International Journal of Sustainable Transportation

ISSN: 1556-8318 (Print) 1556-8334 (Online) Journal homepage: http://www.tandfonline.com/loi/ujst20

The Identification of Hazardous Road Locations: A Comparison of the Blacksite and Hot Zone Methodologies in Hong Kong

Becky P. Y. Loo

To cite this article: Becky P. Y. Loo (2009) The Identification of Hazardous Road Locations: A Comparison of the Blacksite and Hot Zone Methodologies in Hong Kong, International Journal of Sustainable Transportation, 3:3, 187-202, DOI: 10.1080/15568310801915583

To link to this article: https://doi.org/10.1080/15568310801915583

International Journal of Sustainable Transportation, 3:187–202, 2009

ISSN: 1556-8318 print/1556-8334 online
DOI: 10.1080/15568310801915583

The Identification of Hazardous Road Locations: A Comparison of the Blacksite and Hot Zone Methodologies in Hong Kong

Becky P. Y. Loo
Department of Geography, The University of Hong Kong,
Pokfulam, Hong Kong

ABSTRACT
In this paper, the spatial characteristics of road crashes in Hong Kong are analyzed. This paper first introduces the hot zone terminology in road safety research. Then, the methodological issues and difficulties are discussed. Finally, the spatial characteristics of road crashes in Hong Kong in 2004 are analyzed by the hot zone methodology. In particular, the hot zone methodology is compared with the blacksite methodology, which is commonly used by road safety administrations worldwide. The analysis is important both theoretically in enriching our ways of conceptualizing and identifying hazardous road locations and practically in providing useful information for addressing road safety problems.
Key Words: blacksites, crash analysis, GIS, hot zones, road safety.

1. INTRODUCTION

Spatial information about crashes is often included in road crash databases (Shinar, Treat, and McDonald, 1983; Ibrahim and Silcock, 1992; Austin, 1993; Khan, Kathairi, and Grib, 2004). However, some of these spatial data are not accurate and precise enough for scientific crash analysis (Austin, 1995; Loo, 2006). To justify the extra resources required for collecting accurate and precise spatial data, the potential values of having high-quality spatial data should be better understood.

One of the useful ways that researchers can make good use of accurate spatial data is the identification of hazardous locations in a road network (Deacon,

Received 14 February 2007; revised 30 October 2007; accepted 7 November 2007.
Address correspondence to Becky P. Y. Loo, Associate Professor, Department of Geography, The University of Hong Kong, Pokfulam, Hong Kong. E-mail: bpyloo@hkucc.hku.hk

Zegger, and Deen, 1975; Transportation Research Board, 1982; Hoque and Andreassen, 1986; Nicholson, 1989; Ogden, 1996). In fact, many road safety administrations worldwide are doing this with the blacksite[1] or hot spot methodology (Transportation Research Board, 1982; Silcock and Smyth, 1984; Transport Department, 2001). This paper introduces a recent and developing methodology—the hot zone methodology (Black and Thomas, 1998; Flahaut et al., 2003). Because this geographical concept is rather new, the terminology is described first. Next, the methodological issues and difficulties are discussed. Finally, the case of Hong Kong is presented.

2. HOT ZONES: TERMINOLOGY

"Where are the dangerous locations in a road network?" Different road users may have different answers. However, road safety administrations need to identify dangerous road locations in a systematic and scientific manner. Typically, the identification of hazardous road locations involve the key aspects of: defining the site (e.g., site or route or area), setting the criteria of "hazardous" (e.g., using the frequency, rate, rate quality control, or potential crash reduction method), and considering relevant factors like exposure (e.g., the traffic volume and traffic mix), crash severity (e.g., weighting fatal crashes more than injury ones), and time period (e.g., selecting a time interval to establish statistical reliability) (Ogden, 1996). Although all these aspects are important for scientific crash analysis, this paper, which summarizes the preliminary results of a larger-scale project, only focuses on the first aspect, that is, the definition of sites. The definition of sites is important because the choice of the spatial unit of analysis will have implications on all other aspects, such as the setting of the critical crash frequency/rate in defining hazardous road locations.

Consider a simple hypothetical network, as shown in Figure 1. The *existing blacksite methodology* of the Hong Kong Special Administrative Region (SAR) Government, as many other overseas administrations (Silcock and Smyth, 1984; Ogden, 1996; Transport Department, 2001), first identifies the road junctions and their surrounding road network within 70 meters. If the crash rate is considered high, a blacksite is identified. Currently, the crash rate is considered high if there are 9 vehicle–vehicle collisions or 6 pedestrian crashes in a year (Transport Department, 2001). Conceptually, blacksites are a type of hot spots, which refer to hazardous road locations as spots or "point" locations (Transportation Research Board, 1982). Because the locations of potential blacksites are only restricted to road junctions and their surrounding areas (shown by the dotted circles in Fig. 1), the blacksite methodology is relatively limited in scope. However, when the level of precision of the spatial data collected is not high, the blacksite methodology, which only requires the identification of crashes near junctions, could prove to be pragmatic.

If more comprehensive and precise spatial data are available, the entire road network can be analyzed by the *hot spot methodology*. Following this methodology,

[1]Hotsite is preferred. Blacksite is used in this paper because it is an official term in Hong Kong.

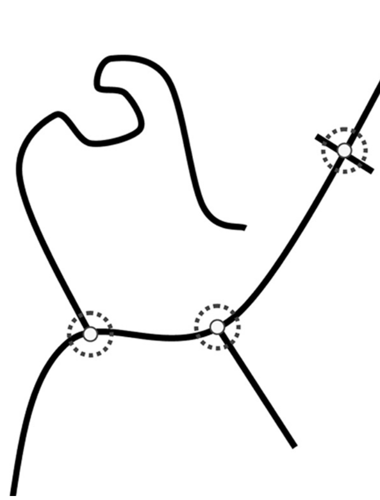


Figure 1. A hypothetical road network.

each road in the entire road network is cut into segments named basic spatial units (BSUs). A BSU, being a road segment (not a network or a subgraph), has to be long enough to allow the identification of crash clusters but short enough to reflect changes in the road environment that have a bearing on road safety. As a general rule, the length of a BSU is defined as 1 hectometer, that is, 100 meters (Black, 1991; Black and Thomas, 1998; Flahaut et al., 2003, Flahaut, 2004).[2] For each BSU, the actual crash rate, AR, is calculated and compared with the critical crash rate, CR. The CR can be defined by various methods, such as the frequency, rate, frequency rate, rate quality control and crash severity methods (Hauer and Persaud, 1984; Hauer, 1996; Stokes and Mutabazi, 1996; Ogden, 1996; Khisty and Lall, 2003). Wherever the AR of a BSU is greater than or equal to the CR, a hot spot or dangerous location is found. If the hypothetical network in Figure 1 consists of 34 BSUs, there are 34 potential hot spots. In comparison, only three potentially dangerous locations could be systematically identified by the existing blacksite methodology.

In addition, the hot spot methodology allows the safety levels of different roads to be compared and analyzed. In Figure 2, the hypothetical crash patterns of two roads with equal length are displayed. If the CR for identifying hot spots is set to be 6 crashes per year, Road 1 has two hot spots at BSUs 1 and 11. Road 2 has only one hot spot at BSU 1. However, is Road 2 safer? On Road 2, the contiguous BSUs with consistently high crash rates from BSU 8 to BSU 12 obviously pose a safety hazard that is worth closer examination. To the road users, the danger posed by these

[2]The author is currently conducting a sensitivity analysis, which includes varying the length of a BSU to 150 meters and 200 meters. Moreover, additional sensitivity tests such as the use of various scientific methods to define CR, the aggregation of crash data over a long period of time, the weighting of crash severity, and a finer analysis by crash type are being conducted.

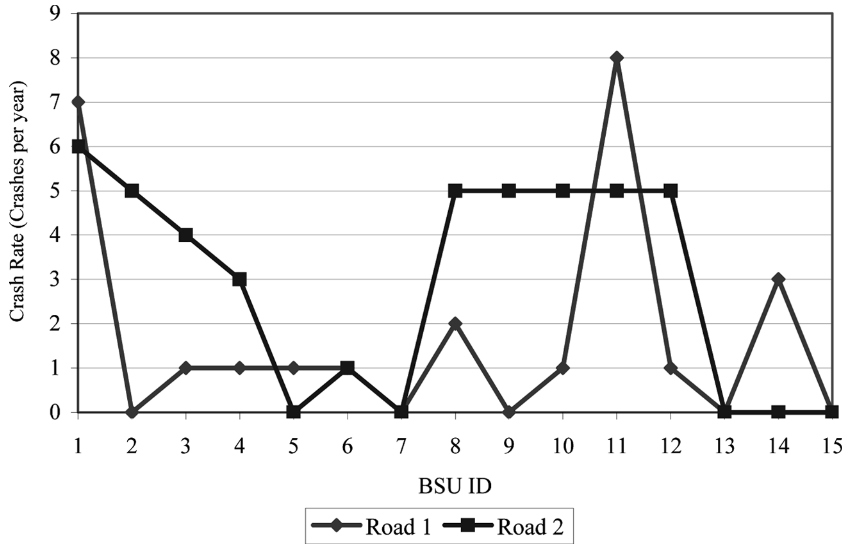


Figure 2. Crash patterns of two roads compared.

contiguous road segments may well be more serious than a hot spot alone. Because the high crash rate at a hot spot may well be the result of randomness rather than any systematic deficiency in the road environment that warrants treatment, the crash rates at these "false-positive" sites are likely to have more significant drops in their crash rates in the ensuring years, even without intervention. This is known as the regression-to-mean problem (Elvik, 2006; Shen and Gan, 2003). Among the different ways of defining CR, Elvik (2006) found the Empirical Bayes (EB) method, which takes into account crash rates at similar sites, to perform best in differentiating the "true-positive" and the "false-positive" sites.

The *hot zone methodology* identifies hazardous road locations consisting of more than one contiguous road segments (defined by BSUs). These hazardous road locations, which may occur at nonjunction locations, are called hot zones. In particular, the spatial contiguity of the road network is emphasized in the identification of hazardous road locations. A most important underlying concept is that road segments are not spatially independent objects. Furthermore, the length of hot zones is not fixed *a priori* but determined by the actual crash records, which may be weighed by the number and severity of injuries or pooled data over many years. With the hot zone methodology, it is worth mentioning that the CR for a hot zone is not used directly for identifying hazardous locations. Instead, the threshold crash rate of a BSU is used. Consider the threshold value of 3 crashes in a BSU in a year. Because a hot zone consists of at least two contiguous BSUs with 3 or more crashes in a year, the hot zone would have a CR of 6 crashes in a year. If the adjoining BSU(s) is (are) also having AR equal to or over the threshold value, it (they) would form part of the same hot zone. Hence, a hot zone consisting of three BSUs would have 9 or more crashes in a year. Nonetheless, one BSU with 9 crashes in a year would not form a hot zone by itself. Statistically, the nature of the "false-positive" problem in the hot zone methodology, which involves the joint probability distribution of two (or more) contiguous BSUs, is therefore different from the hot spot methodology, which only deals with the probability distribution of a single site/BSU independently.

3. METHODOLOGY

Although hot zones are hazardous to road users, there is no systematic methodology to identify them in a road network. Most of the studies conducted to date are based on hypothetical or selected roads rather than the entire road network (Black, 1991; Black and Thomas, 1998; Flahaut et al., 2003; Flahaut, 2004). This paper first develops a three-step geographic information systems (GIS)-based methodology to identify the hot zones in a road network systematically. The methodology is suitable for analyzing real-life complex road networks. In this section, the methodology, together with the major issues and difficulties involved, is outlined.

The first step involves the validation of crash locations. Using Hong Kong as an example, the location of a road crash is identified by the five-figure grid references in the crash database of Hong Kong's Traffic Accident Data Systems (TRADS). The information can be transformed into the x and y coordinates for analysis by GIS. For both technical and nontechnical reasons (see Austin, 1995; Loo, 2006), these crashes are unlikely to intersect with the GIS road network, which is represented by nodes and centerlines. The creation of buffers is one of the solutions to assign the crashes to the GIS road network (Austin, 1995). However, this is likely to cause double-counting, especially in a dense road network. Hence, a GIS-based methodology is required to snap road crashes to the appropriate nodes and line segments. Then, the results should be geo-validated by counter-checking the other spatial descriptors in the crash database, such as street names and crash circumstances, and the data stored in the road network database and the administrative database. Based on the crash database, the road network database, and the administrative database, the locations of 15,026 road crashes in Hong Kong in 2004 were validated. The details are reported in Loo (2006).

The second step is network segmentation and the calculation of crash statistics. This step involves segmenting the network into BSUs and the numbering of all BSUs. Even for a small city like Hong Kong, the road network is 4,249.4 km long. In other words, there are at least 42,494 BSUs. In reality, depending on the spatial configuration of the road network and the segmentation method, the number of BSUs is much higher than that suggested by the total road length. Moreover, many BSUs would be shorter than 100 meters. With GIS, this enormous and highly laborious task can be automated. However, several methodological problems need to be solved. First, double-counting will arise when a crash happened between two BSUs. As the spatial unit of analysis is small, this happens quite frequently. To avoid double-counting, the crash may either be randomly assigned to one of the BSUs or it may be assigned according to a preset rule, such as the BSU with a smaller x coordinate. A thornier methodological problem was the building of the network proximity matrix, which determines the network distance among different BSUs.

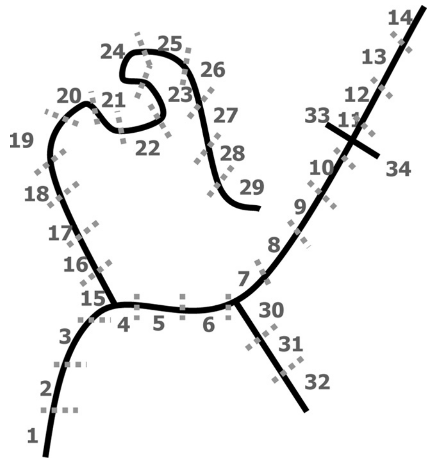


Figure 3. BSUs in the hypothetical network.

This problem can be illustrated in Figure 3. Take the case of BSU 15: which BSUs should be considered as contiguous for the identification of hot zones? The answer is not simply 14 and 16, following the numeric order. Actually, BSU 4 and BSU 16 are contiguous to BSU 15. For BSU 11, for example, its immediate neighbors are 10, 12, 33, and 34. Moreover, BSU 8 is just one BSU away from BSU 30. In a real road network, the problem is much more complicated. This methodological problem can be solved by a GIS algorithm to locate all contiguous road segments. Moreover, the characteristics of the road network, such as the hierarchy of roads, the number of lanes, and the traffic volume (exposure), can be stored in the road attribute table. In addition, the use of GIS allows further spatial analysis by administrative units to shed light on the different crash characteristics in different areas, such as the urban versus the rural areas.

The third step is the identification of hot zones. In relation, a threshold crash rate for a BSU has to be set. For the threshold values of 3, 4, and 5 crashes in a year, the hot zones would have their CRs equal to 6 (HZ_{6+}), 8 (HZ_{8+}), and 10 (HZ_{10+}) crashes in a year, respectively. To reiterate, the primary aim of this paper is to compare the official blacksite methodology with the hot zone methodology. Hence, the values of the CRs have not been set by the more commonly used scientific methods, such as the rate quality method with Poisson distribution or the EB method (Hauer, 1996), but by matching the official blacksite criterion of a minimum of 6 to 9 crashes in a year. Moreover, in order to make a systematic comparison with the annual official list of blacksite locations (such as in Tables 2–4 and Fig. 6 below), the current analysis uses the crash data for 1 year only rather than for 3 (or more) years, which would have minimized the fluctuations caused by randomness (Hauer, 1996; Robert and Veeraragavan, 2004; Cheng and Washington, 2005).

Figure 4 summarizes the GIS-based algorithm involved. The algorithm, which works on small network segments (BSUs) and their crash counts (AR), is based on the assumptions that the precise locations of the crashes be known and that all crashes are geo-validated and snapped to the centerlines of the road network

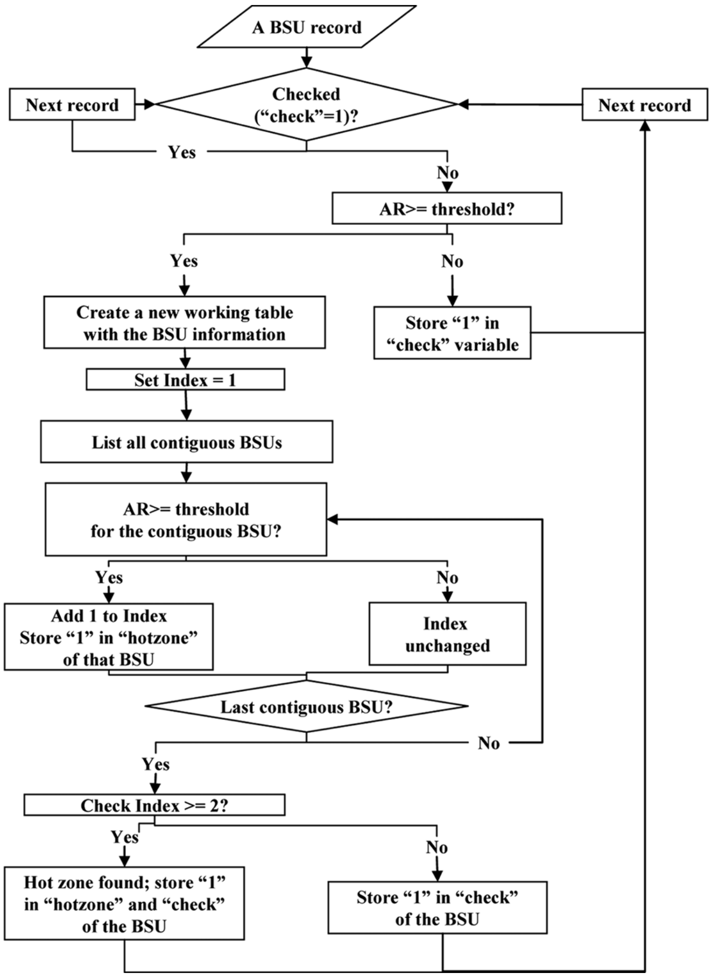


Figure 4. Flowchart showing the steps of identifying hot zones.

(Loo, 2006). To begin with, the first BSU record is examined to see whether its AR is greater than or equal to the threshold crash rate. If the answer is positive, a new working table is created with an index number equal to one. All contiguous BSUs are checked by GIS and listed out in this working table. Then, each contiguous BSU is analyzed. Whenever the AR of a contiguous BSU is also greater than or equal to the threshold crash rate, the index number increases by one. The checking will continue until all contiguous BSUs in the working table have been checked. When this is done, the index number is examined. If the index number is greater than one, a hot zone has been identified and the "hotzone" variable of the BSU in the main table (default = 0) is updated to 1. Because the length of a hot zone is variable, the index number stores the hot zone's length in terms of BSUs. Should the crash severity information be used, the "severity" indices, which record, for example, the cumulative number of traffic crash casualties and the highest level of injury, would also be updated here. The entire process repeats until all BSUs in the road network have been checked. The spatial pattern of hazardous road locations can then be analyzed by plotting all BSUs that form part of a hot zone, that is, BSUs with "hotzone" equal to or greater than one.

4. THE CASE OF HONG KONG

In 2004, there were altogether 15,026 crashes in Hong Kong. Sixty-six blacksites were identified by the Transport Department of the Hong Kong SAR Government for treatment. How serious was the road safety problem if the hot zone methodology is applied? This study focuses on all expressways and the networks of main and secondary roads of 10,000 meters or above.

The results of the hot zone analysis at the territory-wide level are summarized in Table 1. If the threshold values of 5 crashes (HZ_{10+}), 4 crashes (HZ_{8+}), and 3 crashes (HZ_{6+}) in a year are used, the numbers of hot zones identified in 2004 are 24, 61, and 159 respectively. If the individual hot zones are examined, the longest hot zones (Max(n)) consisted of 6 to 8 contiguous BSUs. The maximum crash rate for an individual BSU ($Max(AR_{BSU})$) was 25 crashes in a year for all threshold values. The maximum crash rates of individual hot zones ($Max(AR_{HZ})$) with the threshold values of 5 crashes, 4 crashes, and 3 crashes in a year were as high as 63, 63, and 70 crashes in a year, respectively. These hot zones were obviously dangerous. The information is not only of value to the road safety administration but also road users, including drivers and pedestrians.

If the hot zone methodology is very similar to the existing blacksite methodology in the identification of hazardous road locations, most blacksites would also be hot zones, and vice versa. However, only 5 of the 24 HZ_{10+} were both hot zones and blacksites. In this paper, these dangerous locations are labeled "hot-zones-cum-blacksites" (HZ-cum-BS). The share of HZ-cum-BS was 20.8% for HZ_{10+}. In other words, about 80% of these hot zones were not blacksites. The situation was similar for the other threshold values. Only about 21.3% (13 of 61 HZ_{8+}) and 16.4% (26 of 159 HZ_{6+}) of the hot zones identified contained 1 or more blacksites. Hence, the attempt to introduce the hot zone methodology is not a duplication of the efforts spent on blacksites. The two methodologies are different and they supplement each other.

Next, the hot zones are analyzed by road type. For the road network included in this analysis, the general picture displayed by the 2004 official data—road length, crashes, and blacksites—is shown in Table 2. Moreover, the results of the hot zone

Table 1. Hot zones in Hong Kong, 2004.

	N	Max(n)	$Max(AR_{BSU})$	$Max(AR_{HZ})$	HZ-cum-BS
HZ_{10+}	24	6	25	63	5
HZ_{8+}	61	6	25	63	13
HZ_{6+}	159	8	25	70	26

Table 2. An overview of road safety in Hong Kong by road type, 2004.

	Length, km	Crashes	Blacksites	HZ_{10+}	HZ_{8+}	HZ_{6+}
Expressways	164.5 (10.1%)	868 (9.4%)	0 (0%)	1 (4.2%)	4 (6.6%)	15 (9.4%)
Main roads	550.3 (33.8%)	5314 (57.4%)	38 (57.6%)	16 (66.7%)	38 (62.3%)	102 (64.2%)
Secondary roads	915.2 (56.1%)	3076 (33.2%)	28 (42.4%)	7 (29.2%)	19 (31.1%)	42 (26.4%)

Note. Figures shown in parenthesis are based on the road network included in the analysis.

analysis are juxtaposed. Based on Table 2, a number of important observations can be made. First of all, Hong Kong's expressways were much more dangerous than that portrayed by the blacksites. In fact, crashes on expressways display different spatial characteristics or patterns that can be more effectively identified and analyzed by the hot zone rather than the blacksite methodology. Hot zones on expressways can pose a serious road hazard because of the higher speed of vehicular flows. Hence, when crashes happen on expressways, they tend to cause fatalities and serious injuries. Moreover, in view of the geographical redistribution of population to the new towns and nearby suburb areas, expressways are increasingly important in the local transport system. Many people living in new towns and suburb areas rely on buses running on expressways to commute downtown every day. The hot zone methodology, through identifying the hazardous road locations on expressways, can contribute significantly to the improvement of road safety.

Second, main roads were the most dangerous regardless of the methodology used. In 2004, there were 38 blacksites on the main roads of Hong Kong. The number of hot zones, depending on the threshold crash rates, ranged from 16 to 102. While all these blacksites and hot zones were unsafe to road users, which one(s) should be treated with the highest priority? This is a very complex question that needs to take into account the motive of the road safety administration, the "false-positive" problem, the expected cost, and the potential crash reduction, which can be weighted by crash severity or the number and type of injuries suffered by the traffic casualties (Hauer, 1996; Robert and Veeraragavan, 2004; Brijs et al., 2006). Therefore, the results of the hot zone analysis cannot be used directly to rank hazardous road locations. Nonetheless, the analysis does provide some additional information for improving safety. Using both the official blacksite list and the hot zone analysis, hazardous locations on the road network may be classified into three types. The first type refers to blacksites that are also hot zones. They are called HZ-cum-BS. Next, there are blacksites that do not form part of a hot zone. Finally, there are hot zones that do not contain any blacksite.

Using this typology of classifying hazardous road locations, the general situation in Hong Kong in 2004 is shown in Table 3. For expressways, because there is no blacksite, the number of HZ-cum-BS is also zero. For main roads, there were 2 HZ_{10+}-cum-BS, 7 HZ_{8+}-cum-BS, and 17 HZ_{6+}-cum-BS. For secondary roads, the numbers were 3, 6, and 9 for HZ_{10+}-cum-BS, HZ_{8+}-cum-BS, and HZ_{6+}-cum-BS, respectively. Based on these numbers, the percentages of blacksites that were also hot zones are listed. These percentages can have significant policy implications. For instance, the treatment of the 17 HZ_{6+}-cum-BS on main roads would have

Table 3. Hot zones-cum-blacksites in Hong Kong by road type, 2004.

	HZ_{10+}-cum-BS		HZ_{8+}-cum-BS		HZ_{6+}-cum-BS	
Expressways	0		0		0	
	—	*100%*	—	*100%*	—	*100%*
Main roads	2		7		17	
	5.3%	*87.5%*	*18.4%*	*81.6%*	*44.7%*	*83.3%*
Secondary roads	3		6		9	
	10.7%	*57.1%*	*21.4%*	*68.4%*	*32.1%*	*78.6%*

Notes: 1. Percentages of blacksites in the HZ-cum-BS category are typed in italics.
2. Percentages of hot zones not in the HZ-cum-BS category are underlined.

already included 44.7% of the blacksites identified by the Transport Department of the Hong Kong SAR Government. Similarly, the treatment of the 9 HZ_{6+}-cum-BS on secondary roads would have addressed about 32.1% of the blacksites happening on this road type. In addition, the percentages of hot zones that were not blacksite are also listed. These shares give us an idea about the relative significance of dangerous road locations that were indicated by hot zones only. One can see that the percentages were quite high (ranging from 57.1% to 100%). For main roads, the percentages were above 80% (ranging from 81.6% to 87.5%) for all threshold crash rates.

Last but not least, the official blacksites are analyzed geographically by district. Table 4 compares the distribution of blacksites and hot zones in Hong Kong

Table 4. Analysis of road safety records by district, 2004.

District	**Region**	**Blacksites**	HZ_{10+}	HZ_{8+}	HZ_{6+}
Central and Western	Hong Kong Island	2	0	3	6
Wan Chai	Hong Kong Island	4	1	2	12
Eastern	Hong Kong Island	1	1	5	14
Southern	Hong Kong Island	0	0	0	1
Yau Tsim Mong	Kowloon Peninsular	31	6	16	33
Shum Shui Po	Kowloon Peninsular	11	1	2	9
Kowloon City	Kowloon Peninsular	9	0	2	11
Kwun Tong	Kowloon Peninsular	0	4	4	6
Wong Tai Sin	Kowloon Peninsular	1	2	3	7
Sha Tin	The New Territories	4	0	1	3
Tai Po	The New Territories	0	2	5	10
Kwai Tsing	The New Territories	0	1	3	6
Tsuen Wan	The New Territories	1	0	1	13
Tuen Mun	The New Territories	0	0	0	1
Yuen Long	The New Territories	0	4	8	16
Sai Kung	The New Territories	2	2	4	6
Northern	The New Territories	0	0	2	5
Islands	The New Territories	0	0	0	0
Total	—	66	24	61	159

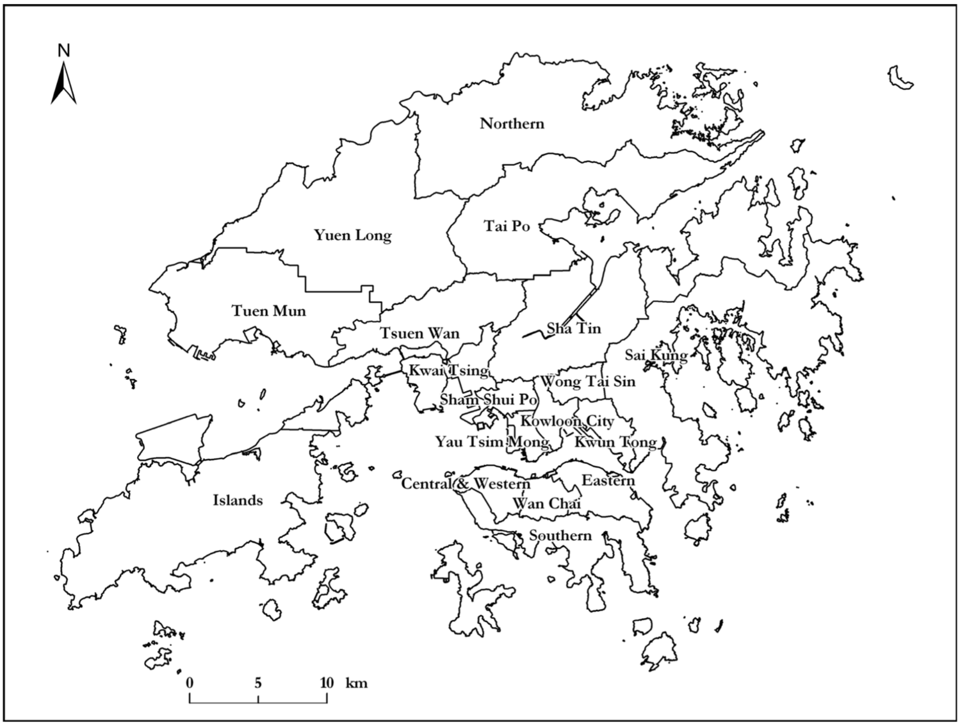
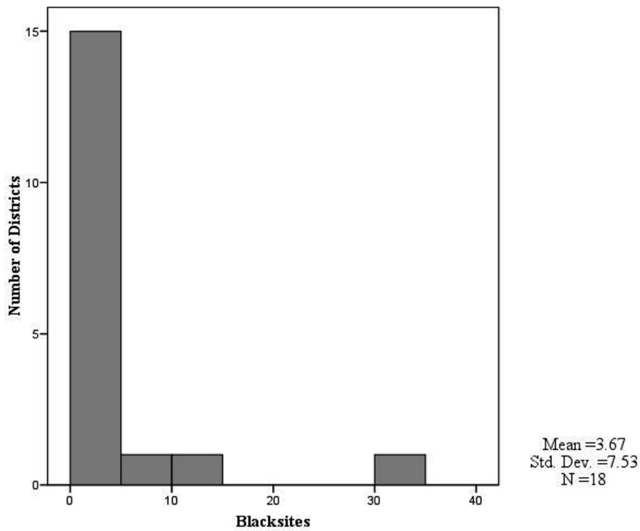


Figure 5. Districts in Hong Kong.

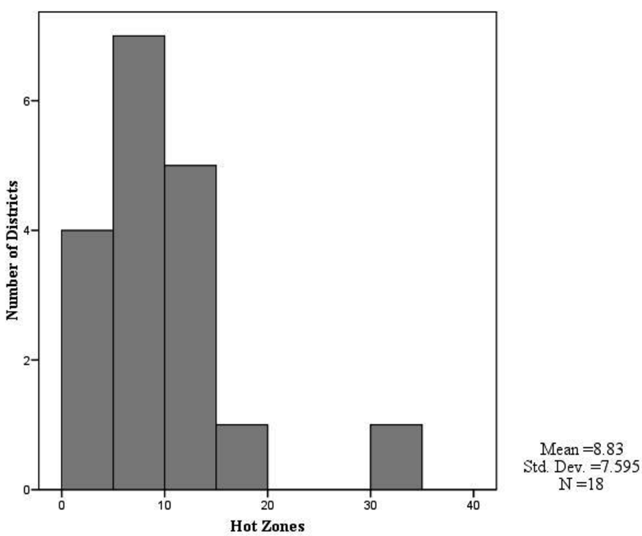
by district. Figure 5 shows the district names and boundaries. In 2004, there were 66 blacksites in Hong Kong. Yau Tsim Mong had the worst record. Nearly half of the blacksites (47.0%, or 31 of 66) were found in this district. Shum Shui Po (11 blacksites) and Kowloon City (9 blacksites) followed. Spatially, the blacksites were highly concentrated. In 2004, the three districts already accounted for 77.3% (51 of 66) of all blacksites in Hong Kong. Moreover, these three districts are all located in the Kowloon Peninsula. Wanchai on the Hong Kong Island and Sha Tin in the New Territories each had 4 blacksites in 2004. There were also a few blacksites in Central and Western, Eastern, Wong Tai Sin, Tuen Wan, and Sai Kung districts. For all the other districts in Hong Kong, there were no blacksites at all. However, was there no systematic spatial clustering of road crashes in these districts?

If one focuses on HZ_{8+} in Table 4, the total number of hot zones identified is comparable with the total number of blacksites identified by the official methodology. Many important spatial features are discernible. First, the hot zone methodology also points to Yau Tsim Mong as the most dangerous district, though the concentration (26.2%, or 16 of 61) was much lower than that suggested by the blacksites (47.0%, or 31 of 66). Moreover, the road safety problem in Shum Shui Po and Kowloon City was much less clustered than that suggested by the blacksites. Both districts had 2 HZ_{8+} in 2004. In contrast, Yuen Long in the New Territories was having a much more clustered pattern. Although no blacksite was identified in Yuen Long, as many as 8 HZ_{8+} were found in the district. In addition, the hot zone methodology was also much more effective in identifying dangerous road locations in the Eastern district on Hong Kong Island and Tai Po district in the New Territories. Both districts had 5 HZ_{8+} in 2004. However, only one blacksite was identified in the Eastern district and no blacksite was found in Tai Po district.

Figure 6 shows the histograms of the blacksites, HZ_{6+}, HZ_{8+}, and HZ_{10+} by district. It can be seen that the frequency distributions of hot zones are much less positively skewed than that of the blacksite. Districts with no hazardous road



(A)



(B)

Figure 6. Histograms of hazardous locations by district. (A) Blacksites; (B) HZ_{6+}; (C) HZ_{8+}; (D) HZ_{10+}.

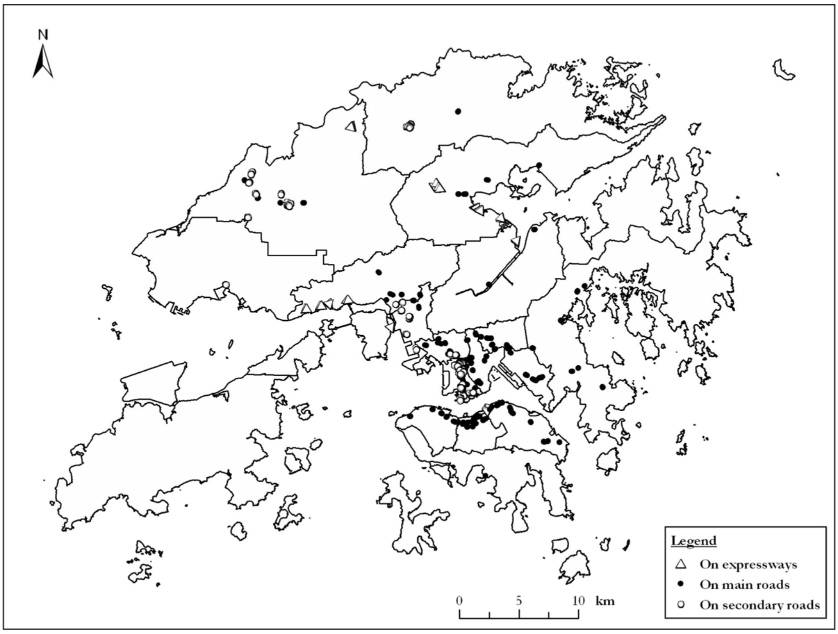


Figure 7. Hot zones with 6 or more crashes in Hong Kong, 2004.

locations were dramatically reduced, when the hot zone methodology (which considers the entire road network rather than just the junction areas) was applied. Moreover, the variability of road safety records among different districts became much smaller. The coefficients of variation (standard deviations over means) were 2.05, 0.86, 1.11, and 1.32 for the blacksite, HZ_{6+}, HZ_{8+}, and HZ_{10+} methodologies, respectively. Other than Yau Tsim Mong District (which clearly performed more poorly than the other districts regardless of the methodologies used), the road safety problems in most other districts have become more apparent with the hot zone methodology. The question of whether all these hot zones were "true-positive" locations requires more vigorous scientific analysis and the pooling of crash data over a longer period of time (Cheng and Washington, 2005). Nonetheless, the current analysis does point to the need of spending more efforts in identifying the hazardous road locations in the non-urban-core areas of Hong Kong.

Figure 7 shows the locations of HZ_{6+} both by road type and by district, with triangles denoting hot zones on expressways, circular dots denoting hot zones on main roads, and squares showing hot zones on secondary roads. It is clear that crashes in different districts exhibited different characteristics (see Figure 5 for district names). For instance, Tuen Wan and Tai Po districts were having many hot zones on expressways. In Tai Po district, 60% (6 of 10) of all HZ_{6+} happened on expressways. Hot zones located on Hong Kong Island, however, mainly happened on main roads. In fact, all hot zones identified in Central and Western, Wanchai and Southern districts were found on main roads. In contrast, hot zones in Yau Tsim Mong district were mostly on secondary roads. Within this district, 19

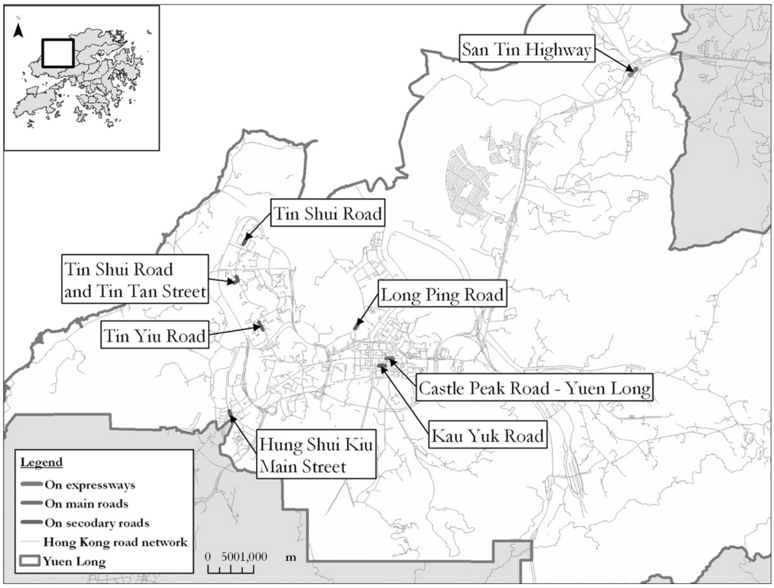


Figure 8. Hot zones with 6 or more crashes in Yuen Long District, 2004.

(57.6%) of the 33 HZ_{6+} occurred on secondary roads. Similar characteristics can also be discerned in the spatial patterns of HZ_{8+} and HZ_{10+}. However, the analysis is more comprehensive with HZ_{6+} because only one district (Islands) had no hot zone in 2004 (see Table 4). The regional characteristics should be useful to district transport offices and police in targeting their road safety improvement measures, including publicity programs.

To round up and to show the usefulness of the hot zone methodology at the local level, Figure 8 shows the hot zone locations (HZ_{6+}) in Yuen Long district. The blacksite methodology suggests that there was *no* systematic spatial clustering of road crashes in Yuen Long in 2004. None of the 66 blacksites were located in this district. However, the hot zone methodology suggests that crashes did systemically cluster in Yuen Long. There were 4 HZ_{10+} in 2004 (Table 4). The number increased to 16 when HZ_{6+} were analyzed. All these hot zones were unidentified hazardous road locations. In particular, Figure 8 shows that road crashes systematically clustered at San Tin Highway on the expressways and Castle Peak Road-Yuen Long on the main roads. For secondary roads, hot zones were found on Long Ping Road, Tin Shui Road, Tin Shui Road-Tin Tan Street, Tin Yau Road, Hung Shui Kiu Main Street, and Kau Yuk Road. These data are valuable to the road safety administration, especially at the local level.

5. CONCLUSIONS AND WAYS FORWARD

To conclude, accurate spatial data are essential and valuable. Based on the analysis in Hong Kong, the spatial characteristics of crashes are likely to vary by road type and by district. Moreover, though the hot zone methodology supplements the blacksite methodology, it is superior and more flexible in many aspects, especially in the identification of hazardous road locations on expressways and in the rural areas. As a step further, two additional sets of research questions need to be addressed. The first set of problems relates to the other aspects of the identification of hazardous road locations. In particular, the setting of the criteria of "hazardous" by taking into account exposure, crash severity, and time period (that is, the definition of CR) and an evaluation of the "false-positive" problem are important. These additional research efforts would shed valuable light on the robustness and sensitivity of the hot zone methodology. The second set of questions relates to the treatment of hot zones. One must realize that the identification of hot zones (even "true-positive" ones) is only just the beginning. Moreover, it is the means rather than the ends. A big challenge ahead is to eliminate the spatial clusters of crashes at hot zones without causing accident migration. For instance, if full pedestrianization is introduced at the hot zones, there will be no vehicles and, hence, no crashes. However, will the number of crashes nearby increase? To handle the road safety problems found at hot zones, there are no simple solutions. Nonetheless, it is certain that multidisciplinary research efforts would be required to address the problems.

ACKNOWLEDGMENTS

This research project is funded by the Research Grant Council (HKU 7296/04H).

REFERENCES

Austin K. 1993. The collection and use of additional sources of road safety data in highway authorities. Traffic Engineering and Control 34:540–543.

Austin K. 1995. The identification of mistakes in road crash records: Part 1, locational variables. Accident Analysis and Prevention 27:261–276.

Black WR. 1991. Highway accidents: A spatial and temporal analysis. In: Transportation Research Board, Transportation Research Record 1318. Washington, DC: Transportation Research Board, National Research Council, pp. 75–82.

Black WR, Thomas I. 1998. Accidents on Belgium's motorways: A network autocorrelation analysis. Journal of Transport Geography 6:23–31.

Brijs T, Bossche FV den, Wets G, Karlis D. 2006. A model for identifying and ranking dangerous accident locations: A case study in Flanders. Statistica Neerlandica 60:457–476.

Cheng W, Washington SP. 2005. Experimental evaluation of hotspot identification methods. Accident Analysis and Prevention 37:870–881.

Deacon JA, Zegeer CV, Deen RC. 1975. Identification of hazardous rural highway locations. In: Transportation Research Board, Transportation Research Record 543. Washington, DC: Transportation Research Board, National Research Council, pp. 16–33.

Elvik R. 2006. New approach to accident analysis for hazardous road locations. Transportation Research Record 1953. Washington, DC: Transportation Research Board, pp. 50–55.

Flahaut B, Mouchart M, Martin ES, Thomas I. 2003. The local spatial autocorrelation and the kernel method for identifying black zones: A comparative approach. Accident Analysis and Prevention 35:991–1004.

Flahaut B. 2004. Impact of infrastructure and local environment on road unsafety: Logistic modelling with spatial autocorrelation. Accident Analysis and Prevention 46:1055–1066.

Hauer E, Persaud BN. 1984. Problem of identifying hazardous locations using accidents data. In: Transportation Research Board, Transportation Research Record 975. Washington, DC: Transportation Research Board, National Research Council, pp. 36–43.

Hauer E. 1996. Identification of sites with promise. In: Transportation Research Board, Transportation Research Record 1542. Washington, DC: Transportation Research Board, National Research Council, pp. 54–60.

Hoque M, Andreassen DC. 1986. Pedestrian accidents: an examination by road class with special reference to accident 'cluster.' Traffic Engineering and Control 27:391–397.

Ibrahim K, Silcock DT. 1992. The accuracy of accident data. Traffic Engineering and Control 33:492–496.

Khan MA, Kathairi AS Al, Grib AM. 2004. A GIS based traffic accident data collection, referencing and analysis framework for Abu Dhabi. Proceeding Codatu XI in 2004 in Bucharest (Romania): Towards More Attractive Urban Transportation, Association CODATU, France. Available at: http://www.codatu.org/english/publications/proceeding/conference/codatu11/Papers/khan.pdf. Accessed July 15, 2005.

Khisty CJ, Lall BK. 2003. Transportation Engineering: An Introduction. Upper Saddle River, NJ: Prentice Hall.

Loo BPY. 2006. Validating crash locations for quantitative spatial analysis: A GIS-based approach. Accident Analysis and Prevention 38:879–886.

Nicholson AJ. 1989. Accident clustering: some simple measures. Traffic Engineering and Control 30:241–246.

Ogden KW. 1996. Safer Roads: Guide to Road Safety Engineering. Avebury Technical, Aldershot.

Robert VR, Veeraragavan A. 2004. Hazard-rating scores for prioritization of accident-prone sections on highways. In: Transportation Research Board, Transportation Research Record 1878. Washington, DC: Transportation Research Board, National Research Council, pp. 143–151.

Shen J, Gan A. 2003. Development of crash reduction factors: methods, problems, and research needs. In: Transportation Research Board, Transportation Research Record 1840. Washington, DC: Transportation Research Board, National Research Council, pp. 50–56.

Shinar D, Treat JR, McDonald ST. 1983. The validity of police reported accident data. Accident Analysis and Prevention 15:175–191.

Silcock D, Smyth AW. 1984. The methods used by British Highway Authorities to identify accident blackspots. Traffic Engineering and Control 25:542–545.

Stokes RW, Mutabazi MI. 1996. Rate-quality control method of identifying hazardous road locations. In: Transportation Research Board, Transportation Research Record 1542. Washington, DC: Transportation Research Board, National Research Council, pp. 44–48.

Transport Department. 2001. Transport Planning & Design Manual, volume 5. Hong Kong: Government Printer, Hong Kong.

Transportation Research Board. 1982. Highway Accident Analysis Systems. Washington, DC: Transportation Research Board, National Research Council.