

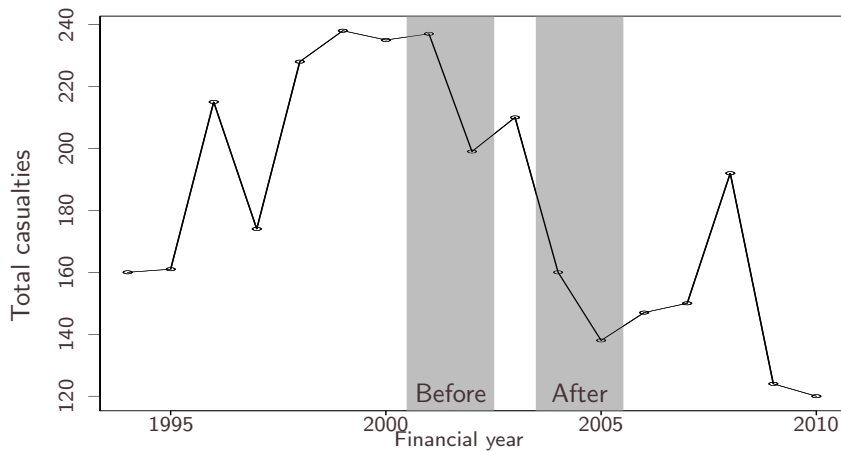
Software for safety scheme evaluation/ hotspot identification

June 2015

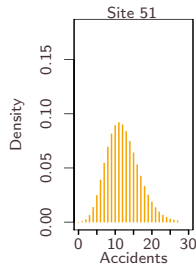
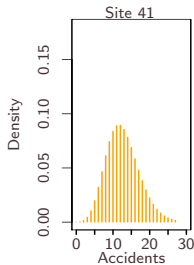
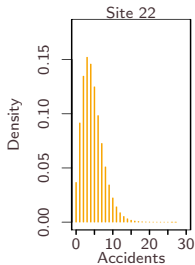
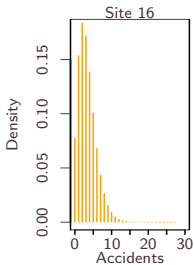
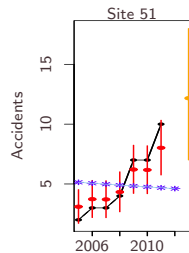
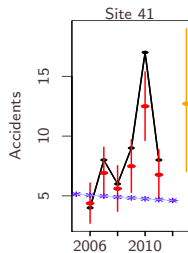
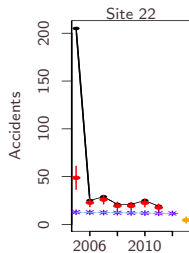
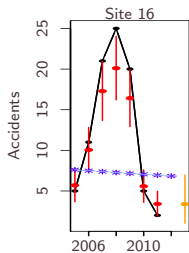
Background

- £5K award to 'pump-prime' research-led impact work
- TADU at Gateshead kindly match-funded
- Builds on work in Thorpe & Fawcett (2012) and Fawcett & Thorpe (2013) which looks at *scheme evaluation*
 - Sites chosen for mobile road safety cameras done so because of their abnormally high casualty record
 - In the 'after' period, might expect to see casualties reduce anyway, even if no scheme implemented
 - Idea: 'clean' the data to separate any casualty reduction into **RTM**, **trend** and **genuine treatment effect**
 - Use cutting-edge statistical techniques
- RTM also a confounding factor in *hotspot identification*

RTM: Scheme evaluation



RTM: Hotspot identification



General methodology: Safety scheme evaluation

1. The data

- Casualty/accident counts in the before/after periods for all treated sites *and* a collection of non-treated (reference) sites
- Information on other variables in the before period for both treated and reference sites; e.g. average speed/85th %-ile speed, road classification, road type, speed limit, flow,...
- Two spreadsheets – in each, rows represent sites and columns represent variables, e.g.:

	Casualties (before)	Casualties (after)	Flow	85th %-ile speed	...
Site 1	10	4	4523	45	...
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮

General methodology: Safety scheme evaluation

2. The method

- Software creates an **Accident Prediction Model** (APM) using the reference data, linking casualties/accidents to the predictor variables (e.g. flow...)
- Applies this APM to predictor information at the treated sites to estimate casualties/accidents in the before period if counts were not abnormally high
- Then uses **Bayes' theorem** to combine the *actual* count in the before period with that from the APM:

	<i>Observed</i>		<i>Model</i>		<i>Difference</i>		
	Before	After	APM	Bayes	Raw	RTM	Treatment
Site 1	12	5	7	9	-7	-3	-4
Site 2	8	3	6	7	-5	-1	-4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
All sites	436	298	275	323	-138	-113	-25

3. The novelty

- Uses state-of-the-art statistical methods: 'Full Bayes' rather than 'Empirical Bayes', recently acknowledged to be the best method (accuracy, precision, reliability)
- Has the ability to try many statistical models for the data and return results based on the 'best model'
- Gives accurate margins of error on RTM/treatment effect estimates
- If general records on 'global' historical accident/casualty counts are available, can also estimate the contribution of trend

General methodology: Hotspot identification

1. The data

- Historical casualty/accident counts at several sites across a geographical region, reported annually (or over some other time unit)
- Information on other predictor variables at these sites – e.g. flow, road type, road classification, average speed, speed limit,...
- No before/after period - perhaps no treatments yet been applied
- Single spreadsheet:

Site	Year	Accidents	Flow	85th %-ile speed	Road type	...
1	2003	18	4532	62.1	Dual	...
	2004	16	4210	64.2	Dual	...
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	2015	8	4091	68.6	Dual	...
	2016	(6,9) / 81.3-82.5%				
	2017	(6,11) / 76.5-84.2%				
<i>Pred/P(exceeds 4)</i>						
2	2003	31	9234	75.3	M	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮

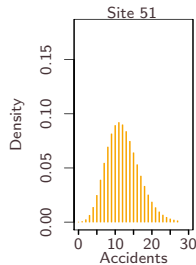
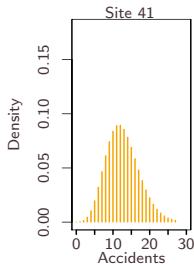
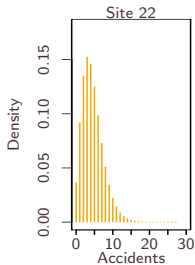
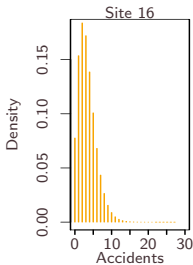
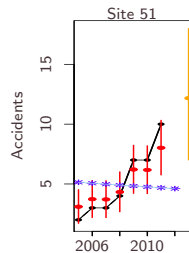
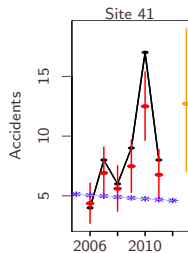
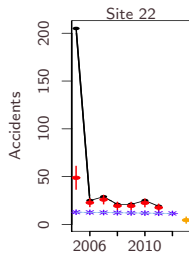
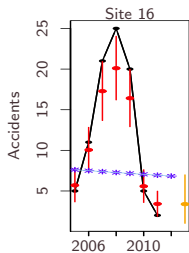
2. The method

- Software creates an APM using information available at *all* sites, or perhaps a tailor-made set of sites using a technique known as **propensity score matching** (PSM)
- This APM can allow for modifications to the road at a particular site (has **crash modification factors** built-in). Also allows for trend, where appropriate
- As before, observed counts (in each year) are ‘modified’ by the APM to clean for RTM; for example, unusually high/low counts will probably be de/inflated
- Exploits the powerful **Bayesian predictive distribution** to accurately forecast short-term future accident counts (with error)

3. The novelty

- Using cutting-edge/innovative modelling techniques developed with colleagues at PTV
- The model 'learns': As more data comes on-line annually, the model adjusts and gives more weight to the data, rather than the APM
- As more data comes on-line, error margins for future predictions decrease
- Can be used as a tool to help guide future investment - employ safety schemes proactively rather than reactively, preventing accidents before they happen
- Predictions 'cleaned' for RTM and trend

General methodology: Hotspot identification



The software

- Exploits a user-friendly interface of the statistical software package R, called 'Shiny'
- Free to download
- Easy to install
- Has the option to change the modelling aspects, if desired; otherwise, standard 'best-fitting' defaults are used
- Very much version 0.000001 at the minute - would like your input into what you'd like this software to give in terms of output