasonal and trend components within the monthly suicide deaths rates time series. For both models, the dependent variable was the average daily suicide death rate by month per 100,000 persons. Both models included months of the year as dummy dependent variables and January as the base case. Model 2 also includes a single lagged suicide death rate dependent variable in order to test explicitly whether the suicide rate observed in the month prior to a given month had an association with that month's rate. These regression models allow for the estimation of robust parameters explaining any variance in the data over our observation period (January 2007 to November 2018).

2.6 Train, test and apply monthly suicide death rate forecasting models

The full January 2007 till November 2018 monthly death rate time series data were split into training and testing datasets at a ratio of 75:25. The training dataset, January 2007 till December 2015 was used to identify and fit various forecasting models. The testing dataset, January 2016 till November 2018, was used to test each of the forecast models. The R 'forecast' package (Hyndman & Athanasopoulos, 2018) was used to train and test five separate forecasting models. The R 'forecast' package is able to apply a range of automatic univariate forecasting models.

For each model, point forecasts for predicted monthly rates of suicide January 2016 till November 2018 were plotted against the observed values for those months. These plots allowed for a visual interpretation of performance for each model. Included with these plots were 80% and 95% prediction intervals, which capture an estimate of an interval within which the test observations will fall with the respective probabilities, given the observed range in the training data.

The accuracy of each model was also more formally assessed using the Mean absolute percentage error (MAPE). The MAPE measure captures the difference between values predicted by the forecast model and observed values, expressed in percentage terms (Hyndman & Koehler, 2006; Ord & Fildes, 2012). Diagnostics of residuals were then checked for models with the highest accuracy. Residuals are the variance remaining in the data after fitting a time series model. A strong forecasting model will yield residuals that are uncorrelated and have a zero mean. Further, the calculation of prediction intervals will be simpler if residuals are additionally normally distributed and have a constant variance. A time series plot of residuals, histogram of residuals, and auto-correlation plots were used to examine these attributes for the residuals of models with highest accuracy.

The most accurate and best-fitting model was used to forecast monthly suicide death rates for December 2018 till December 2020. Point estimates with 80% and 95% confidence intervals were generated for predicted values.

The following five time series models were trained and tested:

2.6.1 Exponential Smoothing State Space Model.

The exponential smoothing state space (ETS) Model (Gardner, 1985, 2006; Hyndman et al., 2008) allows for seasonal and/or trend component to be incorporated into the model. The model handles additive and multiplicative trend and/or season components.

2.6.2 Autoregressive Integrated Moving Average Model.

Using an autoregressive integrated moving average (ARIMA) model (Box et al., 2011; Brockwell & Davis, 2016; Pena et al., 2011) is necessary to capture the effects of autocorrelation in a time series dataset. The ARIMA model includes parameters for seasonality, trend, as well as autoregressive and/or moving average terms to account for any autocorrelation. ARIMA time series analysis assumes that the underlying time series is stationary or can be made stationary by differencing it 1 or more times. This is known as the ARIMA (p, d, q) model where d denotes the number of times a time series has to be differenced to make it stationary. The procedure that was used to estimate the ARIMA model automatically chooses the best model based on the performance metrics such as the AIC value.

2.6.3 Exponential Smoothing State Space Model with Box-Cox Transformation, ARMA Errors, Trend and Seasonal Components Model.

Exponential smoothing state space model with Box-Cox transformation, ARMA errors, trend and seasonal components (TBATS) Model (De Livera et al., 2011) model is designed to model complex seasonal patterns such as when there are multiple seasonal periods or cyclic patterns (e.g. daily, weekly monthly or yearly patterns) within a single time series.

2.6.4 Seasonal Decomposition of Time Series by Loess Model.

Seasonal decomposition of time series by Loess decomposes a time series into seasonal, trend and remained components using loess model (Bergmeir et al., 2016). Forecasts of STL objects are obtained by applying a non-seasonal forecasting method to the seasonally adjusted data and re-seasonalising using the last year of the seasonal component.

2.6.5 Neural Network Autoregression Model.

Neural network autoregression (NNAR) model (Crone et al., 2011), is a feed-forward neural networks model with a single hidden layer and lagged inputs for forecasting univariate time series. Artificial neural networks are forecasting methods based on simple mathematical models of the brain. They allow complex nonlinear relationships between variables. The NNAR model used has three components: p, P & k. p denotes the number of lagged values used as inputs, P denotes the number of seasonal lags, and k denotes the number of hidden nodes present.

3 Results – Describing monthly suicide rates

As shown in Figure 1, there appeared to be no meaningful differences between the National Preliminary, Revised, and Final monthly death count data versions for the time periods between January 2007 and November 2018 when such data was available. Consequently, the latest available data were used for all further analyses. The latest available data is preliminary for 2017 and 2018, revised for 2016, and final for 2007 to 2015 inclusive. Figure 2 presents the time series plot of all latest available monthly Australian suicide deaths count data, for the same period.

Figure 1 Available Preliminary, Revised, and Final Monthly Australian Suicide Death Count Data: January 2007 to November 2018

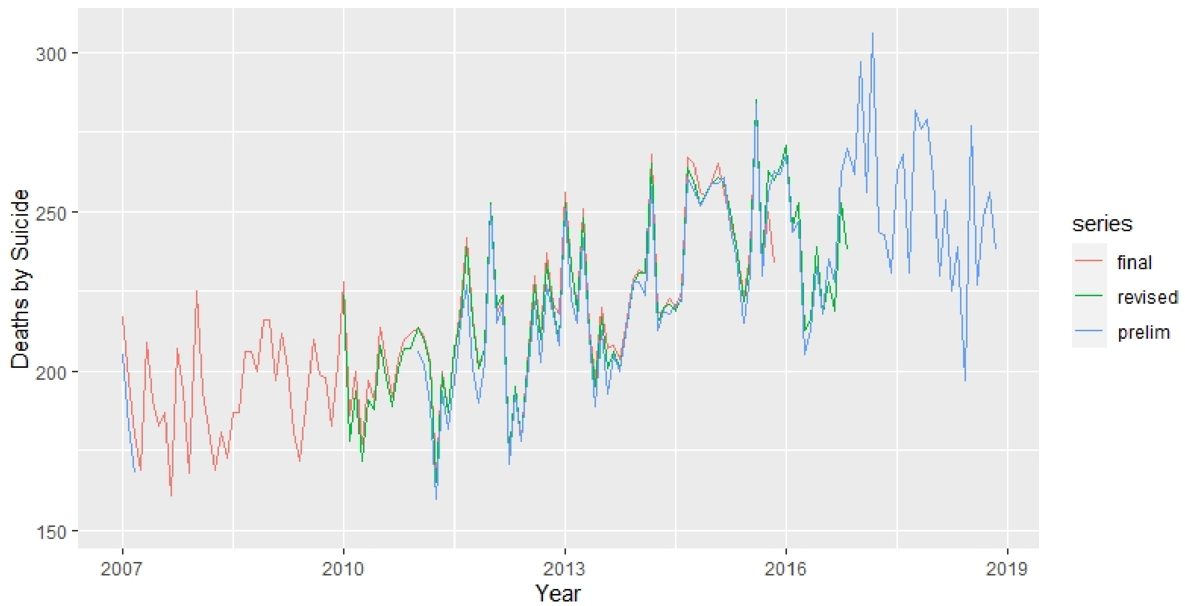
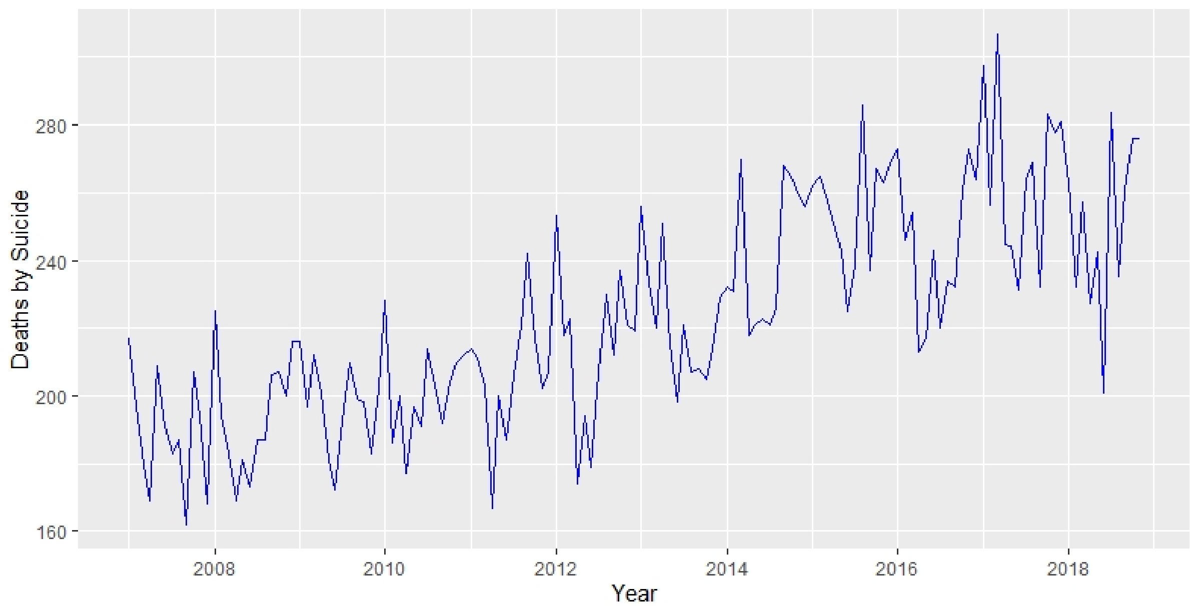


Figure 2 Monthly Counts of Suicide Deaths in Australia: January 2007 to November 2018



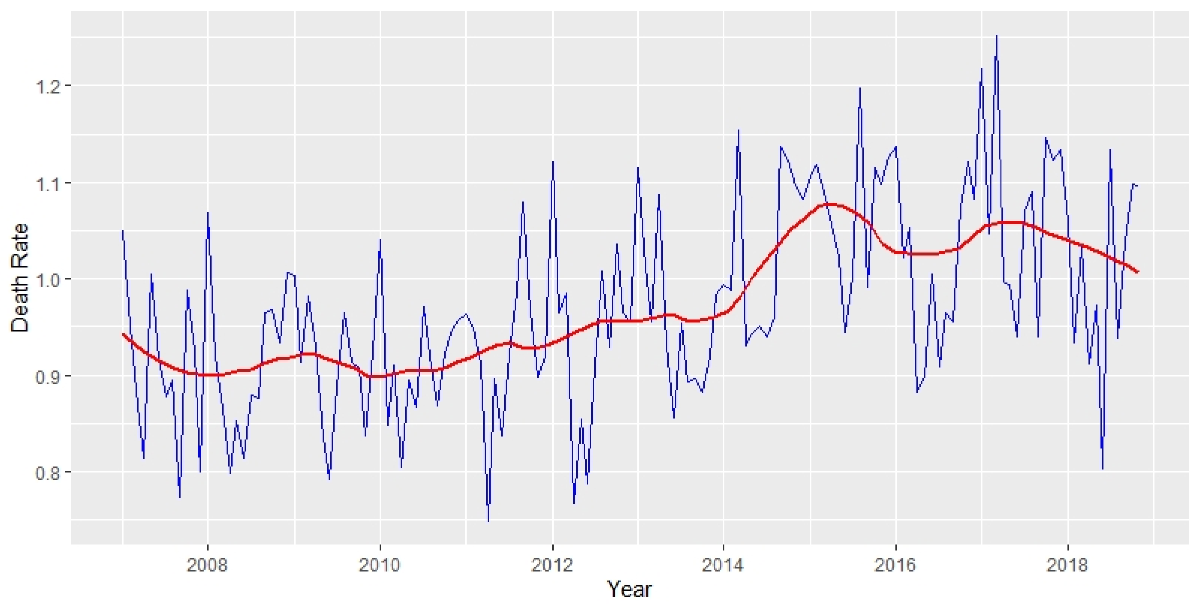
3.1 Exploring monthly Suicide Death Rates per 100 000 Persons

Figure 3 presents a time series plot of the calculated Australian suicide death rates per 100,000 persons, for the period January 2007 to November 2018 inclusive. Figure 4 presents a time series plot of the calculated Australian suicide death rates per 100,000 persons by gender for the same time period. See Appendix A for additional time series plots presenting the calculated monthly death rates separately for each state and territory jurisdiction. For all of the time series, the red line is the fitted Loess regression line.

A visual analysis of Figure 3 highlights the possible trend component of the data. Death rates appear to have increased from 2010 till 2015, and then fluctuated thereafter with a potential decline between mid-2017 and late-2018. An examination of Figure 4 suggests that, overall, this trend is quite similar for males and females. However, death rates for males appear to have been approximately three times greater than for females at any given point in time.

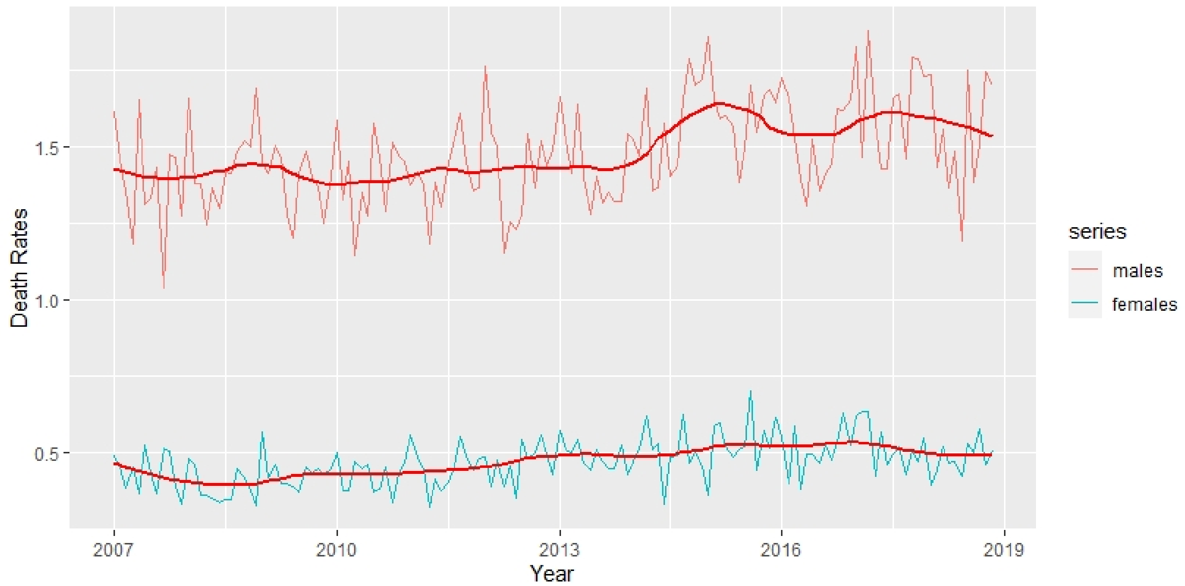
By the end of the period, suicide rates were highest in the Northern Territory, Queensland, and Western Australia. Looking at the State/Territory specific plots, there appears to have been a positive trend in suicide rates in NSW, Queensland and, to a lesser extent, Western Australia. The other States/Territories had a relatively flat profile, with no jurisdictions exhibiting a noticeable downward trend.

Figure 3 Monthly Australian Suicide Death Rates per 100, 000 Persons: January 2007 to November 2018



Note. The red line is an overlayed loess regression curve.

Figure 4 Monthly Australian Suicide Death Rates per 100,000 Persons by sex: January 2007 to November 2018



Note. The red lines are overlayed loess regression curves

The time series data set was assessed to be an additive model, because the magnitude of season fluctuations do not appear to vary with the level of the series. Taken together, a visual inspection of Figures 5, 6, and 7 provide information about possible seasonal components of the data. Figure 5 presents a time series plot of monthly suicide death rates separately for each year, using a more conventional horizontal axis. Figure 6 presents the same information using polar coordinates, which make the time series axis circular rather than horizontal. Figure 7 presents a seasonal subseries plot, where data for each month are displayed as a separate mini time series.

Figures 5, 6, and 7 also show the presence of likely outliers for most years of the time series. During 2018 the highest monthly suicide death rate occurred in July, whereas for other years July had a relatively low rate of suicide deaths. For 2007 one of the lowest observed monthly suicide death rates occurred in December, a month which tended to have higher relative rates of suicide deaths within the other years of the time series. We did not explore any potential causal models for these outliers, but note that they may be due to either random variation, or external factors specific to that month and year. Outliers notwithstanding, Figures 5, 6, and 7 do show possible seasonal patterns, with the lowest suicide rates having occurred in April and June (around 0.85 deaths per 100,000 people) and the highest rates in January (around 1.07 deaths per 100,000).

Figure 5 Seasonal Plot of Monthly Australian Suicide Death Rates per 100,000 Persons: January 2007 to November 2018

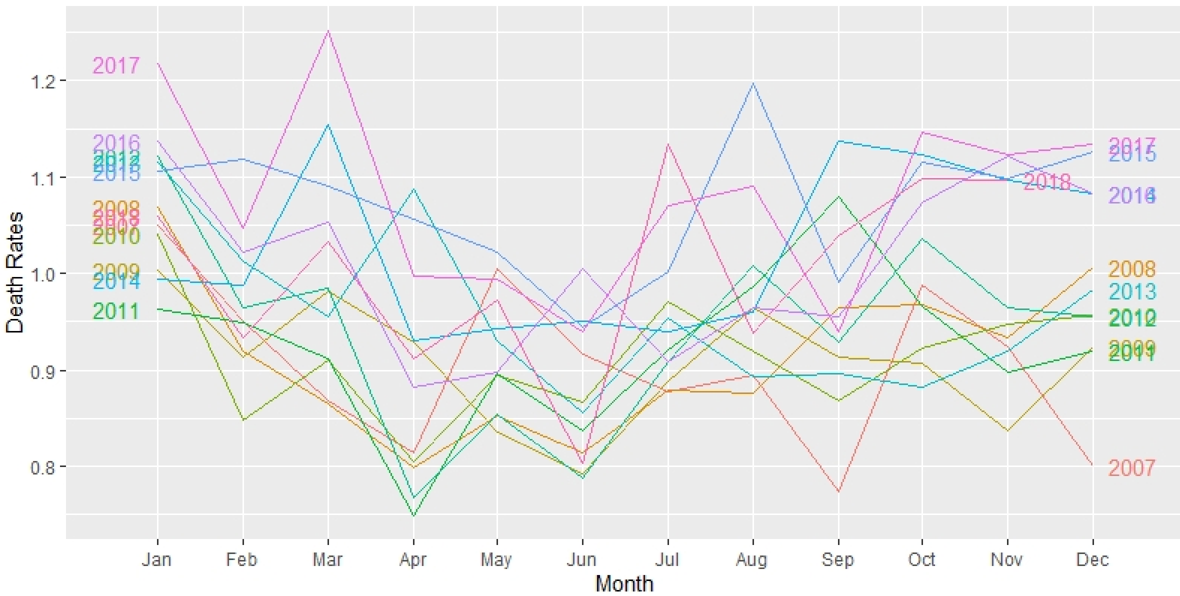


Figure 6 Polar Seasonal Plot of Monthly Australian Suicide Death Rates per 100,000 Persons: January 2007 to November 2018

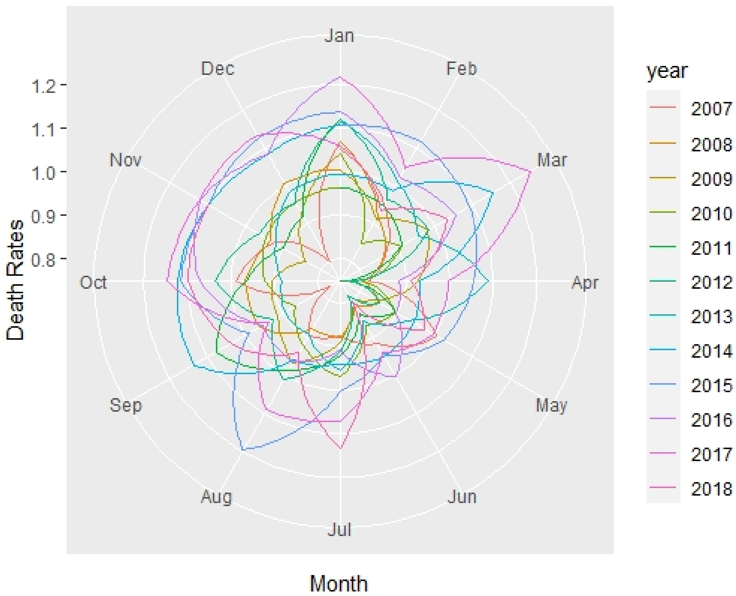
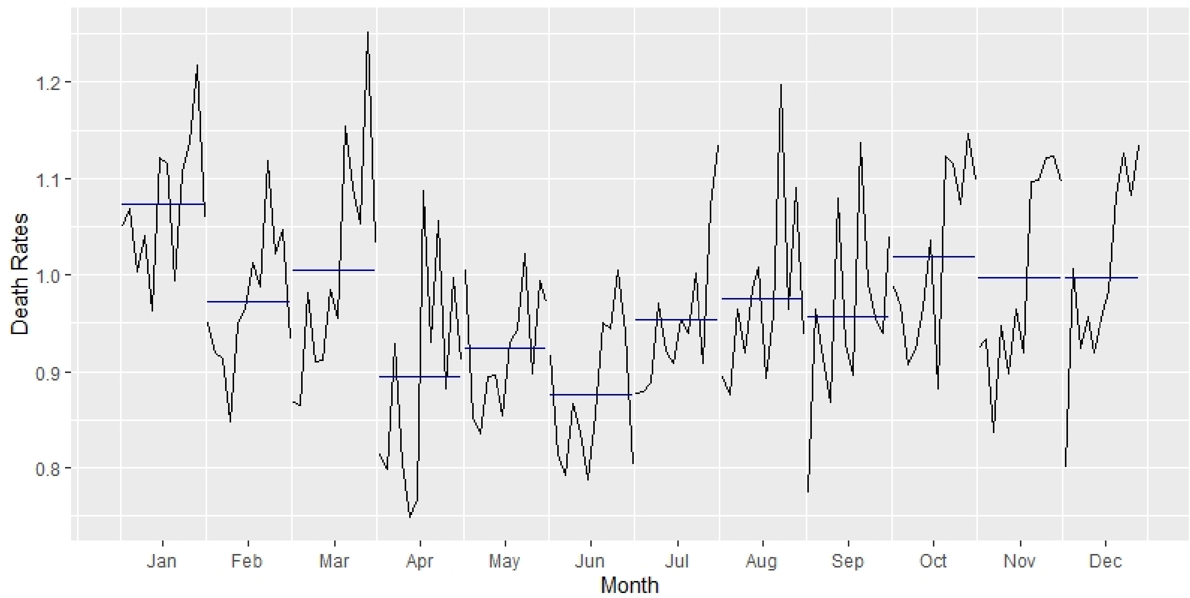


Figure 7 Seasonal Subseries plot of Monthly Australian Suicide Death Rates per 100,000 Persons: January 2007 to November 2018



Note. The blue horizontal lines indicate the average death rate per 100,000 persons, for each month across the 12 observed years (except for the month of December which has observations for only 11 years)

3.2 Modelling seasonality and trend in monthly suicide

In our next set of descriptive analysis, time-series regression models were implemented as a means of identifying possible seasonal and trend components within the monthly suicide deaths rates time series.

The first step in any time series analysis is to test for stationarity or the presence of a unit-root. Specifically, a time-series has a unit root if the association between a value in time *t* and value in time *t-1* (known as the lagged dependent variable) has a value of one. If the time series does have a unit-root, then standard linear regression models can produce invalid estimates as the assumptions around the independence of error terms are violated and the effect of a ‘shock’ to the system does not decay through time. In the absence of a unit-root, however, standard models are more robust.

In order to test for stationarity, we undertook the Augmented Dickey-Fuller test (Becketti, 2013) with and without a trend term and with up to three lags. For all estimations, we are able to reject the null hypothesis of a unit root.[2]

We then re-estimated the Augmented Dickey-Fuller test with 12 lags. That is, explicitly allowing for the suicide rate in a given month to be correlated with the rate in the same month, one year prior. We were not able to reject the null hypothesis with 12 lags (test statistic = -2.501, p-value = 0.3274) strongly suggesting that seasonal factors are likely to be leading to a non-stationary series, and that any time series modelling of suicide needs to account for this. Confirming this, we then differenced the suicide rate by 12 months (subtracting the rate in time *t-12* from the rate in time *t*), and found that this was a stationary series (test statistic = -4.622, p-value = 0.0009). This strongly suggests that any modelling of

[2] The most detailed model, with a trend and three lags, gave a test statistic of -5.439 and a p-value of 0.0000

suicide rates should either adjust for seasonal effect by differencing, or through seasonal dummies, as outlined below.

The results of the regression models with a trend and seasonal dummies are consistent with the trend and seasonal time series plots presented in the previous section.[3] Beginning with the seasonal factors, the results of the regression analyses suggest that January had the highest rate of suicide deaths (though there was no significant difference between the months of January and February). The lowest suicide death rates occurred during the months of April through to July (and June in particular). The death rate for other months (March, August, September, October, November, and December) fell somewhere in between that of January and the period April through to July. The (linear) trend term is positive, showing that controlling for seasonal factors, suicide death rates are higher on average at the end of the period compared to the start of the period.

We can quantify the size of these associations in terms of the overall Australian population (using Model 1). The month of June had an average of 0.006 fewer deaths each day per 100,000 persons compared to January, or 1.4 fewer deaths per day for a population of 25 million persons. By the end of the time series, we predict an average of 43.2 additional people died by suicide per 31-day month for a population of 25 million Australians, when compared to the number of suicide deaths at the beginning of the time series, and controlling for growth in the size of the Australian population.

Results of model 2 found a significant positive association between the rate of suicide deaths and single lagged rate of suicide deaths. A comparison of Model 1 and Model 2, shows that this positive association did not change the direction or statistical significance of the other variables in the model. This result suggests that controlling for trend and seasonality, the rate of suicide deaths in a given month was positively correlated with the rate of suicide deaths in the subsequent month.

[3] The model results also hold when we estimate a Negative Binomial model, with the Australian Estimated Resident Population incorporated as an exposure variable. That is, the Negative Binomial regression finds a significant positive trend, a correlation with the lagged dependent variable, and the same months being significantly different from the omitted category (January).

Table 2 Time Series Model of Estimated Average Daily Suicide Death Rates by Month: January 2007 to November 2018

Independent variable	Model 1		Model 2	
	Coeff.	Signif.	Coeff.	Signif.
Lagged adjusted death rate			0.20103	**
February	-0.00077		-0.00111	
March	-0.00229	**	-0.00248	**
April	-0.00494	***	-0.00483	***
May	-0.00496	***	-0.00431	***
June	-0.00562	***	-0.00497	***
July	-0.00408	***	-0.00330	***
August	-0.00347	***	-0.00300	***
September	-0.00305	***	-0.00270	***
October	-0.00214	**	-0.00187	*
November	-0.00182	*	-0.00174	*
December	-0.00267	***	-0.00264	***
trend (month)	0.00004	***	0.00004	***
Constant	0.03174	***	0.02566	***
Adjusted R-Squared	0.4828		0.5016	
Number of observations	143		142	

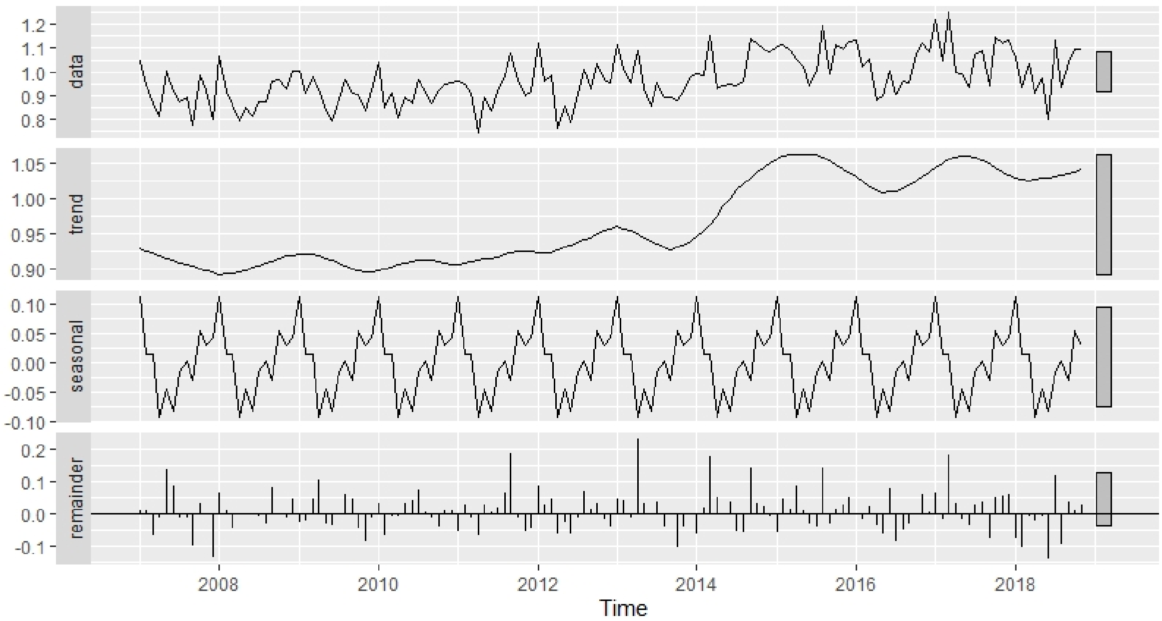
Notes: *p<.01, **p<.05, ***p<0.01

3.3 Decomposing seasonal and trend components

Seasonal and Trend decomposition using Loess (STL) was used as a more formal means of identifying and isolating trend, seasonal, and remainder components of the monthly suicide death rates time series using a more flexible model than that provided in the previous sub-section. The data plot presents full time series data for monthly suicide death rates per 100,000 persons. The trend plot presents the overall trend component of the time series data, given by the loess curve. The seasonal plot presents the estimated seasonal component of the time series data. The remainder plot shows the component of the data not explained by the trend or seasonality. The grey bars on the far right of each plot are range bars and indicate the relative magnitude of each decomposed component (and the amount of variance in the original data each component accounts for). The results of this STL analysis are shown in Figure 8.

This more formal analysis supports the visual description provided earlier, controlling for both random variation and seasonal factors. It also extends the econometric analysis by allowing for a far more flexible trend term. Doing so shows that death rates appear to have been reasonably steady up until 2010, increased from 2010 till 2015, and then fluctuated thereafter. The small decline into 2018 implied in the visual description earlier does not appear to have been maintained in the more formal decomposition presented below.

Figure 8 Seasonal and Trend decomposition using Loess of Monthly Australian Suicide Death Rates per 100,000 Persons: January 2007 to November 2018



4 Results – Building a projection model for monthly deaths

From a policy perspective, an important question regarding suicide is what the likely rate will be over subsequent periods if current trends and observed patterns continue. This allows for resource planning, but perhaps more importantly, allows for observed rates to be compared to a data-driven counterfactual. For example, using the linear model presented in Table 2 or some of the more flexible models outlined in this section, we could create a projection of what the suicide rate would be in April, or May of 2020 (or beyond) if pre-COVID-19 trends continued. If the observed suicide rate was significantly higher than what was projected, then this gives *prima facie* evidence that the spread of COVID-19 has led to increases in suicide. Furthermore, by adding plausible values around these projections, we can also estimate whether the increase is within the variation observed in previous periods.

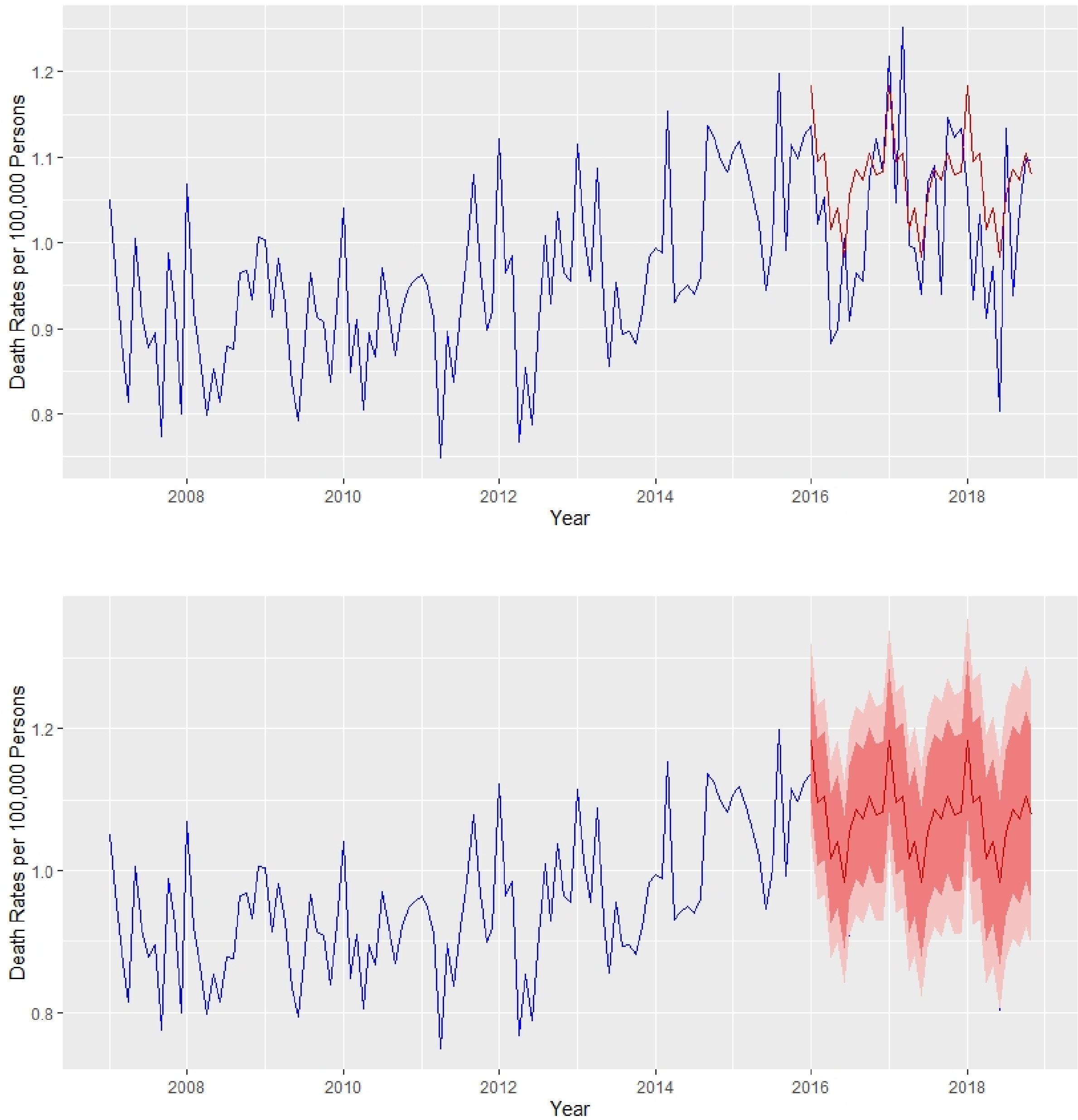
As described earlier, in order to build and test different projection models, we split our data into two periods using the ratio of 75%:25% as training: testing datasets. Specifically, the training data (from 2007 till the end of 2015) is used for generating or learning the model and the testing data (2016 till November 2018) is used to test the model's accuracy. We outline five models below, and then give diagnostic tests of their ability to fit the data. Unlike the econometric model described in Table 2 above where we were interested in parameter values rather than prediction accuracy, in this section, we allow for more flexible, data-driven specifications in our projection model (nonparametric forecasting models).

4.1 Building projection models

4.1.1 Model 1: Exponential Smoothing State Space Model

Figure 9a fits an exponential state space smoothing (ETS) (A,N,A) with additive error, no trend and additive seasonality model which automatically optimizes the choice of model parameters. Point forecasts for the next 36 months are shown with a red line, which roughly matches the historical pattern of the data. Figure 9b gives an 80% (dark red) and 95% (light red) prediction intervals where the 95% interval covers the test data well, which means the point forecasts with a 95% prediction interval are good, i.e., the point forecasts contain a range of values which should include the actual death rates in the test data with a 95% probability.

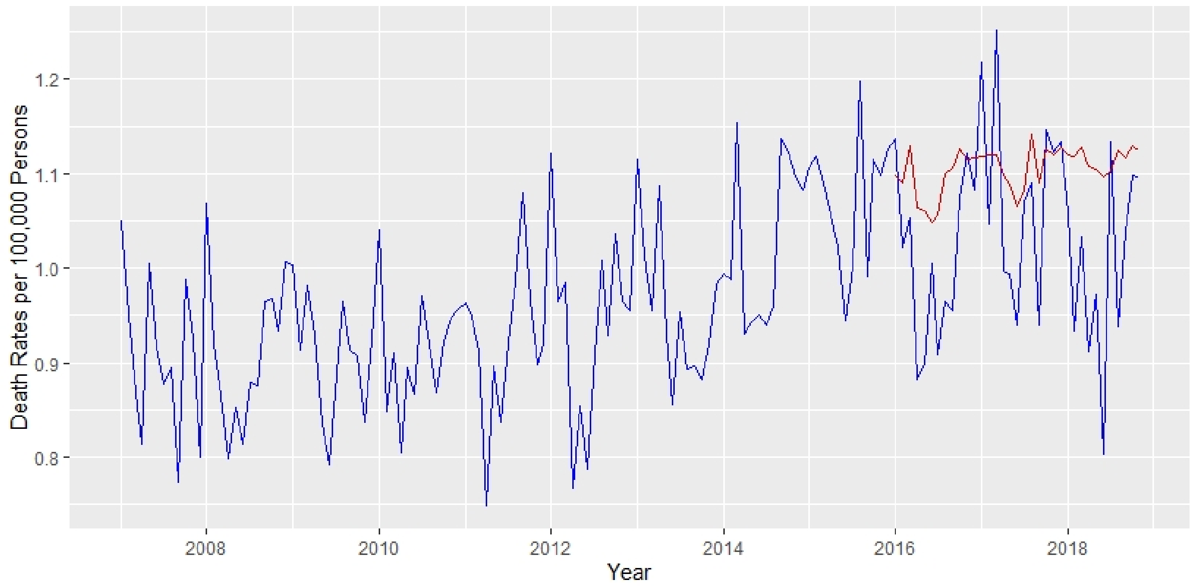
Figure 9a and 9b ETS forecasts of monthly death rates per 100,000 persons

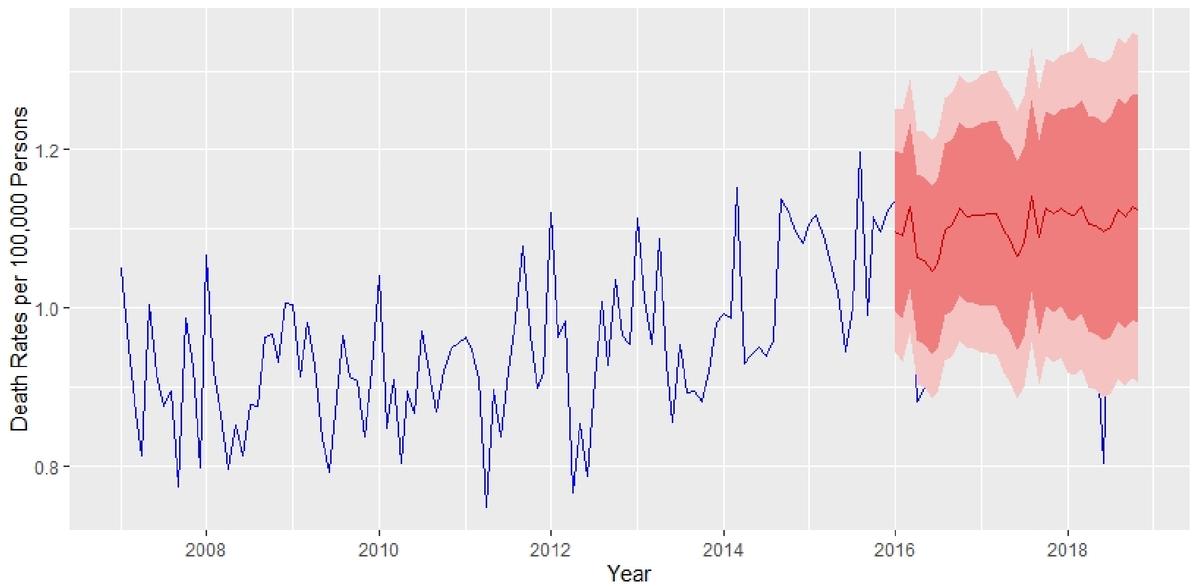


4.1.2 Model 2: Autoregressive Integrated Moving Average Model

Figure 10 fits an Autoregressive Integrated Moving Average Model (ARIMA)(1,1,1)(2,0,0)[12] model, which denotes a monthly data with non-seasonal differencing by 1, AR order 1 and MA order 1 and seasonal differencing by 2, AR order 0 and MA order 0 for the seasonal parts. A forecast for the next 36 months is shown in red with 80% and 95% prediction confidence intervals in Figure 10b. Notice that the forecasts roughly match the historical pattern of the data but it probably does not look as good as the ETS model, especially when some test data appear to be out of the 95% prediction interval. The ETS model may be a result of a better fit to the data, but that decision can be made visually.

Figure 10a and 10b ARIMA forecasts of monthly death rates per 100,000 persons

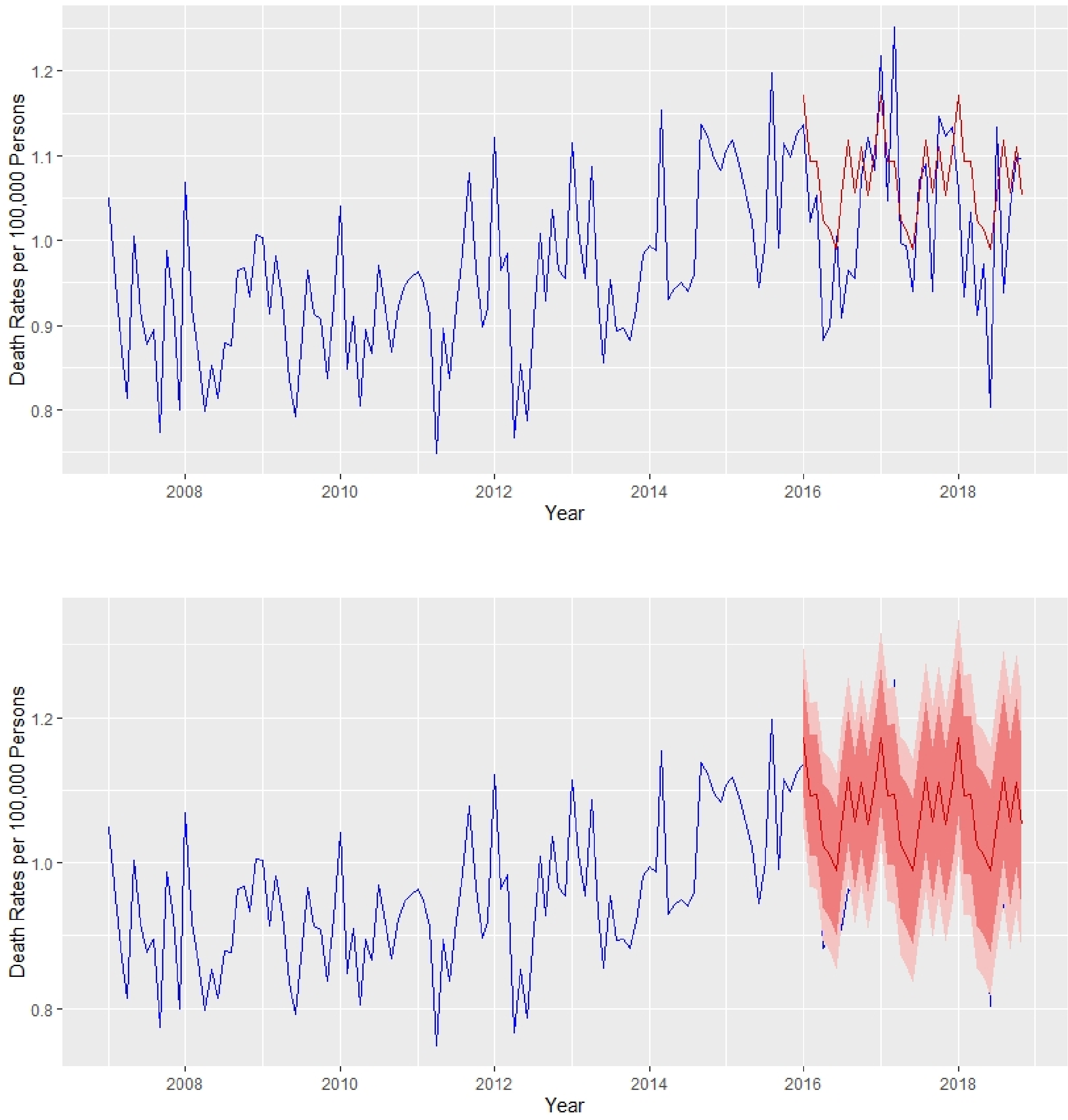




4.1.3 Model 3: Exponential Smoothing State Space Model with Box-Cox Transformation, ARMA Errors, Trend and Seasonal Components Model

The third model trained is a versatile exponential smoothing state space model with Box-Cox transformation, ARMA errors, trend and seasonal components (TBATS), given by Figure 11a. Point forecasts for the next 36 months are shown with a red line, which roughly matches the historical pattern of the data. Figure 11b gives an 80% (dark red) and 95% (light red) prediction intervals where the 95% interval covers the test data well, which means the point forecasts with a 95% prediction interval are good, i.e., the point forecasts contain a range of values which should include the actual death rates in the test data with a 95% probability. The results are somewhat closer to the ETS model.

Figure 11a and 11b: TBATS forecasts of monthly death rates per 100,000 persons

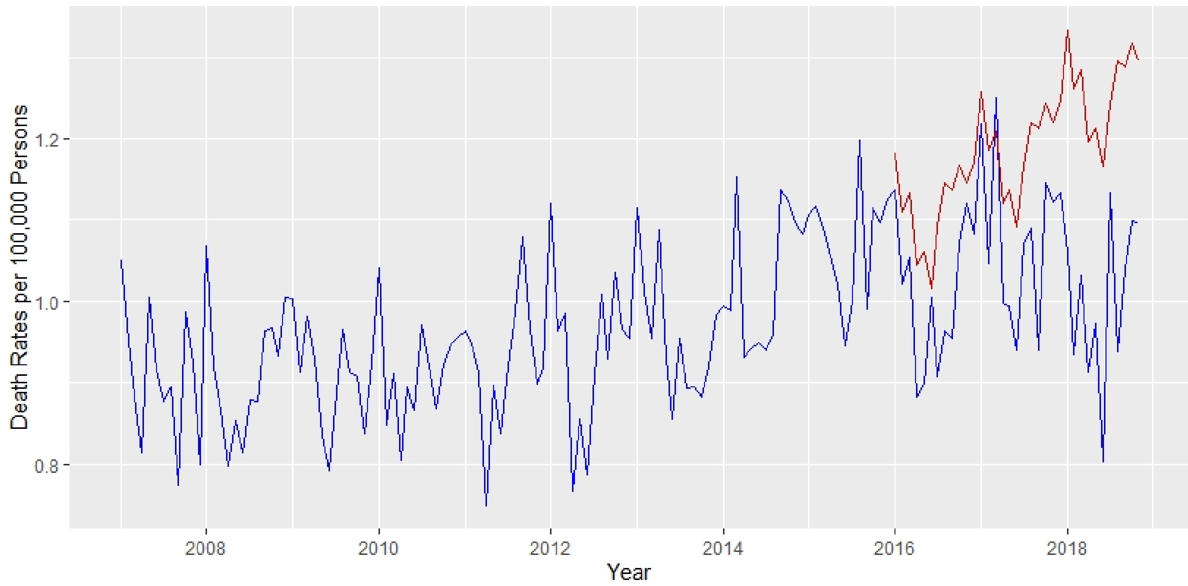


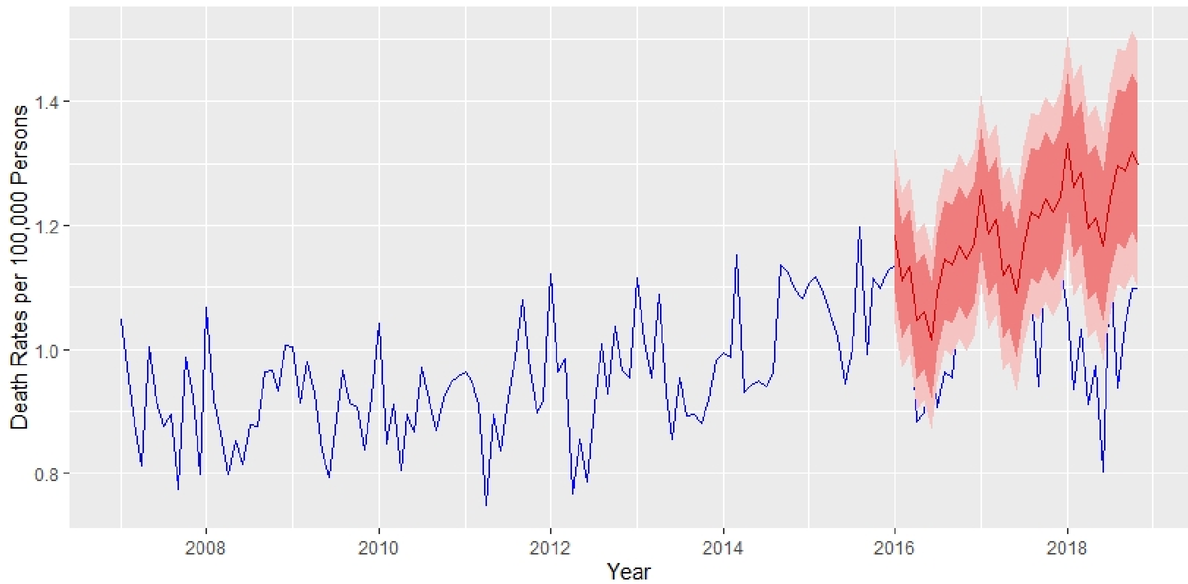
4.1.4 Model 4: Seasonal Decomposition of Time Series by Loess Model STL

The fourth forecasting model is seasonal decomposition of time series by loess (STL), which decomposes a time series into seasonal, trend and irregular components using loess. Forecasts of STL objects are obtained by applying a non-seasonal forecasting method to the seasonally adjusted data and re-seasonalizing using the last year of the seasonal component.

In the forecasts given by Figure 12a and Figure 12b, we can see the figures for point forecasts and with prediction intervals respectively. STL appears to over-predict with larger death rates that increase with time.

Figure 12a and 12b STL forecasts of monthly death rates per 100,000 persons

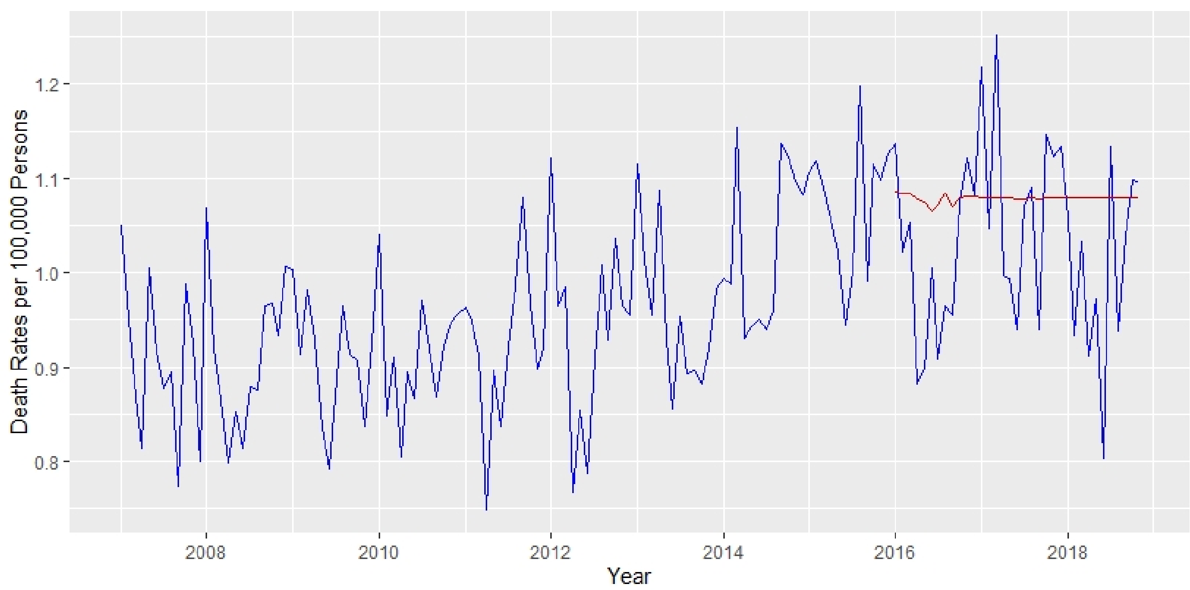


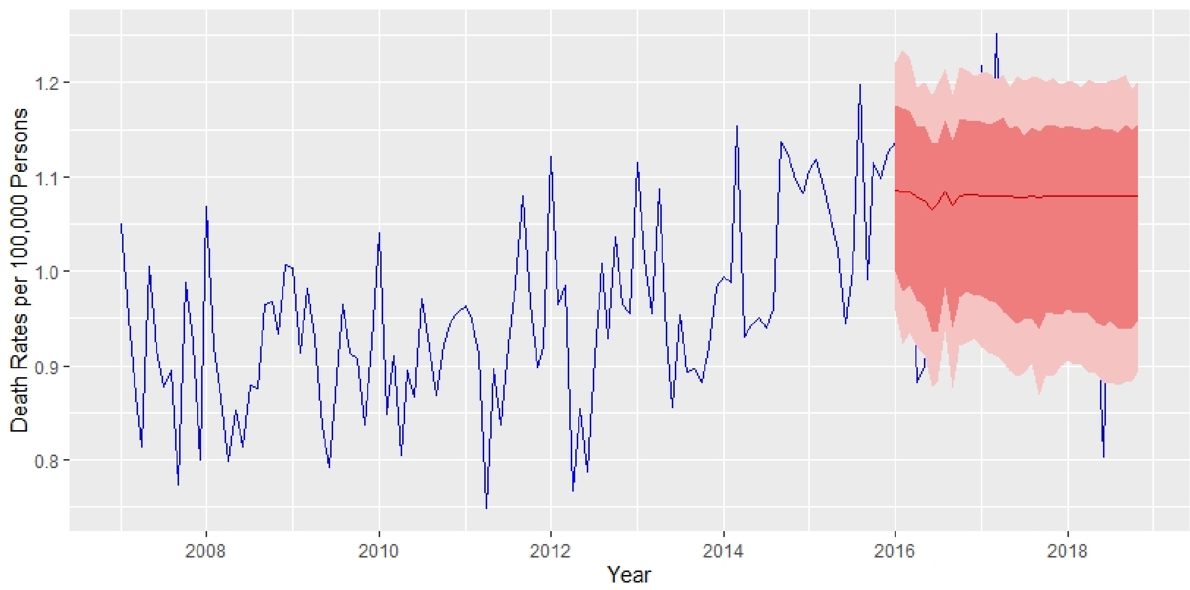


4.1.5 Model 5: Neural Network Autoregression Model

In the final forecasting model, the neural network autoregression (NNAR) model is trained as a feed-forward neural networks model with a single hidden layer and lagged inputs for forecasting univariate time series. NNAR(p,P,k) model has three components: p, P & k: p denotes the number of lagged values that are used as inputs, P denotes the number of seasonal lags and k denotes the number of hidden nodes that are present. In this case, the procedure automatically generates NNAR(3,1,2)[12] model. Figure 13a presents the point forecasts while Figure 13b presents the prediction intervals.

Figure 13a and 13b NNAR forecasts of monthly death rates per 100,000 persons





4.2 Choosing the best forecasting model

Across the five models there is significant variation in the accuracy of the forecasts and projections. One way to choose amongst the models is to compare their residuals to see how adequately a model has captured the information in the data that was excluded from the training data. The residuals are equal to the difference between the observed values and the corresponding fitted values: $e_t = y_t - \hat{y}_t$.

A good forecasting method will yield residuals that are uncorrelated over successive time periods. If there are correlations between residuals, then there is information left in the residuals which should be used in computing forecasts. Also, the residuals should have zero mean. If the residuals have a mean other than zero, then the forecasts are biased. In addition, it is also useful for the residuals to be approximately normally distributed and have constant variance.

Many models may meet the above criteria. MAPE (mean absolute percentage error) is the average of the absolute percentage errors of forecasts. It is the most commonly used measure of forecast accuracy since it is easy to understand conceptually and it has the advantages of scale-independency and interpretability (Kim & Kim, 2016). However, MAPE becomes highly skewed when observed values in the time series are close to zero and infinite when observations equal to zero, making it unsuitable for some time series that have low report values. The death rates data does not fall into this category, which makes MAPE an ideal measure.

The ETS (MAPE=7.44%) and TBATS (MAPE=7.52%) models have the lowest average error values. The ARIMA and NNAR models also perform reasonably well by this measure (MAPE=9.70% and 9.35% respectively) whereas the STL model performs relatively poorly (MAPE = 16.33%).

Looking at the two preferred forecast models based on the lowest MAPE the figures below produce time plots, ACF plots and histograms of the residuals (with an overlaid normal distributions for comparison).

Figure 14 ETS model diagnostic plots

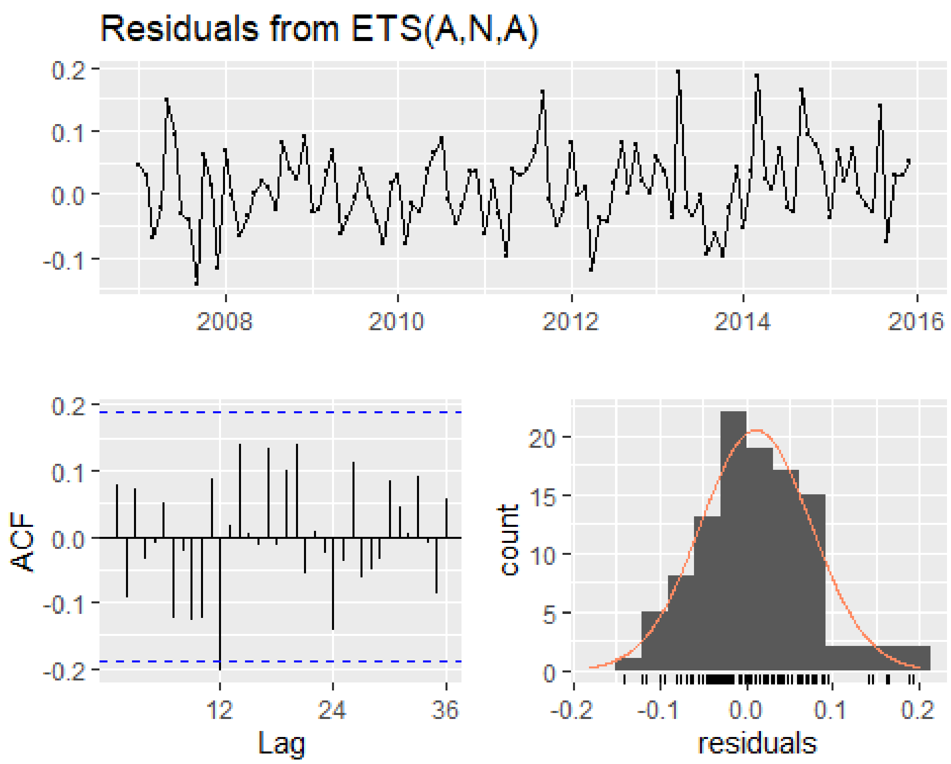
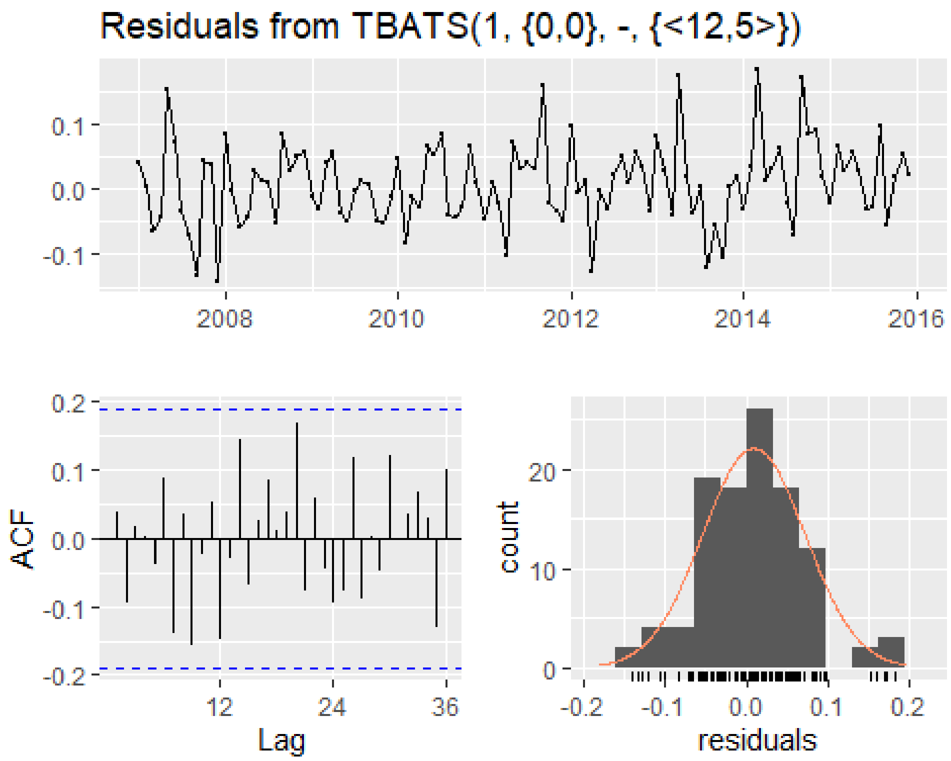


Figure 15 TBATS model diagnostic plots

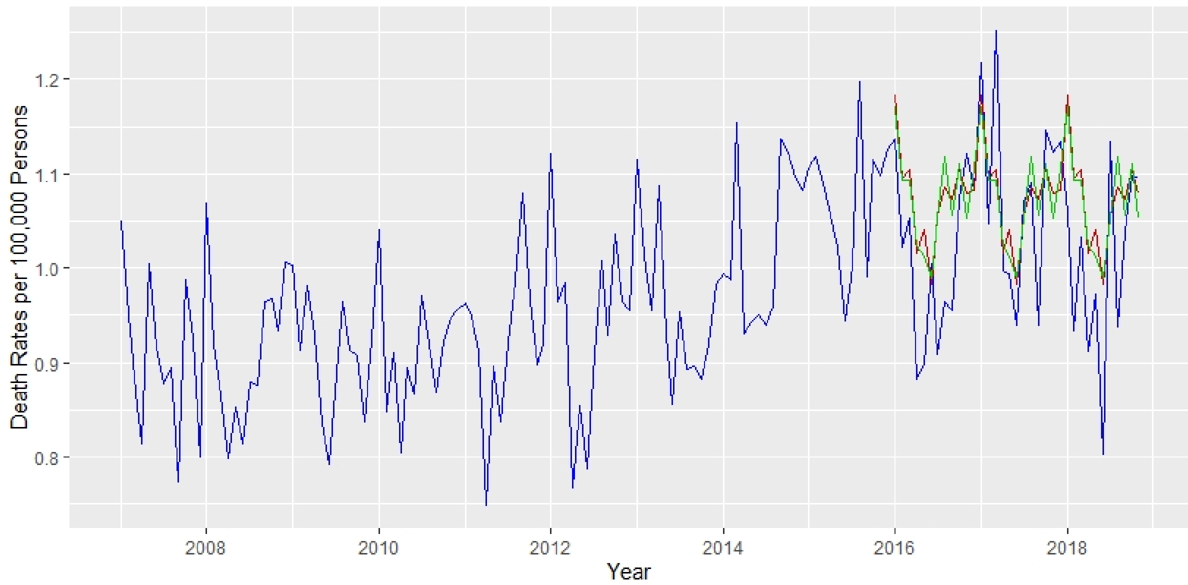


The above graphs show that the ETS and TBATS methods produce forecasts that appear to account for all available information. The means of the residuals in these two methods are close to zero and there is no significant correlation in the series of residuals. In these two methods, we see that in the ACF plots, the spikes are within the required limits except for one spike in ETS model. Since the autocorrelations are small, we conclude that the model does not exhibit a significant lack of fit. The time plot of the residuals shows that the variation of the residuals stays much the same across the historical data and therefore the residual variance can be treated as constant. This can also be seen on the histogram of the residuals. The histogram suggests that the residuals appear to be approximately normal for both models.

4.3 Comparing forecast accuracies to choose the final model

With reasonable predictions in both models, they can be graphically compared to see which one performs best visually, or how close they are in the forecasts for the 36 months of the test data, given in Figure 16.

Figure 16 Comparing ETS and TBATS model forecasts



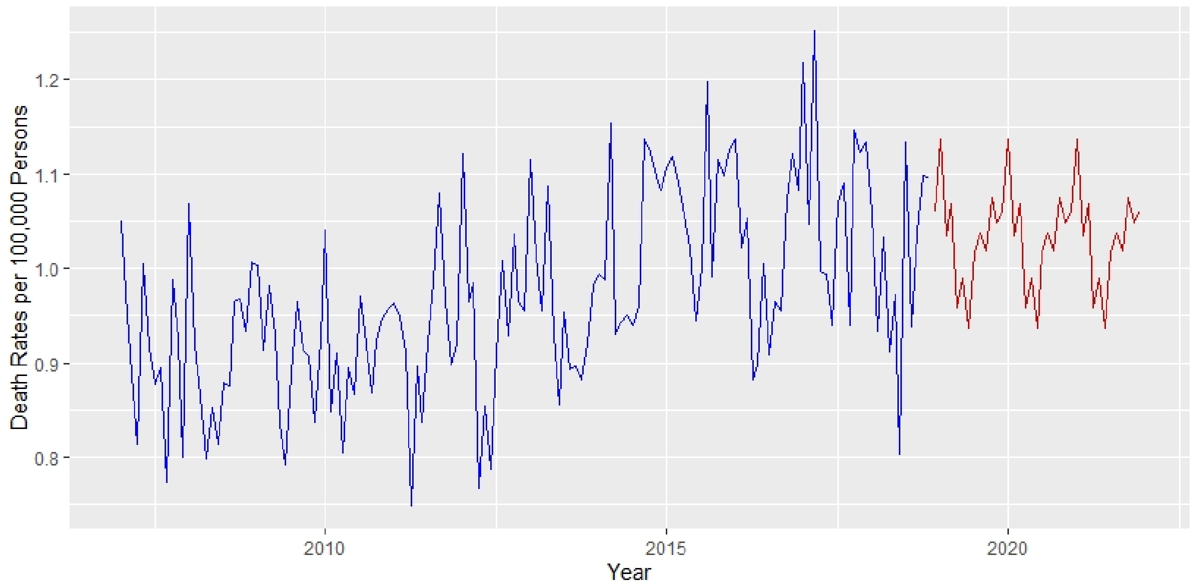
Note. The red line indicates the predicted point forecasts of the ETS model while the green one indicates the same for the TBATS model.

4.4 Forecasting monthly suicide death rates for December 2018 till December 2020

The ETS model (due to the slightly smaller MAPE) was re-estimated using the complete monthly suicide death rates time series, January 2007 to November 2018, and forecast projected monthly suicide death rates for December 2018 to December 2020. The MAPE for the training dataset was 5.49%.

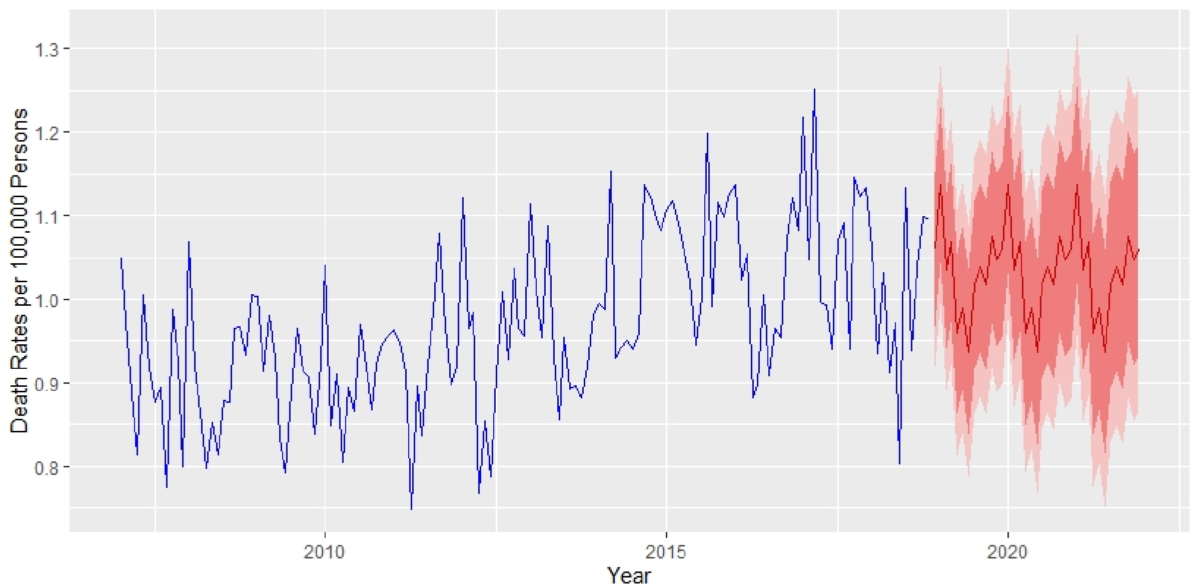
Figure 17a and 17b present the results of fitting this model. Refer to Appendix B for the predicted point estimates with 80% and 95% confidence intervals.

Figure 17a Exponential State Smoothing Model: Point Forecasts for December 2018 to December 2020



Note. The blue line indicates the observed dataset for monthly suicide death rates. The red line indicates the predicted point forecasts of the model.

Figure 17b Exponential State Smoothing Model: Point Forecasts with Confidence Intervals for December 2018 to December 2020



Note. The blue line indicates the observed dataset for monthly suicide death rates. The red line indicates the predicted point forecasts of the model. The dark red shaded area indicates the 80% confidence interval around the point forecasts. The pink shaded area indicates the 95% confidence interval around the point forecasts.

5 Discussion and concluding comments

Suicide is a major cause of death in Australia, resulting in significant loss of expected years of life. Better understanding trends in suicide, and developing Australia's suicide forecasting capability has the potential to strengthen our public health planning and responses and improve evaluation of prevention or intervention initiatives.

Increased robustness of suicide monitoring additionally provides opportunities to measure the impact of large events such as the recent Black Summer bushfires and the COVID-19 pandemic. The aim of this paper was to describe, model, and project suicide rates based on historic, monthly count data provided by the AIHW. The combined impact on data quality of late death registrations and the revisions processes were also accounted for within our methodology.

Based on our descriptive and econometric analysis, we conclude that there has been an increasing rate of suicide over the time period (2007 to 2018). A linear trend model estimates that an average of 43.2 additional people are likely to have died by suicide per 31-day month for a population of 25 million Australians at the end of the time period, when compared to the number of suicide deaths at the beginning of the time period, even after controlling for population growth. However, a more flexible trend analysis suggests that death rates appear to have been reasonably steady up until 2010, increased from 2010 till 2015, and then fluctuated thereafter.

Overall, the observed trend is quite similar for males and females. However, death rates for males appear to have been approximately three times greater than for females at any given point in time. By the end of the period, suicide rates were highest in the Northern Territory, Queensland, and Western Australia. There appears to have been a positive trend in suicide rates in NSW, Queensland and, to a lesser extent, Western Australia with all other States/Territories having a relatively flat profile.

Furthermore, we observe that there is a consistent monthly variation with January and February having the highest rates of suicide, and April to July having the lowest rates. Across the period, the month of June had an average of 0.006 fewer deaths each day per 100,000 persons compared to January, or 1.4 fewer deaths per day for a population of 25 million persons.

Based on the observed data, we were able to test a number of projection models, concluding that an ETS model (exponential smoothing within a state space framework) best fits the observed data when 75 per cent of our observations are used to train the model, and the other 25 per cent of our observations are used to test it. Having built and tested this model, we were able to project the likely suicide rates beyond the current observation window, alongside plausible upper and lower bounds based on the existing observations.

An important area of future work is to test our projection model (for December 2018 and beyond) with observed data once it becomes available. While this data will only be preliminary, it will allow us to measure whether a model built on data up until the end of 2018 can accurately project suicide rates beyond that period. Once data becomes available for the COVID-19 period, we will also be able to test whether observed suicide rates are above what would have been projected in the absence of this external shock.

A second area of future analysis is to replicate the model building and projections for population subgroups within Australia. We have shown that males and females have different rates of suicide, and somewhat different trends. However, other yearly data suggests that there is also variation by age group, State/Territory, and ethnicity. While there is likely to be more random variation in the data the smaller the population size, our methodology allows us to capture this in the projection intervals.

One of the benefits of re-estimating the projections for smaller population sub-groups is that they may be affected differently by exogenous shocks. For example, if one jurisdiction, age group, or small area was exposed to an external shock and other wasn't, then if the observed suicide rates for that sub-group was outside of the projection range whereas this does not occur for the rest of the country (essentially a difference-in-difference model), then this provides even stronger evidence for an effect of that exogenous shock on suicide.

Ultimately, to identify effective policy responses to minimize deaths from suicide in the community, it is important for high-quality data to be available, for that data to be analysed and understood using rigorous data techniques, and for conclusions on future or observed changes to take into account the significant variation and uncertainty in suicide by month and by other characteristics. The forecasting model that we have developed will hopefully contribute to this monitoring and analytical endeavour.

References

Australian Bureau of Statistics. (2012, July 2). *Suicides, Australia, 2010* (catalogue no. 3309.0). https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3309.0Main+Features12010?OpenDocument

Australian Bureau of Statistics. (2019, September 25). *Causes of death, Australia,* 2018 (catalogue number 3303.0). https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/3303.0~2018~Main%20Features~Intentional%20self-harm,%20key%20characteristics~3

Australian Bureau of Statistics. (2020, June 18). *Australian demographic statistics, Dec 2019* (catalogue no. 3101.0). https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3101.0Main+Features1Dec%202019?OpenDocument

Australian Institute of Health and Welfare. (2020, July 23). Suicide and intentional self-harm. Snapshot. https://www.aihw.gov.au/reports/australias-health/suicide-and-intentional-self-harm

Australian Institute of Health and Welfare: Henley G & Harrison JE (2019). *Injury mortality and socioeconomic influence in Australia, 2015–16*. (catalogue number INJCAT 208). https://www.aihw.gov.au/reports/injury/injury-mortality-and-socioeconomic-influence-in-au/contents/table-of-contents

Australian Institute of Health and Welfare: Harrison, J.E., Pointer, S. & Elnour, A. A. (2009, July 28). *A review of suicide statistics in Australia* (catalogue number INJCAT 121). https://www.aihw.gov.au/reports/injury/a-review-of-suicide-statistics-in-australia/contents/table-of-contents

Becketti, S. (2013)*. Introduction to time series using Stata, revised edition*. Stata Press.

Bergmeir, C., Hyndman, R. J., & Benítez, J. M. (2016). Bagging exponential smoothing methods using STL decomposition and Box–Cox transformation. *International Journal of Forecasting,* 32(2), 303–12. https://doi.org/10.1016/j.ijforecast.2015.07.002

Bickley, H., Hunt, I. M., Windfuhr, K., Shaw, J., Appleby, L., & Kapur, N. (2013). Suicide within two weeks of discharge from psychiatric inpatient care: A case-control study. Psychiatric Services, 64(7), 653-659. https://doi.org/10.1176/appi.ps.201200026

Biddle, N., Edwards, B., Herz, D., & Makkai, T. (2020). *Exposure and the impact on attitudes of the 2019-20 Australian Bushfires.* ANU Centre for Social Research and Methods. https://csrm.cass.anu.edu.au/research/publications/exposure-and-impact-attitudes-2019-20-australian-bushfires-0

Biddle, N., Edwards, B., Gray, M., & Sollis, K. (2020a) *Mental health and relationships during the COVID-19 pandemic.* ANU Centre for Social Research and Methods. https://csrm.cass.anu.edu.au/research/publications/mental-health-and-relationships-during-covid-19-pandemic

Biddle, N., Edwards, B., Gray, M., & Sollis, K., (2020b, May 7) *Hardship, distress, and resilience: The initial impacts of COVID-19 in Australia.* ANU Centre for Social Research and

Methods. https://csrm.cass.anu.edu.au/news/hardship-distress-and-resilience-initial-impacts-covid-19-australia

Biddle, N., Edwards, B., Gray, M., & Sollis, K., (2020c) *Alcohol consumption during the COVID-19 period: May 2020.* ANU Centre for Social Research and Methods. https://csrm.cass.anu.edu.au/research/publications/alcohol-consumption-during-covid-19-period-may-2020

Box, G.E.P, Jenkins, G.M., & Reinsel, G.C. (2011). *Time series analysis: Forecasting and control* (4th ed.). John Wiley & Sons.

Brockwell, P. J., & Davis, R.A. (2016). *Introduction to time series and forecasting* (3rd ed). Springer.

Calati, R., Ferrari, C., Brittner, M., Oasi, O., Olié, E., Carvalho, A. F., & Courtet, P. (2019). Suicidal thoughts and behaviors and social isolation: A narrative review of the literature. *Journal of Affective Disorders*. 245, 653-667. https://doi.org/10.1016/j.jad.2018.11.022

Case, A., & Deaton, A. (2017, Spring). Mortality and morbidity in the 21st century. B*rookings Papers on Economic Activity*. 397-476. https://doi:10.1353/eca.2017.0005

Case, A. & Deaton, A. (2020). *Deaths of despair and the future of capitalism*. Princeton University Press.

Cheung, Y. T., Spittal, M. J., Pirkis, J., & Yip, P. S. (2012). Spatial analysis of suicide mortality in Australia: Investigation of metropolitan-rural-remote differentials of suicide risk across states/territories. *Social Science & Medicine*, 75(8), 1460-1468. https://doi.org/10.1016/j.socscimed.2012.04.008

Chitty, K. M., Schumann, J. L., Schaffer, A., Cairns, R., Gonzaga, N., J., Raubenheimer, J., E., Cater, G., Page, A., Person, S-A, & Nicholas, A. B. (2020). Australian suicide prevention using health-linked data (ASHLi): Protocol for a population-based case series study. *BMJ Open*, 5(10), e038181. https://doi:10.1136/bmjopen-2020-038181

Choo, C.C., Harris, M. H, & Ho, R.C. (2017). Prediction of lethality in suicide attempts: Gender matters. *OMEGA - Journal of Death and Dying*, 80(1), 87-103. https://doi.org/10.1177/0030222817725182

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403), 596-610.

Cleveland, R. B., William, S. C., Jean E. M., & Terpenning, I. (1990). STL: A seasonal-trend decomposition. *Journal of Official Statistics.* 6(1), 3–73.

Coroners Court of Victoria. (2020, August 27). *Coroners court monthly suicide data report* (Report 1). https://www.coronerscourt.vic.gov.au/sites/default/files/2020-08/Coroners%20Court%20Monthy%20Suicide%20Data%20Report%20-%20Report%201%20-%2027082020.pdf

Crone, S. F., Hibon, M. & Nikolopoulos, K. (2011). Advances in forecasting with neural networks? Empirical evidence from the Nn3 competition on time series prediction. *International Journal of Forecasting.* 27(3), 635–60. https://doi.org/10.1016/j.ijforecast.2011.04.001

Deady, M., Tan, L., Kugenthiran, N., Collins, D., Christensen, H. & Harvey, S.B. (2020). Unemployment, suicide and COVID-19: Using the evidence to plan for prevention. *Medical journal of Australia*. 213(4). https://doi-org.virtual.anu.edu.au/10.5694/mja2.50715

Deaton, A. (2013). *The great escape: Health, wealth, and the origins of inequality*. Princeton University Press.

De Livera, A. M., Hyndman. R.J., & Snyder, H.R. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association.* 106(496), 1513–27. https://doi.org/10.1198/jasa.2011.tm09771

De Leo, D., Sveticic, J., & Milner. A. (2011). Suicide in Indigenous people in Queensland, Australia: Trends and methods, 1994-2007. *The Australian and New Zealand Journal of Psychiatry*. 45(7), 532-538. https://doi.org/10.3109/00048674.2011.570310

Gardner, E. S. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting.* 4(1), 1–28. https://doi.org/10.1002/for.3980040103

Gardner, E. S. (2006). Exponential smoothing: The state of the art-Part II. *International Journal of Forecasting* 22(4), 637–66. https://doi.org/10.1016/j.ijforecast.2006.03.005

Hunt, I. M., Kapur, N., Webb, R., Robinson, J., Burns, J., Shaw, J., & Appleby, L. (2009). Suicide in recently discharged psychiatric patients: A case-control study. *Psychological medicine*, 39(3), 443–449. https://doi:10.1017/S0033291708003644

Hyndman, R.J. & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd edition). OTexts. https://otexts.com/fpp2/

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting* 22(4), 679–88. https://doi.org/10.1016/j.ijforecast.2006.03.001

Hyndman, R.J, Koehler, A.B., Ord, J.K., & Snyder, R.D. (2008). *Forecasting with exponential smoothing: The state space approach*. Springer.

Jacoby, W. G. (2000). Loess: A nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies*. 19(4), 577-613. https://doi.org/10.1016/S0261-3794(99)00028-1

Kim, S & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting* 32(3), 669–79. https://doi.org/10.1016/j.ijforecast.2015.12.003

Kõlves, K., Potts, B., & De Leo, D. (2015). Ten years of suicide mortality in Australia: Socio-economic and psychiatric factors in Queensland. *Journal of Forensic and Legal Medicine.* 36, 136-143. https://doi.org/10.1016/j.jflm.2015.09.012

Korda, R.J., Biddle, N., Lynch, J., Eynstone-Hinkins, J., Soga, K., Banks, E., Priest, N., Moon, L. & Blakely, T. (2020). Education inequalities in adult all-cause mortality: First national data for Australia using linked census and mortality data. *International Journal of Epidemiology*, 49(2), 511-518. https://doi.org/10.1093/ije/dyz191

Mitchell, R., & Cameron, C. (2017). Self-harm hospitalised morbidity and mortality risk using a matched population-based cohort design. *Australian and New Zealand Journal of Psychiatry*, 52(3), 262-270. https://doi.org/10.1177/0004867417717797

Naghavi, M. (2019). Global, regional, and national burden of suicide mortality 1990 to 2016: Systematic analysis for the Global Burden of Disease Study 2016. *BMJ*, 364, l94. https://doi.org/10.1136/bmj.l94

National Mental Health Commission. (2019). *Annual report 2018–19*. https://www.transparency.gov.au/annual-reports/national-mental-health-commission/reporting-year/2018-2019

Ord, K., & Fildes, R. (2012). *Principles of business forecasting (1st ed.)*. Nelson.

Pena, D., George C. T., & Tsay, R.S. (2011). *A course in time series analysis*. Wiley.

Phillips, A., & Hempstead, K. (2020). Differences in U.S. Suicide Rates by Educational Attainment, 2000-2014. *American Journal of Preventative Medicine*. 53(4), e123-e130. https://doi:10.1016/j.amepre.2017.04.010

Productivity Commission. (2019). *Mental health, draft report*. https://www.pc.gov.au/inquiries/current/mental-health/draft

Qi, X., Tong, S. & Hu, W. (2010). Spatial distribution of suicide in Queensland, Australia. *BMC Psychiatry.* 10(106). https://doi.org/10.1186/1471-244X-10-106

Robinson, J., Too, L. S., Pirkis, J., & Spittal, M. J. (2016). Spatial suicide clusters in Australia between 2010 and 2012: A comparison of cluster and non-cluster among young people and adults. *BMC Psychiatry*, 16, 417. https://doi.org/10.1186/s12888-016-1127-8

Singhal, A., Ross, J., Seminog, O., Hawton, K., & Goldacre, M. (2014). Risk of self-harm and suicide in people with specific psychiatric and physical disorders: Comparisons between disorders using English national record linkage. *Journal of the Royal Society of Medicine*, 107(5), 194-204. https://doi:10.1177/0141076814522033

Stene-Larsen, K., & Reneflot, A. (2019). Contact with primary and mental health care prior to suicide: A systematic review of the literature from 2000 to 2017. *Scandinavian Journal of Public Health*, 47(1), 9-17. https://doi.org/10.1177%2F1403494817746274

Too, L. S., Pirkis, J., Milner, A., & Spittal, M. J. (2017). Clusters of suicides and suicide attempts: Detection, proximity and correlates. *Epidemiology and Psychiatric Sciences*, 26(5), 491-500. https://doi.org/10.1017/S2045796016000391

Too, L. S., Law, P., Spittal, M., Page, A., & Milner, A. (2018). Widening socioeconomic inequalities in Australian suicide, despite recent declines in suicide rates. *Social Psychiatry and Psychiatric Epidemiology*, 53(9), 969-976. https://doi:10.1007/s00127-018-1527-9

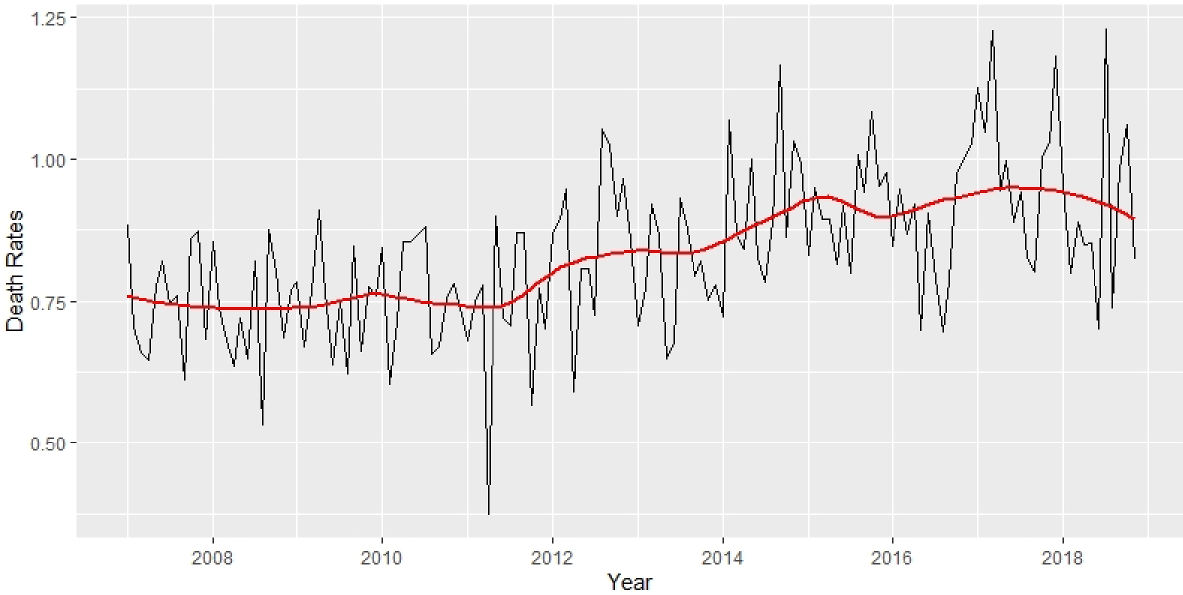
Torok, M., Konings, P., Batterham, P. J., & Christensen, H. (2017). Spatial clustering of fatal, and non-fatal, suicide in New South Wales, Australia: Implications for evidence-based prevention. *BMC Psychiatry*, 17,339. https://doi.org/10.1186/s12888-017-1504-y

World Health Organization (2004). *International statistical classification of diseases and related health problems, 10th revision*. https://icd.who.int/

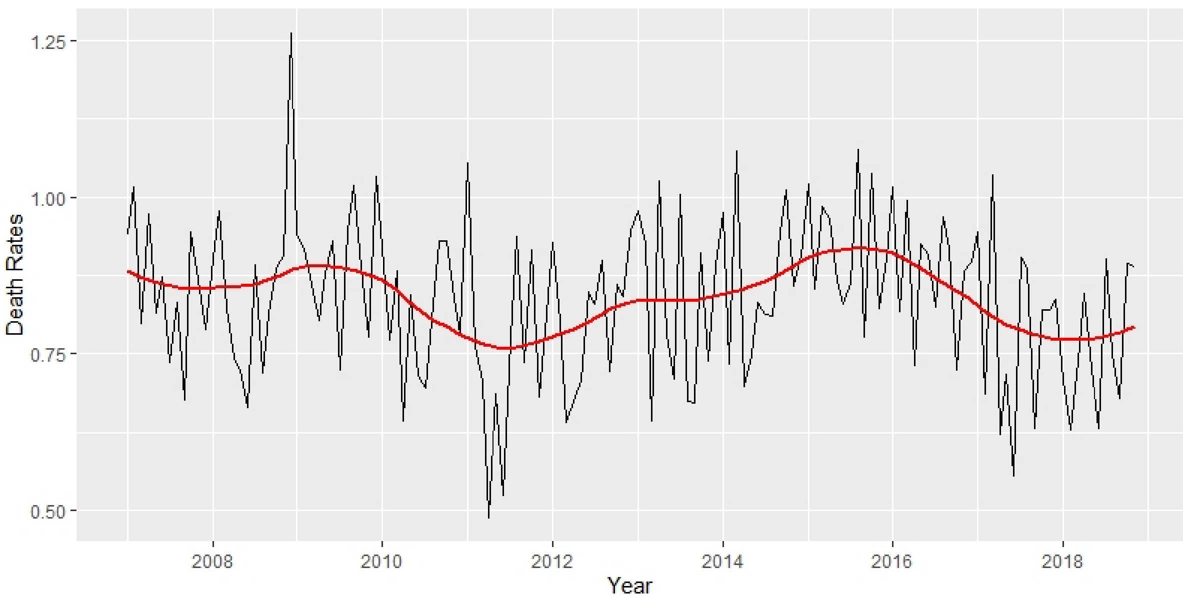
World Health Organization. (2020). *Suicide in the world: Global health estimates* https://www.who.int/publications/i/item/suicide-in-the-world

Appendix A Time series figures for states and territories

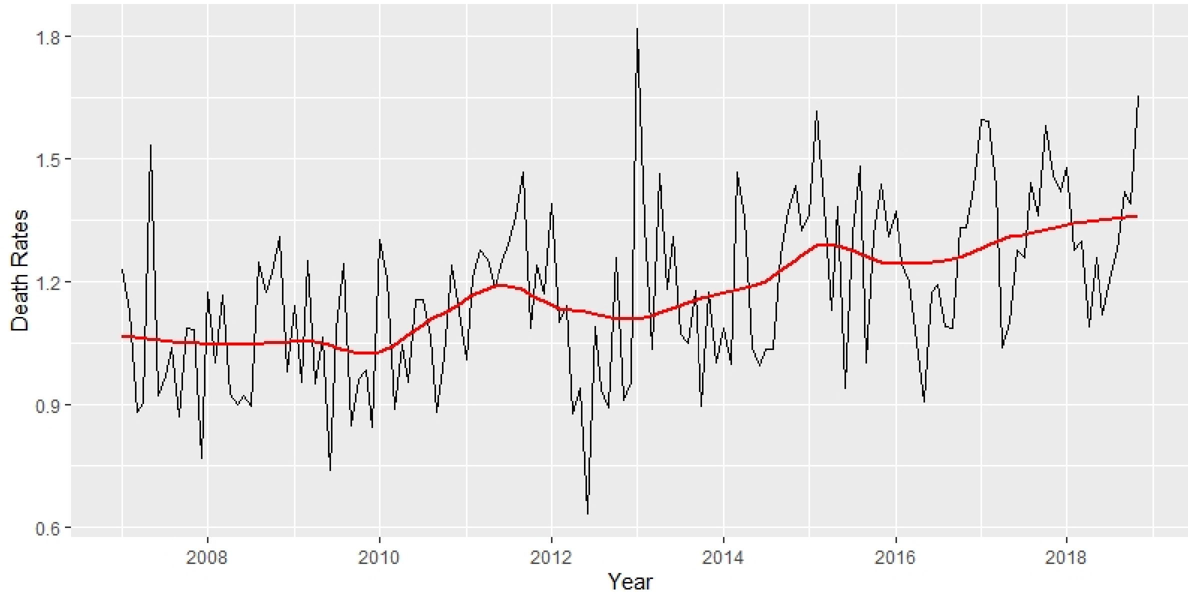
Appendix Figure 1 Monthly Counts of Suicide Deaths in NSW: January 2007 to November 2018



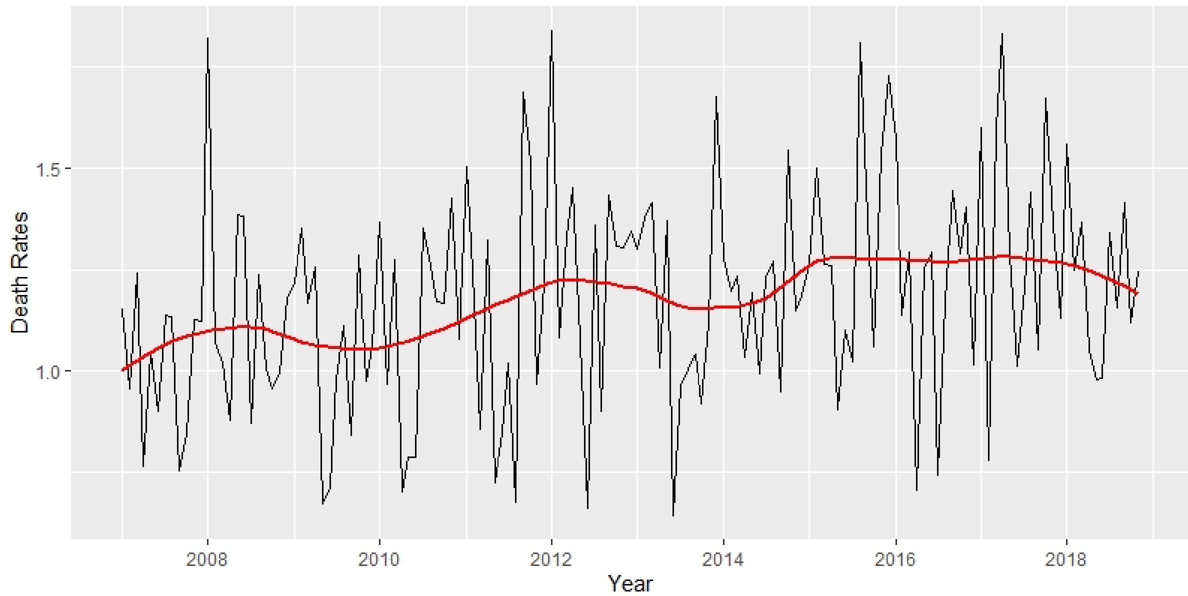
Appendix Figure 2 Monthly Counts of Suicide Deaths in VIC: January 2007 to November 2018



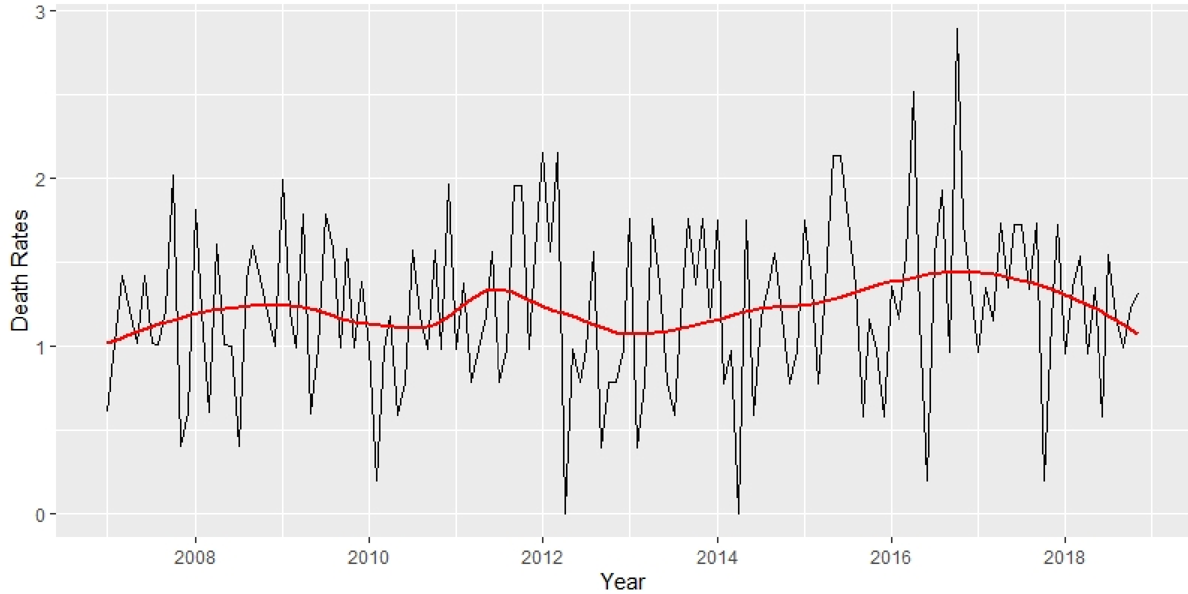
Appendix Figure 3 Monthly Counts of Suicide Deaths in QLD: January 2007 to November 2018



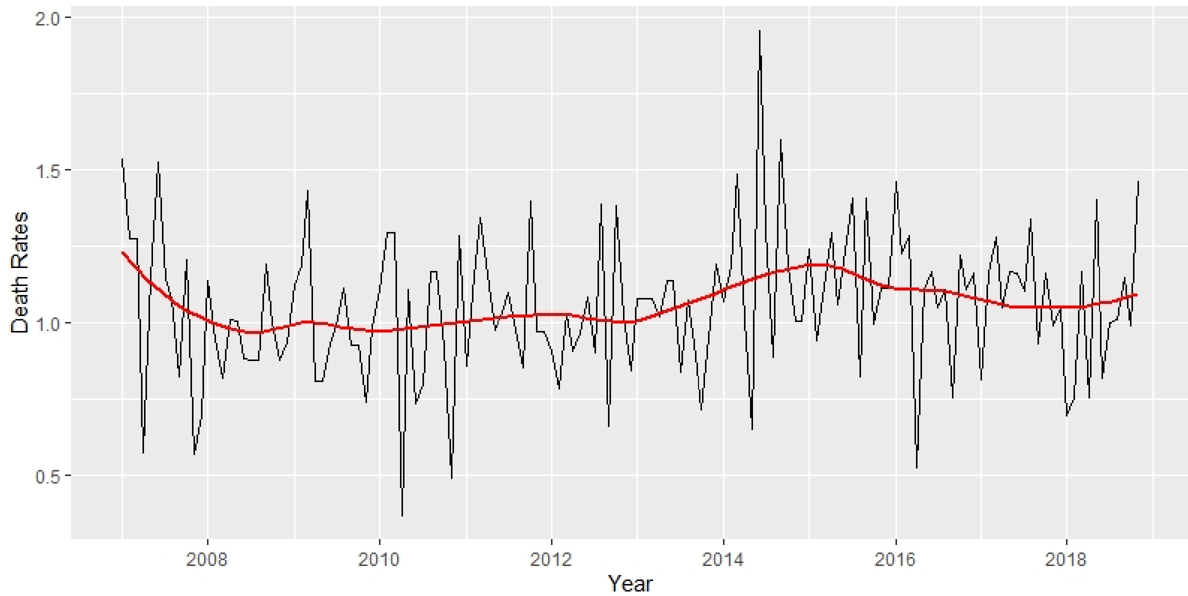
Appendix Figure 4 Monthly Counts of Suicide Deaths in WA: January 2007 to November 2018



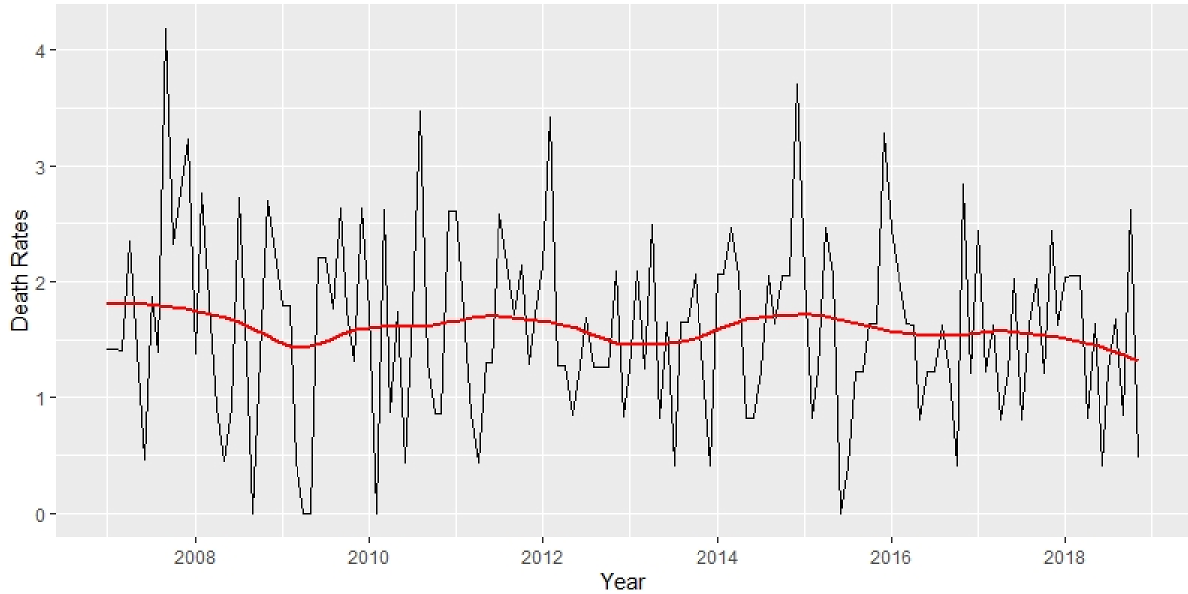
Appendix Figure 5 Monthly Counts of Suicide Deaths in TAS: January 2007 to November 2018



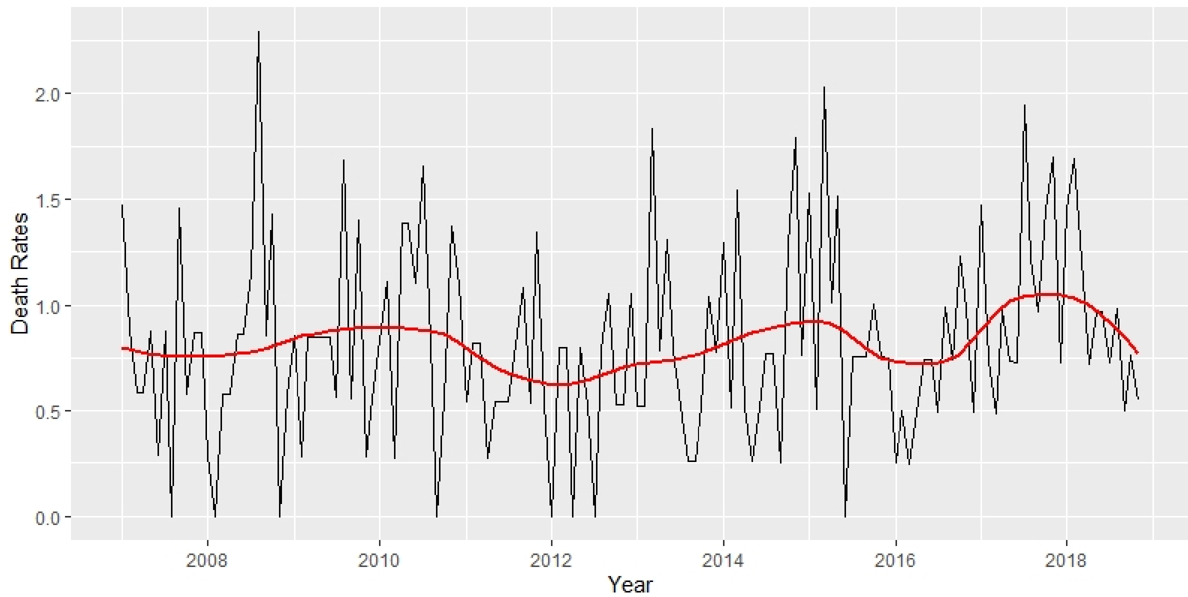
Appendix Figure 6 Monthly Counts of Suicide Deaths in SA: January 2007 to November 2018



Appendix Figure 7 Monthly Counts of Suicide Deaths in NT: January 2007 to November 2018



Appendix Figure 8 Monthly Counts of Suicide Deaths in ACT: January 2007 to November 2018



Appendix B Projection values

Appendix Table 1: Projected monthly suicide death rates per 100,000 population – December 2018 to December 2020

Month	Point projection	80% forecasting range		95% forecasting range	
		Lower	Upper	Lower	Upper
Dec-2018	1.05884720	1.05885	1.05885	1.05885	1.05885
Jan-2019	1.13741520	1.13742	1.13742	1.13742	1.13742
Feb-2019	1.03484820	1.03485	1.03485	1.03485	1.03485
Mar-2019	1.06864660	1.06865	1.06865	1.06865	1.06865
Apr-2019	0.95882740	0.95883	0.95883	0.95883	0.95883
May-2019	0.98875780	0.98876	0.98876	0.98876	0.98876
Jun-2019	0.93740550	0.93741	0.93741	0.93741	0.93741
Jul-2019	1.01724020	1.01724	1.01724	1.01724	1.01724
Aug-2019	1.03774720	1.03775	1.03775	1.03775	1.03775
Sep-2019	1.01818340	1.01818	1.01818	1.01818	1.01818
Oct-2019	1.07515640	1.07516	1.07516	1.07516	1.07516
Nov-2019	1.04787830	1.04788	1.04788	1.04788	1.04788
Dec-2019	1.05884720	1.05885	1.05885	1.05885	1.05885
Jan-2020	1.13741520	1.13742	1.13742	1.13742	1.13742
Feb-2020	1.03484820	1.03485	1.03485	1.03485	1.03485
Mar-2020	1.06864660	1.06865	1.06865	1.06865	1.06865
Apr-2020	0.95882740	0.95883	0.95883	0.95883	0.95883
May-2020	0.98875780	0.98876	0.98876	0.98876	0.98876
Jun-2020	0.93740550	0.93741	0.93741	0.93741	0.93741
Jul-2020	1.01724020	1.01724	1.01724	1.01724	1.01724
Aug-2020	1.03774720	1.03775	1.03775	1.03775	1.03775
Sep-2020	1.01818340	1.01818	1.01818	1.01818	1.01818
Oct-2020	1.07515640	1.07516	1.07516	1.07516	1.07516
Nov-2020	1.04787830	1.04788	1.04788	1.04788	1.04788
Dec-2020	1.05884720	1.05885	1.05885	1.05885	1.05885