|  |
| --- |
| Cycling safety in Scotland: Cycle collision hotspots |
| Part of the Scottish Research Programme |
| June | 2016 |



##### About Sustrans

Sustrans is the charity making it easier for people to walk and cycle.

We are engineers and educators, experts and advocates. We connect people and places, create liveable neighbourhoods, transform the school run and deliver a happier, healthier commute.

Sustrans works in partnership, bringing people together to find the right solutions. We make the case for walking and cycling by using robust evidence and showing what can be done.

We are grounded in communities and believe that grassroots support combined with political leadership drives real change, fast. Join us on our journey. www.sustrans.org.uk

Head Office

Sustrans

2 Cathedral Square

College Green

Bristol

BS1 5DD

© Sustrans 2016

Registered Charity No. 326550 (England and Wales) SC039263 (Scotland)

VAT Registration No. 416740656

Table of contents

[Summary 1](#_Toc455393167)

[1 The Scottish Research Programme 2](#_Toc455393168)

[2 Introduction 3](#_Toc455393169)

[3 Background 3](#_Toc455393170)

[4 Methodology 3](#_Toc455393171)

[4.1 Identifying collision clusters 3](#_Toc455393172)

[4.2 Scoring location variables 4](#_Toc455393173)

[5 Results 6](#_Toc455393174)

[6 Conclusion 7](#_Toc455393175)

# Summary

This paper attempts to identify hotspots for cycle collisions on the roads of Scotland. In identifying hotspots, a list of potential locations to invest resources in for the improvement of cyclist safety can be generated.

The collision clusters were formed using data from the STATS19 database of all reported road traffic collisions. From this data it was possible to obtain the geographic location of each collision (northing and easting), the year of the collision and the severity of any associated cyclist casualties.

An algorithm (Density Based Clustering of Applications with Noise) was used to group together closely packed sets of points in spare (i.e. grouping neighbouring points). The parameters for cluster formation were that the collisions were within 20 metres of the centre of the cluster and that a cluster had to be formed of at least three collisions.

* Based on the size of the cluster (i.e. the how many collisions there were in a cluster), the top 10 hotspots were all in Edinburgh.

However, this paper sought to focus on locations where the number of collisions is high relative to the amount of cycling being done, as well as accounting for the severity of cyclist casualties at the location. These could potentially be considered to be truer hotspots.

The population ‘gravity’ of the cluster location and the proportion of commuters who reported that they cycle to work (Census 2011) in the vicinity of each cluster, were used to give an indication of the usage. The number of slight, serious and fatal cyclist casualties in each cluster was also calculated.

For each of these variables the values were ranked, and the sum of the rank scores across the variables for each cluster provided a final score for the cluster.

* Based on this method of identifying hotspots, only four of the top 20 hotpots were in Edinburgh, however most of the clusters were in urban areas
* The majority of the top 20 hotspots were at either T or staggered junctions (7), or roundabouts (8)
1. The Scottish Research Programme

Sustrans Scotland works in partnership with a wide variety of stakeholders to support the vision of the Scottish Government that by 2020, 10% of everyday trips will be made by bicycle; that a culture of functional walking will be embedded throughout the country and that choosing to walk and cycle through high quality public space will be made easy and predictable.

With our partners, we help to demonstrate how travel behaviour in Scotland can change by delivering practical interventions, making the built environment more conducive to walking and cycling, and engaging with people to present the options available to them to travel actively.

We do this work by managing funds for Transport Scotland and delivering a suite of projects, all match-funded by recipients: projects include the challenge-fund infrastructure schemes Community Links and Community Links Plus, the 2,500 mile National Cycle Network, the community led design intervention Street Design.

We also run behaviour change interventions such as I Bike in schools, Active Travel Champions in workplaces, Active Travel Hubs and support for education and young people, including the annual travel census, the Hands up Scotland Survey (HUSS).

In 2014, with agreement from Transport Scotland’s Sustainable and Active Travel Team we decided to increase investment in research and monitoring by establishing a small research programme. The aim was to provide insight and analysis of the investment in active travelling so a more nuanced understanding of the effectiveness of interventions for walking and cycling in Scotland could be gained.

We discussed this outcome with our Research and Monitoring Unit and commissioned them to produce a set of outputs in the form of a portfolio of some 20 separate but related projects. This report represents the output of one of these projects.

The programme helps to examine and model the scope for change and identify the most appropriate interventions. It includes research on aspects of behaviour change, analysis of infrastructure use, modelling of investment requirements and examination of economic impacts.

These outputs will be used by Sustrans Scotland and our partners to reflect on the projects we deliver, the focus for future investment in walking and cycling as well as research and insight.

1. Introduction

Using data from the STATS19 dataset, and accounting for the volume of cycling in an area, this paper identifies 20 of the most dangerous locations for cyclists across Scotland, based on the volume and severity of collisions over the past decade.

The paper seeks to help identify locations where an investment in cycle safety is most needed.

It must be stressed that this is only one approach to ranking dangerous locations. The principle underlying any ‘hotspot’ identification is to understand, in detail, the local characteristics of cycle safety.

1. Background

The STATS19 dataset holds a record of all reported road traffic collisions where a casualty has occurred. There are a number of weaknesses with this dataset (particularly underreporting of collisions and near misses), but it is nevertheless the most comprehensive dataset of its kind in Scotland.

This paper shows how the dataset can be used to identify locations where there is a high rate of cycle collisions. In so doing locations benefitting from investment in improved cycle safety can be identified.

1. Methodology

There are three main steps which have been taken to identify these ‘hotspots’:

1. Locations where there are ‘clusters’ of collisions are identified
2. These locations are scored using a number of variables
3. The locations are ranked by the sum of these scores
	1. Identifying collision clusters

There are a number of spatial clustering methodologies available, but the one used here is Density Based Clustering of Applications with Noise (DBSCAN). This algorithm groups together closely packed sets of points in some space – that is, points with many nearby neighbours. It is a commonly used method and used in mainstream scientific literature.

DBSCAN requires two parameters: ε (eps (the distance parameter)) and the minimum number of points required to form a dense region (minPts)[[1]](#footnote-1). For the purposes of this exercise, a distance of 20m was used for ε, and minPts was designated as 3. These values were assumed to represent the size of a typical junction and an appropriate size of cluster.

This algorithm was applied to the Easting and Northing co-ordinates of all cycle collisions in Scotland included in the STATS19 dataset (2005-2014).

This resulted in the identification of 253 collision clusters across the Scotland, ranging from a single cluster of 22 collisions, to 134 clusters of the minimum required three collisions.

* 1. Scoring location variables
		1. Variables used

When the list of collision clusters were ranked by the number of collisions in the cluster, it became apparent that this approach was too simplistic to accurately rank the different levels of risk at each location.

The clusters with the greatest number of collisions were all located in Edinburgh, the area in Scotland with the second highest population size[[2]](#footnote-2) and with the joint highest proportion of residents who cycle as their main mode of transport[[3]](#footnote-3) (see Table 3-1). This paper concentrates on identifying locations that have a disproportionately large number of collisions **relative** to the cycling activity at that location.

These are the true cycle safety ‘hotspots’.

Previous research by Sustrans has indicated there are two variables which have the largest impact on the levels of cycling at any one location:

* The population ‘Gravity’ of a location. That is, the size of the population divided by the square of the distance from the location
* The proportion of commuters in the immediate vicinity who reported they cycle to work in the 2011 Census

These variables have been calculated for each of the cluster locations. In the case of population gravity, this was calculated at 1km intervals up to 10km from the cluster location. In the case of the Census data, the value used was the maximum proportion reported of the Lower Super Output Areas (LSOAs – a geographical unit used by the Census) within 1km of the cluster location.

In addition, the number of collisions in each cluster was recorded and weighted according to how recently each collision had occurred. Collisions occurring in 2014 were weighted at 100% and collisions in 2005 weighted at 10%. Intervening years were weighted according to a curved line drawn between these values.

This was done to account for any improvements in cycle safety in the past decade with lower values ascribed to collisions up to 2011, and greater prominence given to collisions that occurred after 2011, due to investment in cycling with an increased focus on safety following the publication of the Cycling Action Plan for Scotland (CAPS)[[4]](#footnote-4).

Finally, the number of serious and fatal collisions in each cluster was also recorded. This was to make sure that locations with a smaller volume of more severe collisions were not unfairly excluded from the ranking. These variables were also weighted according to the date of the collisions, using the same method outlined above.

* + - * 1. Top 10 ranked locations (using number of cycle collisions)

|  |  |  |  |
| --- | --- | --- | --- |
| Rank | Location | Area | Number of cycle collisions 2005-2014 |
| 1 | A7 / Chambers Street | Edinburgh | 22 |
| 2 | B901 / London Street roundabout | Edinburgh | 15 |
| 3 | Dalmeny Street / Easter Road | Edinburgh | 15 |
| 4 | Home Street / A700 | Edinburgh | 14 |
| 5 | Mayfield Road / West Savile Terrace | Edinburgh | 13 |
| 6 | Leith Walk / London Road roundabout | Edinburgh | 11 |
| 7 | Picardy Place roundabout | Edinburgh | 9 |
| 8 | Fountainbridge / Earl Grey Street | Edinburgh | 9 |
| 9 | Drumbrae roundabout | Edinburgh | 8 |
| 10 | Corstorphine Road / Murrayfield Road | Edinburgh | 8 |

* + 1. Scoring method

To allow for weighting to be assigned to the different variables (so the clusters can be ranked according to different priorities[[5]](#footnote-5)), each variable was scored using the same system.

The percentile rank of each cluster was calculated for each variable[[6]](#footnote-6), and the cluster received a score out of 10 according to which decile it fell into, with the minimum score being 1, the maximum score 10. Deciles were selected to give the greatest segmentation of the data without losing usability.

Across the clusters it was apparent that there were no fatal collisions recorded within any of the clusters and that only a couple of dozen clusters contained collisions in which there had been a serious injury to a cyclist. This meant that fatal collisions were irrelevant for the purpose of cluster comparison and so the ‘fatal collision’ variable was not included in the calculation of clusters’ final scores.

The presence of only a small number of clusters which included collisions involving a serious cyclist casualty led to disproportionate differences, between clusters, in serious casualty collision scores relative to the number of serious casualty collisions. In order to reduce the impact of this on the final cluster scores, this variable received a lower weighting then the cluster size, population size and proportion of residents that cycle to work variables[[7]](#footnote-7)

Ascending scoring was used for the weighted cluster sizes with larger clusters receiving larger scores. However, as this paper concentrates on identifying collision clusters that are disproportionately large **relative** to the level of cycling in the area, population and cycle to work proportions were scored using a descending method.

The scores for each variable were summed for each cluster, and the clusters ranked accordingly, with high scoring clusters at the top of the ranking.

1. Results

Table 5-1 shows the top 20 high scoring locations identified by this method, along with the total number of cycle collisions that have occurred at each location between 2005 and 2014.

* + - * 1. Top 20 ranked locations (using scoring method)

|  |  |  |  |
| --- | --- | --- | --- |
| Rank | Location | Area | Number of cycle collisions 2005-2014 |
| 1 | A761 / Arkleston Road | Paisley | 3 |
| 2 | A726 / Parkway roundabout | Near Erskine | 4 |
| 3 | A199 / B1361 / A6094 roundabout | Wallyford | 3 |
| 4 | B959 / Robertson Street | Dundee | 4 |
| 5 | Barrhead Road / Peat Road / Braidcraft Road / Brockburn Road roundabout | Glasgow | 3 |
| 6 | Glasgow Road / Viewlands Road | Perth | 3 |
| 7 | A8 / Cathedral street | Glasgow | 3 |
| 8 | High Street | Tranent | 3 |
| 9 | Main Street / B6482 / Newbattle Road mini roundabouts | Newtongrange | 3 |
| 10 | Mearns Road / A727 | Clarkston | 3 |
| 11 | Drumbrae roundabout | Edinburgh | 8 |
| 12 | A7 / Craigmillar Castle Road | Edinburgh | 7 |
| 13 | A70 / Juniper Avenue | Edinburgh | 3 |
| 14 | Burdiehouse Road / Straiton Road / Lang long roundabout | Edinburgh | 3 |
| 15 | A823 / Laburnum Road roundabout | Dunfermline | 3 |
| 16 | Spiersbridge roundabout | Thornliebank | 3 |
| 17 | A8 / Springhill Parkway | Glasgow | 3 |
| 18 | B9006 / Tower Road | Inverness | 5 |
| 19 | University Road West / Hermitage Road | Stirling | 7 |
| 20 | Starlaw Road / Boghall roundabout | Bathgate | 5 |

The top 20 locations are not widely spread across the country and are typically within urban areas. In particular there was a high representation in and around Glasgow. Although Glasgow has a high population (which would contribute to a lower total score), the proportion of the population that cycle to work is relatively low (which would contribute to a higher score).

As Table 4-2 shows, the majority of the top 20 cluster locations are at either T or staggered junctions (7), or at roundabouts (8). There is a noticeable divergence between the types of location represented by the top 20 clusters and the types represented across all collision clusters.

In particular, 40% of the top 20 clusters were located at roundabouts, whereas 18.5% were at roundabouts across the total selection of clusters identified.

* + - * 1. Cluster locations

|  |  |  |
| --- | --- | --- |
| Location type | Total clusters | Clusters in top 20 |
| T or staggered junction | 100 | 7 |
| Crossroads | 68 | 1 |
| Roundabout | 47 | 8 |
| Not at junction or within 20 metres | 16 | 1 |
| Mini-roundabout | 9 | 1 |
| More than four arms (not roundabout) | 8 | 1 |
| Other junction | 5 | 1 |
| Total | 253 | 20 |

1. Conclusion

This paper has shown that it is possible to identify cycle safety ‘hotspots’ using a methodology that attempts to account for different levels of cycling across the country. It also highlights that the main risk to cyclists occurs when they directly interact with other vehicles – at junctions and roundabouts – rather than when simply sharing the carriageway.

It does not claim to be the definitive list of dangerous locations, nor that the variables included are the only ones that could be used. As ever when using the STATS19 dataset it is important to recognise the limitations of the data in representing cycle safety. A large volume of low severity incidents go unreported and no attempt has been made to measure the ‘near misses’ which contribute so much to feelings of insecurity.

The fact there were no fatal cycle collisions contained within the collision clusters identified also highlights the difference in the nature of collision severities in as much as fatal collisions do not occur at the same locations as serious and slight collisions.

In addition, due to their infrequency, it is not possible to define fatal collision ‘hotspots’ by combining the collision frequency in a limited location size with other relevant variables, in the way that serious and slight collision ‘hotspots’ can be. A different approach would be required to prioritise the dangerousness of locations based on fatal cycle collisions.

1. It is important to note that this algorithm starts from an arbitrary point – different clusters may be identified if the process starts from a different point. [↑](#footnote-ref-1)
2. Mid-year population estimates 2014 [↑](#footnote-ref-2)
3. Scottish household survey 2014 [↑](#footnote-ref-3)
4. http://www.transport.gov.scot/system/files/uploaded\_content/documents/tsc\_basic\_pages/Environment/CAPS\_2013\_-\_final\_draft\_-\_19\_June\_2013\_0.pdf [↑](#footnote-ref-4)
5. In the ranking reported in this paper, a flat weighting across the variables was applied. [↑](#footnote-ref-5)
6. The percentile rank of a value is the percentage of values in its frequency distribution that are equal to or lower than it. [↑](#footnote-ref-6)
7. The weights applied to each variable before summing for a final score were: serious accidents (0.1), collision cluster size (0.3), population size (0.3) and proportion who cycle to work (0.3). [↑](#footnote-ref-7)