# Analysis of death data during the Morwell mine fire

## Summary

The updated analyses gives a 79% to 82% probability of an increase in deaths during the two months of the fire. This is similar to the 80% to 89% probability from the previous analysis. The reduction in probability is because the two additional postcodes (3869 and 3870) showed a slight reduction in death risk.

Allowing the effect of the fire to vary by postcode gives a 94% probability of an increase in deaths in postcode 3844. The highest risks of death were in postcodes 3842 and 3844. There was little to choose statistically between a model with a fixed or varying fire effect across postcodes.

## Introduction

This document contains my analysis of the Morwell mine fire data. This is an updated analysis using data from more postcodes for the years 2004 to 2014. Details on the methods can be found in my original analysis available here: `http://eprints.qut.edu.au/76230/`.

I am happy for this document to be freely shared. I am also happy to answer questions via e-mail: a.barnett@qut.edu.au.

## Methods

### Data

The data were monthly numbers of deaths from 2004 to 2014 for the months of January to December (December data were not available for 2014). The deaths were split by six postcodes (3840, 3842, 3825, 3844, 3869 and 3870) according to usual place of residence. The 11 years, 12 months (11 in 2014) and six postcodes gives 786 observations. There were 6,421 deaths in total.

The previous data were: from 2009 to 2014; only included the months January to June; only included 4 postcodes (3840, 3842, 3825 and 3844); and had 1,811 deaths in total.

### Statistical methods

The statistical methods were as per the previous analysis (`http://eprints.qut.edu.au/76230/`) except for the following differences:

1. Fitting temperature as a non-linear effect using a linear and quadratic term. This is because the new data includes all 12 months (the previous data had just six months) and hence we need to model an increased risk of death during both high and low temperatures.

2. An additional analysis using a model that allowed the effect of the fire to vary over the six postcodes. This was based on qualitative evidence about some differences between postcodes in exposures and evacuations. Hence we might expect the effect of the fire to vary over the six postcodes.

3. Using the deviance information criterion (DIC) to compare the models. The DIC compares the fit of the model (that is, how well it explains the observed number of deaths) and includes a penalty for more complex models. Hence it will hopefully find the most simple explanation that best fits the data.

# Results

## Plots

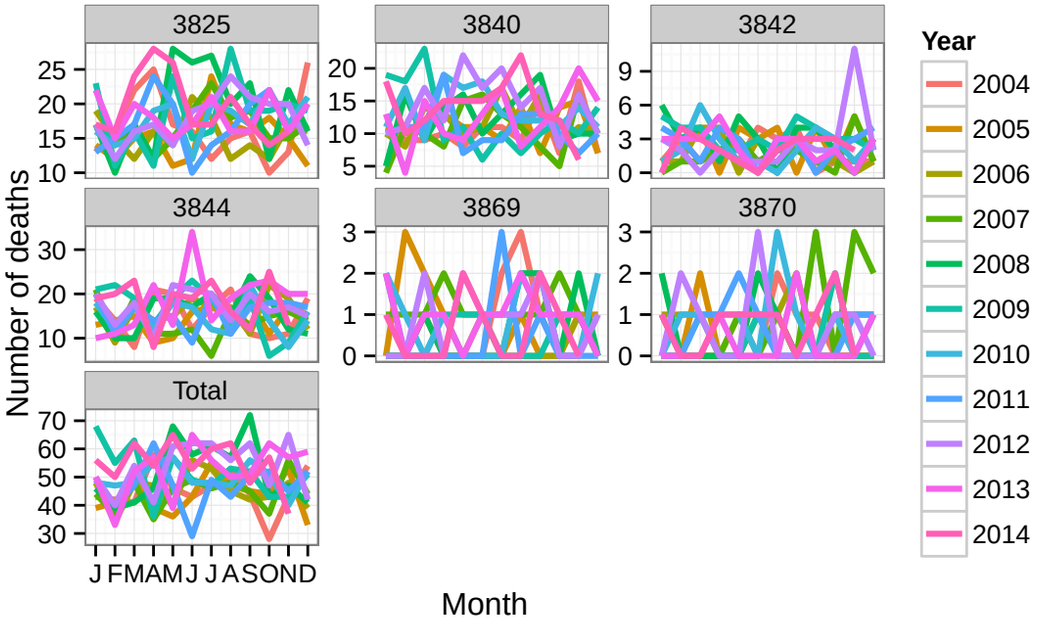


Figure 1: Death numbers by month and year in each postcode and the total number of deaths across the six postcodes. The scales on the $y$-axes differ between postcodes.

There were relatively large spikes in deaths in June 2013 in postcode 3844 and November 2012 in postcode 3842 (Figure 1). The differences in numbers on the $y$-axes between panels in Figure 1 are because some postcodes are larger than others.

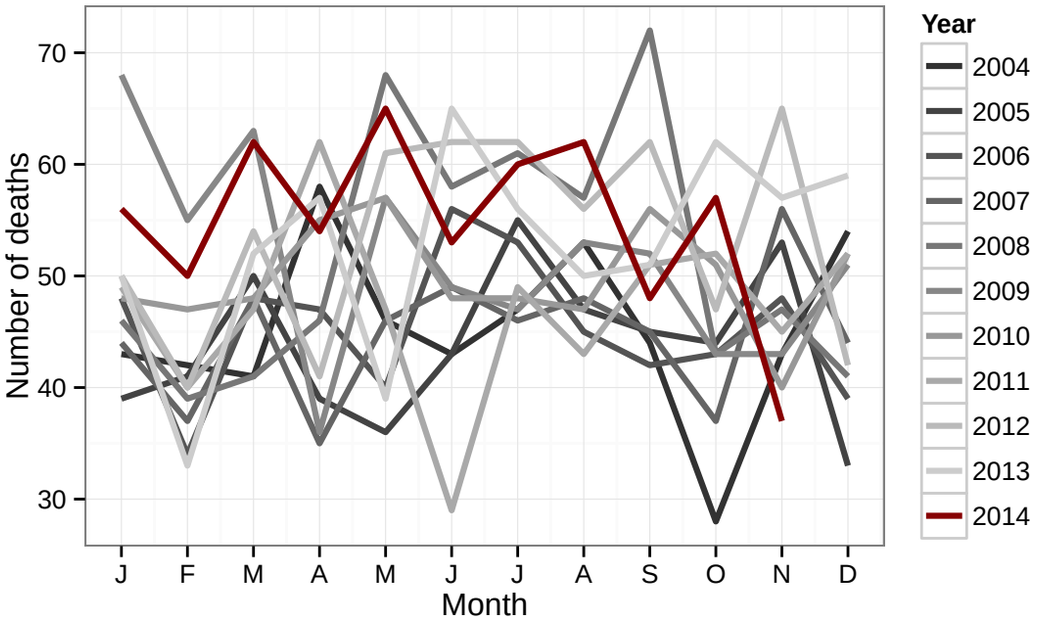


Figure 2: Total deaths across all six postcodes by month and year. The results for 2014 are highlighted in dark red.

Looking at the totals (Figure 2), the deaths in 2014 in February and April do appear to be high. Another year with high deaths rates is 2009 and this may be due to bushfires and extreme heat that summer.

## Statistical model results

Table 1: Estimates without adjusting for temperature. Statistics are the mean and lower and upper 95% credible interval. For the effect of the fire the P-value column gives the probability that the risk of death was increased. Estimates are on a log scale except for the relative risks and absolute number of deaths.

| | Mean | Lower | Upper | P-value |
|---|---|---|---|---|
| Intercept | −0.633 | −0.691 | −0.577 | |
| Trend | 0.016 | 0.008 | 0.024 | |
| Postcode 3825 | 1.524 | 1.459 | 1.590 | |
| Postcode 3840 | 1.162 | 1.093 | 1.231 | |
| Postcode 3842 | −0.504 | −0.613 | −0.396 | |
| Postcode 3844 | 1.431 | 1.366 | 1.498 | |
| Postcode 3869 | −1.740 | −1.922 | −1.564 | |
| Postcode 3870 | −1.873 | −2.070 | −1.688 | |
| Season, cos | −0.058 | −0.093 | −0.023 | |
| Season, sin | 0.005 | −0.030 | 0.039 | |
| Fire | 0.082 | −0.117 | 0.275 | 0.79 |
| Fire, relative risk | 1.090 | 0.890 | 1.316 | |
| Absolute deaths | 0.739 | −0.899 | 2.583 | |

The probability that the death rate was higher than the average during the fire is 0.79. This means that the probability that the death rate was not higher than the average during the fire is 0.21. The mean increase in deaths is as a relative risk is 1.09, or 9 as a percentage. The absolute number of deaths per postcode per month is 0.7, which over 6 postcodes and 2 months is 8.4.

The results after adjusting for temperature are in Table 2. The probability that the death rate was higher than the average during the fire is 0.82. The mean increase in deaths is as a relative risk is 1.1, or 10 as a percentage. The absolute number of deaths per postcode per month is 0.8, which over 6 postcodes and 2 months is 9.6.

Table 2: Estimates after adjusting for monthly temperatures. Statistics are the mean and lower and upper 95% credible interval. For the effect of the fire the P-value column gives the probability that the risk of death was increased. Estimates are on a log scale except for the relative risks and absolute number of deaths.

| | Mean | Lower | Upper | P-value |
|---|---|---|---|---|
| Intercept | −0.650 | −0.717 | −0.584 | |
| Trend | 0.016 | 0.008 | 0.024 | |
| Postcode 3825 | 1.524 | 1.458 | 1.589 | |
| Postcode 3840 | 1.161 | 1.093 | 1.230 | |
| Postcode 3842 | −0.503 | −0.613 | −0.397 | |
| Postcode 3844 | 1.431 | 1.365 | 1.498 | |
| Postcode 3869 | −1.739 | −1.924 | −1.565 | |
| Postcode 3870 | −1.873 | −2.066 | −1.688 | |
| Season, cos | 0.010 | −0.119 | 0.139 | |
| Season, sin | 0.005 | −0.031 | 0.040 | |
| Fire | 0.093 | −0.111 | 0.291 | 0.82 |
| Fire, relative risk | 1.103 | 0.895 | 1.337 | |
| Absolute deaths | 0.843 | −0.857 | 2.754 | |
| Temperature, linear | −0.010 | −0.028 | 0.009 | |
| Temperature, quadratic | 0.001 | −0.001 | 0.002 | |

## Varying effect of the fire over postcodes

Table 3: Estimates of a varying effect of the fire. Including adjustment for monthly temperatures. Statistics are the mean and lower and upper 95% credible interval. For the effect of the fire the P-value column gives the probability that the risk of death was increased. Estimates are on a log scale except for the relative risks and absolute number of deaths.

| | Mean | Lower | Upper | P-value |
|---|---|---|---|---|
| Intercept | −0.648 | −0.714 | −0.581 | |
| Trend | 0.016 | 0.008 | 0.024 | |
| Postcode 3825 | 1.521 | 1.457 | 1.587 | |
| Postcode 3840 | 1.163 | 1.095 | 1.232 | |
| Postcode 3842 | −0.509 | −0.616 | −0.399 | |
| Postcode 3844 | 1.426 | 1.359 | 1.493 | |
| Postcode 3869 | −1.734 | −1.916 | −1.559 | |
| Postcode 3870 | −1.867 | −2.062 | −1.683 | |
| Season, cos | 0.008 | −0.121 | 0.138 | |
| Season, sin | 0.005 | −0.031 | 0.041 | |
| Fire 3825 | 0.079 | −0.253 | 0.386 | 0.69 |
| Fire 3840 | −0.157 | −0.594 | 0.243 | 0.24 |
| Fire 3842 | 0.228 | −0.542 | 0.906 | 0.74 |
| Fire 3844 | 0.246 | −0.069 | 0.552 | 0.94 |
| Fire 3869 | −0.810 | −3.057 | 0.573 | 0.16 |
| Fire 3870 | −0.766 | −3.054 | 0.611 | 0.18 |
| Fire, relative risk 3825 | 1.097 | 0.777 | 1.471 | |
| Fire, relative risk 3840 | 0.874 | 0.552 | 1.275 | |
| Fire, relative risk 3842 | 1.343 | 0.581 | 2.473 | |
| Fire, relative risk 3844 | 1.295 | 0.933 | 1.737 | |
| Fire, relative risk 3869 | 0.614 | 0.047 | 1.773 | |
| Fire, relative risk 3870 | 0.642 | 0.047 | 1.843 | |
| Absolute deaths 3825 | 1.687 | −3.894 | 8.208 | |
| Absolute deaths 3840 | −1.528 | −5.441 | 3.334 | |
| Absolute deaths 3842 | 0.788 | −0.962 | 3.385 | |
| Absolute deaths 3844 | 4.693 | −1.064 | 11.719 | |
| Absolute deaths 3869 | −0.256 | −0.633 | 0.514 | |
| Absolute deaths 3870 | −0.207 | −0.553 | 0.489 | |

The estimated risks of the fire in each postcode are in Table 3 and Figure 3. Three of the six postcodes had a decreased mean risk of death. Postcode 3844 had the largest probability that the death rate was increased of 0.94.

The relative risk was similar in postcodes 3842 and 3844 (Figure 3), but there was more certainty in postcode 3844 as the credible intervals were narrower. This is because 3844 has a larger population than 3842.

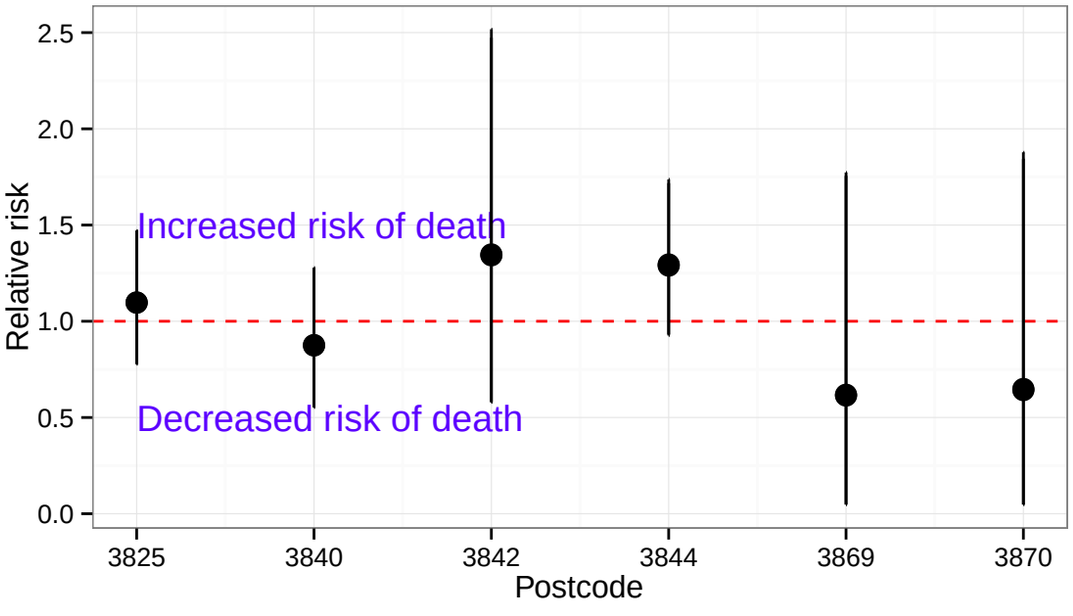


Figure 3: Mean relative risk of deaths due to fire and 95% credible intervals by postcode.

## Choosing the best model (deviance information criterion)

Table 4: Deviance information criterion (DIC) and estimated number of parameters (pD). The lower the DIC the better the model.

| model | pD | DIC |
|---|---|---|
| No weather adjustment, fixed effect of fire across postcodes | 10.0 | 3253.1 |
| Weather adjustment, fixed effect of fire across postcodes | 11.9 | 3255.6 |
| No weather adjustment, varying effect of fire across postcodes | 13.5 | 3253.9 |
| Weather adjustment, varying effect of fire across postcodes | 15.6 | 3256.7 |

The best model according to the DIC is with no weather adjustment and with a fixed effect of the fire across postcodes (Table 4). However, a difference in the DIC of less than 1 is small, and hence there is little to choose between a model with a fixed and varying effect of the fire.

## Residual plots

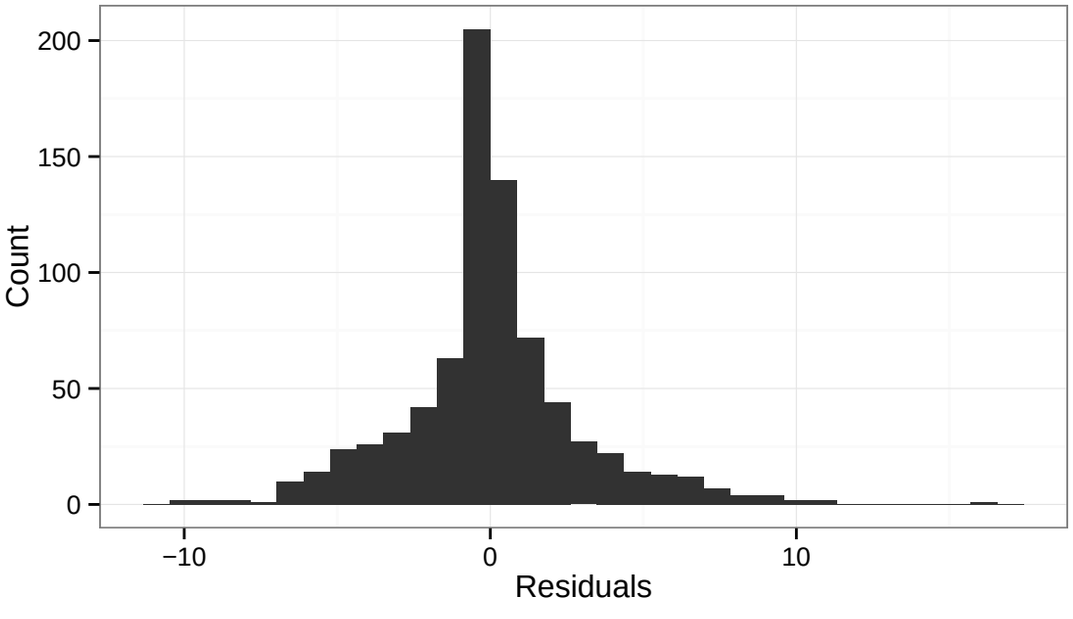


Figure 4: Residuals from the model with no weather adjustment and with a fixed effect of the fire across postcodes.

The residuals are approximately normally distributed (Figure 4). The was one large positive residual which was in postcode 3844 in June 2013 and was 34 deaths when the model predicted only 18. This large spike in deaths was identified in the plots and may have been due to a cold spell and/or flu outbreak.

### Better data

A more accurate analysis could be provided by using more accurate data. This would include:

- Using daily death numbers rather than monthly numbers
- Knowing the cause of death
- Knowing the age of death

Having this information would increase the certainty of any association between the fire and death.