FARS

FARS time series are comprised on the monthly number of deaths in the hourly windows 11PM, 12AM, 1AM, 2AM, 3AM and week groups SunWed, ThuSat. There are ten series in all and their code names are given in the table below.

#	Code	Definition
1	PM11S	11PM SunWed
2	AM12S	12AM SunWed
3	AM1S	1AM SunWed
4	AM2S	2AM SunWed
5	AM3S	3AM ThuSat
6	PM11T	11PM ThuSat
7	AM12T	12AM ThuSat
8	AM1T	1AM ThuSat
9	AM2T	2AM ThuSat
10	AM3T	3AM ThuSat

The fars database includes a blood alcohol variable but in 38% of the records this was missing or not available. These records comprised 45% of the deaths. Until more is known about why so large a portion of the data is incomplete, it was decided that for the purpose of modelling we will use the entire fars dataset, ie. we will ignore the BAC variable.

The purpose of the intervention analysis is to compare our findings with those of the TIRF and MTO datasets.

May 1996 corresponds to observation #49 in this series. A step intervention model defined by,

$$z_t = \mu + \delta_1 \xi_t + N_t \tag{1}$$

where N_t is the error term. Based on the pre-intervention data we assume initially that N_t is normal and independent, so ordinary multiple linear regression can be used. The intervention series are defined by,

$$\xi_t = \begin{cases} 0 & t < 49 \\ 1 & t \geq 49 \end{cases}$$

The pre-intervention series is fairly short so the following two-step approach to the ARIMA identification of the error term N_t will be used. In the first stage the model in eqn. (1) is fit using standard regression. Then the residuals, the estimated values, $\hat{N}_t$, are obtained and the residual autocorrelation of these residuals is examined. This approach should be expected to work well and is indeed theoretically superior to the alternative approach of basing the model identification on the pre-intervention residuals. The reason for this is that in the estimates of the autocorrelations of $\hat{N}_t$ are first-order efficient as are the estimates in the pre-intervention approach. However since the sample size is larger, this approach provides better estimates.

Regression Results

There are no statistical significant interventions. None of the lag one autocorrelations is significant at 5%.

		Estimate	SE	TStat	PValue
PM11S	1	1.8125	0.316267	5.73091	1.31234×10^{-7}
	ξ	−0.244318	0.457322	−0.534237	0.594495

		Estimate	SE	TStat	PValue
AM12S	1	1.6875	0.479426	3.51983	0.000679735
	ξ	0.903409	0.693249	1.30315	0.195847

		Estimate	SE	TStat	PValue
AM1S	1	1.91667	0.349493	5.48414	3.77136×10^{-7}
	ξ	0.0833333	0.505365	0.164897	0.869395

		Estimate	SE	TStat	PValue
AM2S	1	1.85417	0.358629	5.17015	1.40167×10^{-6}
	ξ	0.0549242	0.518577	0.105913	0.915887

		Estimate	SE	TStat	PValue
AM3S	1	1.16667	0.272091	4.28777	0.0000452363
	ξ	0.0378788	0.393444	0.096275	0.923516

		Estimate	SE	TStat	PValue
PM11T	1	11.875	0.832699	14.2609	0.
	ξ	−1.23864	1.20408	−1.0287	0.306378

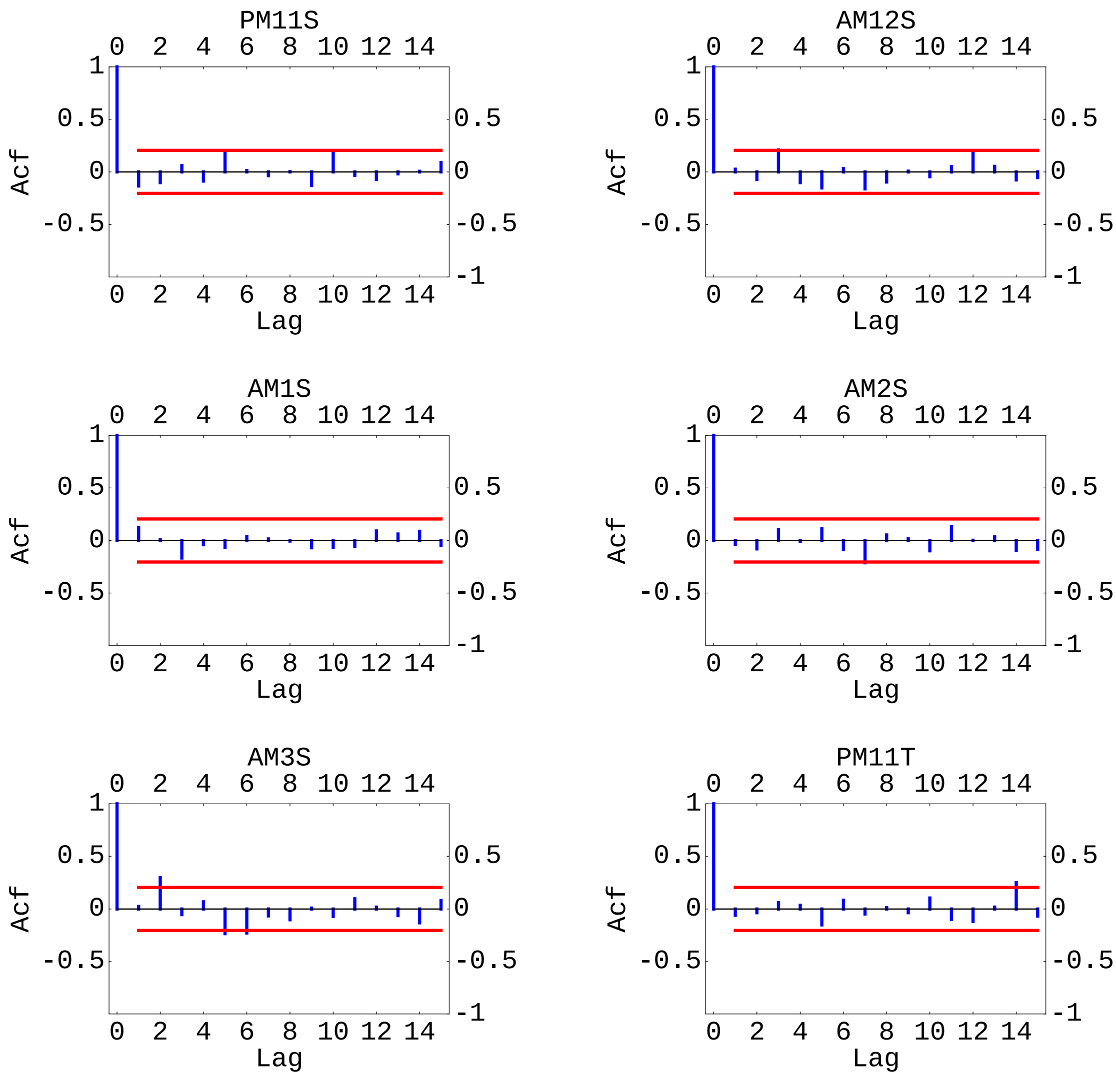
		Estimate	SE	TStat	PValue
AM12T	1	10.3125	0.845908	12.191	0.
	ξ	0.232955	1.22318	0.19045	0.849386

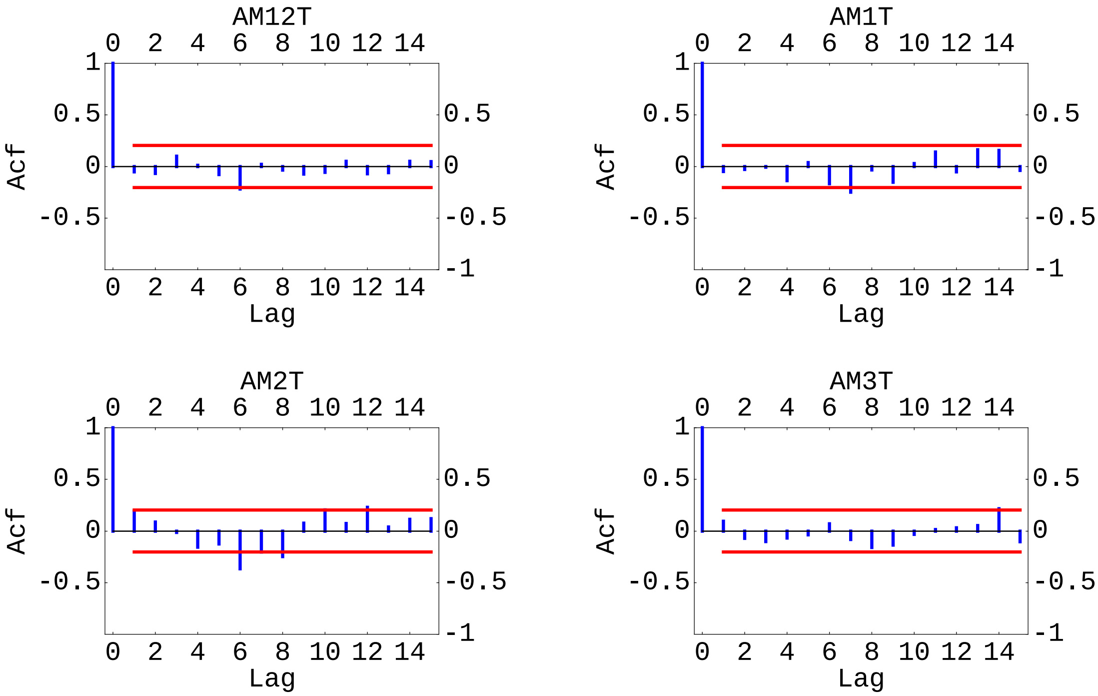
		Estimate	SE	TStat	PValue
AM1T	1	10.5625	0.964922	10.9465	0.
	ξ	−0.448864	1.39528	−0.321703	0.748424

		Estimate	SE	TStat	PValue
AM2T	1	14.125	1.09225	12.932	0.
	ξ	−1.30682	1.57939	−0.82742	0.410188

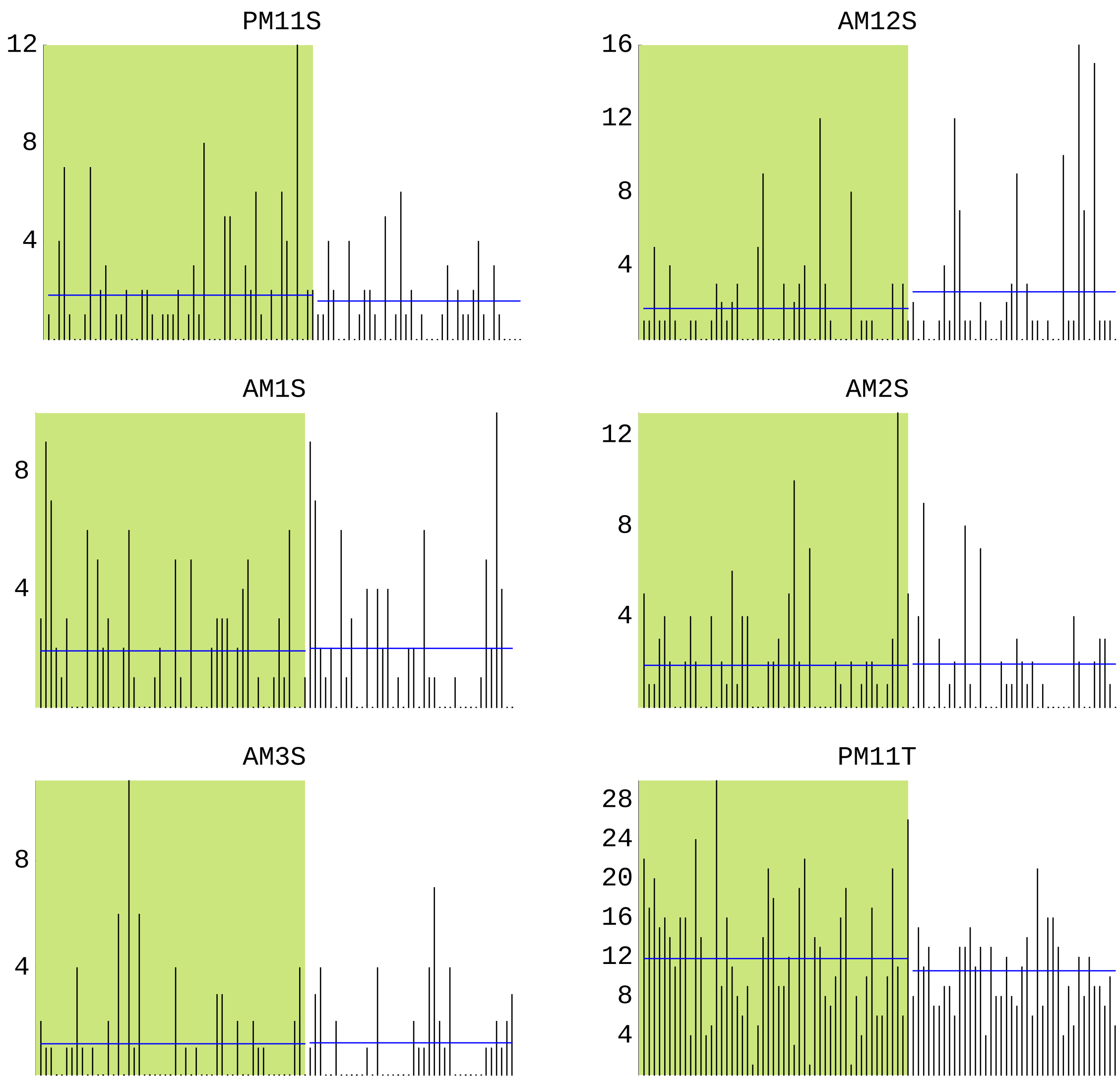
		Estimate	SE	TStat	PValue
AM3T	1	8.89583	0.791846	11.2343	0.
	ξ	−0.645833	1.14501	−0.564043	0.574128

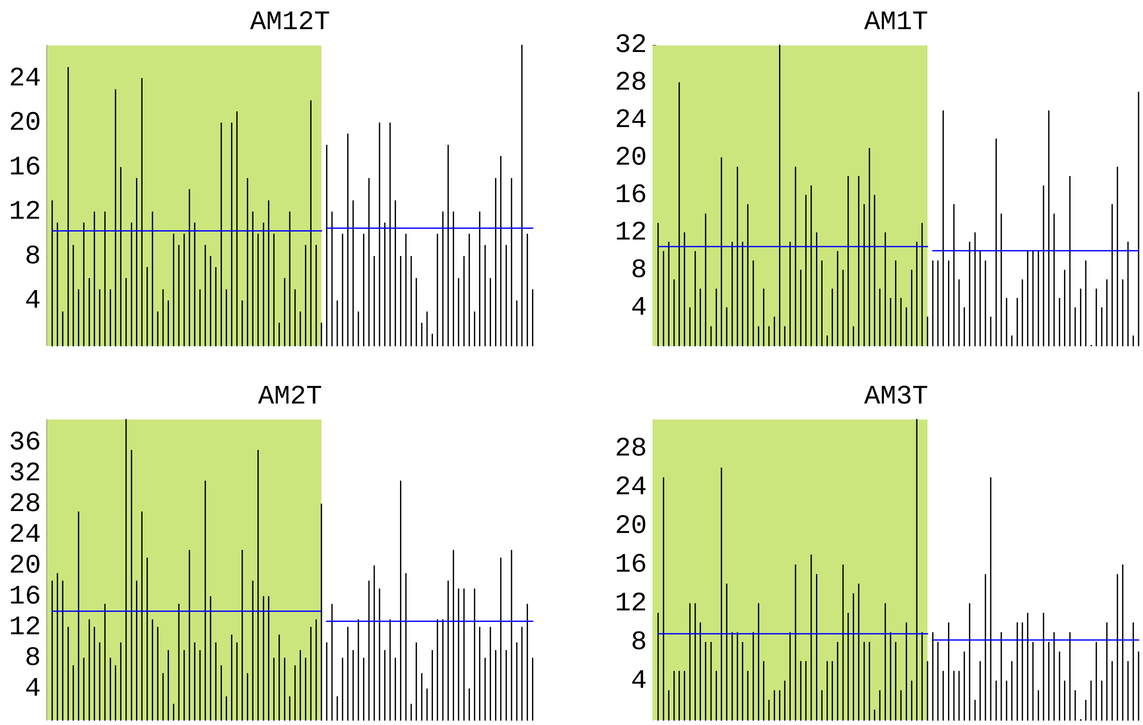
Residual Autocorrelation Analysis




AM12T
Acf
Lag
AM1T
Acf
Lag
AM2T
Acf
Lag
AM3T
Acf
Lag

Data Visualization




AM12T
24
20
16
12
8
4
AM1T
32
28
24
20
16
12
8
4
AM2T
36
32
28
24
20
16
12
8
4
AM3T
28
24
20
16
12
8
4