Online Supporting Information for the article: "Suicide and Drought in NSW, Australia, 1970-2007". Unabridged

Ivan C. Hanigan[1,2] Colin D. Butler[1] Philip N. Kokic[2]
Michael F. Hutchinson[3]

[1]National Centre for Epidemiology and Population Health, Australian National University
[2]Commonwealth Scientific and Industrial Research Organisation
[3]Fenner School of Environment and Society, Australian National University.

Contents

1 Introduction

This document accompanies the R code at this website `https://github.com/ivanhanigan/SuicideAndDroughtInNSW` to calculate the Hutchinson Drought Index and fit the regression models for the paper 'Suicide and Drought in New South Wales (NSW), Australia, 1970-2007'. The calculation of the Drought Index is demonstrated using free data from the Australian Bureau of Meteorology. The suicide mortality data are not publicly available due to confidentiality restrictions. The R code we ran the regressions with is included but the original data are only available for authorised users approved by the Australian Bureau of Statistics and the NSW Registrar of Births Deaths and Marriages.

2 Drought Index

The R code includes a demonstration of the Hutchinson Drought Index [1]. This climatic drought index is shown graphically for a location in the 'Central West' SD of NSW in Figure 1.

Instructions for using R to download and analyse the spatial data from the Australian Bureau of Statistics (`http://www.abs.gov.au`) and the weather data from the Australian Bureau of Meteorology (`http://www.bom.gov.au`) websites are included.

2.1 Calculate the Drought Index

The Drought index is shown in Figure 1 for the SD of 'Central West NSW' during a period which includes a strong drought (1979-83). The raw monthly rainfall totals are integrated to rolling 6-monthly totals (both shown in first panel) which are then ranked into percentiles by month and this is rescaled to range between -4 and +4 in keeping with the range of the Palmer Index [2] (second panel). Mild drought is below -1 in the Palmer index and so consecutive months below this threshold are counted. In the original method 5 or more consecutive months was defined as the beginning of a drought, which continued until the rescaled percentiles exceed -1 again (third panel). The enhanced method imposes a more conservative threshold of zero (the median) to break a drought (fourth panel).

There was also an alternative method devised by Hutchinson where the rescaled percentile values are integrated using conditional cumulative sums. That method is included in the R code however we decided not to use it in this study because the counting method is simpler and gives similar results.

2.2 The Summation Method

When the index is calculated using the sum of each consecutive month's rainfall deficiency score the resulting measure addresses the question of how intense the drought is, rather than just the duration which is provided by the counting method. This version of the index is shown in Figure 2.

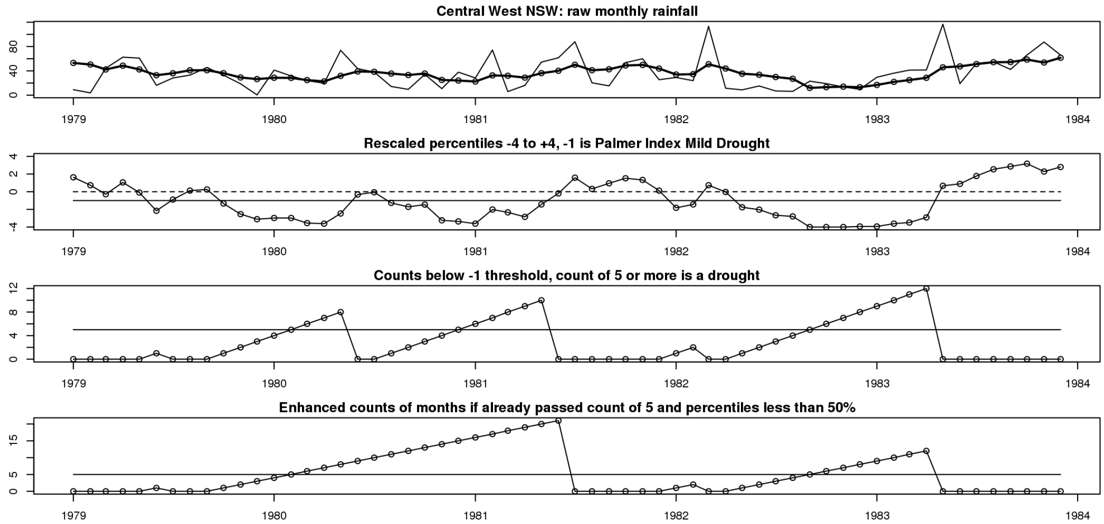

Figure 1: The Drought index in Central West NSW with the enhanced method shown in the fourth panel.

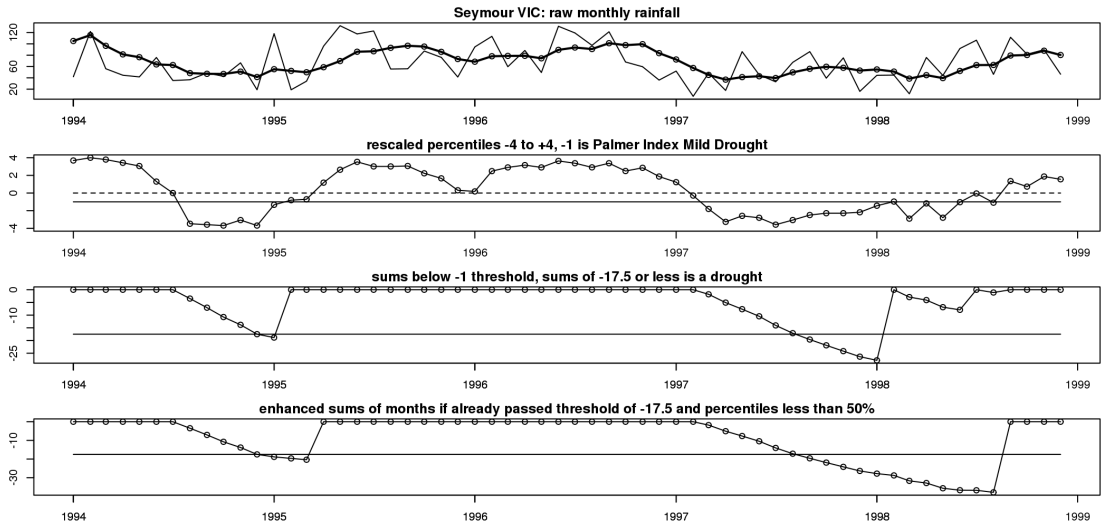

Figure 2: SeymourDrought9499enhanced.png

3 Suicide and Drought Modeling

3.1 Descriptive Statistics of Drought and Suicide

Descriptive statistics for the Drought Index are shown in Table 1. Summary statistics for Suicide rates are shown in Table 2.

Table 1: Descriptive statistics for the drought index

SD group	N droughts	Avg Duration	Max Duration
1 Central West	9	8	12
2 Hunter	11	7	15
3 Illawarra	7	9	16
4 Mid-North Coast	8	8	15
5 Murray	7	8	11
6 Murrumbidgee	10	7	11
7 North and Far Western	8	7	12
8 Northern	5	8	11
9 Richmond-Tweed	13	8	17
10 South Eastern	8	8	11
11 Sydney	9	9	20

Table 2: Descriptive statistics for suicide (PYL = Person Years Lived)

SD group	Avg Death/Month	Avg Pop	Rate/100000 PYL
1 Central West	2	138202	13
2 Hunter	5	430403	13
3 Illawarra	3	280037	13
4 Mid-North Coast	2	183521	12
5 Murray	1	86221	14
6 Murrumbidgee	1	118778	13
7 North and Far Western	2	114460	16
8 Northern	2	146465	14
9 Richmond-Tweed	2	139356	14
10 South Eastern	2	135091	14
11 Sydney	34	3040952	13

3.2 Correlation between Temperature and Drought

We found that monthly maximum temperature variables are not strongly correlated with the drought index in our dataset. Correlation coefficients for the variables are shown in Table 3.

Table 3: Correlations

Variables	Correlation
cor(logDroughtCount,tmax)	0.05
cor(tmax,tmaxanomaly)	0.23
cor(logDroughtCount,tmaxanomaly)	0.35

3.3 Core Model Diagnostics and Variable Selection

We initially fitted age stratified time series Poisson Generalized Linear Models (GLMs). We identified a Core Model that included age, sex, region, season and long term trend. We assessed standard model diagnostics for this. Then we used Generalized Additive Models (GAMs) with the automatic estimation of the optimal amount of smoothing on the drought index using penalised regression splines from the R package: mgcv [3]. These estimated smooths were then explored in GLMs. Many models were fitted to test different combinations of variables. The models are ranked by their Bayesian Information Criterion (BIC) scores in Table 4 (AIC is shown for interest).

Diagnostic plots of the core model are shown in Figure 3.

The effect estimates for initial models of climate are shown in Figure 4. The drought effect was found to be complicated by the countervailing effects in men and women during subsequent modelling.

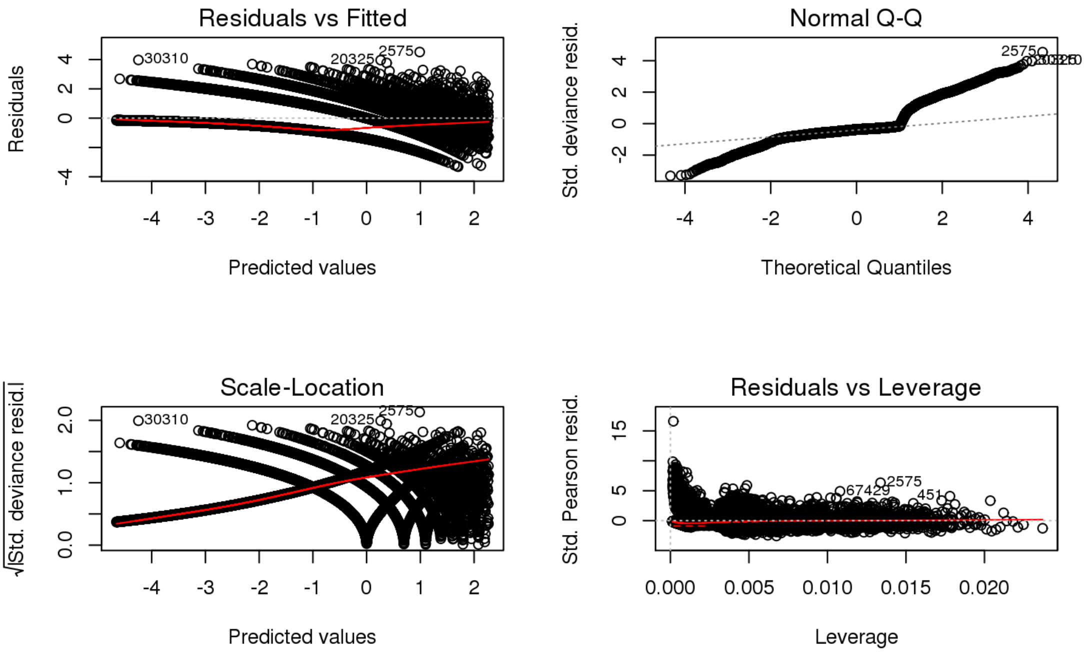

Figure 3: coreModelDiagnosticPlot.png

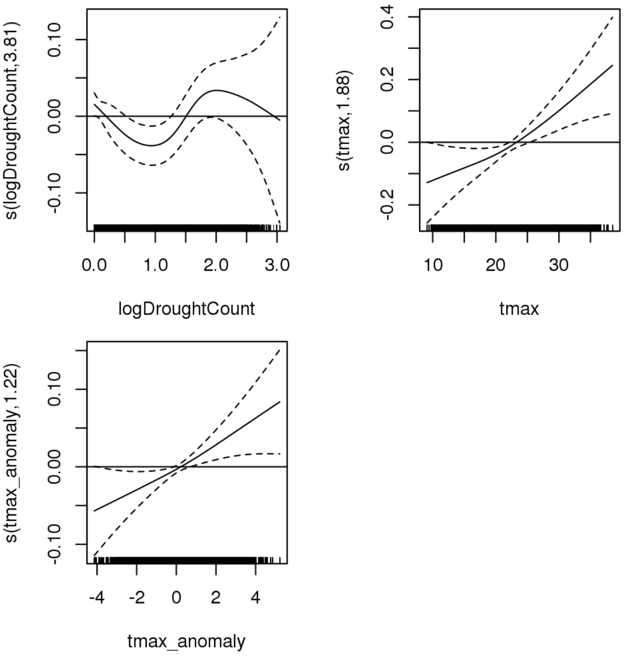

Figure 4: droughtTmaxAnomGAMS.png

Table 4: Models ranked by Bayesian Information Criterion (BIC).

Model	Parameters	BIC	AIC
sd_group*sex	78	69715	69001
age*sex*ns(time,df=3)	68	69814	69191
tmaxanomModel	69	69816	69184
tmax_anomaly*sex	70	69827	69186
tmaxModel	71	69830	69180
ns(tmax,3)*sex	74	69831	69154
ageSexTrendSineXtra	70	69835	69194
droughtModel	73	69845	69186
tmax_anomaly*ns(time,3)	72	69847	69188
sd_group*ns(time,3)	98	69859	68962
With Rural 30-49 Sex Strata	98	69869	69127
Without Rural 30-49 Sex Strata	97	69873	69141
ns(tmax,3)*tmax_anomaly	75	69874	69187
tmax_anomaly*agegp	75	69877	69191
ns(logDroughtCount,5)*tmax_anomaly	79	69884	69179
ns(logDroughtCount,5)*sex	78	69885	69189
tmax_anomaly*sd_group	79	69904	69181
ns(tmax,3)*ns(time,3)	80	69909	69176
interactionDrtAgeSexRuralModel2	188	69937	69091
ns(logDroughtCount,5)*ns(time,3)	88	69963	69194
ns(logDroughtCount,5)*ns(tmax,3)	91	69968	69171
ns(tmax,3)*agegp	89	70012	69197
sd_group*sex*ns(time,3)	138	70075	68812
ns(logDroughtCount,5)*agegp	103	70080	69201
ns(tmax,3)*sd_group	101	70117	69192
ns(logDroughtCount,5)*sd_group	123	70250	69225
agegp*sd_group	128	70347	69175
sd_group*age*sex*ns(time,df=3)	618	74801	69143

3.4 Suicide and Drought Model by Age, Sex and Region

Our final GAM estimated curved response functions for drought and suicide by age, sex and region are shown in Figure 5. This model is labelled 'interactionDrtAgeSexRuralModel2' in Table 4. It included drought effects for each age/sex/region subgroup:

$$
\begin{aligned}
log(O_{ijk}) \quad = \quad & s(Drought \times Sex \times AgeGroupBy20years \times RuralOrUrbanRegion) \\
& + AgeGroupBy10years_i \times Sex_j \times s(Time, df = 3, basis = NaturalCubicSpline) \\
& + StatisticalDivision_k \\
& + s(Month, df = 4, basis = CyclicCubicSpline) \\
& + s(tmaxAnomaly) \\
& + offset(log(Pop_{ijk}))
\end{aligned}
$$

Where:

O_{ijk} = monthly suicide counts by AgeGroupBy10years$_i$, Sex$_j$ and StatisticalDivision$_k$

s(Drought × Sex × AgeGroupBy20years × RuralOrUrbanRegion) are interaction effects

Time = the month number in the sequence from Jan-1970 until Oct-2007

Month = the months of the year ranked from 1 to 12

s( ) = penalized regression splines, degrees of freedom (df) may be specified

tmaxAnomaly = monthly averaged temperature maxima anomalies from long term averages

Pop_{ijk} = interpolated population by month in each group

The eleven regions were classified as rural or urban based on the locations of the three major cities of NSW: Sydney, Newcastle and Wollongong. All other regions were classed as rural.

The estimated degrees of freedom from the GAM were then used with parametric splines in a GLM to estimate the effect sizes. A key drought effect reported in the paper was for rural males aged 30-49 where an Interquartile Range (IQR) rise in drought index gave a Relative Risk (RR) of 1.15 (95CI 1.08 to 1.22). The IQR for the drought index is about 2 months. For the temperature anomaly term there was a RR of 1.03 (95CI 1.01 to 1.05) per IQR rise (1.6 degrees C).

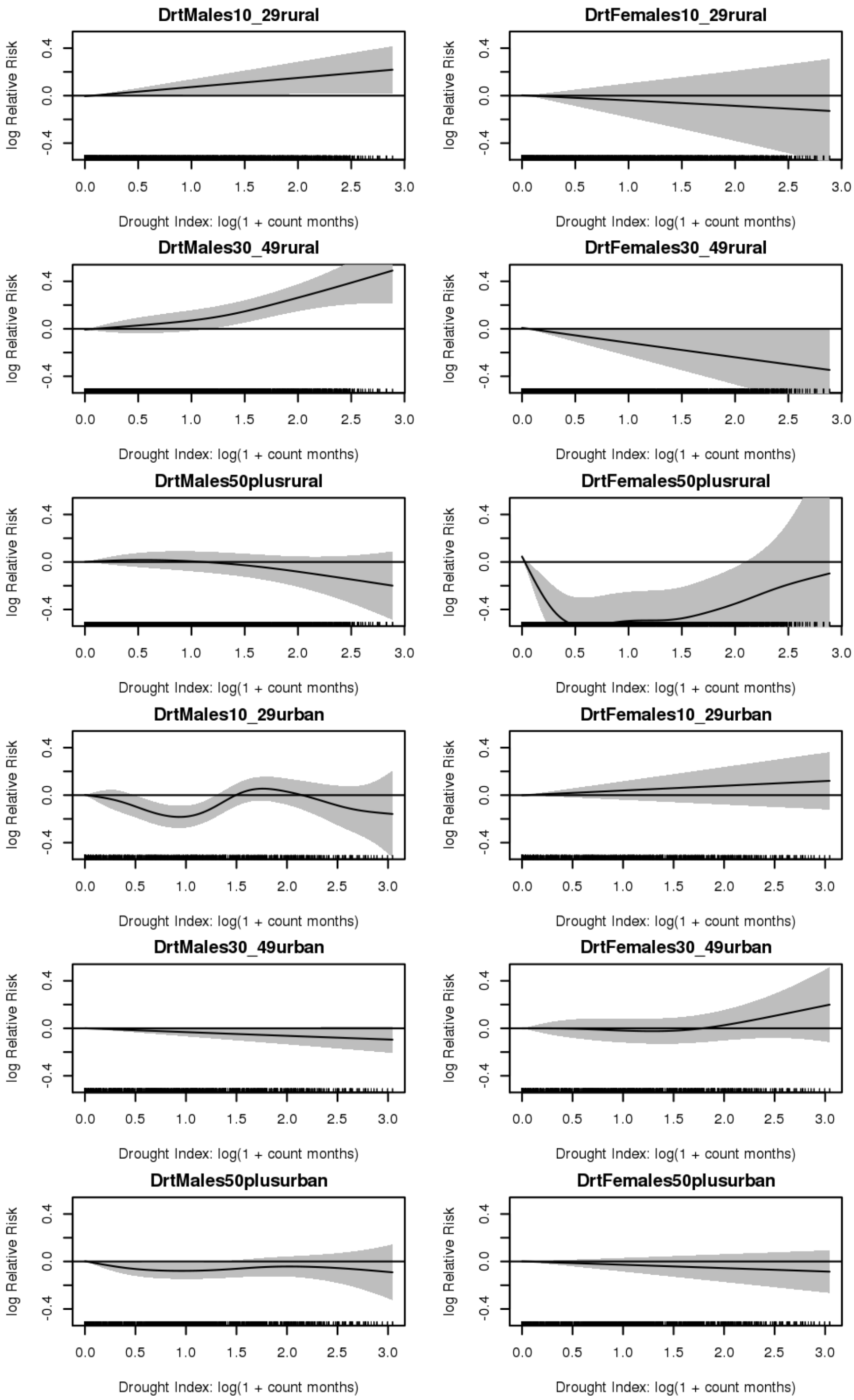

Figure 5: Estimated response functions for suicide and drought in each of the subgroups.

3.5 Attributable Number of Deaths

The predicted number of rural male suicides aged 30-49 per annum associated with droughts over our study period was 4.01 (95%CI 2.14 to 6.05, p = 0.000015), accounting for 9% of the total in 38 years.

However this effect only applies in the months that were in drought, and to a greater extent depending on the intensity of the drought. As drought is a rare and episodic event this estimate is obviously an underestimate of the real impact in terms of numbers of deaths during droughts and potential years of life lost.

The predicted number of rural female suicides aged 30-49 per annum associated with droughts are estimated for comparison with the figure for males. The decreased number of rural female suicides aged 30-49 per annum associated with droughts over our study period was -0.72 (95%CI -1.32 to -0.01, p = 0.041787).

3.6 Test the Sex Stratification

To find out if the inclusion of a seperate term for Rural Males and Rural Females aged 30-49 is warranted we performed a likelihood ratio test with an alternative model where the drought effect was not stratified by sex. The model was significantly better when including the Rural 30-49 sex stratification (likelihood ratio test p = 0.000077).

4 Sensitivity Analyses

4.1 Enhanced Drought Index

We conducted sensitivity analyses for the drought exposure variable. The drought index was enhanced with the threshold needed to end a drought made more stringent. For example in Figure 1 the drought in 1980 would not have ended in the middle of that year given the new threshold but would have continued into 1981 (the fourth panel).

The drought effects estimated were similar to those from our previous modeling.

The key effect estimates are shown for the enhanced drought index in Figure 6.

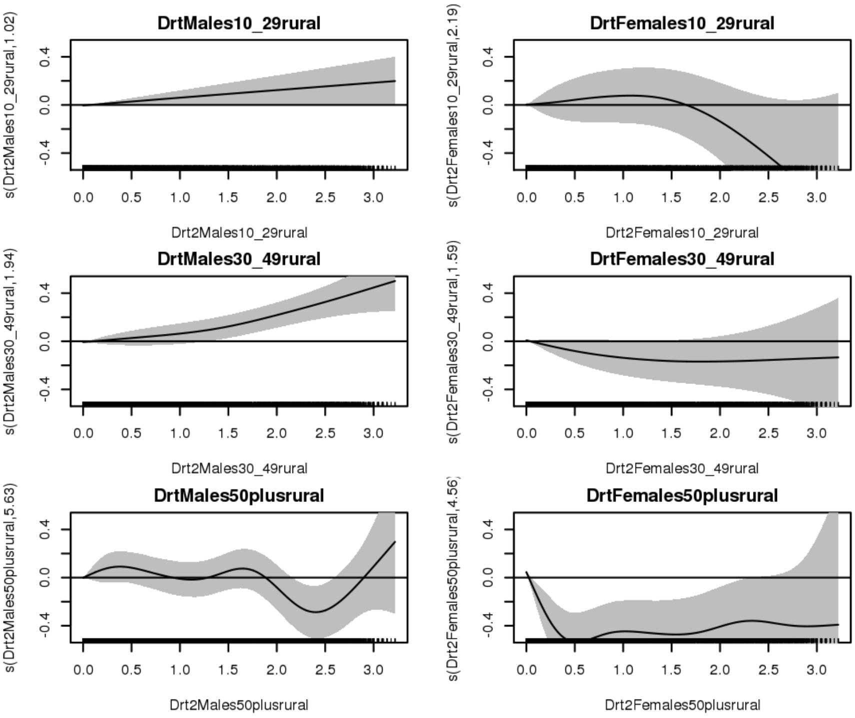

Figure 6: interactionDrtAgeSexRuralModel2enhanced.png

4.2 Self-harm Coded as Undetermined

A sensitivity analysis was conducted that combined the suicide deaths with deaths coded as 'Self inflicted injury, undetermined if intentional'. This analysis agreed with our previous modelling.

The key effect estimates for the drought index effect on Suicides Plus Undetermined are shown in Figure 7.

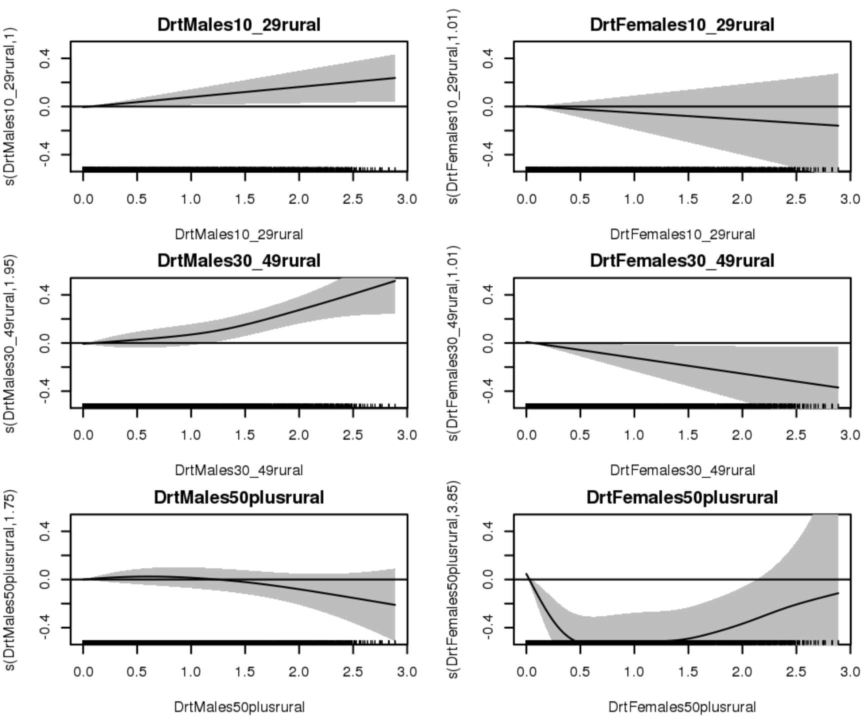

Figure 7: interactionDrtAgeSexRuralModel2SuicidePlusUndetermined.png

4.3 Drop High Leverage Points

A sensitivity analysis was finally conducted that dropped any observations identified as having high leverage. Dropping these observations from the final model produced effect estimates that also agreed with our prior modeling results

A diagnostic plot of the leverage and residuals is shown in Figure 8. Dropping observations with high leverage produced effect estimates that also agreed with our prior modeling results (Figure 9).

4.4 plot check

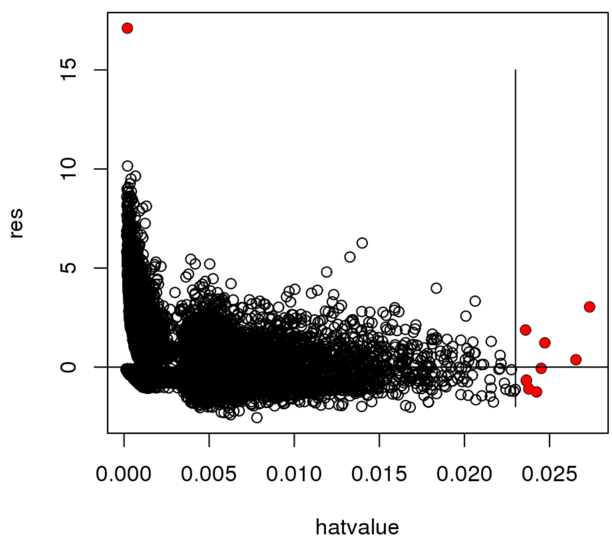

Figure 8: interactionDrtAgeSexRuralModel3checkLeverage.png

4.5 plot do

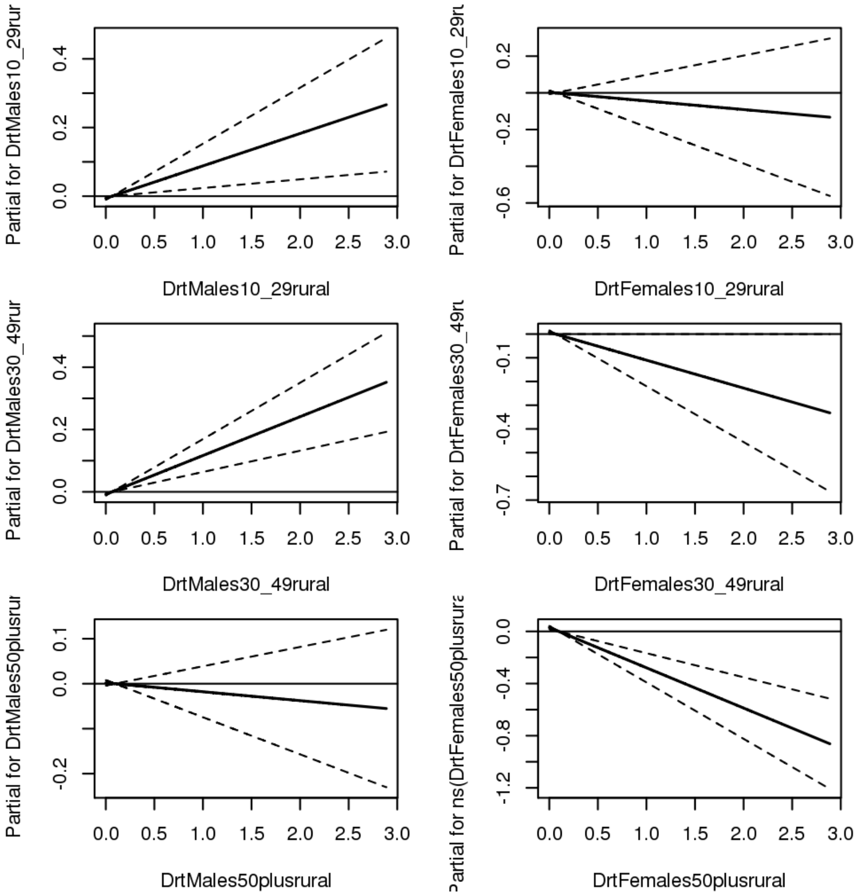

Figure 9: interactionDrtAgeSexRuralModel3noLeverage.png

References

[1] Smith, D. I, Hutchinson, M. F, & McArthur, R. J. (1992) *Climatic and Agricultural Drought: Payments and Policy.* (Centre for Resource and Environmental Studies, Australian National University, Canberra, Australia).

[2] Palmer, W. (1965) *Meteorological drought. Research paper No. 45.* (U.S. Department of Commerce Weather Bureau, Washington, D.C.).

[3] Wood, S. (2008) Fast stable direct fitting and smoothness selection for generalized additive models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **70**, 495–518.

5 Code for Figures 1 and 2

The R codes to fit the model and display the exposure-response relationships for the drought index on rural suicides (Figures 1 and 2 of the paper) are shown below using Sweave.

```
> ######################
> #do,  show model fig1 and 2
> ######################
>
>
> # first fit the model
>         interactionDrtAgeSexRuralModel2 <- gam(deaths ~  s(mm, k=3, fx=T, bs = 'cp')
+         + s(DrtMales10_29rural)
+         + s(DrtMales30_49rural)
+         + s(DrtMales50plusrural)
+         + s(DrtFemales10_29rural)
+         + s(DrtFemales30_49rural)
+         + s(DrtFemales50plusrural)
+         + s(DrtMales10_29urban)
+         + s(DrtMales30_49urban)
+         + s(DrtMales50plusurban)
+         + s(DrtFemales10_29urban)
+         + s(DrtFemales30_49urban)
+         + s(DrtFemales50plusurban)
+         + s(tmax_anomaly)
+         + agegp2
+         + rural
+         + sd_group
+         + sex
+         + agegp
+         + agegp*sex*ns(time,df = 3)
+         + offset(log(pop)), data=data,family=poisson)
>
```

The code to create this graph is shown next:

```
> ######################
> #do,  show plot fig 1 and 2
> ######################
>
>
>
>         # now make a plot of each group effects
>
>         png('RuralMales20.png',res=200,width = 600, height = 1000)
>         layout(matrix(c(1:4),ncol=1),heights=c(1,1,1,0.2))
>         par(mfrow=c(4,1), mar=c(0.1,4,1.5,0.5), cex=.7)
>         plot(interactionDrtAgeSexRuralModel2,select=2,se=T, ylim = c(-0.8,0.8), shade=TRUE,shad
>         abline(0,0)
>         title('Rural Males aged 10-29', cex=.5, font.main = 1)
>         plot(interactionDrtAgeSexRuralModel2,select=3,se=T, ylim = c(-0.8,0.8), shade=TRUE,shad
>         abline(0,0)
>         title('Rural Males aged 30-49', cex=.5, font.main = 1)
>         plot(interactionDrtAgeSexRuralModel2,select=4,rug=F,se=T, ylim = c(-0.8,0.8), shade=TRU
>         abline(0,0)
>         title('Rural Males aged 50 plus', cex=.5, font.main = 1)
>         par(mar=c(1,4,6,0.5))
>         plot(1,1,type = 'n', xaxt = 'n', yaxt='n',ylab='',xlab='', axes = F)
>         title(main = 'Drought Index: log(1 + count months)',  font.main = 1,cex.main=.9)
>         dev.off()
>         png('RuralFemales20.png',res=200,width = 600, height = 1000)
>         layout(matrix(c(1:4),ncol=1),heights=c(1,1,1,0.2))
>         par(mfrow=c(4,1), mar=c(0.1,4,1.5,0.5), cex=.7)
>         plot(interactionDrtAgeSexRuralModel2,select=5,se=T, ylim = c(-0.8,0.8), shade=TRUE,shad
>         abline(0,0)
>         title('Rural Females aged 10-29', cex=.5, font.main = 1)
>         plot(interactionDrtAgeSexRuralModel2,select=6,se=T, ylim = c(-0.8,0.8), shade=TRUE,shad
>         abline(0,0)
```

```
>        title('Rural Females aged 30-49', cex=.5, font.main = 1)
>        plot(interactionDrtAgeSexRuralModel2,select=7,rug=F,se=T, ylim = c(-0.8,0.8), shade=TRU
>        abline(0,0)
>        title('Rural Females aged 50 plus', cex=.5, font.main = 1)
>        par(mar=c(1,4,6,0.5))
>        plot(1,1,type = 'n', xaxt = 'n', yaxt='n',ylab='',xlab='', axes = F)
>        title(main = 'Drought Index: log(1 + count months)',  font.main = 1,cex.main=.9)
>        dev.off()
>
>
```