

Road Safety Data Analytics: Prediction and Evaluation

Joining The Dots 2026

Lee Fawcett, Joe Matthews

Who Are We?

- Lee: Reader in Applied Statistics, School of MSP
- Joe: Senior Lecturer in Data Science, School of MSP
- Applied Statisticians – lucky to be able to work on many projects!
 - **Environmental Extremes:** UK Met Office, Sellafield, EDF Energy
 - **Road Safety:** Northumbria Safety Camera Partnership, North Yorkshire CC, Lisbon Municipality, Estonian Traffic Bureau, Transport Association Canada
- Projects in Road Safety
 - **Speed cameras:** Before/After studies, evaluations, value for money?
 - **Predictive analytics:** “Proactive versus Reactive”
 - Contributory Factors
 - Route Analysis
- Data Science Leads: RAPTOR

What Do We Do?

We develop statistical methods to analyse road safety data, mainly focused around two problems:

1. Hotspot prediction

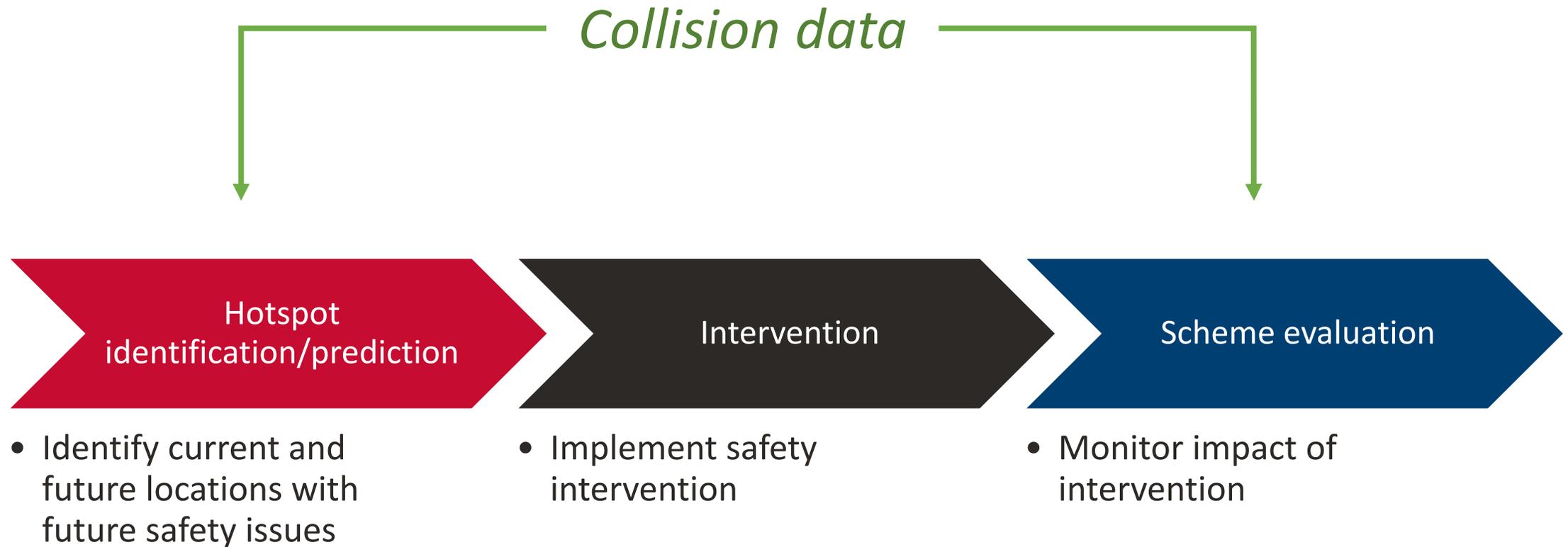
- a) Where are the highest collision rates on a network?
- b) Where are they going to be?
- c) How can we factor in the harm caused by a collision?

2. Scheme Evaluation

- a) Do treatments, e.g. speed cameras, significantly improve safety?
- b) How can we robustly estimate their effectiveness?

To allow practitioners to use our methods, we developed **RAPTOR**, a free-to-use suite of software apps that implement our methods

Collision Data and the Road Safety Practitioner



Scheme Evaluation

Use Before/After data on collisions, casualties, or some other metric of interest, to determine if a road safety intervention has been successful – after successfully accounting for confounding effects.

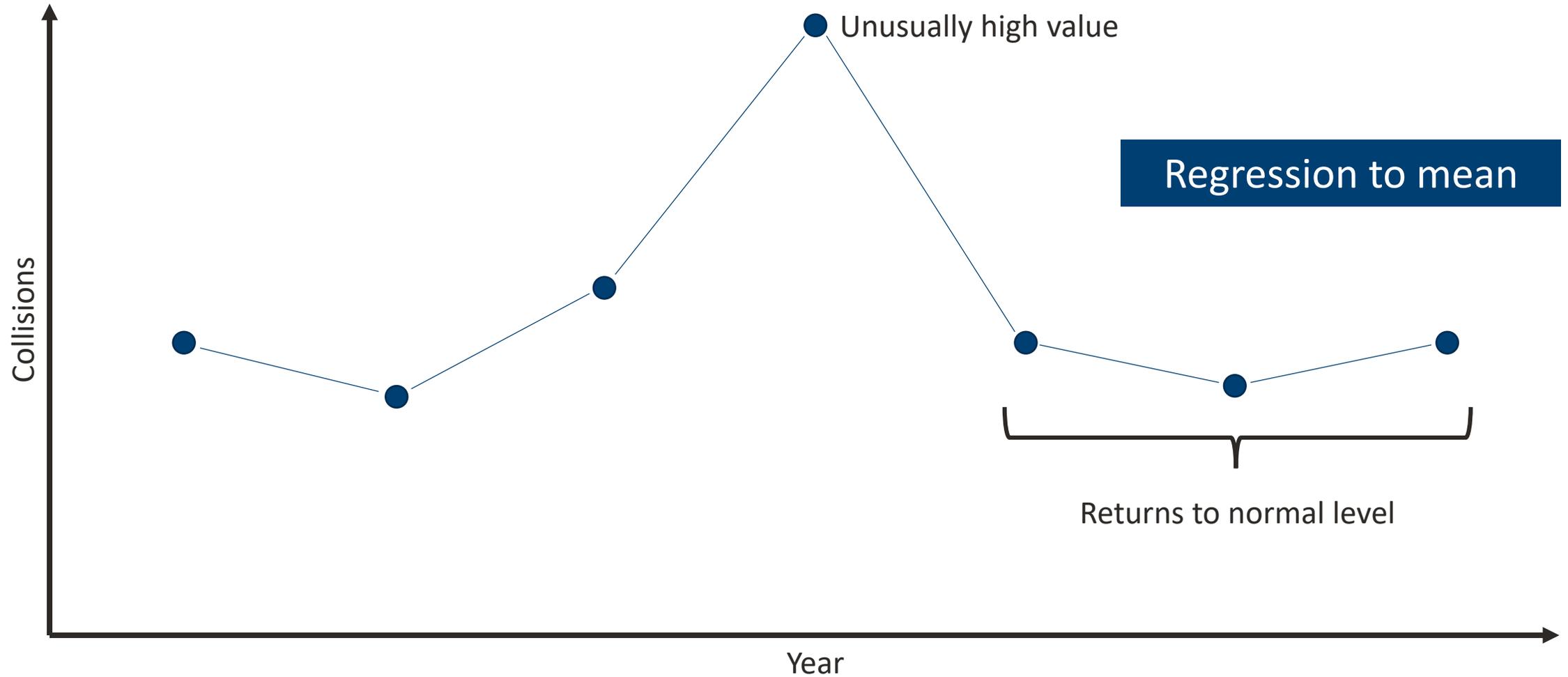
Scheme Evaluation: The Problem

The main issue: Confounding effects, including regression to the mean (RTM)

Occurs when unusual blips correct themselves naturally over time

Failing to account for this effect can lead to overinflated estimates of treatment effects

RTM Effect



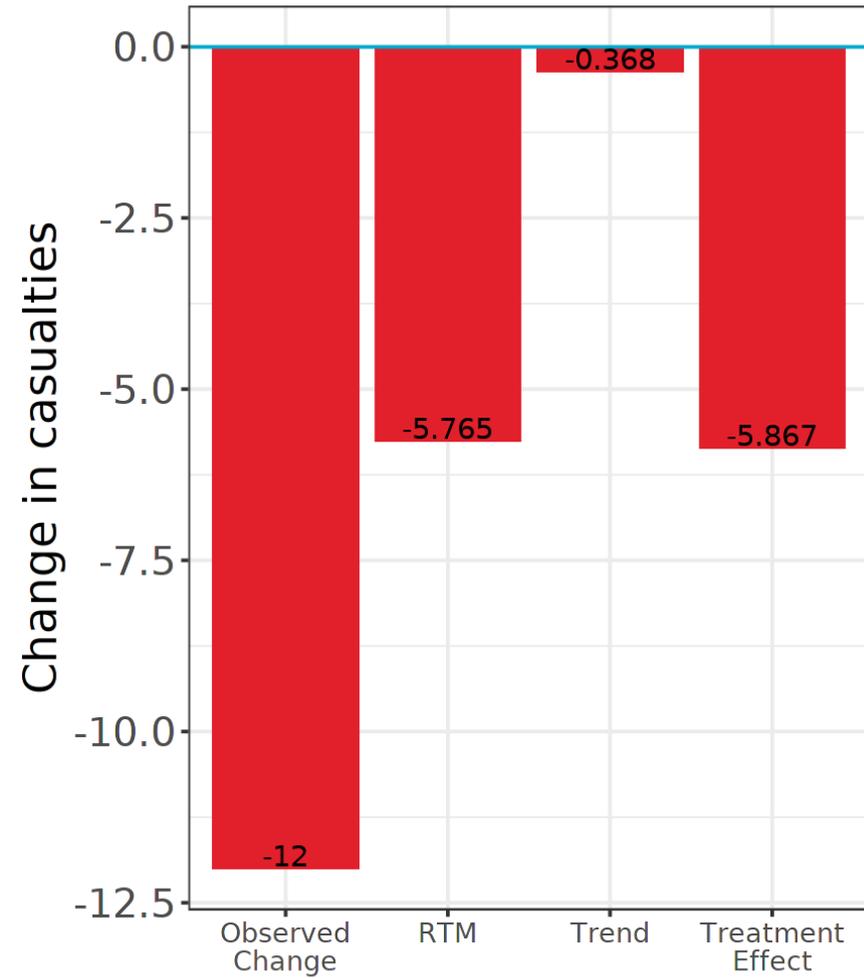
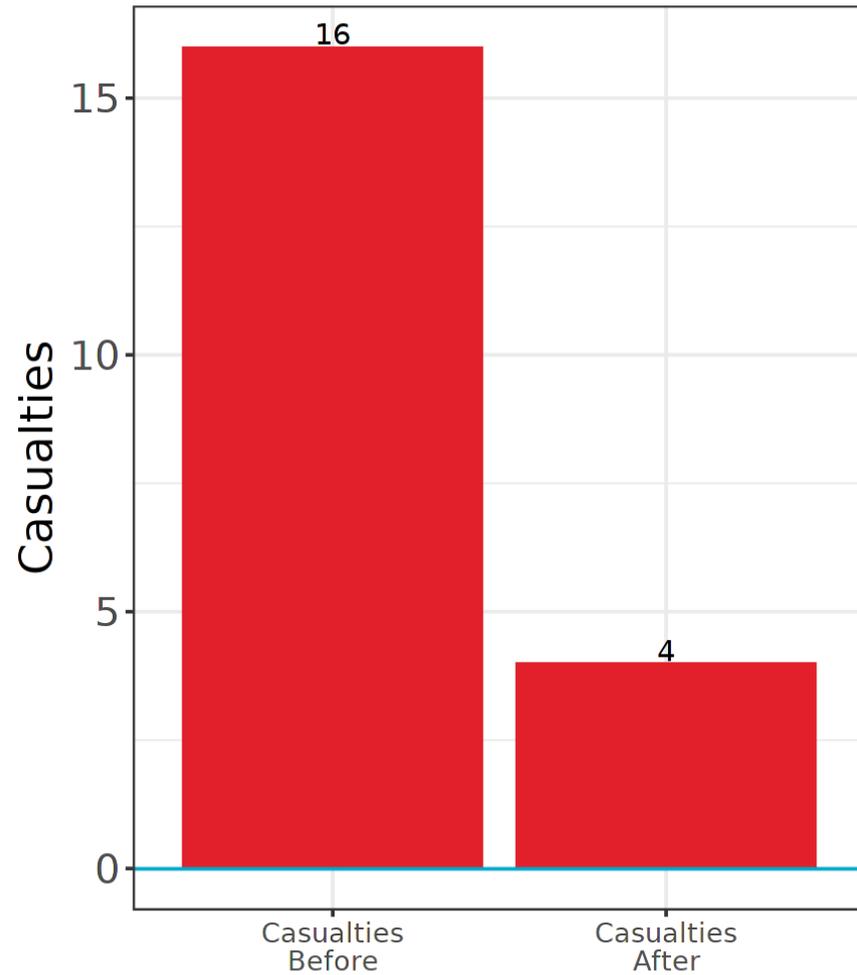
Scheme Evaluation: The Solution

We analyse treated sites, by using other non-treated sites as a control group

This allows us to identify unusual blips in our treated data, and separate the RTM effect from the true treatment effect

We can also identify the effect of temporal trend in the data, which can also interfere with results

RAPTOR Example Output



Scheme Evaluation: Part 1

Foreword



Police and Crime
Commissioner
North Yorkshire



Making North Yorkshire's Roads Safer

Overview & Impact – 2017



Northumbria

- **56 mobile safety**

The results are encouraging as independent research by academics at Newcastle University shows an estimated 20% reduction in casualties owing specifically to the presence of mobile safety camera vans. To put that into some context, the number of people

No

Constabulary

- B/A study on **Cumbria Road Harm Index**
- CRHI accounts for collision severity
- Significant reduction in CRHI after RTM and other confounding factors, and interventions at some sites

Scheme Evaluation: Past Projects



- **Leading Pedestrian Interval** at 8 junctions in downtown Vancouver
- Does not require collisions or casualties – cameras provide info on near misses (“Post Encroachment Times”)
- Requires a different statistical modelling framework: Extreme Value Theory



- **Fixed safety cameras at 30 locations across the city of Tallinn: B/A study**
- Moving windows for B/A periods
- Significant trend and RTM effects identified: Accounted for ~65% of casualty reduction
- Informed rollout of fixed and average speed cameras across Estonia



- **New, and highly controversial, fixed safety cameras in Manhattan and Brooklyn**
- B/A study, incorporating spatial effects
- Significant spatial effects, some trend and RTM
- Significant reduction in collisions and near misses as a result of the cameras

Hotspot prediction

Use the historic collision counts at sites on the network to predict collision counts at these sites in the future to guide future investment in road safety

Hotspot Prediction

Collision hotspots are often determined by comparing historic counts against predetermined thresholds

“Why do we have to wait for something bad to happen before we intervene?”

We use site and network data to forecast into the future, allowing for proactive hotspot identification

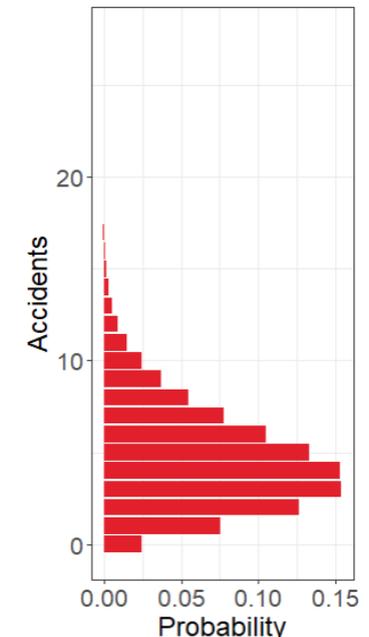
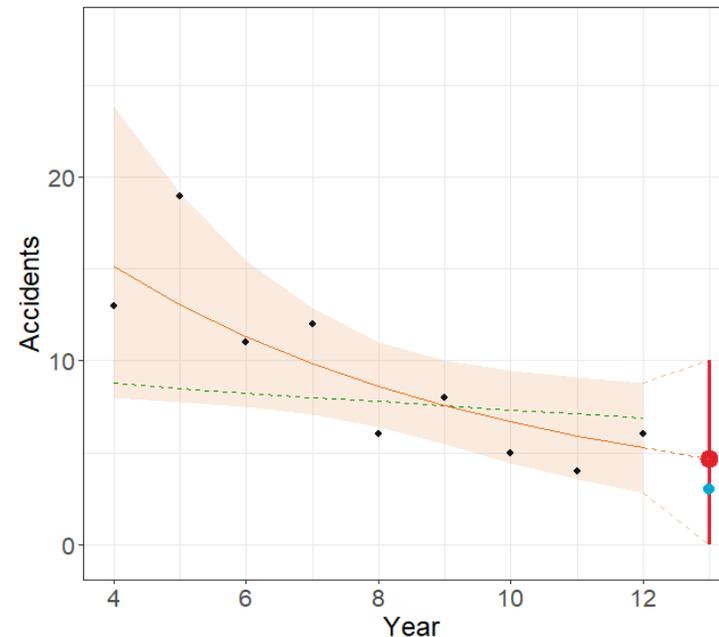
Predicted number of accidents at each site

Show entries

Search:

	Mode ▲	Median ▼	Mean ▼	Lower Bound ▼	Upper Bound ▼
Site.10000664	3	4	4.692	0	10
Site.502	5	5	5.626	1	10
Site.3560	5	5	5.371	1	10
Site.936	38	39	38.943	22	57

Site ID: 10000664



Predicted collision counts

Not just a single “most-likely” value, but the chances for all possible collision count outcomes.

Gain insight into the **underlying level of safety** at each hotspot, letting you see “through the noise”

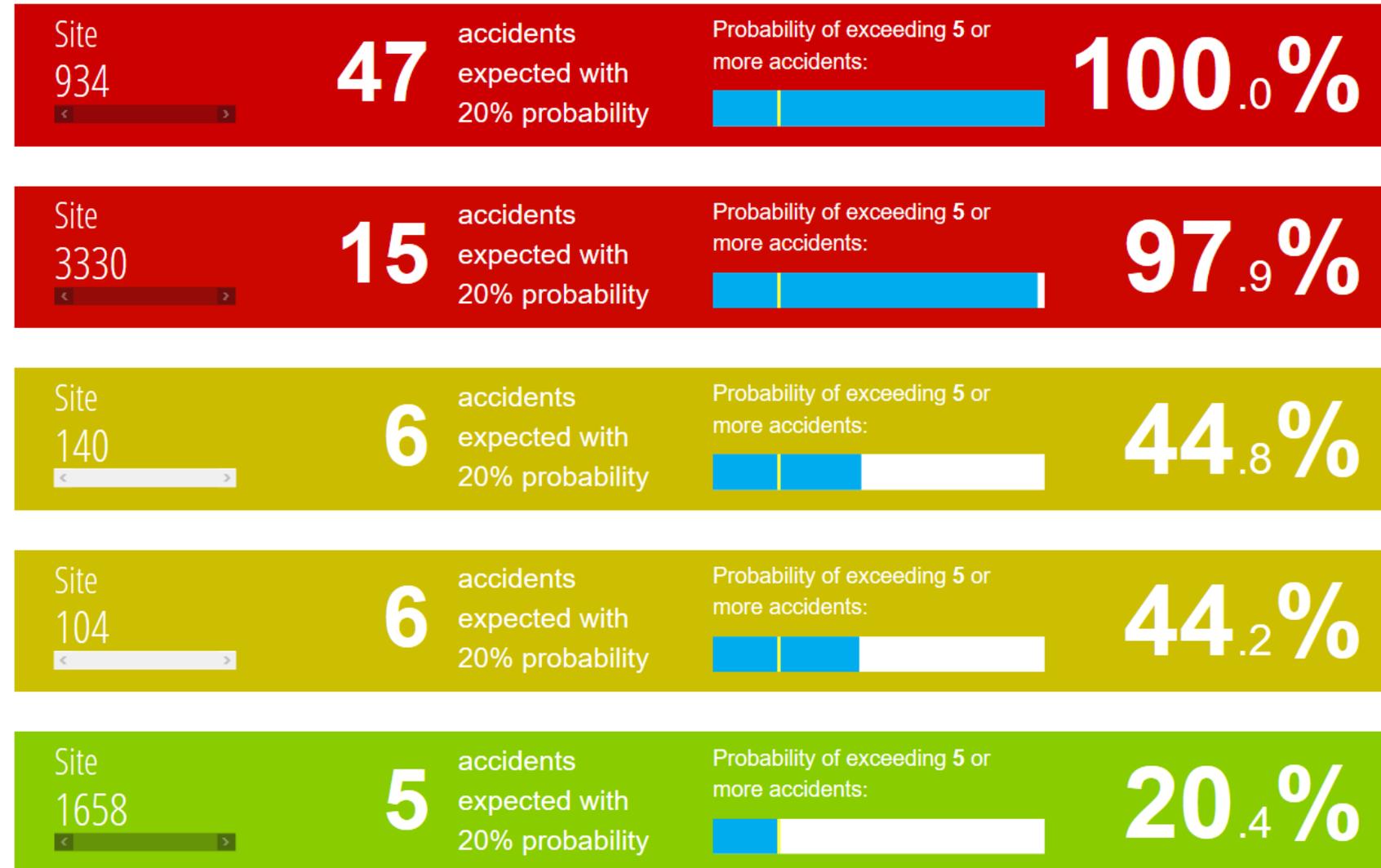
Site Warnings

RAPTOR finds the probability each site will exceed your chosen collision warning threshold.

Sites with warnings

Showing sites that are predicted to have 5 accidents next year with 20% probability or higher.

There are 5 sites with warnings.



Hotspot Prediction: Past Projects

Northumbria
RSCP ($R^2 = 85\%$)

- **Casualty prediction at hundreds of prospective hotspots**
- Partnership uses RAPTOR output to inform decision-making process re: safety scheme implementation

National Highways
($R^2 = 95\%$)

Route Analysis: Collision prediction on A1

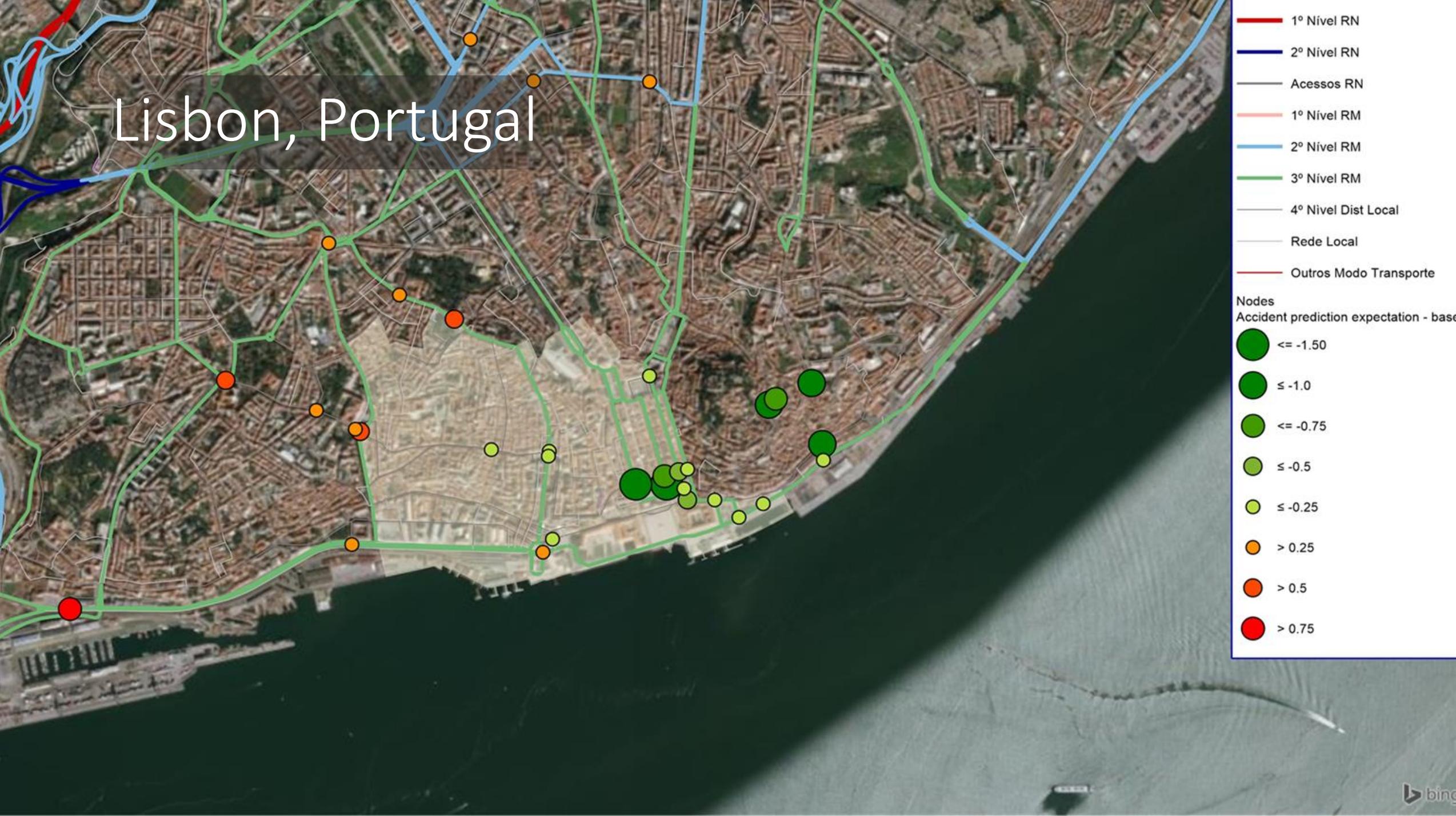
New York City
DoT ($R^2 = 75\%$)

Involved in a dispute between state officials over speed cameras in **Brooklyn and Manhattan**

Lisbon Municipality

- **Proposed pedestrianised zone in central Lisbon**
- Prediction of collision hotspots under this, and several other, “what if” scenarios

Lisbon, Portugal



Hotspot Prediction: Road Harm Index

Partnered with Cumbria Constabulary (DI Jack Stabler and team) on their Road Harm Index project

Locations are identified as harm spots based on the harm caused by collisions there

We used scheme evaluation to evaluate their harm spot interventions, and predicted future harm spots

Road harm index is now being more widely used by police forces across the UK

Conclusion

Statistical methods are necessary for effective road safety scheme evaluation and hotspot prediction

The RAPTOR tool allows practitioners to use cutting edge research methods to analyse their data

We are looking for more organisations to work with on road safety data projects

If you're interested please get in touch lee.fawcett@ncl.ac.uk + joe.matthews@ncl.ac.uk