

# **RESEARCH@LOCATE'15**

## **Proceedings of Research@Locate 15 in conjunction with Locate 15**

Brisbane, Australia  
March 10-12, 2015

Bert Veenendaal, Allison Kealy (Eds.)

## **Editors**

### **Bert Veenendaal**

Department of Spatial Sciences

Curtin University

Bentley, WA 6102

Australia

b.veenendaal@curtin.edu.au

### **Allison Kealy**

Department of Infrastructure Engineering

University of Melbourne

Parkville, VIC 3010

Australia

a.kealy@unimelb.edu.au

## **Preface**

Locate is the annual conference on spatial information in Australia and New Zealand, bringing together for its second year all joint forces of SSSI, SIBA, ANZLIC, OSP, LINZ, CRC SI, PSMA, ASIERA and others. Locate is the meeting point of industry, government and academia in one of the fastest growing areas of IT.

Research@Locate, the academic research stream at Locate, aims to be the premier academic meeting event in the Australasian region. Research@Locate is independently organized by the Australasian Spatial Information Education and Research Association (ASIERA). Research@Locate provides a transparent full-paper peer review process, with carefully selected presentations and papers, and with its own annual, open-access proceedings.

Research@Locate is organised by the Australasian Spatial Information Education and Research Association (ASIERA, [www.asiera.org.au](http://www.asiera.org.au)). ASIERA represents a significant part of the academic segment of the spatial information industry in Australia, with several hundred people in fundamental and applied research and innovation, and with a responsibility for educating and training future generations of spatial professionals.

### Acknowledgements

Research@Locate would not have happened without the support of the institutions behind Locate, SIBA and SSSI. We also wish to thank our colleagues that served on the International Program Committee, and Kate Rampellini for assisting with the review organisation and the production of the proceedings.

Bert Veenendaal, Chair  
Allison Kealy, Co-chair

March 2015

## **Program Committee**

### **Program Committee Chairs 2015**

Bert Veenendaal, Curtin University, Australia (Chair)

Allison Kealy, The University of Melbourne, Australia (Co-chair)

### **Program Committee 2015**

Colin Arrowsmith, RMIT University, Australia

Jagannath Aryal, University of Tasmania, Australia

David Bruce, University of South Australia, Australia

Mark Shortis, RMIT University, Australia

Xiaoli Deng, University of Newcastle, Australia

Allison Kealy, The University of Melbourne, Australia

Ahmed El-Mowafy, Curtin University, Australia

Don Grant, RMIT University, Australia

Ori Gudes, Curtin University, Australia

Eric Guilbert, The Hong Kong Polytechnic University, Hong Kong

John Hayes, Queensland University of Technology, Australia

Mohsen Kalantari, The University of Melbourne, Australia

Samsung Lim, University of New South Wales, Australia

Xiaoye Liu, University of South Queensland, Australia

Kim Lowell, Cooperative Research Centre for Spatial Information, Australia

Feng Lu, State Key Laboratory of Resources and Environmental Information Systems, China

Mahender Kotha, Department of Earth Science, Goa University, India

Kevin McDougall, University of South Queensland, Australia

Antoni Moore, University of Otago, New Zealand

Gerhard Navratil, TU Vienna, Austria

Dev Raj Paudyal, University of South Queensland, Australia

Femke Reitsma, University of Canterbury, New Zealand

Chris Rizos, The University of New South Wales, Australia

Pascal Sirguey, University of Otago, New Zealand

Xiaohua Tong, Tongji University, China

Bert Veenendaal, Curtin University, Australia

Jinling Wang, The University of New South Wales, Australia

Geoff West, Curtin University, Australia

Martin Tomko, The University of Melbourne, Australia

Stephan Winter, The University of Melbourne, Australia

Todd Robinson, Curtin University, Australia

Medhi Ravanbakhsh, RMIT University, Australia

Cecilia Xia, Curtin University, Australia

Wen-zhong Shi, The Hong Kong Polytechnic University, Hong Kong

Matt Duckham, The University of Melbourne, Australia  
Maria Vasardani, The University of Melbourne, Australia  
Abbas Rajabifard, The University of Melbourne, Australia  
Petra Helmholtz, Curtin University, Australia  
Tim Malthus, CSIRO, Australia  
Qin Cheng-Zhi, State Key Laboratory of Resources and Environmental Information Systems,  
China

## **Steering Committee 2015**

Research@Locate is backed by the current members of ASIERA:

Jagannath Aryal, University of Tasmania, Australia  
David Bruce, University of South Australia, Australia  
Mark Shortis, RMIT University, Australia  
Xiaoli Deng, University of Newcastle, Australia  
John Hayes, Queensland University of Technology, Australia  
Kevin McDougall, University of South Queensland, Australia  
Femke Reitsma, University of Canterbury, New Zealand  
Chris Rizos, The University of New South Wales, Australia  
Pascal Sirguy, University of Otago, New Zealand  
Stephan Winter, The University of Melbourne, Australia  
Bert Veenendaal, Curtin University, Australia

## Members

The University of Melbourne  
The University of New South Wales  
RMIT University  
Curtin University  
The University of Newcastle  
University of Tasmania  
University of Southern Queensland  
University of South Australia  
Queensland University of Technology  
University of Canterbury  
University of Otago



Brought to you by:



# Table of Contents

## Session 1: Geospatial analysis and modelling

Maintaining Relational Consistency in a Graph-Based Place Database.....	1-12
<i>Hao Chen, Maria Vasardani &amp; Stephan Winter</i>	
Site Based Data Modelling for Indigenous Land Tenure Systems.....	13-20
<i>Simon Watkinson &amp; Bert Veenendaal</i>	
A Voxel-Based Skewness and Kurtosis Balancing Algorithm for Updating Road Networks from Airborne Lidar Data.....	21-33
<i>Li Liu &amp; Samsung Lim</i>	
A Qualitative Evaluation of a Proposed Metro Map for Melbourne.....	34-45
<i>William Cartwright</i>	
A Simulation Study on Automated Transport Mode Detection in Near-Real Time using a Neural Network.....	46-57
<i>Rahul Deb Das, Nicole Ronald &amp; Stephan Winter</i>	

## Session 2: Position and mapping

AUSGeoid09 Performance in Mountainous Terrain: A Case Study in the Blue Mountains..	58-68
<i>Joseph Allerton, Volker Janssen &amp; Bill Kearsley</i>	
A Spatial Analysis Approach to Evacuation Management: Shelter Assignment & Routing..	69-77
<i>Xuefen Liu &amp; Samsung Lim</i>	
A Data Model for Integrating GIS and BIM for Assessment and 3D Visualisation of Flood Damage to Building.....	78-89
<i>Sam Amirebrahimi, Abbas Rajabifard, Priyan Mendis &amp; Tuan Ngo</i>	
Positioning Eye Fixation and Vehicle Movement: Visual-motor Coordination Assessment in Naturalistic Driving .....	90-99
<i>Qian Sun, Jianhong Xia, Nandakumaran Nadarajah, Torbjörn Falkmer, Jonathan Foster &amp; Hoe Lee</i>	
Effectiveness of DOS (Dark-Object Subtraction) Method and Water Index Techniques to Map Wetlands in a Rapidly Urbanising Megacity with Landsat 8 Data.....	100-108
<i>Sean Gilmore, Ashty Saleem &amp; Ashraf Dewan</i>	
Robust Methods for Feature Extraction from Mobile Laser Scanning 3D Point Clouds...	109-120
<i>Abdul Awal Md Nurunnabi, Geoff West &amp; David Belton</i>	

# Maintaining Relational Consistency in a Graph-Based Place Database

Hao Chen  
hchen@student.unimelb.  
edu.au

Maria Vasardani  
mvasardani@unimelb.  
edu.au

Stephan Winter  
winter@unimelb.  
edu.au

Department of Infrastructure Engineering,  
The University of Melbourne, Australia

## Abstract

People use natural language (NL) descriptions to communicate spatial information, mostly referring to space in qualitative terms. In this research, a graph database is used to store such qualitative spatial information as derived from NL descriptions. It focuses on developing and testing qualitative spatial reasoning mechanisms to maintain relational consistency within the graph database. The study provides a first step into using a graph database to storing and querying qualitative spatial data from NL place descriptions, and provides some insights for the system implementation.

## 1 Introduction

People represent environment in natural language (NL) descriptions typically in a qualitative way of spatial features and their relationships, such as “Building *A* is to the left side of Building *B*” or “The campus is a little bit north from the city”. Such qualitative information could provide an intuitive approach for representing human spatial knowledge. Since this knowledge can be extracted in the form of triplets of locatum-relationship-relatum (as in “campus”-“north of”-“city”) a formal model of two nodes and a linking edge comes to mind. Graph databases, which have already proven useful for modelling qualitative relationships in other areas, present a way to store and query these triplets.

This research focuses on building such a graph database that can store triplets while maintaining relational consistency. The research will go back to Qualitative Spatial Reasoning (QSR) for consistency checking. QSR has been used in other contexts, such as robot navigation, e.g. (Moratz, Nebel et al. 2003), general QSR calculi in AI, e.g. (Frank 1991, Freksa 1992, Zimmermann and Freksa 1996), or for systems modelling NL descriptions, e.g. (Belouaer, Brosset et al. 2013, Basiri, Amirian et al. 2014).

The hypothesis of this study is that a graph-based place database can preserve relational consistency in transactions about triplets representing locata, qualitative spatial relationships, and relata. It is anticipated that, as an outcome, a graph-based place database is able to identify and flag if new input is violating the current consistency and causing logical contradictions. For this purpose the paper studies the relationships and calculi in question, suggests reasoning constraints and algorithms for consistency checking, and tests these algorithms.

The rest of this paper is structured as follows: Section 2 discusses related work and tools. Section 3 explains the methodology for data modelling, as well as reasoning configurations to preserve relational consistency. Section 4 explains the experiments and provides a demonstration of the implemented system. The experiment results presented in Section 5 and discussed in Section 6. Section 7 includes conclusions and suggestions of future work.

---

*Copyright © by the paper's authors. Copying permitted only for private and academic purposes.*

In: B. Veenendaal and A. Kealy (Eds.): Research@Locate'15, Brisbane, Australia, 10-12 March 2015, published at <http://ceur-ws.org>

## 2 Related work

### 2.1 Graph database as a NoSQL database

The NoSQL database family (NoSQL means ‘not only SQL’) includes a variety of database management systems that are not restricted to relational SQL models. It is not in contrast to but rather complements relational databases (Basiri, Amirian et al. 2014). Graph databases are one of the major databases in the NoSQL family other than key-value, document and column databases (Tiwari 2011). Graph databases are based on graphs (Biggs, Lloyd et al. 1986) and employ nodes and edges with properties to store data and their relations.

Graph databases have two advantages in modelling rich-connected data, compared to relational databases (Güting 1994, Wiebrock, Wittenburg et al. 2000). First, they provide a natural and efficient way for network-based data modelling using nodes and edges, while relational databases are more suitable for modelling set-based data. Second, querying connected data can be cheaper in a graph database by traversing paths rather than by a join operation (or even recursive join) in a relational database.

Graph databases are today used in modelling information from social networks, such as Facebook, as well as general knowledge, for example the Google Knowledge Graph. In this study, a graph database is used to model place information since it provides a natural, suitable, efficient and flexible way for modelling spatial objects and the qualitative spatial relationships among them, with no need for pre-defining a tabular schema as in a relational database.

### 2.2 Modelling spatial information in a graph structure

The concept of triplets has been widely utilized (Wiebrock, Wittenburg et al. 2000, Sproat, Coyne et al. 2010, Kordjamshidi, Frasconi et al. 2012). One specific triplet, that of locatum-qualitative spatial relationship-relatum, has been introduced as the *spatial property graph* (Vasardani, Timpf et al. 2013). It consists of the spatial object to be located (the *locatum* L), the reference object (the *relatum* R) and the qualitative spatial relationship between them (r). So far, the spatial property graph has been used for constructing plausible sketch maps (Vasardani, Timpf et al. 2013). Belouaer et al. introduced another rule-based approach to generate a spatial semantic network inferred from verbal descriptions (Belouaer, Brosset et al. 2013). Both approaches focus on identifying semantic similarity among stored nodes. Our study, on the other hand, focuses on the relationships instead of spatial objects.

### 2.3 Qualitative Spatial Reasoning and Calculi

QSR is a constraint-based process that enables spatial intelligence to reasoning about space (Cohn and Hazarika 2001, Dylla, Frommberger et al. 2006). QSR algorithms can be used to solve constraint-based problems by reasoning on different types of qualitative spatial relations.

In this context, Frank suggested QSR with *cardinal directions* and *qualitative distance* relationships (Frank 1991, Frank 1992). For cardinal directions he suggested a cone model (Figure 1a), half-plane models (Figure 1b), and a neutral zone model (Figure 1c). The neutral zone model with its composition table is applied in this study, with additional rules to increase the flexibility of natural language.

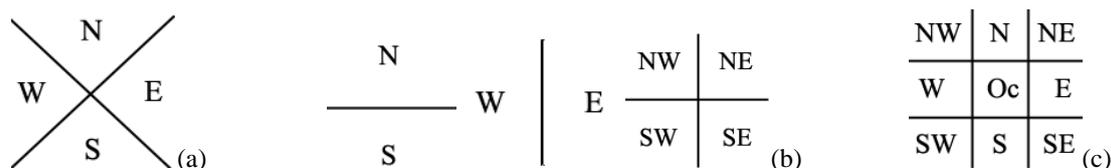


Figure 1: Cardinal direction systems

For qualitative distance relationships Frank discussed two-step (close, far), three-step (close, middle, far) and multi-step systems (Frank 1992). Those systems consider qualitative distance as mappings into intervals that form a partition of the positive real numbers, such as (for a three-step system) close =  $[0, 1)$ ; middle =  $[1, 3)$ ; far =  $[3, \infty)$  (Figure 2a). In a linear space such relationships can be added, and  $\text{dist}(A,B) + \text{dist}(B,C) \geq \text{dist}(A,C)$ . The resulting interval will then be mapped back to the corresponding symbols (close, middle, far) again. In place databases, however, stored place configurations are from two-dimensional space, where nothing else than the usual triangle inequality holds. Figure 2b shows for example a situation where ‘A is middle to B’ and ‘B is middle to C’ means that A and C can be in any relationship, for example near.

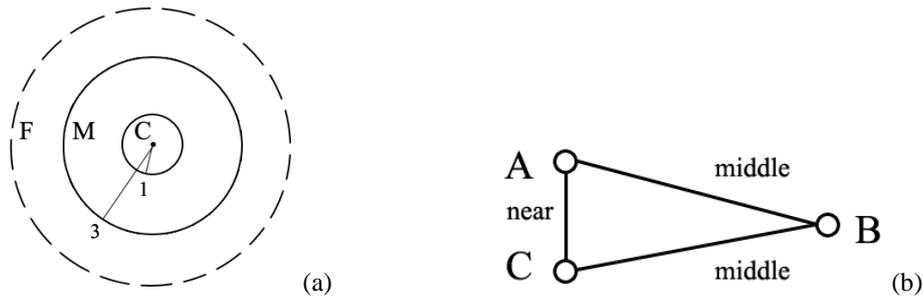


Figure 2: Deriving qualitative distance relationship on a plane

Freksa provided an approach for reasoning with *relative direction* relationships (Freksa 1992). Later, Zimmermann and Freksa discussed composition cases (Zimmermann and Freksa 1996). However, relative direction relations are not considered in this study, for reasons discussed in Section 3.

Egenhofer et al. discussed a methodology for modelling and reasoning with *topological* relations (Egenhofer and Herring 1990, Egenhofer and Franzosa 1991), known as the 4-intersection and 9-intersection models. The two models consider whether the interior ( $A^\circ$  and  $B^\circ$ ), the boundary ( $\partial A$  and  $\partial B$ ) and the exterior ( $A^c$  and  $B^c$ ) of two spatial objects intersect, and map the different cases to topological relationships. A composition table was also introduced for reasoning (Egenhofer 1991). An alternative model, based on first order logic instead of point-set topology, is the Region Connection Calculus (RCC) introduced by Randell, Cui and Cohn in 1992 (Randell, Cui et al. 1992). The 4-intersection model and the RCC both distinguish the same eight relations between two simple connected spatial objects. In this study, Egenhofer’s 4-intersection model and composition table are used.

The different calculi of QSR, each of them applicable on one type of spatial relationships, have been packed into toolboxes. For example, SparQ (Wallgrin, Frommberger et al. 2007) is an integrated system which provides a series of Qualitative Spatial Calculi functionalities and aims at supporting tasks related to constraint-based reasoning and qualitative consistency-checking. In this research, however, the flexibility that comes with the use of natural language requires a critical discussion and extension of such calculi.

### 3 A relationally consistent graph-based place database

Triples of place information, as extracted from NL description, can be integrated and stored in a graph database. The graph database should only store relationships that are consistent with each other in order to avoid contradictions in querying or reasoning, even if the information is not necessarily true according to geographic reality. In this section, methods will be provided to maintain relational consistency.

#### 3.1 Modelling place information using graph database

A triplet consisting of {locatum, relation, relatum}, such as {Building A, to the right of, Building B} is represented in a graph database using two distinct nodes for locatum and relatum, and a directed edge from the locatum to the relatum, as shown in Figure 3. In the following sections, L-r->R will be used to represent a triplet, such as “Building A-to the right of->Building B”.

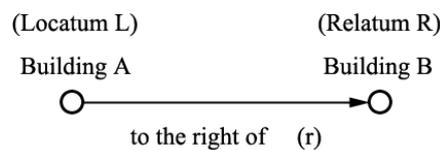


Figure 3: A stored triplet of two nodes and one edge

Each time a new triplet is to be added, the names of the locatum and relatum will be queried to verify whether already nodes of these names exist in the database. If the nodes already exist, the triplet is integrated in the existing graph, otherwise new nodes are created and connected with directed edges. The graph can become a multi-graph when more than one triplet are added with the same locatum and relatum (e.g., A-right of->B; A-west of->B).

#### 3.2 Relational consistency

Relational inconsistency occurs in a graph database when two pieces of information can be derived from the stored data that are logically contradicting. For example, if a real world situation as shown in Figure 4 is described by “A-west of->B” and then by “B-west of->A” there is a contradiction. Contradictions can also occur over longer cycles, such as in “A-west of->B”, “B-west of->C” and then “C-west of-A”.

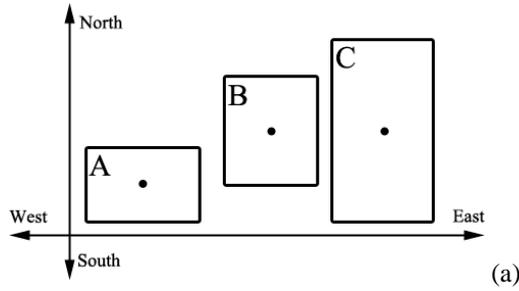


Figure 4: Three buildings in the cardinal reference frame

**Definition:** Two triplets are relationally-consistent if, without any additional knowledge, no contradicting information can be derived according to the provided reasoning rules.

**Definition:** A graph database is relationally-consistent if, at any state, it is free from relational inconsistency within the stored data.

### 3.3 Composition of relationships

A compositional inference is a deduction based on two relationships  $R_1(A, B)$  and  $R_2(B, C)$ , which results in  $R_3(A, C)$  (Cohn and Hazarika 2001). Qualitative spatial calculi, with a limited set of expressions, may require  $R_3(A, C)$  to be a set of possible relations  $S_3(A, C)$ . Let us denote  $\circ$  for the composition operation between two relationships. Composition operation can (but not necessarily) follow the following rules:

Commutative law:  $A \circ B = B \circ A$

Associativity law:  $(A \circ B) \circ C = A \circ (B \circ C)$

Distributive law:  $A = \{A_1..A_i\}$ ,  $B = \{B_1..B_j\}$ ,  $A \circ B = \sum_{m=1}^i \sum_{n=1}^j A_m \circ B_n$

A table can be constructed to represent the composition results  $S_3(A, C)$ . For a calculus of  $n$  possible relationships, an  $n*n$  composition table consists of each composition result between any two relationships in  $n$ .

### 3.4 General reasoning configuration

For the time being, any new triplet  $L-r \rightarrow R$ ,  $r$  must be from one of the three categories: cardinal direction, qualitative distance, or topological relationship. All other relationships are not yet covered; they might be stored without consistency checking, or conservatively just rejected.

In order to identify relational inconsistency, any new input triplet  $L-r \rightarrow R$  will be compared to existing knowledge between  $L$  and  $R$  derived from the stored data.

**Definition:** Given a triplet  $L-r \rightarrow R$ , an *existing path* (EP) is a sequence of nodes and edges in the graph database that starts from  $L$  and ends with  $R$ , while all the relationships in the path belong to the category of  $r$ .

**Definition:** An *existing knowledge* (EK) is a set of derived relationships between  $L$  and  $R$  from an EP. If the length of EP equals to one, then EK is the single relation in the EP. If the length of EP is greater than one, then for all relationships with  $1..n$  in EP in sequence order,  $EK = r_1 \circ r_2 \circ .. r_j$ , which could be either a single relationship or a set of relationships.

A new triplet  $L-r \rightarrow R$  is regarded consistent with a database if it is relationally consistent with all EKs between  $L$  and  $R$ . If the triplet is consistent, it will be added to the graph database; otherwise an inconsistency will be flagged.

### 3.5 Reasoning configuration for cardinal direction relationships

Table 1 below shows the implemented cardinal direction relationship set (left), and possible verbalizations (right).

Cardinal Directions	Example of NL Descriptions
north of	north, to the north of, northern, north from, N, etc.
northeast of	northeast, to the northeast of, northeast from, NE, etc.
east of	east, to the east of, eastern, east from, E, etc.
southeast of	southeast, to the southeast of, southeast from, SE, etc.
south of	south, to the south of, southern, south from, S, etc.
southwest of	southwest, to the southwest of, southwest from, SW, etc.
west of	west, to the west of, western, west from, W, etc.
northwest of	northwest, to the northwest of, northwest from, NW, etc.

For cardinal directions this research adopts the neutral zone system (Figure 1c), as it is assumed that all spatial features have a two-dimensional extend. As for any other system, each cardinal direction can be associated with a region such that these regions form a JEPD partition of the Euclidian plane. The size of the neutral zone is defined by the relatum R, while the locatum L in a triplet L-r->R belongs to one of the eight zones (NW, N, NE, E, SE, S, SW and W).

Table 2 shows Frank’s composition table for the neutral zone model. ‘Any’ in the table means again that any of the eight relationships is possible. However, the composition table is not directly applied in this work. This is because the semantics of cardinal directions in NL descriptions are more vague than a logical model suggests. For example, when people say North, they could mean anywhere within zones of W, NW, N, NE and E—in the neutral zone model—instead of only N, because if the half-planes of cardinal directions intersect in a valid relation, it seems that people consider them as conceptually consistent (i.e., non contradicting), as they can generalize from the more specific relations, to the more general and broader application region of each relation (half-planes). Therefore, without more detailed knowledge, a certain level of uncertainty should be accepted and not be regarded as inconsistency. For instance, in reasoning N and NW should be considered as consistent while N and S should be regarded as inconsistent.

Table 2: Frank’s composition table for the neutral zone model

	N	NE	E	SE	S	SW	W	NW
N	N	N	NE	E	Any	W	NW	N
NE	N	NE	E	E	E	Any	N	N
E	NE	E	E	E	SE	S	Any	N
SE	E	E	SE	SE	S	S	S	Any
S	Any	E	SE	S	S	S	SW	W
SW	W	Any	S	S	S	SW	W	W
W	NW	N	Any	S	SW	W	W	W
NW	N	N	N	Any	W	W	N	NW

**Definition:** An *extended cardinal direction set* (ECD) is a set of consistent cardinal direction relationships given a cardinal direction relationship r in a NL description.

The ECD for any cardinal direction relationship is shown in Table 3. The table provides a set of relationships that are regarded as relationally consistent with r. For instance, “Building A-north of->Building B” and “Building A-east of->Building B” are considered as relationally consistent, and their half-plane intersection exists in the form of a valid relation north-east (NE); while “Building A-north of->Building B” and “Building A-south of->Building B” are relationally inconsistent, as obviously an intersection of their half-planes does not exist.

Table 3: Extended cardinal direction set for r

	ECD set
N	W, NW, N, NE, E
NE	N, NE, E
E	N, NE, E, SE, S
SE	E, SE, S
S	E, SE, S, SW, W
SW	S, SW, W
W	S, SW, W, NW, N
NW	W, NW, N

ECDs provide a ‘tolerant’ mechanism that first stores triplets which are considered to be relationally consistent. Then, when more triplets with new knowledge between the same two spatial objects are stored, and the number of EP (and EK) increases, the system becomes less tolerant, as a new input triplet will need to be consistent with every EK in order to be consistent with the graph database.

**Rule 1:** Given an input triplet L-r->R and r is a cardinal direction relationship, if for every EK between L and R, r belongs to the ECD of at least one of the elements in that EK, then the triplet is relationally consistent with the graph database.

### 3.6 Relative direction relationships

Relative direction relationships require a reference frame for relative directions. The reference frame can be linked to the orientation of the observer (egocentric, to *my* left), the orientation of the relatum (allocentric, to the *left* of the front side), or be a projection of the speaker to the reference frame of the recipient (*your* left). Unfortunately triplets do not provide this spatial reference frame. For instance, when someone says “A is in front of C” then this sentence’s ambiguity with respect to the spatial reference frame cannot be resolved without further context (which is not accessible). This is the reason why in this research reasoning mechanisms for relative direction relationships are not provided.

### 3.7 Reasoning configuration for qualitative distance relationships

Table 4 below shows the implemented qualitative distance relationship set, followed by possible verbalizations.

Table 4: Qualitative distance relationships

Qualitative Distance Relation	Examples of NL descriptions
Near	beside, next to, close to, near, around, etc.
Middle	not far from, middle, etc.
Far	far from, furthest, distance from, very far, end, etc.

The only reasoning rule for the qualitative distance relationship between A and C is the triangle inequality. Frank's three-step intervals are applied: (N = [0, 1); M = [1, 3); F = [3, ∞)). Table 5 shows the corresponding composition table, using the distance computations of Algorithm 1.

Table 5: Composition table of qualitative distance intervals

	N = [0, 1)	M = [1, 3)	F = [3, ∞)
N = [0, 1)	[0, 2) N, M	[0, 4) any	[2, ∞) M, F
M = [1, 3)	[0, 4) any	[0, 6) any	[0, ∞) any
F = [3, ∞)	[2, ∞) M, F	[0, ∞) any	[0, ∞) any

#### Algorithm 1: Distances for qualitative distance interval composition

```

1:  a = [xa, ya), b = [xb, yb)           # a ∘ b = [xc, yc)
2:  yc = ya + yb
3:  if xb ≥ xa:
4:      xc = |xb - ya|
5:  else:
6:      xc = |xa - yb|
7:  return [xc, yc)

```

**Rule 2:** Given an input triplet L-r->R and r is a qualitative distance relationship, if r is element of EK between L and R, then the triplet is relationally consistent with the graph database.

### 3.8 Reasoning configuration for topological relationships

Table 6 shows the implemented topological relationship set (left), followed by possible verbalizations (right).

Table 6: Topological relationships

Topological relation	Example of NL descriptions
Disjoint	Building A is away from building B
Meet	Building A is adjacent to building B
Equal	A and B are actually the same building
Inside	The office is inside Building A
coveredBy	The pond is at the north end of the park
Contains	Building A contains the office
Covers	The park covers the pond at its north end
Overlap	Part of Building A is on the neighbour's parcel

Egenhofer's composition table for the 4-intersection model (Egenhofer 1991) is shown in Table 7. Here, 'any' means no more specific information can be deducted: any topological relationship is possible. The commutative law does not apply for the composition operation between topological relationships. For instance, *disjoint* ∘ *inside* = {d, m, i, cB, o} while *inside* ∘ *disjoint* = d.

Table 7: Composition table of topological relationship

	disjoint	meet	equal	inside	covered By	contains	covers	overlap
disjoint	any	d,m,i, cB,o	d	d,m,i, cB,o	d,m,i, cB,o	d	d	d,m,i, cB,o
meet	d,m,ct, cv,o	d,m,e, cB,cv,o	m	i,cB,o	m,i,cB,o	d	d,m	d,m,i, cB,o
equal	d	m	e	i	cB	ct	cv	o
inside	d	d	i	i	i	any	d,m,i, cB,o	d,m,i, cB,o
covered By	d	d, m	cB	i	i,cB	d,m,ct, cv,o	d,m,e, cB,cv,o	d,m,i, cB,o
contains	d,m,ct, cv,o	ct,cv,o	ct	e,i,cB, ct,cv,o	ct,cv,o	ct	ct	ct,cv,o
covers	d,m,ct, cv,o	m,ct, cv,o	cv	i,cB,o	e,cB,cv,o	ct	ct, cv	ct,cv,o
overlap	d,m,ct, cv,o	d,m,ct, cv,o	o	i,cB,o	i,cB,o	d,m,ct, cv,o	d,m,ct, cv,o	any

**Rule 3:** Given an input triplet  $L-r \rightarrow R$  and  $r$  is a topological relationship, if  $r$  is element of EK between  $L$  and  $R$ , then the triplet is relationally consistent with the graph database.

## 4 Implementation

The data used to test the algorithms are 731 triplets derived from 42 place descriptions collected by Vasardani et al. (Vasardani, Timpf et al. 2013). The triplets were extracted from NL descriptions of the University of Melbourne’s Parkville campus, given by a group of graduate students with different levels of familiarity with the campus. These descriptions had already been pre-processed by NL parsing rules, synonym detection and relationship classification. Of these 731, 325 triplets have a cardinal direction, a qualitative distance, or a topological relationship.

### 4.1 System overview

The Neo4j community version is used in this research as local host visualization platform. The Neo4j Python-embedded is the main module used to interact with the local database host. The system structure is shown in Figure 5.

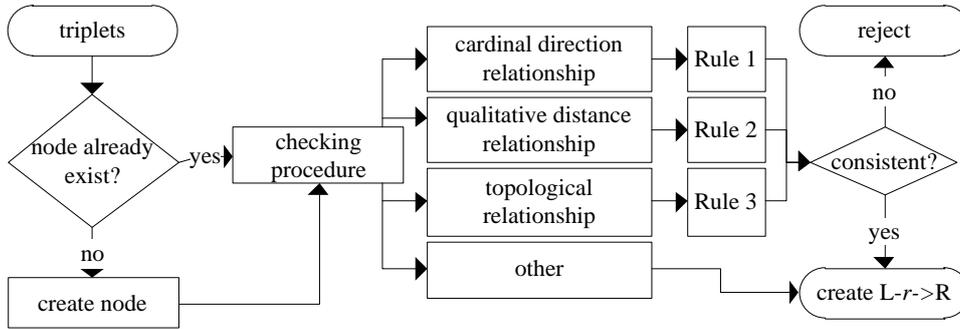


Figure 5: Diagram of reasoning process

### 4.2 Experiment design

The experiment consists of two parts. In Part 1, during the import stage, the reasoning algorithms were applied to identify inconsistent triplets from the set of 731. Each flagged triplet is then manually verified of whether it really causes relational inconsistency or not. Thus, rate  $P_1$  calculated by Formula 4.2.1, is the precision of the algorithm. For instance, if the system detects 10 inconsistent triplets, and 9 of them are verified, then  $P_1$  equals to 0.9 (maximum 1).

$$P_1 = \frac{\text{number of verified inconsistent triplets}}{\text{number of inconsistent triplets identified}} \quad (4.2.1)$$

In Part 2, after all triplets are imported, an additional set of 60 made-up consistent and inconsistent triplets of the three categories of relationships were input to the database. In this case it was known in advance which of the triplets are consistent or inconsistent. Again it was tested whether the system is able to correctly accept or reject each of them accordingly. The rate  $P_2$  calculated by Formula 4.2.2 provides another insight on the precision of the algorithm.

$$P_2 = \frac{\text{number of correctly accepted or rejected triplets}}{\text{number of all made-up triplets}} \quad (4.2.2)$$

In line with the hypothesis it is anticipated that both  $P_1$  and  $P_2$  equal to 1, in which case the graph database is proven to be able to maintain relational consistency.

### 4.3 Maximum query depth

When querying for EPs that start from node L and end with R, a maximum path length has been set, since computing all the paths between two nodes in a graph draws on computation time. This theoretical drawback can only be justified by pragmatics: First, it can be observed that after just a few steps of reasoning, composition tables in many places end with ‘any’ anyway. The second pragmatic argument is based on the First Law of Geography: Near things are more related than far things, which raises the expectation that relevant spatial information is provided in short cycles. The parameter for the maximum path length, Maximum Query Depth (MQD), has been tested in order to observe when the overall computation time becomes unacceptable. This test was done by adding all the 731 triplets into an empty graph database and recording the overall computation time of reasoning for each triplet, over a range of MQD.

## 5 Observed Results

For Part 1, all 731 triplets were added from the descriptions, and no inconsistent triplet occurred. All triplets were then verified, and  $P_1=1$ . For Part 2, 60 triplets were tested. 60 out of 60 were processed correctly—both all consistent and all inconsistent cases were identified—by the graph database, thus  $P_2=1$ .

Figure 6a shows the overall number of EP computed in Part 1 for every triplet from the set of 325 in the considered relation categories. Figures 6b, 6c, and 6d show the number of EP for triplets for each respective category (MQD = 3).

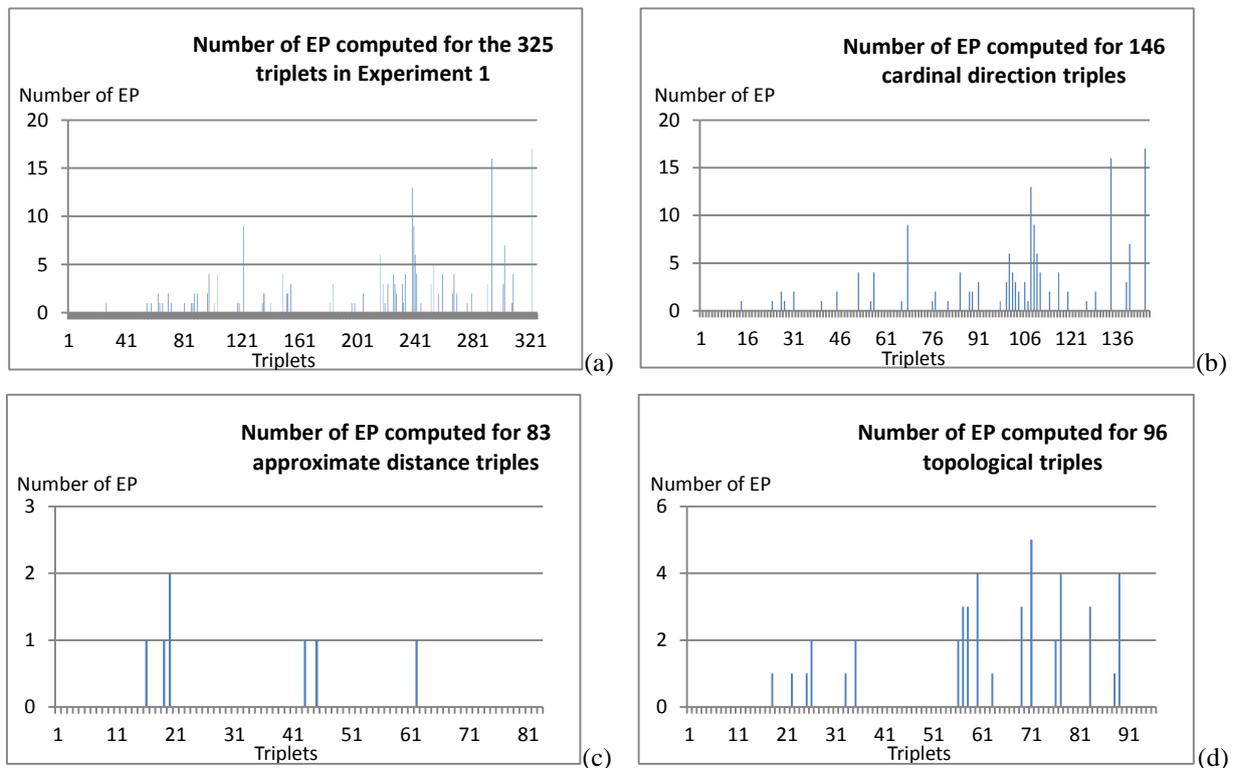


Figure 6: Number of EP computed in Part 1

Figure 7 shows that once a made-up consistent triplet was accepted and added, the following inconsistent triplet has at least one more EP queried than the previous one.

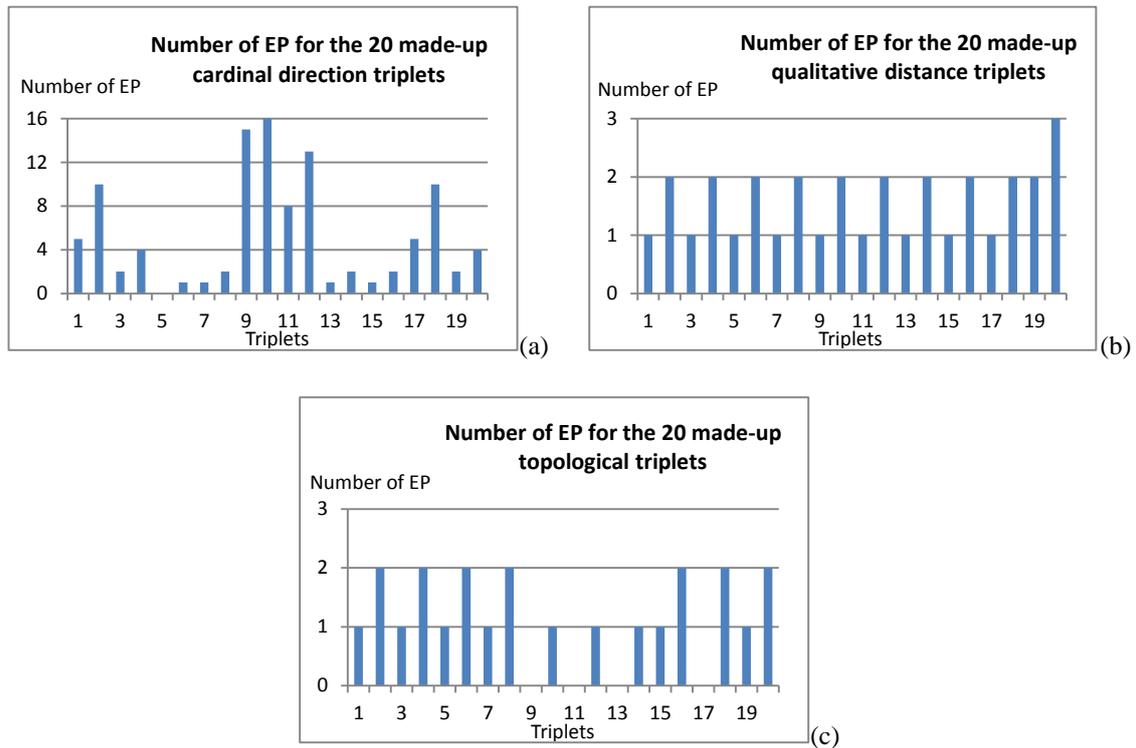


Figure 7: Number of EP computed in Part 2

Computation time for MQD that equals to 1, 2, 3 and 4 respectively is shown in Figure 8. When MQD is set to 1, 2 and 3, the computation time is rather stable (around 0.1s). But when it is set to 4, the computation time suddenly begins to soar as the size of the graph database grows.

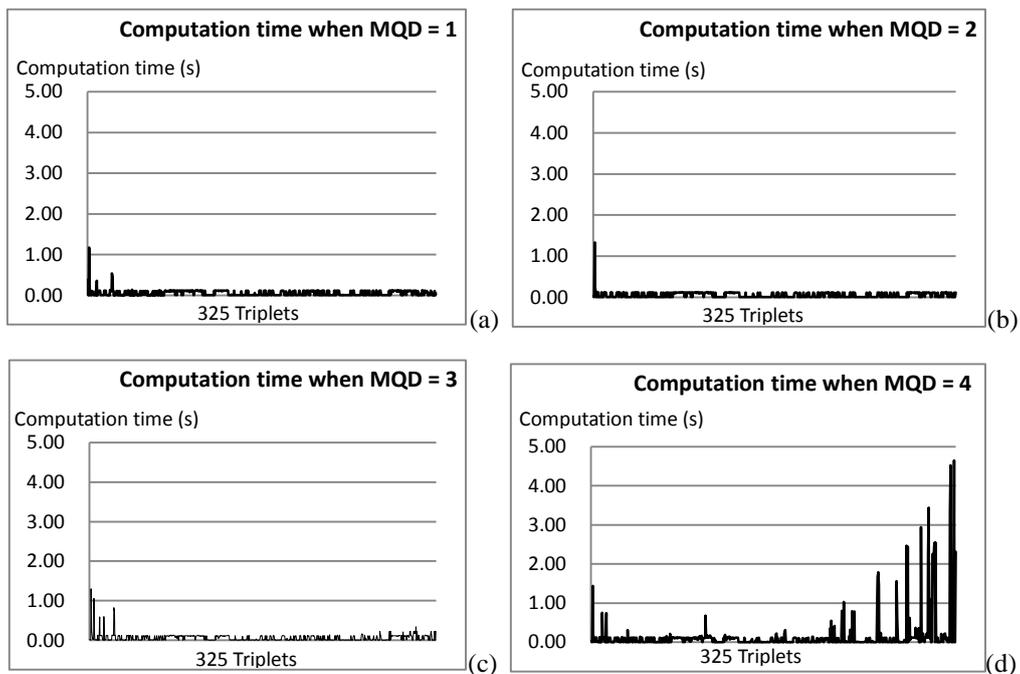


Figure 8: Computation time for MQD equals to 1, 2, 3 and 4

## 6 Discussion

The implementation of relational consistency rules is the first step of automatically reason with qualitative spatial relationships from NL descriptions in a graph database. The implemented mechanisms are flexible and robust to preserve relational consistency in a graph database.

In Part 1 of the experiment, the 731 triplets extracted from the 42 place descriptions are relationally consistent. This result is not surprising, since the descriptions are from students that are already, to various extents, familiar with the environment of the University of Melbourne's Parkville campus, therefore the possibility of giving misleading descriptions is low (but still possible, such as by mistake, and which makes the consistency checking mechanisms necessary). Alternatively, also malicious users could try to add inconsistent knowledge and should be prevented. However, the system does only flag an inconsistency; it has no mechanisms to decide whether the already stored or the added relations are the truer representation of geographic reality.

According to the result given by Part 2, the system is capable of distinguishing consistent and inconsistent triplets. Additionally, as shown in Figure 6, the numbers of EP found for qualitative distance relationships are overall small. This indicates that there are not as many neighbouring relationships that belong to the qualitative distance relationship category, therefore it is not as rich-connected as cardinal direction and topological relationships.

The system is flexible and accommodates certain degrees of ambiguity and uncertainty. The graph database is robust for maintaining relational consistency with given rules and formalized triplets, even though the reasoning algorithms are flexible.

Among the input spatial objects, 'Campus' is a special one. It is rather a container of most other spatial objects in these (campus) place descriptions. Among all 731 triplets, 144 of them (approximately 19.7%) related to 'Campus'.

Although relative direction relationships were not reasoned, it was observed that some relative direction relationships appear 'inconsistent' in expression when considered without their individual spatial reference frame. For instance, in one triplet, the Engineering Department is "right of" the South Lawn, while in another triplet, the relationship between them is "left of". Triplets of relative directions occupy approximately 25% of the total number of triplets; hence they must be addressed in future work by ways of capturing and maintaining their spatial reference frame.

Additionally, each relationship is only checked for consistency within its own category. Relational consistency checking among different categories is not considered in this research; therefore the system is not yet able to reason relationships from multiple categories. For example, "A-inside->B" and "A-far->B" are contradicting since A and B must be disjoint in order to be far from each other. Due to the fact that 'inside' and 'far' do not belong to the same category, they will not be regarded as relationally-inconsistent by the implemented database. Thus, reasoning rules among different categories of relationship should be developed in future research.

It was anticipated that by setting different MQD, the number of EP queried for any triplet  $L-r->R$  would be different. Also, it is expected that the reasoning results would be different since cycles longer than MQD can still be inconsistent with a triplet to be added. When implemented, no difference in reasoning result was detected by setting different MQD; however, the difference of number of EP detected is significant.

Although more EP could be queried by setting larger MQD (and consequently more EK can be used for consistency checking), it is suggested to set MQD to 3 in order to limit the computation time. As shown in Figure 8, it can be predicted that with a larger dataset, the computation time would soon become unacceptable with larger MQD.

## 7 Conclusions and future work

This research proves that a graph database can preserve relational consistency when modelling qualitative place information derived from NL. The hypothesis is verified by the outcomes of the experiments described in earlier section. The implemented system is overall flexible enough to accommodate NL descriptions and robust for maintaining relational consistency.

Data triplets in this research include spatial relationships from categories such as cardinal directions, qualitative distances, topological relations, relative directions and other (non-categorized) relationships. Reasoning rules were developed for the first three categories. When a new input triplet violates a consistency rule, it will be flagged by the graph database to prevent logical contradiction and relational inconsistency. Currently, the system can only process reasoning based on relationships that belong to the same category. It is not yet able to reason on triplets with relationships from multiple categories.

The path length of reasoning cycles has to be limited for computational reasons. Different MQD were tested for composition reasoning, and an MQD of 3 still delivered acceptable computation times.

This research is the first step in attempting to build a relationally consistent database for place information from NL. Such a graph database, built and maintained in relational consistency, can then be used to solve relevant decision making tasks, such as landmark navigation; or be used for other applications, such as plausible sketch map drawing.

The major limitations of this research are the limited size and context of the test dataset (all campus descriptions of the same campus). More interesting observations are expected if the graph database is used for testing datasets referring to different locations. Second, the robustness of mapping NL descriptions of qualitative distance

relationships into intervals needs further research. It is known that qualitative distance judgements change with the context in a conversation, do not follow the commutative law (from “A is near B” does not necessarily follow that, within the same context, “B is near A”), and negation within the three-step model is not addressed. For instance, “not far” could mean either ‘near’ or ‘middle’ in a description, however it is mapped to ‘middle’ in this research. Third, as discussed, the system is not yet able to reason over relationships from different categories, and also it is not able to reason with relative direction relationships.

Apart from the limitations discussed above, other work can also be done in future studies. First, semantic reasoning could be introduced to automatically identify synonyms of relationship labels, thus helping a graph database further to maintain relational consistency over a larger variety of relationship names. Secondly, due to the ambiguity of natural language descriptions, relationships cannot be categorized automatically currently. For instance, ‘in’ and ‘on’ in NL descriptions could stand for not only topological ‘inside’ (i.e. room in the building; clock-tower on the lawn), but also other relations (i.e. station on the road). Mechanisms could be developed by considering the type of the referenced objects as well as the context, in order to categorize relationships automatically.

## References

- Basiri, A., P. Amirian and A. Winstanley (2014). Use of Graph Databases in Tourist Navigation Application. *Computational Science and Its Applications–ICCSA 2014*, Springer 8583: 663-677.
- Belouaer, L., D. Brosset and C. Claramunt (2013). Modeling spatial knowledge from verbal descriptions. *Spatial Information Theory*, Springer 8116: 338-357.
- Biggs, N., E. K. Lloyd and R. J. Wilson (1986). *Graph Theory*, 1736-1936, Clarendon Press.
- Cohn, A. G. and S. M. Hazarika (2001). "Qualitative spatial representation and reasoning: An overview". *Fundamenta informaticae* 46(1-2): 1-29.
- Dylla, F., L. Frommberger, J. O. Wallgrün and D. Wolter (2006). SparQ: A toolbox for qualitative spatial representation and reasoning. *Qualitative Constraint Calculi: Application and Integration*, Workshop at KI.
- Egenhofer, M. J. (1991). Reasoning about binary topological relations. *Advances in Spatial Databases*, Springer 525: 141-160
- Egenhofer, M. J. and R. D. Franzosa (1991). "Point-set topological spatial relations". *International Journal of Geographical Information System* 5(2): 161-174.
- Egenhofer, M. J. and J. Herring (1991). Categorizing binary topological relations between regions, lines, and points in geographic databases. 9: 94-91.
- Frank, A. U. (1991). Qualitative spatial reasoning with cardinal directions. *Österreichische Artificial-Intelligence-Tagung/Seventh Austrian Conference on Artificial Intelligence*, Springer 287: 157-167.
- Frank, A. U. (1992). "Qualitative spatial reasoning about distances and directions in geographic space." *Journal of Visual Languages & Computing* 3(4): 343-371.
- Freksa, C. (1992). Using orientation information for qualitative spatial reasoning, *Spatio-Temporal Reasoning*, Springer 639: 162-178.
- Güting, R. H. (1994). GraphDB: Modeling and Querying Graphs in Databases. In *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB)*, Morgan Kaufmann: 297-308.
- Kordjamshidi, P., P. Frasconi, M. Van Otterlo, M.-F. Moens and L. De Raedt (2012). Relational learning for spatial relation extraction from natural language. *Inductive Logic Programming*, Springer 7207: 204-220.
- Mackaness, W., P. Bartie and C. S.-R. Espeso (2014). Understanding Information Requirements in “Text Only” Pedestrian Wayfinding Systems. *Geographic Information Science*, Springer 8728: 235-252.
- May, A. J., T. Ross and S. H. Bayer (2005). "Incorporating landmarks in driver navigation system design: An overview of results from the REGIONAL project." *Journal of Navigation* 58(1): 47-65.
- Moratz, R., B. Nebel and C. Freksa (2003). Qualitative spatial reasoning about relative position. *Spatial cognition III*, Springer 2685: 385-400.

- Randell, D. A., Cui, Z. and Cohn, A. G. (1992). A Spatial Logic based on Regions and Connection. 3rd Int. Conf. on Knowledge Representation and Reasoning, Morgan Kaufmann: 165-176.
- Sproat, R., R. E. Coyne and J. B. Hirschberg (2010). Spatial Relations in Text-to-Scene Conversion, Computational Models of Spatial Language Interpretation Workshop at Spatial Cognition 2010.
- Tiwari, S. (2011). Professional NoSQL, John Wiley & Sons.
- Vasardani, M., S. Timpf, S. Winter and M. Tomko (2013). From descriptions to depictions: A conceptual framework. Spatial information theory, Springer 8116: 299-319.
- Wallgr ün, J. O., L. Frommberger, D. Wolter, F. Dylla and C. Freksa (2007). Qualitative spatial representation and reasoning in the SparQ-toolbox. Spatial Cognition V Reasoning, Action, Interaction, Springer 4387: 39-58.
- Wiebrock, S., L. Wittenburg, U. Schmid and F. Wysotzki (2000). Inference and visualization of spatial relations. Spatial Cognition II, Springer 1849: 212-224.
- Zimmermann, K. and C. Freksa (1996). "Qualitative spatial reasoning using orientation, distance, and path knowledge." Applied Intelligence 6(1): 49-58.

# SITE BASED DATA MODELLING FOR INDIGENOUS LAND TENURE SYSTEMS

Simon Watkinson<sup>1,2</sup>  
simon.watkinson@nlc.org.au

Bert Veenendaal<sup>1</sup>  
b.veenendaal@curtin.edu.au

<sup>1</sup>Department of Spatial Sciences, Curtin University, Perth, Western Australia

<sup>2</sup>Anthropology Branch, Northern Land Council, Darwin, Northern Territory

## Abstract

The Native Title Act (1993) and the Aboriginal Land Rights Act (NT) (1976) add an extra layer of consideration to land tenure interests in Northern Australia. Traditional Indigenous perspectives follow a site based ontology which is still practiced across most parts of Northern Australia. This paper presents preliminary research into a site based spatial data model for tracking Indigenous Land Tenure.

## 1 Introduction

Research into Aboriginal land tenure systems in Australia have been largely confined to the anthropological disciplines (for example, see Peterson & Rigsby, 2014), which generally deal in qualitative and theoretical analysis. Ethnographers working within the legal framework of Aboriginal Land Rights Act (1976) (ALRA) and the Native Title Act (1993) (NTA) use Geographic Information Systems (GIS) to produce maps of Indigenous Land Tenure, however much ethnographic information is decoupled from the map and must be accessed through related ethnographic reports. The spatial sciences have thus far produced very little in the literature to assist GIS practitioners working in this field. Specialised published GIS toolsets are typically focused around the practicalities of western interests, such as the management or the exploitation of natural resources and offer little assistance to the problem space.

This paper presents core spatial elements of indigenous land tenure established from a review of ethnographic materials in the study region. The elements are mapped to a relational structure using Unified Modelling Language (UML) modelling techniques. The data model presented here forms the spatial foundation for a much larger, integrated ethnographic GIS which can incorporate and map linguistic, ecological and social organisation factors.

## 2 Review of Ethnographic Literature

Selection of appropriate ethnographic material had two basic requirements, firstly, that the material covered a broad enough geographic area that it could capture variations in land tenure systems in the research area and secondly, the material had to be relevant to the legal framework of ALRA and NTA. The selection process thus consisted of a spatial review of ALRA land claim material. This material was preferable as most ALRA claims were undertaken in a hostile political environment, hotly contested by the Northern Territory Government. Thus the land tenure models presented in these reports were considered to be sufficiently robust and

---

*Copyright © by the paper's authors. Copying permitted only for private and academic purposes.*

In: B. Veenendaal and A. Kealy (Eds.): *Research@Locate'15*, Brisbane, Australia, 10-12 March 2015, published at <http://ceur-ws.org>

reliable, primarily on account that the claims were successful. Most researchers who contributed to this body of literature were from the academic disciplines and in their anthropologists reports provided references to the underlying theoretical body of knowledge in this field, notably Radcliffe-Brown's monograph *The Social Organisation of Australian Tribes* (1931). Figure 1 provides a geographic summary of the claim material in this literature review.

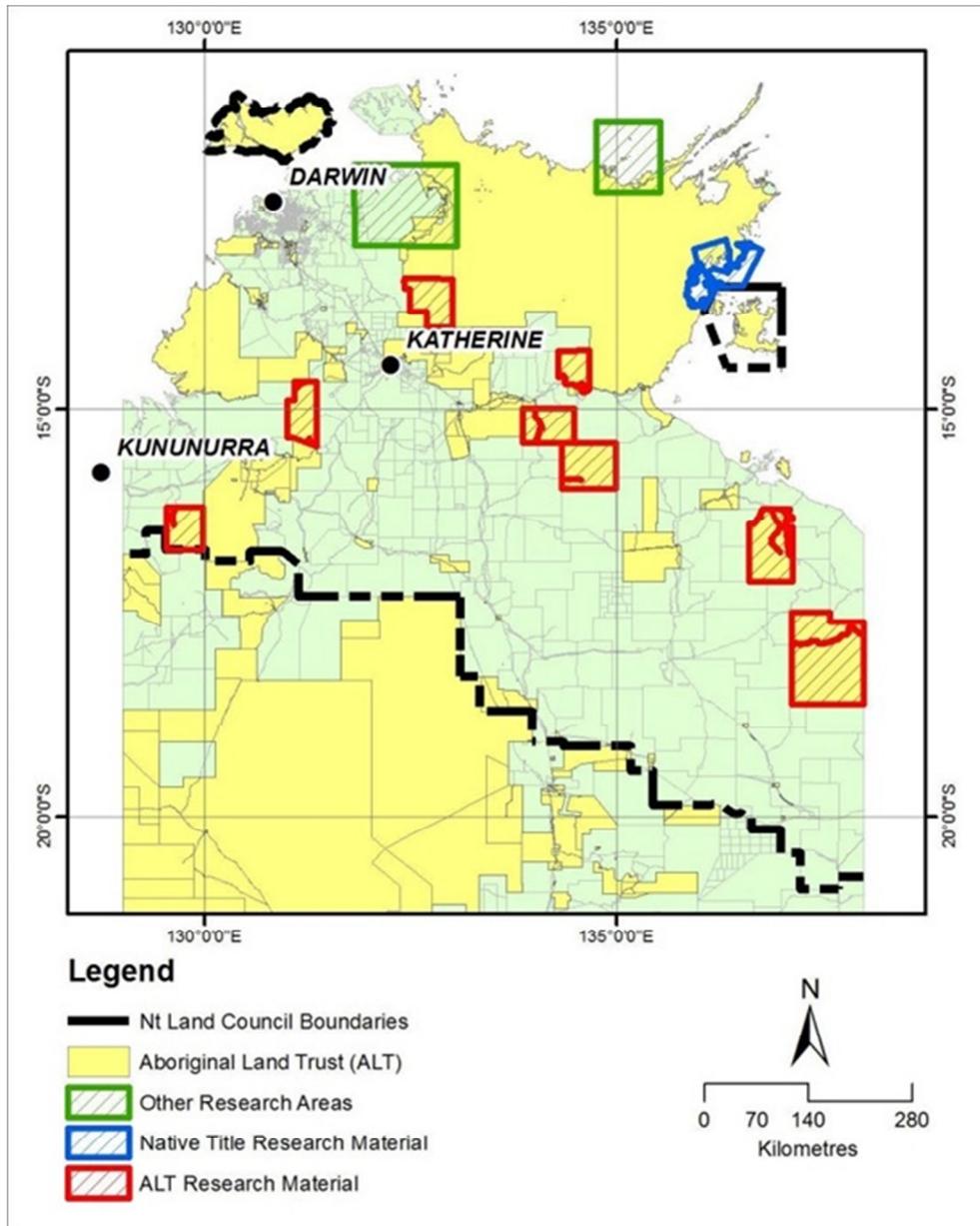


Figure 1 Geographic spread of literature review

## 2.1 Summary of findings of the review

The most remarkable feature of Aboriginal land tenure systems in this review is the profound interrelatedness of a range of factors, together which form an encyclopaedic knowledge base on which the system operates. The five most prevalent factors in this system are location, religion, social organisation, language and ecology. To undertake complete data modelling for all these factors represents a colossal undertaking and is out of scope for this study; however, a brief description will be presented and the basic spatial elements defined.

## 2.2 Basic social organisation

Radcliffe-Brown (1931) provides the baseline research into group organisation of Australian aboriginal groups, and it's that basic structure which formed a common reference point for research throughout the reviewed literature. Radcliffe-Brown (1931) defines the important local group as the *horde*, that being a small group of persons owning territory with known boundaries, and possessing common proprietary rights over the land, and everything on or in it. This primary land owning group is often named the clan or the Local Descent Group (LDG) and it is the latter acronym that will be used here.

Membership of the group is primarily determined by descent through the father's line (the patriline). Male members enter the group by birth and remain in it forever. Women are born into the group, or incorporated through exogamous marriage from another group, with the former case having to join another group upon marriage.

Radcliffe-Brown (1931) provides a description of a larger territorial unit, the *tribe*. The unity of the tribe is primarily linguistic, that being the tribe consists of people who speak a common language or dialects of a common language. The name of the tribe and the language are usually the same, and with the common linguistic affiliation comes a unity of custom throughout the tribe (Radcliffe-Brown, 1931). Radcliffe-Brown (1931) clearly states the tribe, while occupying the combined territory of its member descent groups, is *not* a land owning unit. Land ownership is by the LDG and the LDG is independent and autonomous.

There is often some blurring between tribes and local descent groups. For example, a local descent group could be a subdivision of a tribe, or large enough that it *is* a tribe. Palmer and Brady (1991) grouped all claimants of the Amanbidji Land Claim together as a single group based on linguistic affiliation, however more recent analysis of the claim genealogies reveals there are in fact several LDG's who consider themselves autonomous from each other. In the Alligator Rivers Stage II Land Claim, Keen et al (1980) found that groups in the East of the claim area followed Radcliffe-Brown's basic model, with distinctly named LDG's occupying distinctly named country areas, affiliating with each other through a common language, forming Radcliffe-Brown's *tribe*. However Keen *et al* (1980) found the Western groups identified themselves through language only, referring to themselves and their country by the language name. Regardless of this differing strategy of identification, Keen et al (1980) reported that the LDG existed in effect as there were several patrilineal descent groups occupying differing areas of the language country, but these groups do not seek to identify themselves as separate at the LDG level.

The *tribe*, the LDG (the clan or *horde*) and descent membership are useful starting points in building a model of indigenous land tenure. While the discussion of social organisation ends here, it is important to note the land tenure model remains simplistic without further exploration and inclusion of kinship. Some kin relations are very important as individuals inherit rights and responsibilities to land through them. For Example Layton and Bauman (1996) explain that men who relate to country through their mother and father's mother, while not identified as men with primary spiritual responsibility (i.e. land owners), hold ceremonial and ritual responsibilities for that country (land managers). Morphy and Morphy (2003) identify kin relations manifest as both ego-centric (individuals relating to individuals), and socio-centric (groups relating to groups) both of which include relationships denoting rights and responsibilities to country.

## 2.3 The religious perspective

The problem space in this paper is scoped around legal definitions of indigenous land tenure. The practitioners of anthropology and mapping of land tenure operates from definitions of land ownership in the ALRA and the NTA. The earliest of these definitions comes from the Woodward report, quoting Ronald Berndt, who asserts that 'there are for aborigines two levels of ownership, the primary or religious level and the secondary or economic level' (Woodward, 1974). During the Gove land rights case, Mr Justice Blackburn remarked '...the fundamental truth about the aboriginals' relationship to the land is that whatever else it is, it is a religious relationship' (Northern Territory Supreme Court *et al*, 1971).

The term *dreaming* used in this paper is the generalised term for ancestral beings which created the landscape. Taking from Devitt & Martin (1996), who describe the Wardaman tribal traditions, the dreamings were human-like figures active as ancestors in the past. They travelled and created the landscape, seasons, animals, plants and people in it. The ancestral beings did not die, but they remain present in geographic locations or in species of plants and animals. People trace their origins to specific dreamings, and from those dreamings they inherit land, identity and spiritual power. With that inheritance comes spiritual responsibilities for the places in which they travelled and created, and a topology of spiritual connection to other groups and places can be drawn. James and Baymarrwangga (2013) aptly describe this religious landscape as a site based ontology.

## 2.4 Spatial primitives of Indigenous Land Tenure Systems

### 2.4.1 Sites and named places

The *site* represents the finest grain of land interest in the literature. The site is a named place, which often (but not always) represents a distinct topographic feature such as a hill or a creek. Sometimes it is a grove of trees or even an individual tree. The size of a named place (like anywhere) can vary considerably, sometimes occupying an area large enough to engulf other places, and with a core centre somewhere within it. These named places can be of value because of their carrying capacity for game and other food sources, and of course for the spiritual power imbued within it from the ancestors.

The site can be mapped into a coordinate space as a simple primitive, the point. This of course reveals nothing about the size of the place or its shape. For example, the site “i” shown in Figure 2 is actually a tributary to the Victoria River. It’s assumed full extent is shown in shading. When originally mapped from a Northern perspective, Site “i” was mapped as the headwaters of the tributary. However, travelling along the river, it was identified and named at the mouth. Further interrogation of the matter by Anthropologist John Laurence with senior informants reveal the focal point is usually the river mouth however the place name extends and includes areas up to the headwaters (John Laurence, personal communication, August, 2014).

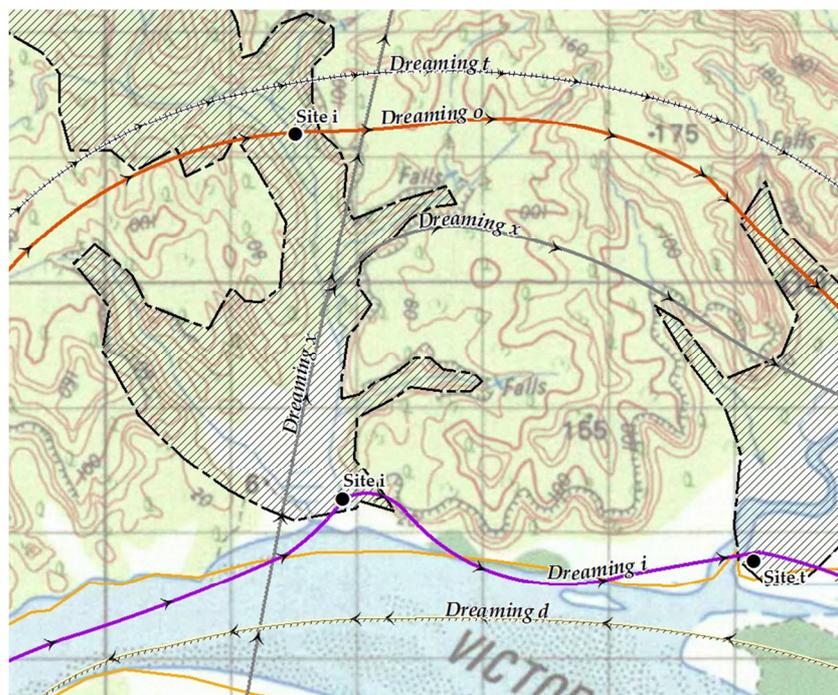


Figure 2 Site example

### 2.4.2 Site vs Estates

As was discussed above, the LDC owns in company a common territory with known boundaries. This territory is often referred to in the reviewed literature as the estate. The concept of boundedness creates some difficulties in mapping. Indeed sometimes Aboriginal land owners remark on a natural feature which forms a distinct end of one estate and the beginning of another. Also, there are ‘handover sites’, where ceremonial management of a travelling dreaming is transferred from one group to another which enables the establishment of a boundary (at least in part). Very rarely can these factors provide enough data to form reliable enclosed estate boundaries. Keen & Lewis (2007) remark the estate boundary is a concept which received more attention from western researchers than indigenous subjects when thinking about land ownership. From the experience of the author, TAOs give less attention to the boundary of an estate; rather, it is the loci of spiritual power within the estate which carries importance.

Sutton (1983) describes the estate as being a collection of sites in which the LDG have primary rights, thus the site collection is a more desirable option as it presents a finer grained representation of land and provides a list of named places which are talking points with informants and more accurate mapping of land interests. Nonetheless,

it can still be very useful to produce maps which represent land interest as a collection of estates (especially for larger scale maps). The site based approach can still cater for this kind of representation using convex hulls and minimum bounding polygons to build estate representations (if needed), and even language or tribal areas.

### 2.4.3 Dreamings

Usually the dreamings visited places in a sequential fashion, creating a topology of events forming what is often referred to as a dreaming track. The logic of the system does not limit dreamings to being present and contained within the estates of a particular LDC, or even the areas of a particular language district. For example, the stormbird dreaming (*Scythrops novae-hollandiae*) migrating from South East Asia into the pastoral belt of the Top End (around 16 degrees south) travels as far south as Port Augusta in South Australia (located at around 32 degrees south), transforms to the Native Cat, and travels North again through the MacDonnell ranges in central Australia and onward to the Northern Coast (Nash, et al. 1993). Nash et al. (1993) goes on to explain that while the dreamings establish ceremonial links between LDG's, and some outside individuals are considered ritual experts in some matters, primary spiritual responsibility for sites remain with the LDG. There was no deviation from this understanding throughout the material reviewed.

## 3 Data Modelling

The review of ethnographic materials enabled the definition of the meaningful entities in indigenous land tenure and the relationships between them. These entities were then modelled into a data structure using the entity relationship (E-R) methodology as explained by Rigaux et al. (2002) using UML principles. Before implementing the model in an RDMBS, the high level conceptual data model was presented to a forum of field anthropologists from the Northern Land Council as well as senior academic researchers from the Australian National University.

The ethnographic data model (Figure 3) focuses on the structural components, and is intended as a guide for GIS practitioners who may be required to assist in ethnographic mapping. Ideally it is best implemented using RDBMS or ORDBMS technology, however this may not be available, and thus a modified data structure for a flat file implementation must be devised. Spatial entities in the diagram are represented as varbinary, which is a MS SQL Server binary data type. This data type is only representational and would vary depending on the target database implementation.

The relational model here presents a basic structural guide to mapping Aboriginal interests in land. The site and the dreaming form the two most important spatial entities, with the Group\_Site table offering a link to derive estate and even language area maps such as those produced by Horton and AIATSIS (1996), or the Tribes of Australia (Tindale, 1974). It is considered that this is the baseline structure required for coupling more complicated facets of Aboriginal social organisation to a mapped space and forming an accessible data repository for ethnographic research and analysis.

### 3.1 Sites

Site representation allows for both a point and area representation, with the Site table holding the primary key (ie Site areas are optional, Site points are compulsory). Traditional owner interest is mapped to sites via the Group\_Site table. Obviously some LDGs become extinct and others move to succession of their estates. For example, Keen and Lewis (2007) reported this in the Kakadu Repeat Claim; thus, for mapping that is intended for long term data recording, the inclusion of time stamps marking the period for which a group holds interest in a place is required.

### 3.2 Language and Local Descent Groups

In this ethnographic data model the Language entity is synonymous with the *tribe* and the Group entity with the LDG or *clan*. The People and People\_Group tables enable individuals to be listed under groups. Strictly speaking, people only belong to a single group through their fathers line, however this was modelled as a M:N relationship to enable easy listing of group membership through the patriline or matriline. The two recursive relationships in the people table enable for multi-generation descent tracking and form a base structure for mapping kinship between individuals (which will be discussed elsewhere).

Both the Language and Estate entities have the capacity for a spatial expression (as polygons). In some instances, researchers will map country at this coarse scale, however, these polygons can also be determined algorithmically from the Sites table. The exact nature of the algorithm obviously would vary depending on the deployment strategy of the GIS. For example, PostGIS offers a concave hull operation which can be used inline with standard SQL statements ([http://postgis.net/docs/ST\\_ConcaveHull.html](http://postgis.net/docs/ST_ConcaveHull.html)), whereas desktop solutions may require a combination of operations such as Voronoi polygons clipped by Delaunay triangulations.

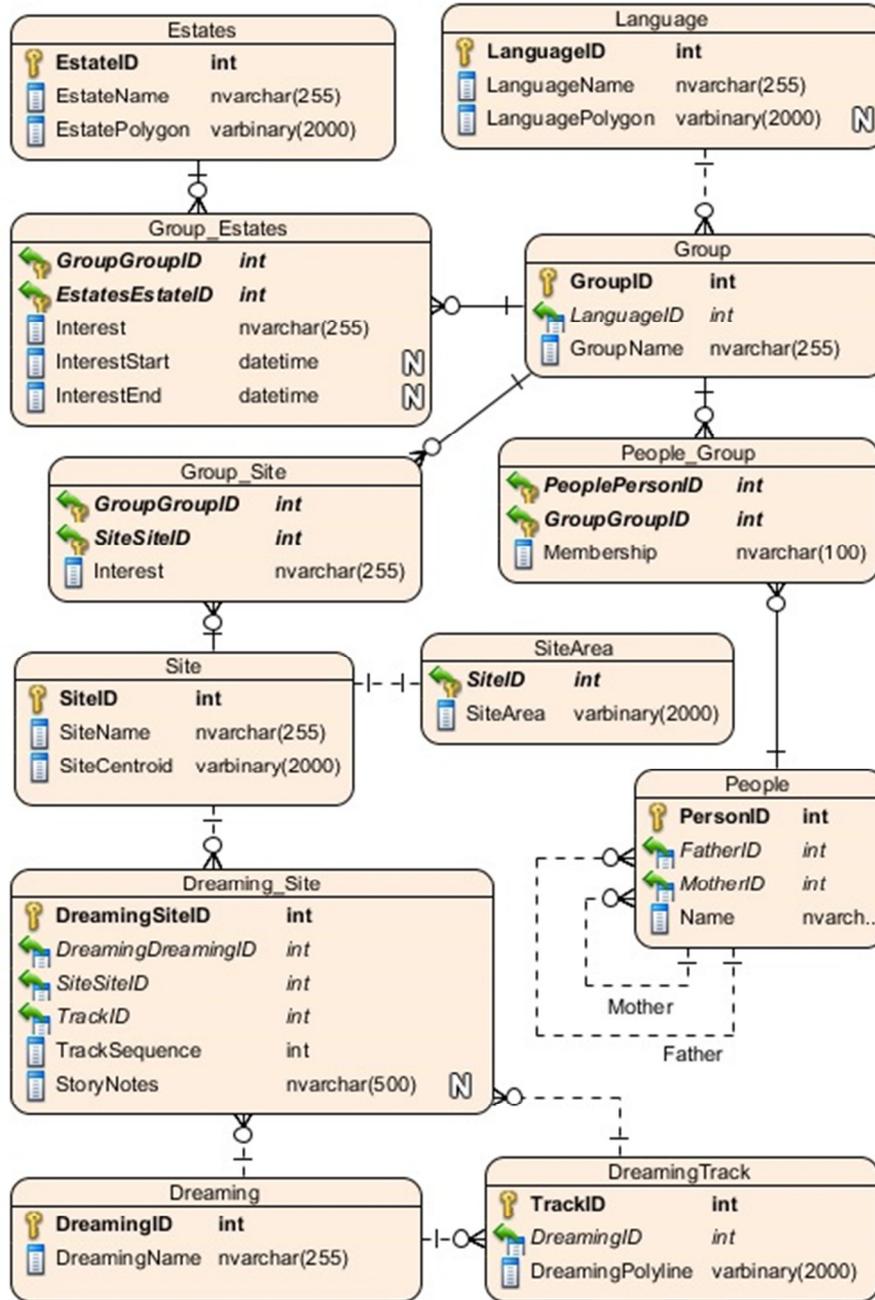


Figure 3 Ethnographic data model

### 3.3 Dreamings

To properly model how dreaming relates to place, four entities are required. The Dreaming table simply acts as an index of individual dreamings which can be linked to any Site via the Dreaming\_Site table. It's often that a dreaming may visit a site, continue to other places and then return to that site later on, thus the composite Primary Key consisting of SiteID and DreamingID must be disregarded and replaced with a simple identity column. In the case that a researcher wishes to list sites in a specific sequence (ie, form a track) then the TrackID and SequenceID fields must be populated. The TrackID and SequenceID combinations must be unique in these instances. These sequenced Dreaming\_Site records can be computed into polylines and stored in the DreamingTrack table. Multiple tracks per dreaming are possible. Alternatively, DreamingTrack entities can be directly drafted into the GIS software as polylines, and the Dreaming\_Site table populated via linear referencing techniques. It is also common that some dreamings do not travel at all, and remain at a single locus. In these instances, the TrackID and SequenceID must be set to a null value, which removes the possibility of the TrackID and the SequenceID being concretely defined as a unique constraint using DDL.

## 4 Expanding the Data Model

The expansion of this data model to include kinship is the first obvious step toward enabling more sophisticated research potential. For example, Trigger (1982) found the subsection system, along with totemic affiliation formed a conspicuous part of land and group identification. Morphy and Morphy (2003) report in Eastern Arnhem land, that marriage preferences are flagged within the kinship system, that exogamous marriage is not symmetrical (groups do not directly swap wives), and that some kin relations are manifest on a group to group level. James & Baymarrwangga (2014), adapting Keen (1978) provide very clear insights as to how these kin relations might be modelled.

A second direction for expansion is the incorporation of a dictionary structure. There are two main reasons for this, firstly, while language changes from one place to another, the social and spiritual connections do not. The same entity can have a name in many different languages. Without the ability for this data model to properly model when a thing is truly different, as opposed to only being different in name, then any one instance of the data model has limited geographic applicability, and broadscale ethno-geographic analysis becomes compromised.

The second reason for dictionary modelling is to assist in data collection. It is important to understand that there is always a divide between the ethnographer and the culture he/she researches. The divide is not only cultural, but it is also linguistic. It is highly likely that the research ethnographer is liaising with people who speak English as a second or third language, and due to their historical circumstances, probably do not have a high level of schooling. Including fields for the linguistic interpretation of named things in the database prompts the researcher to understand the meaning of what is being said to them (adding a failsafe to mapping too many places called "what there?"), and then also provides an inroad into uncovering possible connections that may have otherwise been overlooked.

## 5 References

- Aboriginal Land Rights (Northern Territory) Act. (1976). (Cth).
- Devitt, J. and Martin, D. (1996). Wardaman Claim to Innesvale pastoral lease: Anthropologists report. Darwin, Northern Land Council.
- Graham, R., Holcombe, S., Stead, J., Bern, J., and Wells, S. (2002). Urapunga Station Land Claim: Anthropologists Report. Darwin: Northern Land Council.
- Horton, R.D., Australian Institute of Aboriginal and Torres Strait Islander Studies, Auslig/Sinclair, Knight, Merz. (1996). Aboriginal Australia. Canberra: Aboriginal Studies Press.
- James, B. and Baymarrwangga, L. (2014). The Yan-nhangu Atlas and Illustrated Dictionary of the Crocodile Islands. Singapore: Tien Wah Press.
- Keen, I. (1978). One Ceremony, One Song: An Economy of Religious Knowledge among the Yolngu of northeast Arnhem Land, PhD thesis. Canberra: Australian National University.
- Keen, I., Barker, G., Chaloupka, G., Fraser, R., Haritos, A. and Mobbs, R. (1980). Alligator Rivers Stage II Land Claim. Darwin: Northern Land Council.
- Keen, I., and Lewis, G. (2007). Kakadu Repeat Claim: Anthropologists' Statement. Darwin: Northern Land Council.
- Morphy, H., and Morphy, F. (2003). An Anthropological report on the Yolngu people of Blue Mud Bay, in relation to their claim to native title in the land and sea. Darwin: Northern Land Council.
- Nash, D., Sutton, P., and Morel, P. (1993). Muckaty Pastoral Lease Land Claim: Anthropologists Report. Darwin: Northern Land Council.
- Native Title Act (1993). (Cth).
- Northern Territory Supreme Court., Blackburn, R.A., Milirrpum, and Nabalco Pty. Ltd. (1971). Milirrpum v. Nabalco Pty. Ltd and the Commonwealth of Australia (Gove land rights case) : a claim by Aborigines that their interests in certain land had been invaded unlawfully by the defendants. Sydney : Law Book Co.
- Palmer, K. and Brady, M. (1991). Amanbidji Land Claim. Anthropologists report on behalf of the claimants. Darwin. Northern Land Council.
- Peterson, N., & Rigsby, B. (Eds.). (2014). Customary marine tenure in Australia. Sydney University Press.
- Radcliffe-Brown, A. R. (1931). The Social Organisation of Australian Tribes. The Oceania Monographs. No 1. Melbourne. Macmillan and Co. Limited.
- Rigaux, P., Scholl, M.O., and Viosard, A. (2002). Spatial Databases: With applications to GIS. Morgan Kaufmann Publishers.
- Tindale, N. B. (1974). Aboriginal Tribes of Australia. Berkeley, Calif. : University of California Press.
- Trigger, D. (1982). Nicholson River (Waanyi/Garawa) Land Claim Submission: Land Claim Book. Northern Land Council.
- Woodward, E. (1974). Australian Aboriginal land rights commission: reports, Aboriginal Land Rights Commission, viewed 05 December 2014, (<http://apo.org.au/node/36136>).

# A Voxel-Based Skewness and Kurtosis Balancing Algorithm for Updating Road Networks from Airborne Lidar Data

Li Liu

Samsung Lim

School of Civil & Environmental Engineering

School of Civil & Environmental Engineering

The University of New South Wales

The University of New South Wales

Sydney, NSW 2052, Australia

Sydney, NSW 2052, Australia

l.liu@unsw.edu.au

s.lim@unsw.edu.au

## Abstract

In this paper, we proposed a road extraction algorithm that utilises voxel-based skewness and kurtosis balancing in order to update existing road networks accurately and reliably from airborne lidar data. The proposed algorithm consists of initial road extraction followed by a series of refinements including width constraints, spatial interpolations and curve fitting which are successfully implemented to reduce false-positives, join up misclosures and determine centrelines of the extracted roads. A numerous sample lidar datasets are tested in order to assess the proposed algorithm. The quantified completeness and correctness of the classification results are 98.20% and 98.54%, respectively; hence it is concluded that the proposed algorithm works effectively in acquisition of road features from airborne lidar data.

## 1. Introduction

Keeping road network databases up-to-date is important for many applications e.g. emergency handling, car navigation, tourism, traffic management and monitoring, intelligent transportation systems, web map services, location-based services, because of rapid urbanisation in the developing and developed countries (Zhang and Couloigner, 2004; Zhang, 2004; Baltasvias and Zhang, 2005). An ongoing problem faced by geospatial information providers is how to improve the accuracy of the geospatial database with the limited resources. In some countries, road networks in their national database are found to be inaccurate at the level of errors ranging from 4-200 m (Bentabet et al., 2003). Updating road networks sometimes lags behind up to a few decades (Bentabet et al., 2003). Currently, several methods are available in updating the road networks, including field surveying, vector map comparison, image processing and lidar-based mapping. Field surveying is labour-intensive, time-consuming, and vulnerable to manual errors. Vector map comparison is free of manual errors but suffers from the

---

*Copyright © by the paper's authors. Copying permitted only for private and academic purposes.*

In: B. Veenendaal and A. Kealy (Eds.): Research@Locate'15, Brisbane, Australia, 10-12 March 2015, published at <http://ceur-ws.org>

generalization errors. Image processing is sensitive to the image resolution and is not feasible in occluded areas. Lidar-based mapping is often considered accurate, suitable for large-scale scenarios and cost-effective.

Height differences, normal vectors, and contextual information are the most widely used geo-features to extract roads from lidar data. In general, typical feature extraction strategies consist of segmentation and clustering (Choi et al. 2007; Zhu and Mordohai, 2009; Zhao and You, 2012), region growing (Akel et al., 2005; Clode et al. 2004a, 2004b, 2007) and kerb detection (Vosselman and Zhou, 2009). Choi et al. (2007) extracted roads with a series of circle buffers where the maximum possible slope of roads is used to eliminate the erroneous object clusters. However, this technique requires a strict threshold which makes it difficult to deal with mountainous areas. Zhu and Mordohai (2009) generated a road likelihood map with lidar points and extracted the main road regions with a minimum cover set algorithm. Zhao and You (2012) developed the elongated structure templates that detect candidate road regions and introduced a voting scheme to refine the road parameters.

Akel et al. (2005) applied a region growing segmentation method based on elevation and normal vectors to detect road areas. Clode et al. (2007) introduced a hierarchical system to extract road points from airborne lidar data, and convoluted the results with phase-coded discs to generate vectorised road networks. Vosselman and Zhou (2009) detected small height jumps between kerbstones and road surfaces according to height differences and an elevation threshold. Zhou and Vosselman (2012) extended their study and refined the detection process by using a sigmoidal function. However, their results show that it is not well suited to mobile lidar data because of the occlusion by large objects such as vehicles and trees. Having the aforementioned algorithms tested, our conclusion is that the existing algorithms do not perform well in detecting highly occluded roads.

Because of the shadow effect in high-density vegetation areas, a good level of completeness of extracted roads is difficult to achieve. For the purpose of higher accuracy and completeness, some researchers have utilised aerial images in addition to lidar data. Hu et al. (2004) utilised aerial imagery and lidar data together in combination to extract road points. Their results show that the accuracy and completeness are improved. Zhu et al. (2004) introduced the associated road images extracted from lidar points in order to enhance the results from real road images from aerial mapping. However, road extraction from the fusion of optical images and lidar points is still challenging since the co-registration of the images and the lidar points may become a noticeable error source.

Existing geospatial data can provide a good complement to lidar data as it can be served as a priori knowledge about roads (Vosselman, 2003; Hatger and Brenner, 2003; Elberink and Vosselman, 2006). Vosselman (2003) used lidar points and cadastral maps to reconstruct a road network and update the road database. Hatger and Brenner (2003) applied a fast region-growing algorithm to extract road geometry parameters from lidar data and existing road databases. Similarly, Elberink and Vosselman (2006) fused lidar data and 2-dimensional topographic maps to extract road points. However, these methods cannot cope with the map scale difference and the generalization process.

This study aims to utilise airborne lidar data and the national infrastructure database in order to update the existing road networks. We employ a voxel-based skewness and kurtosis balancing algorithm which is based on the assumption that ground points exhibit a normal distribution in order to filter out object points from the lidar data. Since multi-classes and various road materials can distort the normal distribution, the existing road database is utilised to segment the lidar point clouds into buffer-like tiles so that the assumption can hold. One of the advantages in doing so is that only a few parameters i.e. voxel size and buffer size are necessary. Also, this method can deal with different types of roads regardless of the road materials (e.g. asphalt or concrete) if the candidate roads in the same tile are made of a single material. It should be noted that not only shadows of objects e.g. trees and buildings but also open areas e.g. parking lots may affect the road extraction results, hence further refinements are essential to improve the accuracy and completeness, once initial extraction is performed. A series of refinements including width constraints, spatial interpolations and curve fitting are implemented in the proposed method in order to reduce false-positives, join up misclosures and determine centrelines of the extracted roads.

## **2. Skewness and Kurtosis Balancing**

Skewness and kurtosis balancing (Bartels et al, 2006; Bartels and Hong, 2006; Bao et al., 2008; Bartels and Hong, 2010; Crosilla et al., 2013) have been widely used to separate ground points and object points from a lidar point cloud. It has been commonly applied to creation of a Digital Terrain Model

(DTM), segmentation and classification, but has not been used for road extraction purposes. The underlying assumption of skewness and kurtosis balancing is that the measurements of a homogeneous class are supposed to exhibit a normal distribution (Duda et al., 2001), and the presence of other classes would distort the normal distribution. Thus, the removal of the object points from lidar data would lead to a normal distribution of the remaining ground points.

One of the methods to assess the asymmetry of data distribution is skewness, denoted by  $sk$  (Davies and Goldsmith, 1984) in Equation (1) which is also known as the third moment of the mean (David, 1953).

$$sk = \frac{1}{N \cdot \sigma^3} \cdot \sum_{i=1}^N (z_i - \mu)^3 \quad (1)$$

where  $N$  is the total number of points in the dataset,  $z_i$  for  $i \in \{1, 2, \dots, N\}$  are the attribute values of lidar points,  $\sigma$  is the standard deviation of the dataset, and  $\mu$  is the arithmetic mean of the dataset. Kurtosis, known as the fourth moment of the mean (David, 1953) and denoted by  $ku$  (Davies and Goldsmith, 1984) in Equation (2), is another measure. It is a measure of the size of a distribution's tail and a degree of the dominance of peaks in the distribution.

$$ku = \frac{1}{N \cdot \sigma^4} \cdot \sum_{i=1}^N (z_i - \mu)^4 \quad (2)$$

Compared with other algorithms such as Expectation-Maximization (EM) (Dempster, et al, 1977) and  $k$ -means (Rottensteiner, et al., 2005), the classic skewness and kurtosis balancing under the assumption of a normal distribution is parameter free. EM and  $k$ -means need to go through the iteration process, and require a threshold to halt the iteration process and the maximum number of Gaussian components which influence the accuracy of the results.

Intensity and elevation are common attributes used in skewness and kurtosis balancing. To test the effectiveness of the two attributes, we apply elevation, intensity, and a combination of the two attributes to a sample data, respectively. The results are shown in Figures 1-3. As seen in Figure 1, the elevation values of some road points in the blue rectangle exceed those of vegetation in the black rectangle because of the slope of the road. As a result, road points in the blue rectangle can be eliminated while the vegetation points in the black rectangle can remain. Hence the elevation-based skewness and kurtosis balancing algorithm (SKBA) may not perform well in mountainous areas. It is observed in Figure 2 that the intensity-based SKBA can remove most of object points, however, the intensity values of some trees may also satisfy the normal distribution. While SKBA using the combination of elevation and intensity shows a reasonable accuracy, a large amount of highly elevated road points can be omitted. Taken all the factors into account, the intensity-based SKBA is applied in this paper.

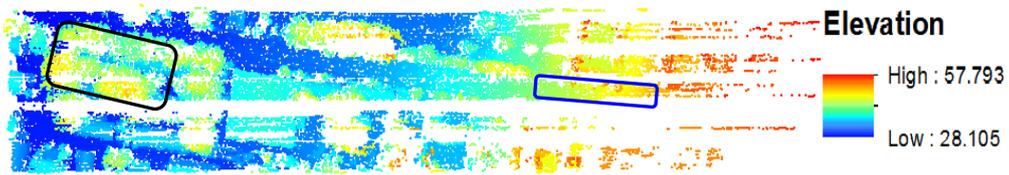


Figure 1. Results of elevation-based SKBA

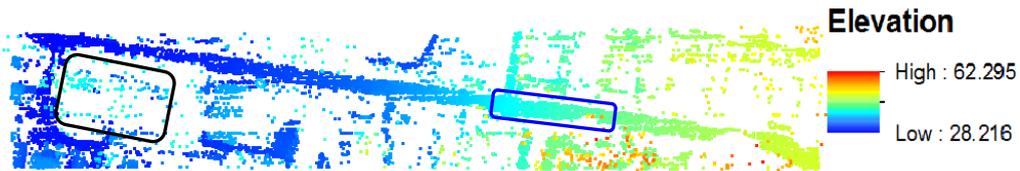


Figure 2. Results of intensity-based SKBA



Figure 3. Results of SKBA based on the combination of elevation and intensity

### 3. Methodology for Road Extraction

A framework of the proposed road network update scheme is illustrated in Figure 4. The first step is to segment lidar data according to the buffers generated by the road network. The second step is to partition the buffered points according to geographic coordinates and reorder the points along the main direction of the existing roads. Voxels are created on the basis of geographic coordinates in order to allocate the reordered points to different voxels and remove object points. Once the object points are removed, the intensity-based SKBA is applied to the remaining points. Some of non-road features such as parking spaces can be possibly classified as roads, and the misclosures among road points may occur because of the blockage by tall trees and buildings. Hence refinements including width constraints, spatial interpolations and fitting curves are applied to the initial extraction results.

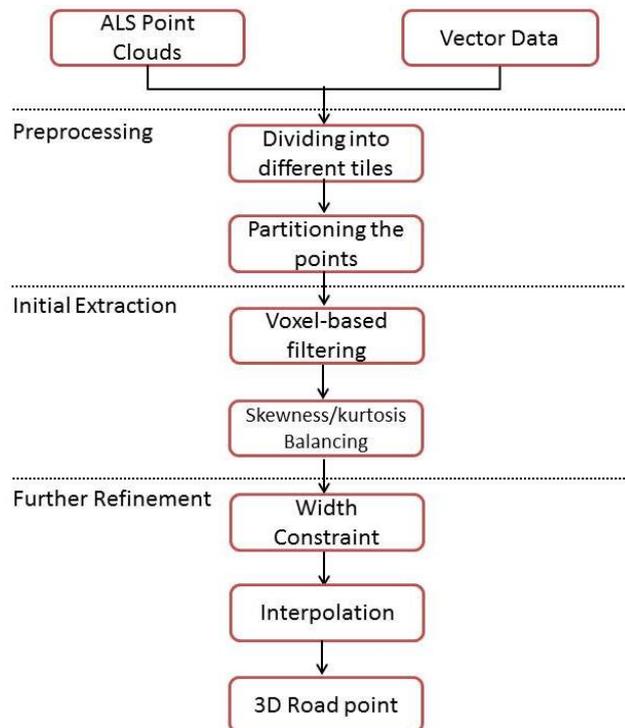


Figure 4. Framework of updating road networks from airborne lidar data.

### 3.1 Tiling of lidar points

Skewness and kurtosis balancing does not consider the material characteristics of the road, therefore different road materials within a road network can disrupt the normal distribution of the intensity values of the whole dataset. It seems that this is the main reason why the researchers have not used skewness and kurtosis balancing for road extraction purposes. Since skewness and kurtosis balancing is largely reliant on the assumption of a normal distribution, preprocessing is required to make the assumption hold. Crosilla et al. (2013) implemented a subdivision method to isolate individual areas so that the assumption of normality can apply. In this study, we take advantage of the road networks to buffer the lidar points which segregates road points of a similar material, making the intensity values in a buffer show a normal distribution. A buffer is generated along each line segment and the lidar points falling into the buffer are regarded as a part of the neighbourhood. Otherwise the points are regarded as outside of the neighbourhood and are removed. To detect and extract the extension of existing roads and new lanes, the buffer distance should be given large enough to cover the expected extension and new lanes. However, if the buffer distance is set too large, roads with different materials may be present in one buffer. Taken this factor into account, the buffer distance along the main direction of the road segment is set to 10 m, whereas the buffer distance perpendicular to the main direction of the road segment is set to 20 m. It turns out that this step accelerates the computation process.

### 3.2 Partitioning of lidar points

Airborne lidar data typically consists of a huge amount of points. Therefore processing the entire points without proper data management is time-consuming. Also, iterations in skewness and kurtosis balancing will take time if the whole dataset is processed at once. To reduce the computation time effectively, our approach is to partition lidar points along the main direction of road segments according to the geographic coordinates. The core steps of partitioning of lidar points are explained as follows:

- a) Calculate the coordinate difference between the starting node and the ending node of the line segment in  $x$ -direction and  $y$ -direction, respectively.
- b) If the absolute value of the coordinate difference between the starting node and the ending node in  $x$ -direction is larger than that in  $y$ -direction, set  $x$ -direction as the main partitioning direction; otherwise set  $y$ -direction as the main partitioning direction.
- c) Search the point with the minimum coordinate value in the main partitioning direction.
- d) For each point, calculate the index number according to the Equation (3):

$$I = INT[(C - C_{\min})/S] \quad (3)$$

where  $I$  is the index number of each point,  $C$  is the coordinate value of a point in the main partitioning direction,  $C_{\min}$  is the minimum coordinate value in the main partitioning direction, and  $S$  is the step value which is equal to the absolute value of the slope. INT is a mathematical operation that rounds the result to the nearest integer.

- e) Cluster the points with the same index number. The partitioned points are indexed for the purpose of efficient data management.

### 3.3 Voxel-based filtering

Popescu & Zhao (2008) used voxels to create vertical profiles of trees. Hosoi and Omasa (2006) used voxels to create canopy profiling. Lim and Suter (2009) generated super-voxels to classify lidar points. As observed in Figure 2, some trees in low-elevation areas remain after the processing since their intensity values also satisfy the normal distribution. Therefore, it is essential to remove object points in low-elevation areas before the SKBA is applied. In this study, we propose a new approach which is voxel-based filtering that requires the following steps:

- a) Set the voxel size.
- b) Calculate the voxel index for each point according to the coordinates and allocate points to the corresponding voxels. The voxel index is the corresponding serial number of a voxel in  $x$ ,  $y$ ,  $z$  direction.
- c) Set the search index to the main direction of the road which can be calculated from the road network.
- d) Check the bins along the search index for each tile. For each bin, find the index of a voxel in  $z$  direction that contains the point of the minimum  $z$  value in the height bin and remove the points whose voxel index in  $z$  direction is larger than the minimum. In addition, if the minimum voxel index in  $z$  direction is greatly larger than that of the neighbourhood, all points in the height bin are removed.

### 3.4 Skewness and Kurtosis Balancing

After the voxel-based filtering, the intensity-based SKBA is applied to the remaining points to filter out object points. In fact, points of the maximum intensity value are removed before applying the intensity-based SKBA since the intensity of road points is overall the lowest in the sample datasets. The process can be described as follows:

- (i) Calculate the skewness and kurtosis values of the dataset.
- (ii) For a normal distribution, the kurtosis value is 3 and the skewness value is 0. If the absolute skewness is larger than the threshold and the absolute difference between kurtosis and 3 is larger than the threshold, go to step (iii); otherwise, go to step (vi).
- (iii) The point with the maximum intensity value is regarded as the object point and is removed.
- (iv) Recalculate the skewness and kurtosis values of the dataset.
- (v) Repeat steps (ii)-(iv) until all the points are processed.
- (vi) If there are some remaining points, label them as candidates for road points.

### 3.5 Refinement of Candidate Road points

Once the initial road extraction is performed, some open areas can be shown as roads since they present similar characteristics as roads. Also, gaps between road segments may occur because of the blockage by trees, buildings and/or vehicles. Therefore, further refinements are needed to improve the results. The refinement process consists of width constraints, interpolations and fitting a curve to the road points. At first, road sections i.e. cross-sectional points perpendicular to the main direction of roads are generated. Then the width of the road is calculated based on the mean values since the widths of roads within a road network may vary. Then the road section whose width is larger than that of the road segment i.e. a part of the road between two nodes is processed. The points far away from the centre point e.g. points of more than a half of the road width away, are removed. After the false positives are removed, a local inverse distance weighting (IDW) interpolation is applied to connect the misclosures caused by the blockage of trees, buildings, or vehicles along the main direction of the road. The local linear interpolation is used to minimise the bias caused by the curves along the road. Connectivity is enforced to join up the misclosures between two different road segments. In case of the jagged edges, symmetric points to the road network are calculated. To obtain a smooth road surface, a curve fitting is adopted, and the centrelines of roads are calculated. These processes are shown in Figure 5.

Because of the buffer distance along and perpendicular to the main direction of the road, the intersections between adjacent roads will appear, especially in conjunction areas. To refine the topology, we force the connecting roads end in the intersection nodes if two roads intersect. In terms of multiple intersection nodes, the mean intersection node is calculated according to the coordinates and we force the connecting roads end in the mean intersect nodes.

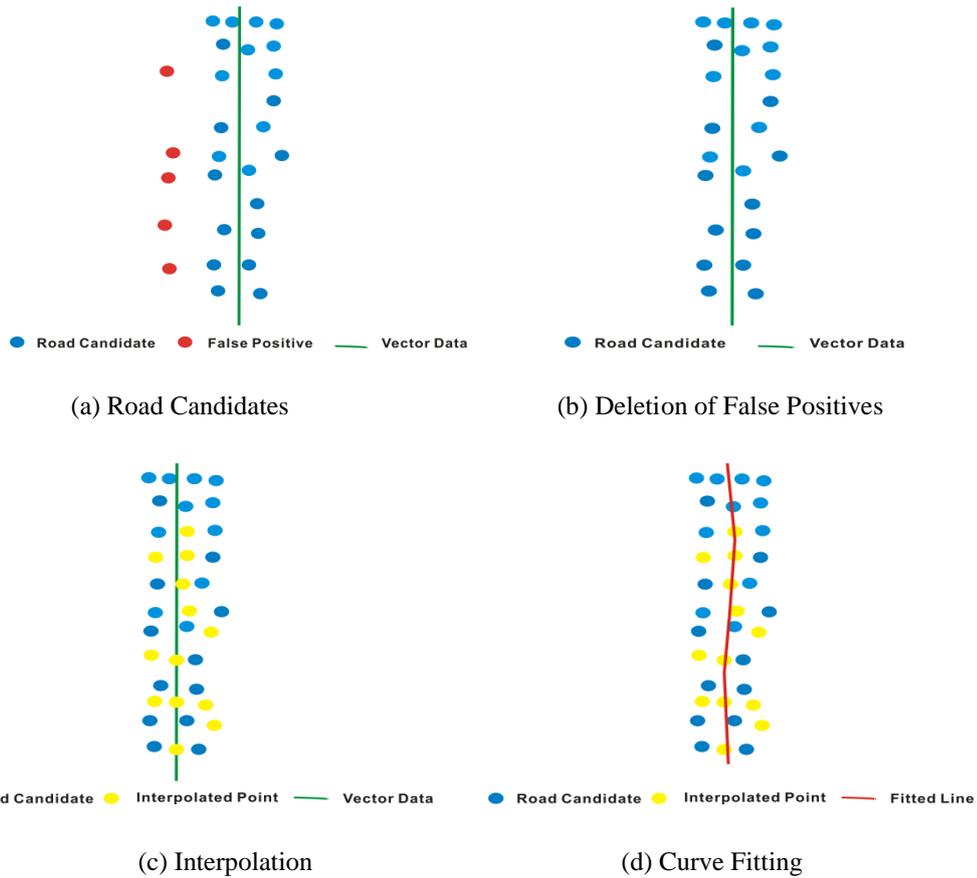


Figure 5. Refinement process: (a) road candidate; (b) deletion of false positives; (c) interpolation; (d) curve fitting

## 4. Road Extraction Results

### 4.1 Test data

To validate the proposed method, tests with various scenes and lidar data samples are conducted. The first sample is a part of Anzac Parade which is an arterial road and surrounding neighbours near the University of New South Wales (UNSW), consisting of 128,847 points with a point density of 0.94 points/m<sup>2</sup>. The surroundings of the dataset are complicated as they contain both high-rise buildings and small residential houses. In addition, different types of roads are present in the sample, including an arterial road, a local lane, an avenue and a local street. Also, Anzac Parade consists of two parallel roads with a road separator in the middle. As the study area is near a university campus, many roads are highly occluded by cars. The second sample is a part of Botany Street which is another arterial road and its surrounding residential areas near UNSW, including 69,064 points. Apart from roads, this sample dataset mainly consists of small houses and dense vegetation. The main challenge for this dataset is its sparse point density of 0.74 pt/m<sup>2</sup>. The third sample is a residential area near UNSW, including 144,255 points. In this dataset, the high slopes along the streets cause huge elevation differences up to 50.73 m, which is a great challenge to distinguish ground points from building points. Another problem which may affect the completeness of the result is that roads are heavily blocked by trees in some areas. The fourth sample is a part of Barker Street which is a main road and its surrounding areas near UNSW, consisting of 124,690 points. Residential houses, trees and roads are main features in the dataset and the challenge of this sample is a high slope along the street resulting in the fact that the elevation values of some houses are even lower than those of some roads.

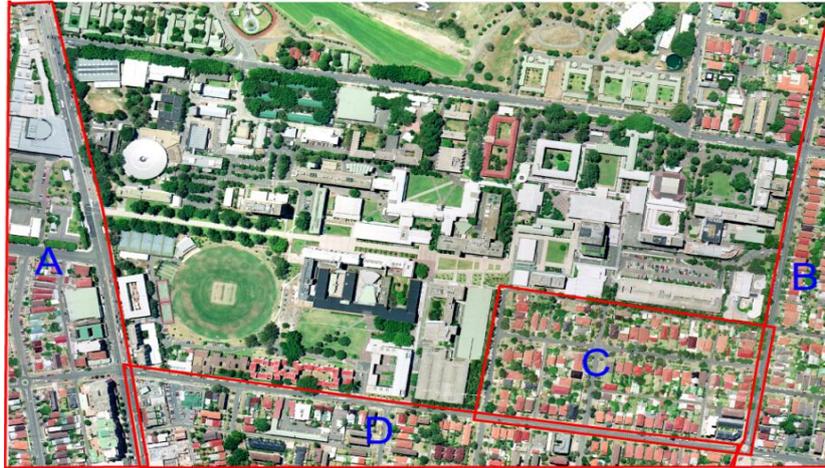


Figure 6. Four airborne lidar datasets: (A) Anzac Parade, (B) Botany Street, (C) a residential area, (D) Barker Street near UNSW.

## 4.2 Results

### 4.2.1 Extraction Results

Based on the proposed algorithm, the voxel sizes in  $x$ ,  $y$ ,  $z$  directions are set to 10 m, 10 m, and 1.5 m, respectively. If the voxel size in  $x$ - or  $y$ -direction is too small, some building rooftops may split into different voxels, which shortens the height difference between adjacent neighborhoods, therefore misclassifies building rooftops as part of ground points. If the voxel size is too large, some road points in mountainous areas will be removed as well. Figure 7 shows the overlay of the centrelines of the four sample datasets and the aerial image. Figure 8 shows the road extraction results and corresponding centrelines of the four sample datasets, which indicate the feasibility of the proposed algorithm.

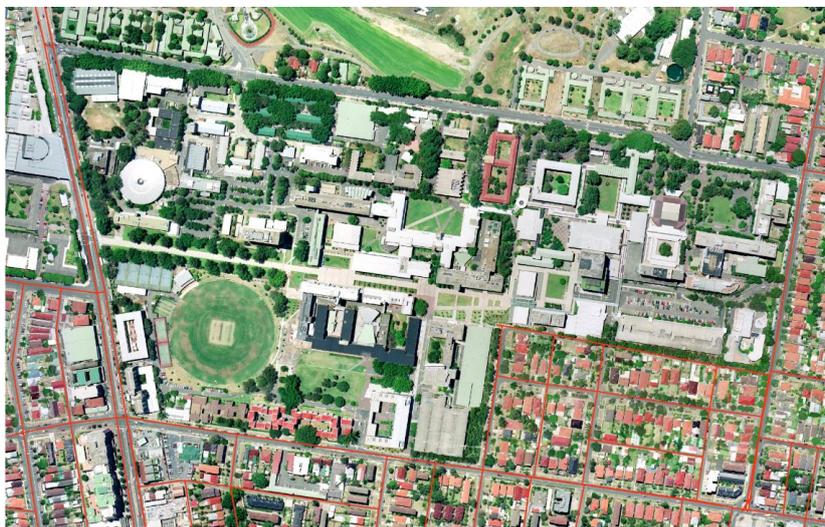


Figure 7. The overlay of the centrelines of four sample datasets and the aerial image

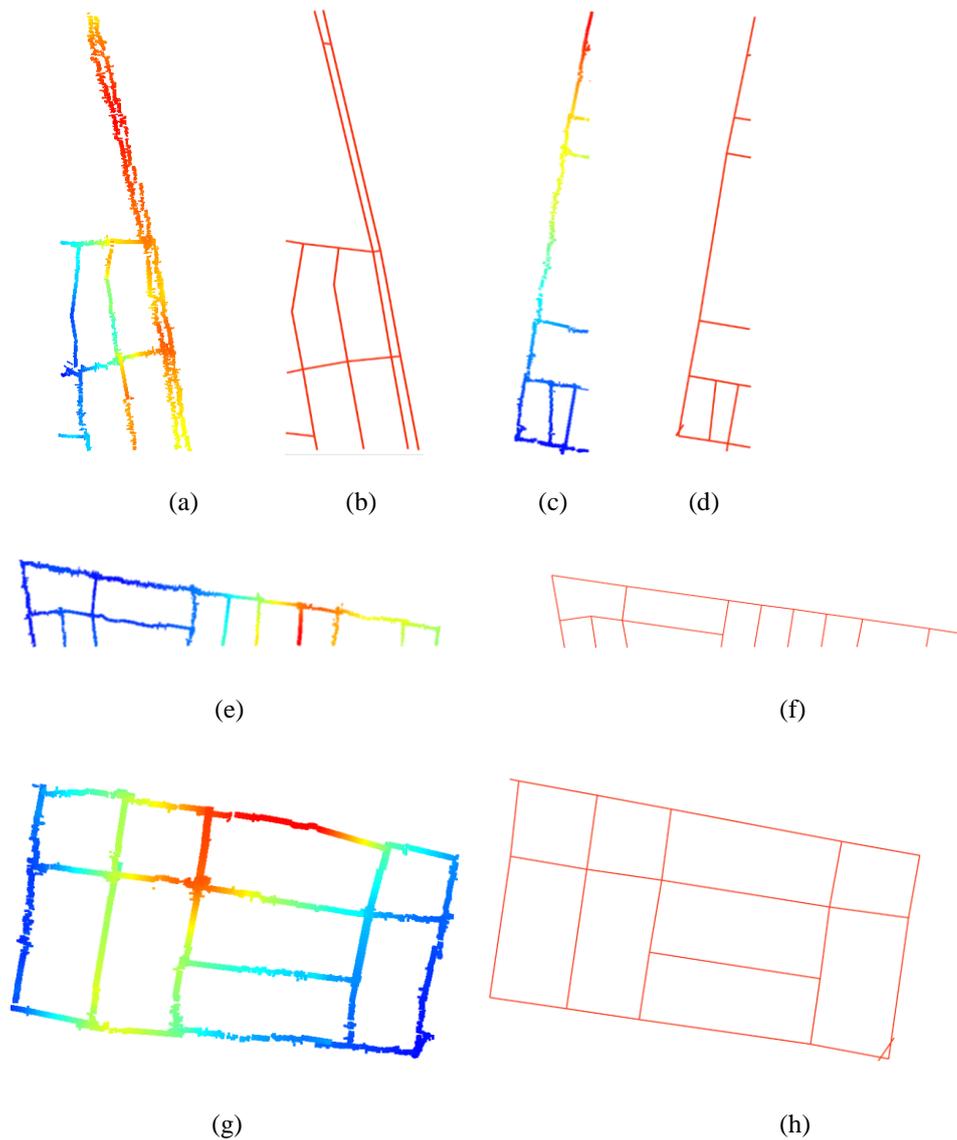


Figure 8. Results of the extracted road network and respective centrelines: (a) extraction results of Anzac Parade, (b) estimated centrelines of Anzac Parade, (c) extraction results of Botany Street, (d) estimated centrelines of Botany Street, (e) extraction results of Barker Street, (f) estimated centrelines of Barker Street, (g) extraction results of a residential area, (h) estimated centrelines of the residential area.

#### 4.2.2 Quantitative evaluation of 3D road results

Heipke et al. (1997) presented that the quantitative evaluation of road extraction results can be measured in terms of completeness, correctness and quality which are also used by Clode et al. (2007) and Yang et al. (2013). Completeness is the ratio of the total length of the matched roads to the total length of the reference roads. Correctness is the ratio of the total length of the matched roads to the total length of the extracted roads. Quality is the ratio of the total length of the matched roads to the total length of the sum of extracted roads and unextracted roads. In this study, the reference data is extracted from aerial images by digitising the road centrelines and the length of reference roads is calculated from the centrelines. The length of extracted roads is calculated by summing up the length of each segment which is achieved by a coordinate-based computation. Then the results are compared with the reference data by computing the differences between the extracted road segments and the reference road segments. If the maximum difference is above the threshold, the whole road segment is regarded as an unmatched road; otherwise it is regarded as a matched road. The given tolerance of differences is set to 2.0 m. We also computed the positional accuracy which estimates the distance between the reference and the extracted roads. That is, each road segment is divided into 10 equal parts and compared the positional distance between the corresponding nodes of the reference and the extracted roads. The root-mean-squared value of the 10 distances is regarded as the positional accuracy

of a road segment. The mathematical model of the positional accuracy of a road segment is shown in Equation (4) and the weighted root-mean-square value of the positional accuracy of all road segments is the positional accuracy of the sample dataset, shown in Equation (5). Table 1 shows the statistics of our quantitative analysis.

$$P_r = \sqrt{\sum_{i=1}^{10} D_i^2 / 10} \quad (4)$$

$$P_s = \sqrt{\sum_{r=1}^N L_r P_r^2 / \sum_{r=1}^N L_r} \quad (5)$$

where  $P_r$  is the positional accuracy of a road segment  $r$ ,  $D_i$  is the distance between the  $i^{\text{th}}$  node of the reference and the extracted road,  $P_s$  is the positional accuracy of a sample dataset,  $N$  is the total number of road segments in one sample dataset,  $r=1,2,3, \dots,N$ , and  $L_r$  is the length of the  $r^{\text{th}}$  road segment.

Table 1. Statistics of quantitative analysis

	Length of reference roads (m)	Length of extracted roads (m)	Length of unextracted roads (m)	Total length of matched roads (m)
Sample 1	2493.20	2484.50	7.13	2448.25
Sample 2	1268.89	1317.50	19.42	1120.01
Sample 3	2227.14	2249.14	0.00	1940.50
Sample 4	1958.33	1974.84	0.00	1547.53

Our quantitative analysis results are summarized in Table 2. Although the majority of the extracted roads match with the actual roads, some are slightly deviated from the reference centrelines, some lie on the edge of the roads. The main reason for the deviation is that the elevation change along the road is large, which can lead to the removal of ground points in highly elevated areas. Since there are no points in these areas to indicate which direction to interpolate, an error may occur and the situation becomes worse when the width of the road is large. Although symmetric points corresponding to the road network are calculated to avoid such a situation, existing positional errors in the road networks may result in the positional errors of the symmetric points. Another vulnerable area is the traffic islands. As shown in the extraction results for Samples 1, 3, 4, the approach fails to obtain the shapes of traffic islands. While Sample 1 has the highest correctness, completeness and quality values, the positional accuracy of the test dataset is the lowest among all samples. The unweighted positional accuracy of all the samples are calculated which results in 6.60 m, 4.62 m, 3.78 m and 6.37 m, respectively. This indicates that the road segments in Sample 1 are deviated from the reference centrelines within the tolerance difference.

Table 2. Results of quantitative analysis

	Completeness (%)	Correctness (%)	Quality (%)	Weighted Positional Accuracy of Dataset (m)
Sample 1	98.20	98.54	98.26	10.14
Sample 2	85.01	88.27	83.78	4.03
Sample 3	86.28	87.13	86.28	4.54
Sample 4	78.36	79.03	78.36	6.76

Figure 9 shows a series of experiments for voxel-based filtering with various voxel sizes. As seen in Figure 9, more object points are removed as the voxel size increases. The optimal voxel size is dependent on each sample dataset, hence it has to be estimated by trial and error.

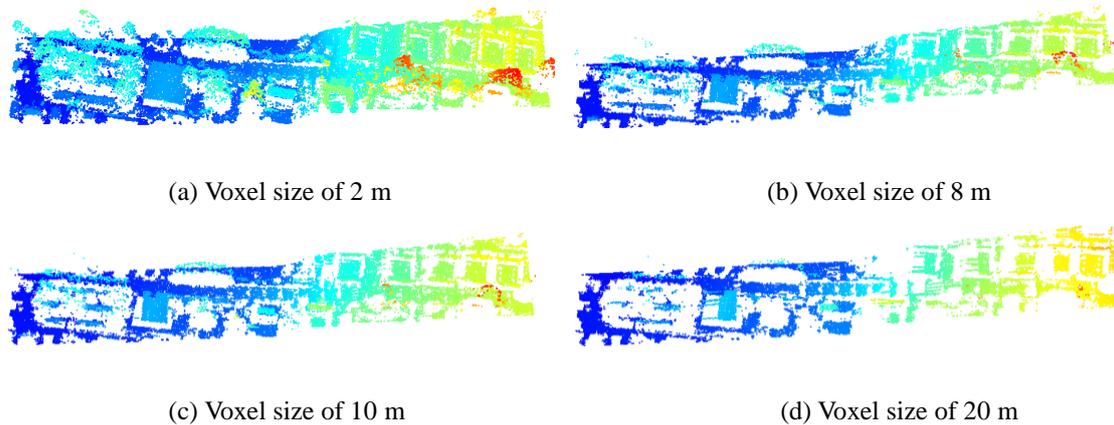


Figure 9. Voxel-based filtering with different voxel sizes

The proposed method may not perform well on the arcs of traffic islands. The method also yields poor results if road networks are not used because of multi-classes in lidar data. If roads in one buffer consist of various materials, omission error of road extraction may increase. The voxel size in  $x$ - or  $y$ -direction is crucial in filtering object points in mountainous areas.

## 5. Concluding Remarks

Updating road networks is critical in many applications e.g. traffic management, traffic monitoring etc. In this paper, we presented a novel method to update existing road databases with airborne lidar data. The proposed method takes advantage of a voxel-based skewness and kurtosis balancing to distinguish road points from the lidar point clouds. To make the intensity of roads suitable to the proposed SKBA, road networks are utilised to divide the lidar points into different tiles since different types of roads within a road network will distort the normal distribution of intensity values of road points. The lidar points are reordered along the main direction of roads and partitioned into sections according to the coordinates. Voxel-based filtering is applied to filter out remaining object points. Skewness and kurtosis balancing is used to classify the candidate road points. In order to refine the results, width-constraints are applied to eliminate false positives, and misclosures within a road segment are filled by a local interpolation method. Calculating road width automatically based on mean values in the refinement step avoids the problem of different road widths within a road network. The proposed method extracts roads from airborne lidar data with an acceptable completeness and correctness. The quantitative results show that it is a feasible method for updating the road databases regardless of the road types. There are some biases in the extracted centrelines that can be larger than the threshold. The main reason for this is that the removal of ground points in highly elevated areas fails to select the interpolation direction correctly. While the algorithm can fill large misclosures caused by trees and building rooftops, it is still difficult to reconstruct road segments if it is totally blocked by trees or building rooftops.

## References:

- Akel N.A., Kremeike K., Filin S., Sester M., and Doytsher Y., (2005). Dense DTM generalization aided by roads extracted from LIDAR data. In: IAPRS 36 (Part 3/W19), 54–59.
- Baltsavias, E., Zhang C.S., (2005). Automated updating of road databases from aerial images. *International Journal of Applied Earth Observation and Geoinformation*. Vol 6, pp 199–213.
- Bao Y.F., Li G.P., Cao C.X., Li X.W., Zhang H., He Q.S., Bai L.Y. and Chang C.Y., (2008). Classification of Lidar Point Cloud and Generation of DTM From Lidar Height and Intensity Data In Forested Area. In: IAPRSIS. Vol. XXXVII. Part B3b. Beijing 2008.
- Bartels, M., and Hong W., (2006). Segmentation of Lidar Data Using Measures of Distribution.
- Bartels, M., and Hong W., (2010). Threshold-free object and ground point separation in LIDAR data. *Pattern Recognition Letters* 31(10): 1089-1099.

- Bartels, M., Hong W, and Mason D.C., (2006). DTM Generation from LIDAR Data using Skewness Balancing. In: ICPR. pp. 566 – 569. Hong Kong.
- Bentabet, L., Jodouin S., Ziou D, and Vaillancourt J., (2003). Road Vectors Update Using SAR Imagery: A Snake-Based Method. IEEE Trans. On Geoscience And Remote Sensing, Vol. 41, pp: 1785-1802.
- Choi, Y.W., Jang Y.W., Lee H.J., and Cho G.S., (2007). Heuristic road extraction. In: International Symposium on Information Technology Convergence, IEEE Computer Society.
- Clode, S., Kootsookos P., and Rottensteiner F., (2004a). The Automatic Extraction of Roads from Lidar Data. In: IAPRSIS, Vol. XXXV-B3, pp. 231 – 236.
- Clode, S., Zelniker E., Kootsookos P., and Clarkson V., (2004b). A Phase Coded Disk Approach to Thick Curvilinear Line Detection. In: 17th European Signal Processing Conference, 6-10 September, 2004, Vienna, Austria, pp.1147-1150.
- Clode, S., Rottensteiner F., Kootsooks P., and Zelniker E., (2007). Detection and Vectorisation of Roads from Lidar Data. PE&RS, Vol. 73(5), 517–536.
- Crosilla, F., Macorig D., Scaioni M., Sebastianutti I., and Visintini D., (2013). Lidar data filtering and classification by skewness and kurtosis iterative analysis of multiple point cloud data categories. Applied Geomatics. Vol. 5, Issue 3, pp 225-240.
- David, F. N., 1953. A statistical primer. London: Griffin.
- Davies, O. L. and Goldsmith P. L., (1984). Statistical methods in research and production, with special reference to the chemical industry. 4th revised. London: Longman.
- Dempster, A.P., Laird N.M., and Rubin D.B., (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society, Series B 39 (1): 1–38.
- Duda, R. O., Hart P. E., and Stork D. G., (2001). Pattern Classification. New York: Wiley.
- Elberink, S.J.O., and Vosselman G., (2006). 3D modelling of topographic objects by fusing 2D maps and lidar data. In: IAPRS, Vol. 36, part 4, Goa, India, September 27-30.
- Hatger, C., and Brenner C., (2003). Extraction of road geometry parameters from laser scanning and existing databases. IAPRS, Vol. 34, 2003, 225-230.
- Heipke, C., Mayer H., Wiedemann C., and Jamet O., (1997). Evaluation of automatic road extraction. IAPRS XXXII (1997), pp. 47–56.
- Hosoi, F., and Omasa K., (2006). Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar. Geoscience and Remote Sensing. Vol.44, Issue:12. pp: 3610 - 3618
- Hu, X., Tao C.V., and Hu Y., (2004). Automatic road extraction from dense urban area by integrated processing of high-resolution imagery and LIDAR data. IAPRS 35 (Part B3), 288–292.
- Lim E.H., and Suter D., (2009). 3D terrestrial LIDAR classifications with super-voxels and multi-scale Conditional Random Fields. Computer-Aided Design. Vol.41. Issue.10. pp:701:710.
- Liu L., and Lim S., (2014). A Novel Algorithm for Road Extraction from Airborne Lidar Data. In: Research@ locate 14. Canberra, Australia. April 4-7, 2014.
- Popescu, S. C., and Zhao K.G., (2008). A voxel-based lidar method for estimating crown base height for deciduous and pine trees. Remote Sensing of Environment. Vol.112, Issue: 3, pp. 767–781.
- Rottensteiner, F., Trinder J., Clode S., and Kubik, K., (2005). Using the Dempster–Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. Information Fusion. Vol.6, Issue:4, pp: 283–300.
- Vosselman, G. and Zhou L., (2009). Detection of curbstones in airborne laser scanning data. In: IAPRS XXXVIII – 3/W8, pp. 111-117, Paris, France, 2009.

- Vosselman, G. (2003). 3-D Reconstruction of Roads and Trees for City Modelling. In: IAPRS, Vol. 34, part 3/W13, Dresden, Germany, pp. 231-236.
- Yang, B. S., Fang L.N., and Li J., (2013). Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds. ISPRS, Vol.79: 80-93.
- Zhang, C.S., (2004). Towards an operational system for automated updating of road databases by integration of imagery and geodata. ISPRS. Vol. 58, Issues 3-4, pp:166-186.
- Zhang, Q.P., and Couloigner I., (2004). A Framework For Road Change Detection And Map Updating. In: IAPRSS 35 (Part B2), pp. 729 - 734.
- Zhao, J.P., and You S.Y., (2012). Road Network Extraction from Airborne Lidar Data Using Scene Context. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
- Zhou, L. and Vosselman G., (2012). Mapping curbstones in airborne and mobile laser scanning data. International Journal of Applied Earth Observation and Geoinformation 18(2012), 293-304.
- Zhu, P., Lu Z., Chen X., Honda K., and Eiumnoh A., (2004). Extraction of city roads through shadow path reconstruction using laser scanning. PE&RS 70(12), 1433-1440.
- Zhu. Q.H. and Mordohai P., (2009). A Minimum Cover Approach for Extracting the Road Network from Airborne Lidar Data. In: IEEE 12th Conference on Computer Vision Workshops.

# A qualitative evaluation of a proposed metro map for Melbourne's underground system.

William Cartwright

School of Mathematical and Geospatial Sciences at RMIT University, Australia

## Abstract

When cities develop a rail network – above or under ground – this is usually accompanied by the publication of a metro map. The map is an essential adjunct, produced and provided to commuters to facilitate informed use of the system.

It is generally agreed that the exemplar metro map is that of the London Underground, designed by Harry Beck, in 1931. Beck's design principles have been copied, modified and adopted by numerous public transport bodies, when producing maps of their own system. His map is regarded as a design classic, and used as the lodestone when judging the effectiveness of other metromaps.

Recently, Public Transport Victoria (PTV) developed a new design for a metromap for the rail system for Melbourne, Australia. This design is greatly based on Beck's design ideas. As part of the press release, PTV sought passenger feedback on its proposed new design.

It is refreshing to see that an organisation like Public Transport Victoria has taken the initiative to implement a new design, and to seek public feedback on this design. However, is this new design effective, or just 'more of the same'? Is the map no different, and perhaps no better, than previous maps of the Melbourne metropolitan rail system, or a great improvement on what was previously published? Are we just going around in circles?

This paper provides the results of an evaluation of this proposal for a new map for the Melbourne metropolitan rail system. It begins by providing a brief history of metropolitan rail maps in Melbourne, to provide a background to what now exists. Then it looks at this recent proposal and outlines the basis for evaluation, which is built around the design principles of Beck's London map. Finally, it provides the results from the evaluation, reports on conclusions from this evaluation and makes recommendations about how the proposed map might be improved.

**Keywords:** underground map, design, Melbourne

## 1. Introduction

Commuters take for granted the mental map they have of how public transportation systems work. Through constant travel on the system, and upgrading their knowledge of the system from time-to-time when system changes are implemented, they maintain a current, usable mental image of what can be considered to be a complex system.

Cartographers and graphic designers design and implement representations of metropolitan transport systems for both regular and new users of the system. These representations – usually referred to as maps (but this description of what is used may vary – see Cartwright, 2014) – must provide a clear 'picture' of the network, its lines and nodes (railway stations, bus and tram stops) and interchanges, so that users of the system can effectively plan journeys and navigate the network once their travel is underway.

Many maps and diagrams have been produced of metropolitan transportation systems to facilitate this. Many are produced as an adjunct to other promotional material produced by transportation authorities or private operators. Here, they generally form part of documentation that serves to promote a certain image. Their main aim in this instance is to reinforce and promote the corporate 'look' desired.

In Melbourne, as noted earlier, a new map design has been proposed to illustrate the rail transportation network. But, will the proposed new Melbourne map work? This is the topic of this paper. It reports, briefly, on the initial analysis of this new design, done with due consideration of the design principles employed by Beck, in his design of a new London Underground map in the early 1930s. Here, the work sought to ascertain whether the design for a new map of Melbourne's rail system did in fact consider Beck's design principles and other information graphics associated with public transport in Melbourne.

## 2. Melbourne's rail system

Melbourne's rail network is based on a commuter rail model centered on the City's Central Business District (CBD) and its main commuter station, Flinders Street Station. It consists of 16 electrified lines, the central City Loop subway, and 207 stations, with a total length of 372 km of electrified lines (Wikipedia, 2014). From its early days Melbourne has been well served with rail transportation, complemented with tramcar and omnibus services. The city established its first railway line 20 years after it was founded (Wikipedia, 2014). In fact Melbourne can claim Australia's first railway line, built between

Flinders Street Station and Sandridge (now Port Melbourne). It opened on 12 September 1854 (Department of Infrastructure and Regional Development, Government of Australia, 2014), and it linked Melbourne's main port of the time with the growing central business area. The Melbourne and Hobson's Bay Railway Company, formed in 1853, operated the line (Wikipedia, 2014). In 1859 the Williamstown railway line opened, connecting Williamstown and Geelong to the new Spencer Street railway station (Wikipedia, 2014).

Looking at a map of Melbourne of 1890 (Figure 1), the extent of the metropolitan rail system in Melbourne can be seen to cover the greater area of the city. The city was already served with an extensive rail system.



Figure 1. Melbourne ca. 1890, showing the extent of the metropolitan rail system.

Source: <http://www.antiqueprints.com.au/sites/antiqueprints.com.au/files/6-158-melbourne-street-map-c1890-1386385569.jpg>

This work is out of copyright.

Private companies developed the railway lines, but this proved not to be economically successful. These companies were gradually taken over by the Victorian Government. Up until the early 1980s the Victorian Railways (later called VicRail) operated the system. The original name of this organization is Victorian Railways Department, and it was created in 1856 (Bau, nd). Later VicRail's country and suburban railway operations were separated, and Metropolitan Transit operated the system (Bau, nd), with V/Line operating intra and inter-state rail services. Later still, in 1998, operation of the metropolitan rail system was fully privatised (Bau, nd).

The railway lines developed further from the centre, associated with the 1880 'Land Boom' in Victoria. The development of the rail system continued, with lines pushed into some suburbs, and some under-patronised lines in inner Melbourne being closed. The lines were electrified and new rolling stock introduced. But, the general configuration of the system was still

focussed around one hub – Flinders Street Station – and it continued to be basically a suburbs-Central Business District radiating system.

The arrival of the railway in the suburbs of Melbourne had a great impact. New railway stations were built, and they became the focus of the provision of intensive development of services and shops that served the commuter.

The Railway Construction Act of 1884 saw the development of numerous lines. An Inner Circle line was established in 1888 and the Outer Circle in 1890 (Danno, nd). However, these circle lines proved not to be economic, as shown in the map of 1895 (Figure 2). The Circle lines were soon closed – the Outer Circle as early as 1897 and the Inner Circle closed in the following century - in 1948 (Danno, nd.).

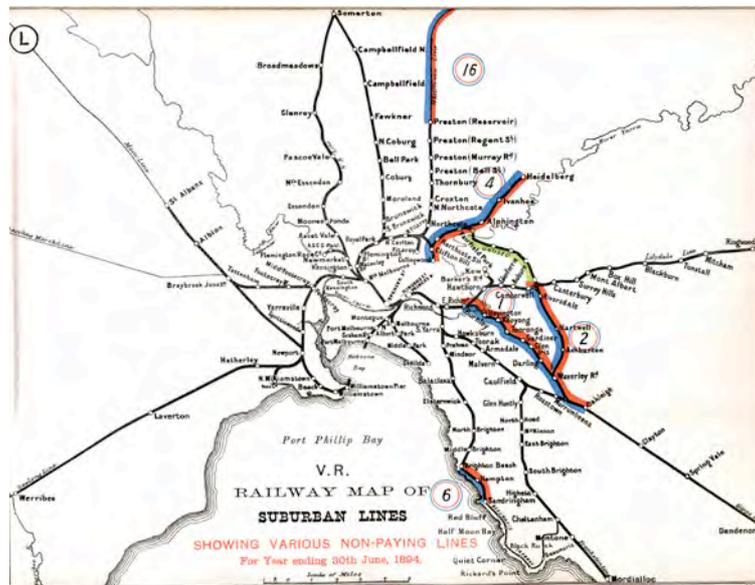


Figure 2. 1895 map of suburban rail lines, showing “Various Non-Paying Lines” for the year ending 30th June 1884.

Source: <http://www.victorianrailways.net/vr%20map/1895submap.html>  
This work is out of copyright.

The next real major change was the construction of the underground City Loop, which opened between 1981 and 1985. This linked Melbourne two main central city stations – Flinders Street and Spencer Street (now Southern Cross station) – with three new underground stations (Parliament, Museum (now Melbourne Central) and Flagstaff) and the suburbs. Rather than simply radiating from Flinders Street Station to the suburbs, the lines now ran under the city, through ‘The Loop’. This is basically the system that operates today – an inner city underground loop linking the original radial railway system.

Before looking at the map being evaluated here, the next section of the paper provides a brief timeline of the maps published to represent the system.

### 3. Mapping Melbourne’s metropolitan rail system

The initial maps of the system were basically produced as overlays atop of existing road network or cadastral plans. They were rudimentary, but showed the system and its relationship to the growing city. The maps shown in figures 3 and 4 are ‘official’ maps from 1884 and 1926.



Figure 3. Melbourne suburban rail system. Victorian Railways, Melbourne : Railway Dept. 1884  
 Source: State Library of Victoria. <http://handle.slv.vic.gov.au/10381/115264>  
 This work is out of copyright.

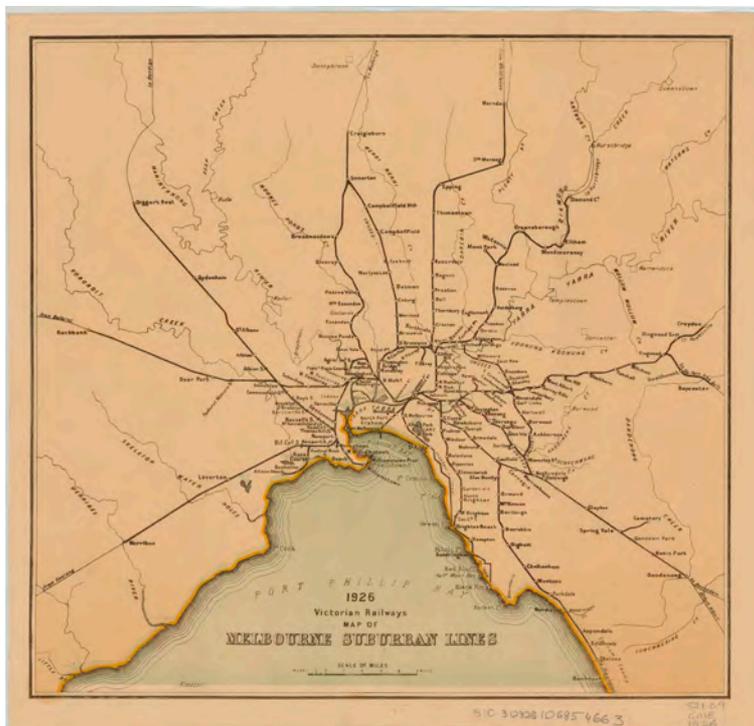


Figure 4. 1926 Victorian Railways map of Melbourne suburban lines. Victorian Railways. Melbourne: The Railways. Source: State Library of Victoria. <http://handle.slv.vic.gov.au/10381/115292>.  
 This work is out of copyright.

As well as these official government maps, a number of private publishing companies also produced their own representations of the system. An example of this type of map, showing railways and tramways, produced by the Melbourne Book Depot, is shown in figure 5.

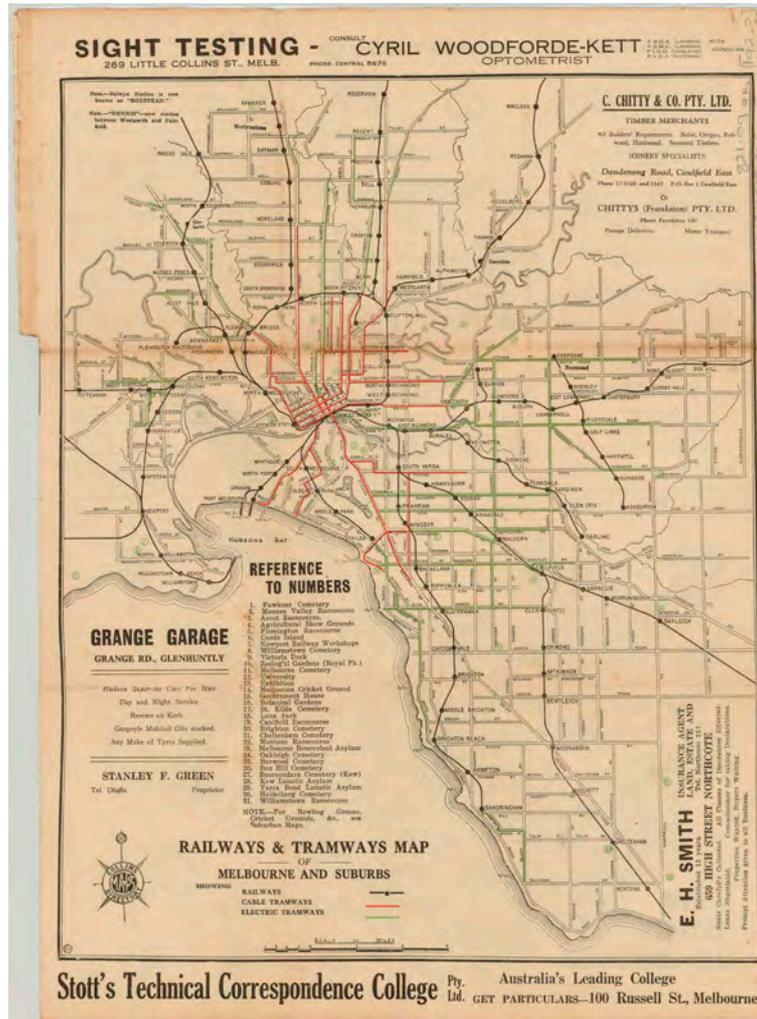


Figure 5. Railways & tramways map of Melbourne and suburbs, 1920. Collins Book Depot. Melbourne: Collins Book Depot? 1920. Source: State Library of Victoria. <http://handle.slv.vic.gov.au/10381/115385>. This work is out of copyright.

Maps fundamentally followed this general design pattern, with both governmental and private sector maps being made available. With the opening of ‘The Loop’ in 1981 a new type of map was developed, and the ‘look’ of the system map was changed to something that promoted this new era of Melbourne’s metropolitan commuter rail system. The map in figure 6 shows the new inner city underground loop and illustrates the 5 electrified lines and connecting non-electrified lines using colour coding.



Figure 6. City Loop 1981

Source: [http://i2.wp.com/railgallery.wongm.com/albums/melbourne-stations/D181\\_8122.jpg](http://i2.wp.com/railgallery.wongm.com/albums/melbourne-stations/D181_8122.jpg)

Later, with the introduction of three suburban fare zones, the maps were changed to illustrate this. The map from 1994 (figure 7) shows the suburban network and the three-zone fare system. Later, the fare zones changed to just two, and the maps were changed to reflect this (figure 8).

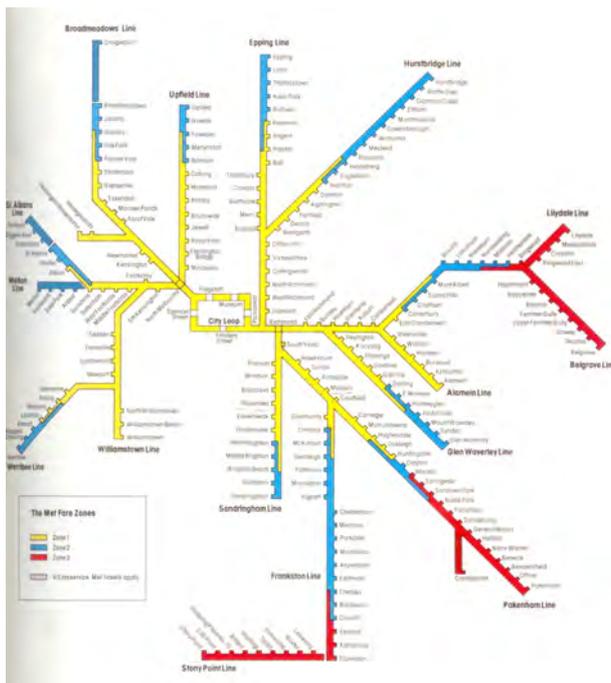


Figure 7. Melbourne suburban rail system, 1994.

Source: <http://melbpt.files.wordpress.com/2008/06/1994-train-map.jpg>



Figure 8. Melbourne metropolitan rail system, 2007.

Source: <http://www.mappery.com/map-of/Melbourne-Train-Map>

As well as rail-specific maps, there also exist maps illustrating the complete metropolitan public transport system. An example is shown in figure 9.

The maps shown in this section of the paper show just a snapshot of the maps produced of the Melbourne rail system. A number of iterations of the current map have been developed, as well as proposals for new maps. A proposal for a new map of the Melbourne rail system is described in the following section.



Figure 9. Melbourne train and tram guide, January 2014.  
 Source: <http://www.railmaps.com.au/melbourn.png>

The next section of this paper describes Beck's map and other system maps that have been developed using his design concepts. Before looking at Beck's map, a general overview of London's transport in the 1930s and the graphics developed to communicate information about the transportation is provided. It is essential to consider Beck's map with due reference to the overall concepts of communication artefact design at The London Passenger Transport Board during that period.

#### 4. London transport 1930s

In the early 1930s London was considering the future of its transportation. In 1931 inner London's population was already a hefty 4,397,003 people and Greater London had a population of 8,110 358 (Exploring 20th Century London, nd). The London Passenger Transport Board pioneered public transport campaign for coherence and efficiency. Part of this move to efficiency was the development of a modern integrated public transport system.

An efficient public transportation system was not guaranteed, with the popularity of the motorcar, and the concepts of changing the city to suit motor transport that was changing the face of cities across the Atlantic. In 1937 the Bressey Report (Bressey, 1938) was released, which looked at how London's transportation system might be changed to implement changes needed if motorized transport was to dominate the city.

However, the London Passenger Transport Board's electrified underground won the day and a modern system continued to be developed. As well as developing this modern system, the London Passenger Transport Board championed the concept of selling the system to the public as a modern system. This was reflected in the design of the London Underground stations, the rolling stock and the graphics that promoted the transportation system's position in contemporary London.

#### 5. Exemplar of metropolitan rail system mapping: Harry Beck and 'followers'. A benchmark for evaluation.

Henry (Harry) Charles Beck (1902 – 1974) worked as an Engineering Draughtsman in the London Underground Signals Office at the London Passenger Transport Board. During a time that he was laid-off, during a time of austerity at The London Passenger Transport Board, he developed his ideas for a different way of portraying the London Underground system. Ken Garland, who wrote an excellent book on Beck's map: *Mr. Beck's Underground Map: A History* (1994), said in an interview in the documentary "The London Underground Map" (BBC, 1987) that Beck's wife told him that Beck was obsessed with the development of his map, and she would find pieces of scrap paper, with design developments on them, scattered throughout the house.

Beck's obsessiveness was not just related to the design of the map, but also to getting his map published. He needed to convince The London Passenger Board's publicity officer, Frank Pick (1878-1941), that his design accorded with the graphics communication concepts championed by Pick. Pick guided the overall public 'look' of The London Passenger Transport Board, and what he chose had to reflect 'the new', modernism and the avant-garde. Finally, Beck's pocket version

design was printed (figure 10) as a trial run and, according to Garland (BBC, 1987), the public loved it. 850,000 copies of the map were in circulation within two months after its introduction (Garland, 1994).



Figure 10. Beck's first published map - the 1933 'pocket' version. Source: <http://www.ltmuseum.co.uk/omnibus/pg/1919b.htm#>  
Image in the public domain (a UK artistic work other than a photograph, made available to the public more than 70 years ago).

Looking at Beck's design certain rules can be noted:

- Except for the Thames River, the geography 'above ground' was removed;
- Only horizontal, vertical or 45° lines were used to demarcate rail lines;
- Each rail line was colour coded (as per the F. H. Stingemore map of 1927 that Beck's map replaced);
- Stations were denoted by 'tickmarks';
- A distinctive symbol was used to show interchange stations;
- The centre of map was enlarged to enable the efficient representation of stations in the central area of London; and
- The outlying extents of the rail system were 'moved' closer to the centre, to facilitate a more comprehensive map.

Following the success of this new design, other maps were developed along the lines of Beck's map. The "Sydney Suburban and City Underground Railway map" (Commissioner for Railways, New South Wales) of 1939 used similar design principles as Beck's representation (Dobbin, 2011). And, Beck's design influenced George Salomon's New York City Transit Authority (NYCTA) map, 1958 and Berlin's current U and S-Bahn maps, designed by Erik Spiekermann.

## 7. Evaluation of the new map for Melbourne

Promotional material and articles circulated in 2014 by Public Transport Victoria (PTV) about the design of a new map for the Melbourne metropolitan rail system were heralded as a process that considered users and provided a map, which was designed according to 'good' design principles (Milman, 2014). PTV piloted a potential version of a new map (figure 11). PTV solicited feedback from the travelling public on this new design.

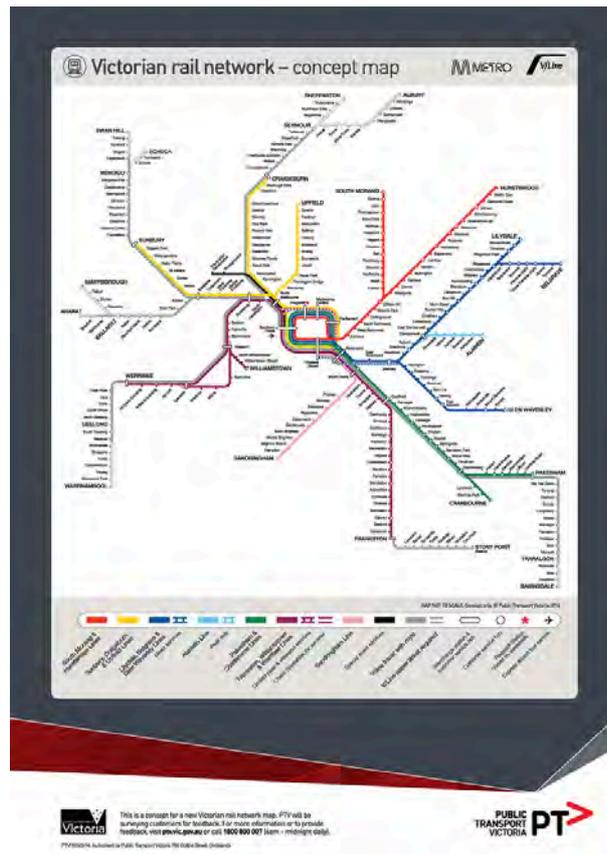


Figure 11. Proposed new map for the Melbourne rail transportation system. Source: [tumblr\\_n434oaHHsA1r54c4oo1\\_1280.jpg](https://www.tumblr.com/post/1280128014340401)

It was developed with due reference to what is considered to an exemplar of designs for representing transportation systems - Harry Beck's 1931 design map of the London Underground map (first published in 1933) (Garland, 1994). The proposed new map for Melbourne uses the 'Beckesque' symbology.

The map also shows the V/Line (intra and-inter state rail system) train connections and where the MYKI (electronic ticketing system) and paper tickets may be used. The different fare zones no longer exist, which is reflected on the map, with no suburban zones shown.

Initial reactions from the travelling public have been quite positive, with Public Transport Users Association spokesman Daniel Bowen saying: "My initial impression: I quite like this" (Carey, 2014, p. 2).

So, the map concept has been launched for comment, initial comments are positive and, from first inspection, the map does accord to Beck's concepts of 'good' design for effective communication of an urban rail network. Closer inspection of the map is warranted, to ascertain whether the map does accord to the design guidelines that Beck proposed. As well, since Beck's map accorded to the design 'suite' of graphics promoted and employed by The London Passenger Transport Board (now Transport for London (TfL)) to promote a modern, efficient, coordinated rail system, this factor also needs to be considered.

Basically, the evaluation compared the proposed new map for Melbourne according to the 'rules' that made the Beck design work. These principles were listed earlier in the paper. The outcomes of the qualitative evaluation are provided in table 1.

Table 1. Considering the proposed new Melbourne map with Beck's design criteria.

Beck design criteria	Proposed new map for Melbourne
Geography 'above ground' removed (except for Thames River)	No geography 'above ground' whatsoever
Only horizontal, vertical or 45° lines for rail lines	Only horizontal, vertical or 45° lines for rail lines
Rail lines colour coded	Rail lines colour coded
Distinctive symbols for interchange stations	Distinctive symbols for interchange stations
Centre of map was enlarged for clarity	Centre of map not enlarged enough
Outlying extents of the rail system 'moved' closer to the centre	No real compression of outlying stations.

Opinion

The map could be improved by adding a stylised Yarra River and Maribyrnong River. This would assist users of the map when deciphering their general location in the city. Also, the centre of the map needs to be enlarged much further than in the current design. (See figure 12 for an enlargement of the central part of the proposed new map for Melbourne) At present this part of the map is much too crowded, and the design would be greatly improved if this were done. Also, the map needs to compress the outlying stations, thus making the map easier to read, and to consider the extent of the services provided by the various lines.

The rail line colours mean nothing – they relate to no previous symbology for different lines.



Figure 12. Central section of the proposed new map for Melbourne. Source: The Guardian, 21 April 2014. (theguardian.com)

The map needs to compress the outlying stations, thus making the map easier to read, and to consider the extent of the services provided by the various lines.

Lastly, the map seems a little ‘remote’, from a design perspective, from its ‘sister’ graphics used by the Met. The map is somewhat of an orphan, when other graphics products used in the Melbourne public transport system. However, having said that, all of the graphics used are really a hodge-podge of various graphics designs. A selection of these can be seen in figure 13.



Figure 13. A selection of Melbourne public transport signs. Photo: William Cartwright, August 2014.

Looking at what had happened in London in the 1930s provides the ‘formula about how, with the appropriate choice of graphics, architecture and vehicle design, a comprehensive, modern image can be promoted. This was London Passenger Transport Board’s Frank Pick’s forte. He insisted that all designs, including graphics should reflect the fact that London, and its transportation system, reflected the modern age of electricity (Garland, 1994).

This type of publicity campaign reflected the British modern consciousness of the time. It promoted the new, the avant-garde.

Beck’s map did accord to the overall ‘look-and-feel’ of other London Passenger Transport Board graphics. The same cannot be said for the new map proposed for Melbourne by Public Transport Victoria.

## 8. Conclusion

The proposal for a new map for the Melbourne rail system is certainly a development by Public Transport Victoria that must be applauded. As well, seeking public comment on the design affords ‘crowd-sourced’ critique and suggestions.

The qualitative evaluation showed that the map generally ‘works’, albeit with work needing to be done to remove the clutter in the inner sections of the map and to foreshorten the extent of the suburban rail lines reach. As well, the map could be improved by the addition of a stylised representation of Melbourne’s two main rivers: the Yarra and the Maribyrnong. This would greatly assist locating and navigation.

Perhaps the problem that cannot be overcome is that the map is somewhat an orphan, within a cluster of other orphan graphics and signage used by Public Transport Victoria. This is a bigger design coordination problem that is beyond just the design of this map. Perhaps a modern-day Frank Pick could help here!

## 9. References

- Bau, M. NY. VR Timeline (nd) <http://www.victorianrailways.net/vr%20history/history.html> Web page accessed 16 June 2014.
- Bressey, C., (1938) “The Bressey Report: A Thirty-year Scheme”, The Times, London.
- British Broadcasting Corporation (BBC), 1987, “The London Underground Map (29 Jun. 1987) (television documentary)
- Carey, A. (2014) “Rail map gets a rainbow update”, The Age, April 21, 2014.
- Cartwright, W. (2014) “Rethinking the definition of the word ‘map’: an evaluation of Beck’s representation of the London Underground through a qualitative expert survey”, International Journal of Digital Earth.

<http://dx.doi.org/10.1080/17538947.2014.923942>

Department of Infrastructure and Regional Development, Government of Australia (2014) History of Rail in Australia. <http://www.infrastructure.gov.au/rail/trains/history.aspx> Web page accessed 16 June 2014.

Danno (2014) Beside the Yarra: Melbourne's Phantom Railways. <http://marvmelb.blogspot.com/2014/04/melbournes-phantom-railways.html>. Web page accessed 16 June 2014.

Dobbin, C. (2012) London Underground Maps: Art, Design and Cartography, Farnham, Surrey: Lund Humphries. Exploring 20th Century London, nd, <http://www.20thcenturylondon.org.uk/server.php?show=nav.41>

Garland, K. (1994) Mr. Beck's Underground Map: A History, Harrow Weald, Middlesex: Capital Transport Publishing.

Grocott, E. (2012) The Underground Group, A Reckless Serenade (blog), 5 June 2012. <http://edwardgrocott.blogspot.com.au/2012/06/underground-group.html>

Milman, O. (2014) Melbourne rail network map gets a splash more colour in overhaul. The Guardian, 21 April 2014. [theguardian.com](http://theguardian.com). Web page accessed 23 April 2014.

Wikipedia (2014) Railways in Melbourne. [http://en.wikipedia.org/wiki/Railways\\_in\\_Melbourne](http://en.wikipedia.org/wiki/Railways_in_Melbourne). Web page accessed 16 June 2014.

# A Simulation Study on Automated Transport Mode Detection in Near-Real Time using a Neural Network

Rahul Deb Das  
rahuld@student.unimelb.  
edu.au

Nicole Ronald  
nicole.ronald@unimelb.  
edu.au

Stephan Winter  
winter@unimelb.edu.au

Department of Infrastructure Engineering  
The University of Melbourne, Australia

## Abstract

Detecting transport modes in near-real time is important for various context-aware location based services and understanding urban dynamics. In this paper we present a simulated study on detecting transport modes in near-real time using a neural network. We have shown how detection accuracy will vary with different temporal window sizes and different combination of modes. Since in urban environment transport modes move slowly due to traffic, considering movement attributes or kinematics alone for mode detection is not sufficient. That is why we investigated how spatial information can improve mode detection accuracy. The model has achieved 82%-95% accuracy using different simulation designs and proves its efficacy over other detection models.

## 1 Introduction

Transport mode detection from trajectories has seen growing interest in research over last few years for its importance in various domains such as context-aware computing, location based services, understanding urban dynamics, travel demand surveys, traffic monitoring, and travel behaviour analysis. Traditionally travel modes have been surveyed in questionnaires, enabling also to capture additional knowledge including purpose of trip. Travel surveys, however, are burdensome, erroneous if made from memory, of low spatial detail, and reach only small sampling rates. Automation should overcome all these issues.

Since the late 1990's, due to advancements in positioning and navigation technology, GPS started being used as a mean to collect travel data and assess its reliability and future possibilities (Wolf, 2000). Eventually, the use of GPS has increased as it has become more precise, portable and ubiquitous. Nowadays people themselves can track their movement trajectories using GPS and potentially other sensors on-board their smart phones (Pieria et al., 2013).

Most of the research on travel mode detection is based on rigid velocity based model. However, a velocity based approach is not always sufficient. For example, low speed conditions, which are nowadays typical in urban traffic due to traffic at capacity, or bad weather, produce mode ambiguities. In low speed traffic conditions, the speed of a bus is similar to a car or bicycle. Therefore, there is a need to consider various non-kinematic attributes along with movement attributes (kinematics) in order to detect different transport modes.

---

*Copyright © by the paper's authors. Copying permitted only for private and academic purposes.*

In: B. Veenendaal and A. Kealy (Eds.): Research@Locate'15, Brisbane, Australia, 10-12 March 2015, published at <http://ceur-ws.org>

Existing transport mode detection research is mostly offline. That means modes are detected once a trip is completed from historical trajectories. Existing methods use the entire trip record in the form of trajectory and then separates the trajectory based on walking based segments into number of meaningful parts that correspond to respective transport modes.

The hypothesis behind this research is that a neural network based model can adjust well in real time with varying movement behaviour and overcomes the mode ambiguity under low speed conditions.

In detecting transport modes, movement characteristics derived from trajectories of the users are the raw data source. In this paper we will concentrate on trajectories of a single sensor, as provided by GPS enabled smart phones. Such GPS trajectories are unlabelled and come in raw format. Other sensors in the phone are neglected for the time being, but can easily be included in the model. A classifier is required that can detect the various transport modes used along each trip in real to near-real time. In this paper, a neural network based classifier has been tested. Our contributions are as follows-

1) We developed a simulated near-real time transport mode detection model based on a multi-layer perceptron neural network.

2) Earlier approaches to neural network based transport mode detection are mostly offline and did not use any spatial information. In this research we show how spatial information can increase the detection accuracy.

3) Selecting a proper temporal window for detecting transport modes in near-real time is critical and context dependent. In this paper we investigate how detection accuracy varies with different temporal window sizes, which helps in selecting a proper window size based on accuracy requirement.

In this paper we also evaluate the performance measure of a multi-layer perceptron neural network in order to detect transport modes in near-real time. A real time model can detect the transport mode epoch by epoch basis (such as second by second). In this research, we simulated queries within short temporal window to detect a given mode instead of second by second basis. Hence, we call this model a near-real time mode detection approach.

Detecting transport mode in near-real time is comparatively an emerging research area. In this paper we have developed a basic but intuitive near-real time mode detection model using a supervised learning approach. Real time mode detection can be useful for a number of applications. Applications include various context-aware location based services where the context could be a given transport mode. A petrol pump can distribute an electronic discounted coupon within its neighbourhood to all the private cars only. Detecting transport modes in real time can also help developing various context-aware mobile applications that can sense the modality and act accordingly. One instant could be developing a mode-dependent auto-answering service on smart-phones. If the mobile senses the owner is in driving mode then the auto-answer can automatically be enabled and helps driver to concentrate on the road rather than receiving any incoming call. Thereby this can help in reducing distractions on the road in order to reduce road accidents. This approach can also be helpful for urban planners or emergency service providers who want to know people's mode choice at a given route or in a given region at a given time window for modeling travel demand or various spatio-temporal events.

The paper is organized as follows. Section 2 discusses related works in transport mode detection from various perspectives. Section 3 discusses some of the basic terminologies and methodology. Section 4 demonstrates data preparation and experimentation. Section 5 shows the experimental results. Section 6 presents the discussion of these results, and Section 7 concludes the paper.

## 2 Related work

Nowadays, smart-phones come with GPS enabled facilities. Since smart-phones are carried by the users almost everywhere and all the time, hence, this positioning facility can be utilized in order to collect trajectories without any external intervention. Once a GPS trajectory has been collected or is in the process of being collected, those trajectories or part there-of can be used for transport mode detection in real time or post-processing mode. In this regard, existing work is mostly based on post-processing of the trajectories or detecting modes offline. Existing literature shows a wide variety of post-processing algorithms and classifiers. Some of the approaches used the classification technique directly without segmenting the GPS trajectories (Byon et al., 2007; Dodge et al., 2009; Reddy et al., 2010). At the same time there are approaches applying segmentation of the entire trajectory into meaningful parts, corresponding to different modes, before classification (Mountain and Raper, 2001; Tsui and Shalaby, 2006; Schussler and Axhausen, 2009; Zheng et al., 2010; Biljecki et al., 2012; Hemminki et al., 2013).

Segmentation is done based on those points that show high probability of mode change. Mountain and Raper used change in speed and direction for segmentation in their work (Mountain and Raper, 2001). However, this approach creates ambiguities in certain cases where the vehicles move slowly and constrained to specific roads or

the rail or tram networks. Liao et al. used proximity to potential change points, such as bus stops or train stops for offline mode detection (Liao et al., 2007). However, GPS accuracy greatly varies in urban environments, depending on the number of satellites in view, time of the day and season, atmospheric conditions and surrounding sources of multipath effects. Other research used change in peaks of acceleration curves in order to segment the trajectory (Hemminki et al., 2013). However this approach also suffers from low ambiguity resolution, typically in low speed condition such as during bad weather or traffic congestion. Another common and intuitive way for segmenting the trajectory is based on detecting walking segments. The rationale behind this approach is the observation that a person generally walks between using two modes of transport. This approach has achieved promising results for segmentation (Tsui and Shalaby, 2006; Zheng et al., 2010; Biljecki et al., 2012). However this approach also fails when there is a quick mode change or walking is negligible.

There have been a number of different algorithms for mode classification used so far. Zheng and colleagues used a decision tree, Bayesian Net, Conditional Random Field (CRF) and Support Vector Machine (SVM) in their work with 75% reported accuracy (Zheng et al., 2010). Gonzalez et al. used neural networks with 91% accuracy (Gonzalez et al., 2010). Some works are solely based on statistical measures (Patterson et al., 2003).

As far as the input parameters or indicators are concerned, prior work mostly concentrated on velocity attributes (Bohte et al., 2008; Schussler and Axhausen, 2009). But in low speed condition velocity and acceleration are not sufficient to resolve the ambiguities. So, more recently, research has incorporated additional movement attributes including heading rate change and stop rates (Zheng et al., 2010). Vibration data has also been tested as an additional attribute with promising results (Ohashi et al., 2013). However, in order to achieve better accuracy and account for GPS signal loss others have used inertial localization and navigation sensors such as accelerometers, along with GPS sensors (Reddy et al., 2010; Hemminki et al., 2013; Ohashi et al., 2013).

Byon and colleagues used GPS trajectories collected by GPS loggers to study detection accuracy in real time. However their focus was mainly on how accuracy varies with different sampling frequencies (Byon et al., 2009). They achieved high detection accuracy at 20 min temporal window. However they observed mainly four modes auto, walk, car, bus. Although Byon and colleagues developed two neural network models, one route specific and another one a universal model, they did not explore how spatial knowledge can help in detecting different modes. Also their approach is limited by their use of GPS loggers: they used instantaneous speed, acceleration, number of satellites in view for a given transport mode to train their classifier. Number of satellites in view depends on particular transport mode. Such as GPS device inside a bus is obstructed by the metallic body and ceiling and vertical windows limiting the number of satellites in view. Whereas a car would have wider front windshield that would allow stronger and multiple GPS signals. However when using smart-phones for detecting modes, instantaneous acceleration, number of satellites in view or horizontal dilution of precision values may not be available.

Gonzalez and colleagues developed a neural network based mode detection model with a core focus on how to reduce streaming of movement data. Earlier work used a static and fixed data transmission procedure but that suffered from high financial costs associated with data transmission as well as computational overhead and storage issues. Gonzalez and colleagues proposed a novel critical point (CP) algorithm to transmit only the relevant GPS points during the trip (Gonzalez et al., 2010).

Since movement states are uncertain and imprecise there are a couple of mode detection approach using fuzzy logic (Tsui and Shalaby, 2006; Biljecki et al., 2010). A fuzzy approach with three criteria and five to ten modes has been tested with an accuracy of more than 90% (Schussler and Axhausen, 2009; Biljecki et al., 2010). However these approaches are rule-based and involve fuzzy antecedents and fuzzy consequents (Zadeh, 1965; Mamdani and Assilian, 1975). This approach cannot adapt with different movement behaviour in real time. Since fuzzy logic based models are developed based on expert knowledge with predefined premise and consequents hence they are not scalable with new parameters and thus pose scalability and flexibility issues. In this paper we present a neural network based model that can learn in real time. A neural network based model is flexible and scalable.

### 3 Theory

In this section we presented some basic definitions, concepts and methodology used in this research.

#### 3.1 Raw Trajectory

A raw trajectory is a set of spatio-temporal points arranged in a chronological order. This can be mathematically expressed as

$$Tr = \{P_i\}; P_i = (x_i, y_i, z_i, t_i); i \in [0, N]; \forall i: (t_i < t_{i+1}) \dots \dots \dots (1)$$

### 3.2 Segment

Any connected part of a raw trajectory with a specific semantics is a segment. For example, if a part of trajectory is extracted with a given annotated mode, then that is a modal segment. Similarly, if certain part(s) of a trajectory is extracted over a given time period that part would be a temporal segment.

### 3.3 Model architecture

The architecture of a multi-layer feed forward back propagation neural network is explained as follows-

A multi-layer feed forward neural network consists of mainly three layers- a) input layer, b) hidden layer, c) output layer. These layers contain one or more than one nodes or neurons. Input nodes are connected to hidden nodes and hidden nodes are connected to output nodes. But nodes of the same layer have to be disjoint and they cannot be connected to each other. Input layer is responsible to get input signals from the external world typically in the form of movement attributes (kinematics) or spatial attributes (non-kinematics) in the context of transport mode detection.

### 3.4 Training

Neural network can learn online and adapt well with given instances. However before using a neural network, it has to be trained to map a given set of inputs to a given output class. The training typically starts from input layer as soon as input stimuli are fed in. Nodes in each layer receive input signal from the preceding layer and send an output signal to the nodes in immediate succeeding layer. Each node multiplies the input signal with a previously established weight, adds a threshold, converts into an output signal through an activation function and sends it to the other nodes in the succeeding layer. The hidden layer is not directly connected to real world. This is the most important layer that processes the information and creates categorizing features for classification which sends signal to the output layer to categorize a given set of feature vectors. Once the output signal produces a response it is evaluated with the actual response. The difference between the predicted response and desired response is the error term of the neural network which is then back propagated to the model in order to adjust the weight and threshold values iteratively. This iterative process goes on in a cyclic way until a prescribed number of cycles (epochs) or a desired error level is achieved during training phase.

The rate at which a neural network learns can be adjusted by changing certain parameters called learning rate (LR) and momentum (M). These parameters control the change in weight and their persistence throughout total number of epochs.

### 3.5 Near-real time simulation

In order to detect transport modes in near-real time, queries will be fetched to a central server with kinematic and non-kinematic information. In this research we used a set of historical trajectories for near-real time simulation purpose. In order to train a neural network model small temporal segments have been extracted from the trajectories. Kinematic and non-kinematic attributes are then estimated over that temporal window in order to capture various movement behaviour of the given mode within that time period. Since a transport mode can exhibit different movement behaviour at different instant hence there is a need to train the classifier with movement behaviours for each mode at different instant of time over different trajectories. In order to extract movement behaviour of each transport mode, temporal segment over a given temporal window of a given mode segment has been extracted at regular interval of time.

### 3.6 Mode segmentation

Each trajectory can be expressed as a set of modal segments. This can be expressed as

$$T = \{SM_j\} \dots\dots\dots (2)$$

Where,  $j \in [1, N]$ ;  $N$ = total number of modes used by the user over the trajectory

$$\text{Each modal segment can be expressed as } SM_j = \{P_{ij}, M_{ij}\} \dots\dots\dots (3)$$

where  $i = i^{\text{th}}$  spatio-temporal index;  $j = j^{\text{th}}$  modal segment index;  $SM_j = j^{\text{th}}$  segment of the trajectory;  $P_{ij} = i^{\text{th}}$  spatio-temporal point in  $j^{\text{th}}$  segment;  $M_{ij} = \text{mode for } i^{\text{th}} \text{ point in } j^{\text{th}} \text{ segment.}$

A modal segment can be divided into a number of overlapping temporal segments. That can be mathematically expressed as -

$$SM_j = \{TW_{kj}\}^1 : k \in [1, M] \dots\dots\dots (4)$$

Where,  $TW_{kj} = k^{\text{th}}$  temporal segment of  $j$  modal segment over time window  $t$ ;  $M = \text{total number of temporal segments over a given modal segment.}$

### 3.7 Temporal segmentation

Once a trajectory is segmented into number of modal segments then each modal segment is segmented in number of temporal segments overlapping by  $(n-1)$  spatio-temporal points, where  $n$  is the number of spatio-temporal points in a given temporal segment. The overlap is chosen as  $n-1$  in order to capture diverse movement behaviour of the given mode at a finer granularity.

Each temporal segment can be expressed as  $TW_{kj}^t = \{P_{ijk}, M_{ijk}\}$  ..... (5)

Where,  $TW_{kj}^t = k^{th}$  temporal segment of 't' time length in  $j^{th}$  modal segment  
 $P_{ijk} = i^{th}$  spatio-temporal point in  $k^{th}$  temporal segment of  $j^{th}$  modal segment  
 $M_{ijk} =$  Annotated mode in  $k^{th}$  temporal segment of  $j^{th}$  modal segment

### 3.8 Kinematics and spatial information

In this research eight kinematic attributes are estimated using Euclidean functions in space-time domain such as average speed, average acceleration, variance of speed, variance of acceleration, maximum speed, maximum acceleration, minimum speed and minimum acceleration. In order to understand how a given mode behaves spatially with respect to different spatial objects (route network or POI), eight spatial relevance measures have been considered (see Table 2). Spatial relevance with respect to different spatial objects is calculated based on spatial proximity of spatio-temporal points to the given spatial object or a part thereof.

### 3.9 POI relevance estimation

In order to estimate POI relevance (proximity to bus stop, train stop, traffic signal or car wash or parking lot) a density-based clustering kernel is ran over each temporal window. Then POI relevance over a given temporal segment is estimated as

$$POIRel_c = POIRel_{c-1} + s * (n/N) \dots\dots\dots (6)$$

Where,  $POIRel_c =$  Relevance measure of a given POI over cluster 'c' over temporal window [t1, t2]  
 $POIRel_{c-1} =$  Relevance measure of a given POI over cluster 'c-1' over temporal window [t1, t2]  
 $s =$  scaling factor ( $s=10$  in this case)  
 $n =$  number of elements in the cluster falling in the search radius of the given POI  
 $N =$  total number of the elements in the cluster

### 3.10 Instance formation

In order to train and test the model using N-fold cross-validation, instances are created in the form of feature vectors which include kinematics and spatial attributes estimated over each temporal segment and fed into the model. A flowchart is given to show the workflow in Figure 1.

### 3.11 Performance measure

In order to evaluate the performance of the model, we used N-fold cross-validation. Since in N-fold cross-validation all the feature space is used using N-1 as training and 1 set of feature vectors as test thus it can capture the state behaviour at a fine granularity. However in hold-back type training, the accuracy of the model depends on the percentage of training instances that can represent all the details and characteristic behaviour of the entire population. Since in real time mode detection instances may vary with a temporal window size and modal movement behaviour, hence performance measure of a N-fold cross-validation strategy has been presented in this research (see experiment and result section), assuming an iterative N-fold cross-validation over growing time can dynamically improve the model in near-real time.

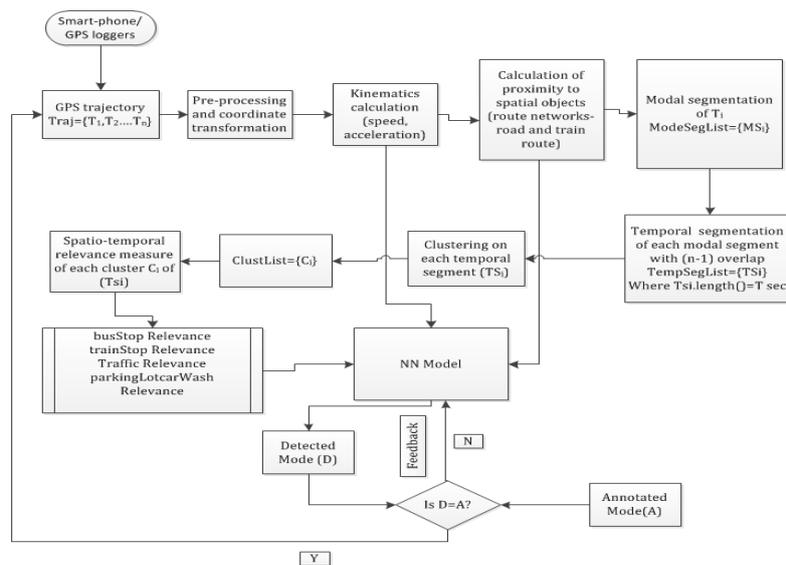


Figure 1: Flowchart of simulated mode detection modeling architecture

## 4 Data preparation and experimental setup

In order to evaluate our hypothesis Microsoft's GeoLife dataset has been used in this experiment (Zheng et al., 2008a; Zheng et al., 2008b; Zheng et al., 2010). The dataset mainly covers Beijing CBD and its surrounding suburb. The dataset was collected by smart-phones and GPS loggers in the form of GPS trajectories. Different sampling intervals ranging from 2-5 seconds have been used in this dataset. Users provided their trajectories and ground truth separately. This dataset contains various transport modes such as car, walk, bus, taxi, train, subway and bike. However, no accuracy measure such as horizontal dilution of precision (HDOP) is provided in this dataset. The dataset also suffers from semantic gaps due to both technical reasons such as urban canyons, indoor environments and misreporting such as time gap between annotating two different mode segments by a user. During pre-processing stage some of the GPS points are found to be outside the study area. There were also some inconsistencies in the annotations, such as walking over unreasonable longer duration or with unreasonable speed. There are also semantic gaps during signal loss and missing annotations in the dataset. The portions of the dataset containing such semantic inconsistencies are discarded. However a future work can use signal loss and contextual information to detect certain modes such as train in a subway.

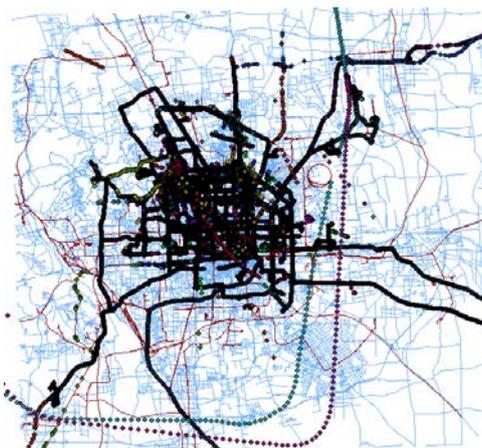


Figure 2: Beijing GeoLife GPS dataset overlaid on the road network (in blue colour) and rail network (in red colour)

In a data filtration stage 2.5 m/s has been set as walking speed threshold (Minetti, 2000), and some of the trajectories are discarded. In this experiment 264 trajectories have been used including training and test trajectories (Fig. 2).

Earlier works used HDOP value and can easily filter noise points (Byon et al., 2009; Gonzalez et al., 2010), but in this dataset we do not have any information that can provide positional accuracy or confidence level for each GPS fix. Hence, we setup two different experimental designs. One with filtered GPS data points where walking speed more than 2.5 m/s have been removed. Another setup was used without any filtering walking speeds. In both cases, raw velocity values are smoothed using an inverse distance weightage (IDW) smoothing kernel.

However technically a segment can also be viewed as a trajectory if it is treated discretely for further analysis. In order to detect transport modes in near-real time, we used a portion of a historical GPS dataset with transport mode annotated to train the mode detection model. Then we generated queries randomly at different instant of the trajectory and fetch the queries to the model to detect the transport mode as if the queries are coming in near-real time. A multi-layer perceptron (MLP) neural network has been realized in this research in order to detect various transport modes. The reason neural network has been investigated in this research because neural network is flexible, easily scalable and most importantly, it can learn online and adapt well.

In this research, five transport modes are considered: car, walk, bus, train and bike. Since car and taxi are difficult to distinguish especially in near-real time hence car and taxi are both grouped as car for time being. However in future car and taxi can be treated separately depending on the availability of contextual information. Similarly, train or light rail and subway are grouped as train. A multi-layer perceptron (MLP) neural network has been modelled using Weka, a Java based open source machine learning package. Since the time window is a critical factor in near-real time mode detection hence different temporal window size has been evaluated such as 120 sec, 180 sec, 240 sec, 300 sec and 600 sec based on subjective judgement. Experiments are also set up using only kinematic information and spatial information along with kinematics. For kinematic information eight movement attributes over a given temporal window have been considered since different modalities may exhibit different movement behaviour (Table 1). When kinematic attributes are used a 8-6-5 MLP was formed, and using spatial and kinematic attributes a 16-10-5 MLP model was used to detect different transport modes (Fig. 3). A popular

approach to select number of hidden nodes can be calculated as the closest integer value of  $\lceil[(\text{input nodes} + \text{target nodes})/2]\rceil$ . Hence we selected 6 hidden nodes when input nodes are 8 and output nodes are 5. Likewise, for 16 input nodes and 5 target nodes, number of hidden nodes are 10.

Table 1: Kinematic attributes

Attribute	Relevance
Average speed (avgSpeed)	Central tendency of a temporal segment in order to approximate a characteristic movement behaviour
Average acceleration (avgAccl)	
Variance of speed (varSpeed)	Spread of movement behaviour over the temporal window
Variance of acceleration (varAccl)	
Maximum speed (maxSpeed)	Upper bound of respective movement attributes within a given temporal window
Maximum acceleration (maxAccl)	
Minimum speed (minSpeed)	Lower bound of respective movement attributes within a given temporal window
Minimum acceleration (minAccl)	

For spatial information eight spatial attributes including proximity to route network and different POIs over a given temporal window have been considered (Table 2).

Table 2: Spatial attributes

Attribute	Relevance
Average road proximity (avgRoadProx)	Central tendency of proximity distribution over a given temporal window
Average railway proximity (avgTrainProx)	
Variance of road proximity (varRoadProx)	Spread of proximity distribution over a given temporal window
Variance of railway proximity (varTrainProx)	
Relevance score for bus stop (busRel)	Relevance measure of each relevant cluster from a given POI based on spatial proximity
Relevance score for train stop (trainRel)	
Relevance score for traffic stop (trafficRel)	
Relevance for parking lot and car wash (plcwRel)	

In order to study the performance of neural network through different training strategies, the model has been realized through N-fold cross-validation (where N=10). Table 3 shows number of instances used in N-fold cross-validation for different time window.

Table 3: Time window vs instances

Time window	Instances
120	15060
180	13735
240	12600
300	11488
480	9166
600	7835

## 5 Experimental results

The performance of the model has been evaluated on five modes against different temporal window size using filtered walking speeds in order to compare with the existing works that used positional uncertainty information. An experiment has also been carried out without filtering walking speeds assuming in real time people may run instead of walking during mode transfer or GPS positions can be subjected to various errors leading to walking speed greater than any threshold value.

The model was trained and tested using 10-fold cross-validation. In the first stage accuracy was tested against different temporal window sizes on trajectories where walking speeds are filtered. In the second stage accuracy was evaluated without filtering the trajectories in order to simulate real time mode detection. In both cases, the model shows that using spatial information can easily outperform the accuracy produced by only kinematics attributes. The reason behind this is that all transport modes may move slowly during traffic congestion or bad weather which leads to mode ambiguities. Figure 4 shows how mode ambiguity may arise using only acceleration measure. In this figure different modes may not be distinguished from their acceleration since they are clustered around the similar acceleration measures (Fig. 4).

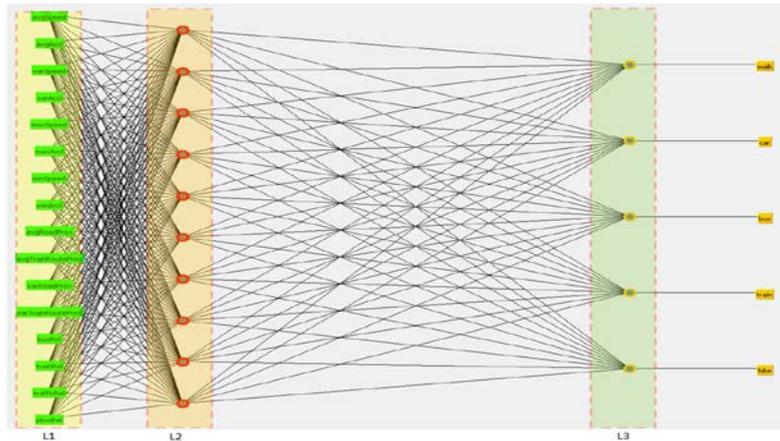


Figure 3: A 16-10-5 MLP neural network based mode detection model. L1 indicates input layer, L2 indicates hidden layer and L3 indicates output layer.

But when spatial relevance, in particular proximity to route network or given POI relevance, is considered the mode has been detected more accurately. In figure 5, the bus mode shows high bus stop relevance and hence bus mode is more prominent from other modes. However since walking can take place anywhere over the footpath near the bus stop hence, some of the walking instances have shown high bus stop relevance owing to false positives (Fig. 5). The rationale is, a car can travel like a train with similar speed and acceleration but the underlying route network would be different and POI relevance will also vary accordingly.

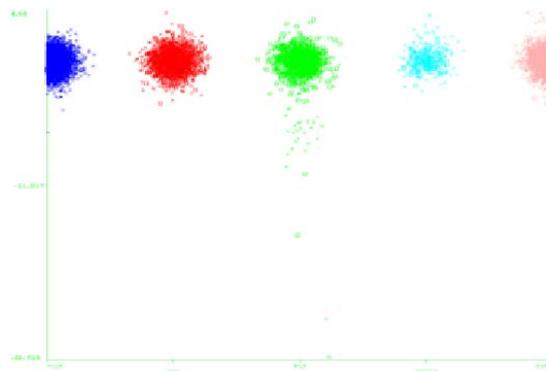


Figure 4: Mode ambiguities from similar acceleration distribution. Class map: {blue: walk; red: car; green: bus; indigo: train; brown: bike}. X-axis: modes; Y-axis: acceleration value

From the accuracy measures it is clear that there is a trade-off between temporal window size and mode detection accuracy. Selecting an optimal window size is context dependent. Overall 300 sec seem to be an optimal window size for near-real time mode detection as the accuracy starts increasing gradually from this point and the accuracy measure is more than 82% for unfiltered trajectories and 86% for filtered trajectories (Fig. 6; Fig. 7).

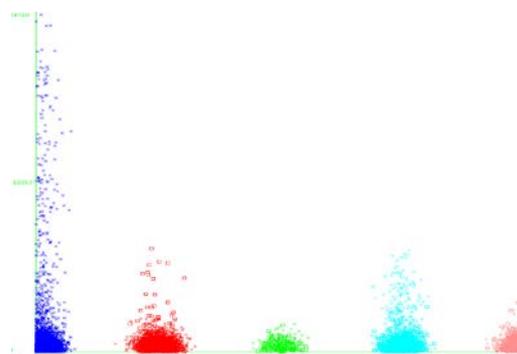


Figure 5: Bus stop relevance vs different modes. Class map: {blue: walk; red: car; green: train; indigo: bus; brown: bike}; X-axis: modes and Y-axis: bus relevance

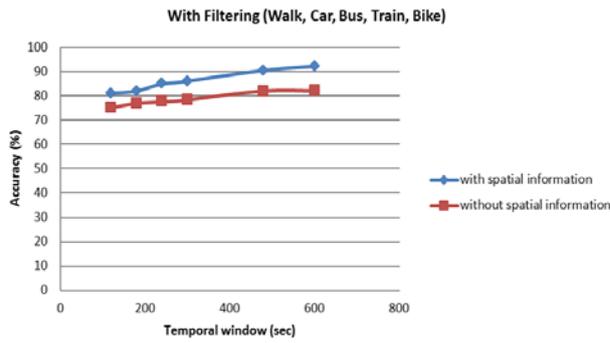


Figure 6: Accuracy measure with filtered walking speed

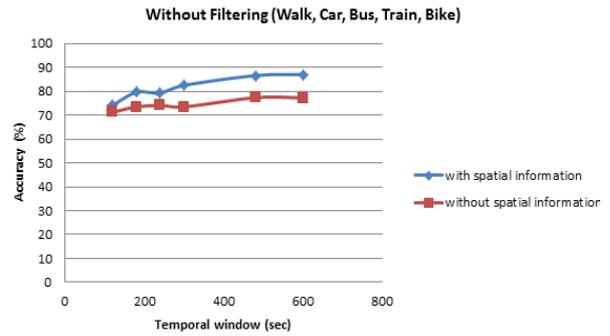


Figure 7: Accuracy without filtered walking speed

The methodology was also tested on car, walk and train assuming these modes will show quite distinct behaviour in terms spatial relevance as well as kinematic relevance. When spatial information was used accuracy reached 95% for filtered trajectories and 93% for unfiltered trajectories (Fig. 9). A spatial visualization is also presented to show the classification accuracy for three modes (train, car and walk). The diagonals are true positives and off-diagonals are false positives (Fig. 8). The figure shows the model can give a high accuracy and less type I and type II error for walking. However due to similar kinematic behaviour some of the car instances are mostly classified as walk owing to type I error. Likewise train instances are sometimes classified as walk and car.

In order to compare state-of-the-art approaches that used only kinematic information, another test was conducted within temporal window of 300 sec, on car, bus and walk modes. It was found there was a small difference in estimated accuracy by using only kinematics inputs, and kinematics and spatial inputs together. However using spatial information and kinematics, the accuracy is certainly more than that of using kinematics alone. The small difference of accuracy can be justified as the bus network has not been used in this research; only the road network was used. A car or bus both can travel on road network and hence it was not easy to distinguish between car and bus. But there is a good chance that car and bus can be easily distinguished by using a bus network. Using spatial information average accuracy for car, walk and bus was achieved 81.24 % whereas without spatial information the accuracy was 79.50 %.

We also compared the performance of our MLP neural network with some of the well-studied machine learning algorithms. The result shows a MLP neural network outperforms other approach. Interestingly the accuracy of a MLP neural network increases as the size of the time window increases whereas other approaches show saturation over growing time window. This clearly shows the ability of a MLP to learn and adapt well in near-real time as more instances come in with fine and varied state behaviour of different modes (Table 4).

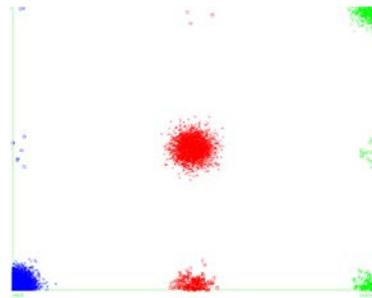


Figure 8: Classification accuracy visualization. Class map: {blue: walk; red: car; green: train}; X-axis indicates actual modes and Y-axis indicates predicted modes

## 6 Discussions

From the result it is evident that spatial information can improve mode detection accuracy significantly, especially in near-real time. In near-real time detecting different transport modes is challenging from their movement attributes only, as the queries are issued for a very short interval and different modes may have similar movement behaviour. Earlier literature did not consider minimum speed and acceleration as all of them are offline and based on segmenting the entire trajectory in each modal segment that normally starts with zero speed and zero acceleration (Gonzalez et al., 2010).

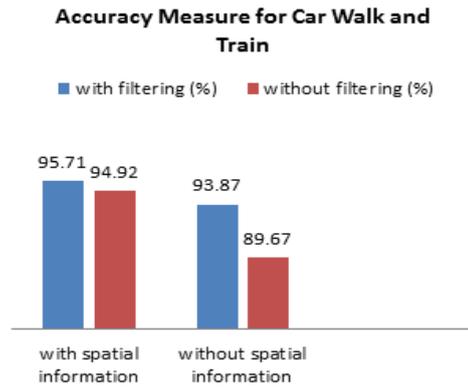


Figure 9: Accuracy measure for three modes

Table 4: Accuracy measures for different classifiers against different time window

Time window (sec)	SVM (%)	MLP Neural network (%)	Logistic regression (%)	RBF Network (%)
120	59.10	74.08	68.51	58.77
180	60.65	79.65	70.32	60.37
240	64.15	79.34	72.51	64.15
300	61.78	82.55	75.20	66.89
480	66.84	86.63	77.86	68.95
600	68.35	87.05	77.76	66.87

But in real time since a query can come at any time and the given mode need not to start or finish from stationary state within the given temporal window, hence minimum and maximum bounds, central tendency and variance of speed and acceleration distribution over the given time window have been considered. In order to supplement the mode detection accuracy various spatial information have been considered assuming the fact that each mode shows distinct spatial behaviour, such as bus and car will travel along the roadway; likewise train will travel on train network only. However, in this research the bus network of Beijing was not used, instead the entire road network consisting of the bus network and roads together was used. In order to distinguish different modes especially bike, car and bus different POI relevance has been considered such as bus stop for bus, traffic for bike, bus and car, car wash and parking lot for car. However, it has been observed that when the time window is short it cannot capture distinct behaviour and thus leads to ambiguities especially for car, bus, walk and bike that share similar movement and spatial relevance at some instants.

From individual classification accuracy for five modes it has been observed walking has the highest false positives as different modes can be slowed down and behave like walking. At the same time this can also give false negative as people can walk on the road, near the bus stop, traffic signal or parking lot and hence respective POI relevance may be higher for a walking segment. That gives a false impression of other modes corresponding to the respective spatial relevance.

There is also a trade-off between the temporal window and the accuracy that raises questions of selecting the proper time window for a given location based service. Say for emergency services, the amount of response time required is less than the time required by an urban planner or traffic engineer to understand travel demand from people's mode choice or location based context-aware advertisements. This also poses challenges of selecting a proper and optimal window size to detect transport modes as accurately as possible. However from this study it is evident that the window size bears an inverse relationship with the accuracy measure, assuming the mode is not changed within the temporal window.

Earlier work using neural network only performed offline analysis considering only three modes. The highest reported accuracy achieved was 91% (Gonzalez et al., 2010). However our approach capable of real time estimation that shows accuracy can reach up to 95% or more when using the road network along with other relevant information for three modes, and 87% when using five modes over 600 sec. However the accuracy depends on number of modes, their spatial relevance and the size of the temporal window. The accuracy also depends on the clustering algorithm used to estimate POI relevance over a given temporal window.

However, at this moment this approach is limited by detecting only a single mode within a given temporal window. But in real time there is a possibility that a quick transfer can take place from one mode to another mode followed by walking. This creates composite mode segments within a same temporal window which is difficult to detect using only GPS signals (Das et al., 2014). In order to distinguish two different modes within a composite segment different inertial sensor information along with GPS is required that can distinctly detect presence of two different modal class within a given temporal segment from their characteristic kinematic signatures.

## 7 Conclusions

Detecting transport modes in near-real time is an emerging research area. This is particularly useful in various context-aware location based services and understanding urban dynamics in near-real time. In order to detect various transport modes Microsoft's GeoLife dataset has been used in this research. In this research a simulated near-real time mode detection classification framework has been developed using a neural network based classifier. We have evaluated the performance of neural network in detecting various modes, since neural networks can adjust well with different input and output parameters online. Neural networks also offer flexibility and scalability in terms of learning ability and accommodating new information from the external world. In this paper, we particularly focused on how real time mode detection accuracy varies with varying temporal window size. This has been figured out within a small temporal window all the transport modes show similar kinematic behaviour. In order to detect different modes more accurately we used various spatial information such as route network information and POI information.

We tested our hypothesis on three sets of modes: two sets containing three modes and one set containing five modes). Our result shows incorporating spatial information can improve mode detection accuracy. We achieved accuracy 95% accuracy on three modes only and 93% accuracy on all five modes. The result also shows a MLP neural network can outperform other machine learning algorithms with growing temporal window size.

Future work will look into distinguishing a composite segment within a given temporal window where a quick transfer has occurred. In order to explore different modes within a temporal segment, different sensor signals such as accelerometer, proximity sensor, gyroscope information are required that can give characteristic movement behaviour of each modes at a very fine granularity. In this research while forming the clusters we only considered spatial relevance of each cluster with respect to given POI. We did not consider temporal relevance as temporal window may vary from as small as 120 sec to as high as 600 sec or more. During smaller temporal window, it is difficult to set a temporal relevance or dwell time. Future research will address spatio-temporal issues while developing potential clusters within a given temporal segment to calculate spatio-temporal relevance for each mode.

## References

- Biljecki, F., Ledoux, H., & Oosterom, P. V. (2012). Transportation mode-based segmentation and classification of movement trajectories. *International Journal of Geographical Information Science*, 17(2), 385-407.
- Bohte, W., Maat, K., & Quak, W. (2008). A method for deriving trip destinations and modes for GPS-based travel surveys. In J. V. Schaick & S. V. d. Spek (Eds.), *Urbanism on Track* : IOS Press, pp. 129-145.
- Byon, Y., Abdulhai, B., Shalaby, A. S. (2007). Impact of sampling rate of GPS enabled cell phones on mode detection and GIS map matching performance. *Transportation Research Board Annual Meeting 7(1795)*, 21.
- Byon, Y., Abdulhai, B., & Shalaby, A. (2009). Real-time transportation mode detection via tracking Global Positioning System mobile devices. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 13(4), 161-170.
- Das., R. D., Ronald., N., & Winter., S. (2014). Clustering based transfer detection with fuzzy activity recognition from smart-phone GPS trajectories, 17<sup>th</sup> IEEE International Conference on Intelligent Transportation Systems, Qingdao, China.
- Dodge, S., Weibel, R., & Forootan, E. (2009). Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects. *Computers, Environment and Urban Systems*, 33(6), 419-434.
- Gonzalez, P., Weinstein, J., Barbeau, S., Labrador, M., Winters, P., Georggi, N., & Perez, R. (2010). Automating mode detection for travel behaviour analysis by using Global Positioning Systems enabled mobile phones and neural networks. *Intelligent Transport Systems, IET*, 4(1), 37-49.

- Hemminki, S., Nurmi, P., & Tarkoma, S. (2013). Accelerometer-based transportation mode detection on smartphones. 11<sup>th</sup> ACM Conference on Embedded Networked Sensor Systems, Rome, Italy.
- Hemminki, S., Nurmi, P., & Tarkoma, S. (2013). Accelerometer-based transportation mode detection on smartphones. 11<sup>th</sup> ACM Conference on Embedded Networked Sensor Systems, Roma, Italy.
- Liao, L., Fox, D., & Kautz, H. (2007). Extracting places and activities from GPS traces using hierarchical conditional random fields. *International Journal of Robotics Research*, 26(1), 119-134.
- Mamdani, E.H. and Assilian, S, (1975). An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies*, 7(1), 1-13.
- Minetti, A. (2000). The three modes of terrestrial locomotion. *Biomechanics and Biology of Movement* B. M. Nigg, B. R. MacIntosh, and J. Mester, Eds., ed: *Human Kinetics*, pp. 67–78.
- Mountain, D., & Raper, J. (2001). Modelling human spatio-temporal behaviour: a challenge for location-based services. *GeoComputation*, Brisbane, Australia.
- Ohashi, H., Akiyama, T., Yamamoto, M., & Sato, A. (2013). Modality classification method based on the model of vibration generation while vehicles are running. 6<sup>th</sup> ACM SIGSPATIAL International Workshop on Computational Transportation Science, Orlando, USA.
- Patterson, D., Liao, L., Fox, D., & Kautz, H. (2003). Inferring high-level behavior from low-level sensors. *Lecture Notes in Computer Science*, Berlin, Germany.
- Pereira, F., Carrion, C., Zhao, F., Cottrill, C. D., Zegras, C., & Ben-Akiva, M. (2013). The Future Mobility Survey: Overview and preliminary evaluation. *Proceedings of the Eastern Asia Society for Transportation Studies*.
- Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., & Srivastava, M. (2010). Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks*, 6(2), 1-27.
- Schuessler, N., & Axhausen, K. (2009). Processing GPS raw data without additional information. 88th Annual Conference of the Transportation Research Board, Washington D.C. USA.
- Tsui, S. Y. A., & Shalaby, A. S. (2006). Enhanced system for link and mode identification for personal travel surveys based on Global Positioning Systems. *Transportation Research Record*, 1972, 38-45.
- Wolf, J. (2000). Using GPS data loggers to replace travel diaries in the collection of travel data. PhD thesis, Georgia Institute of Technology, Atlanta, USA.
- Zadeh, L. A. (1965). Fuzzy sets, *Information and Control* 8(3), 338-353.
- Zheng, Y., Chen, Y., Li, Q., Xie, X., & Ma, W.-Y. (2010). Understanding transportation modes based on GPS data for Web applications. *ACM Transactions on The Web*, New York, USA.
- Zheng, Y., Li, Q., Chen, Y., & Xie, X. (2008). Understanding mobility based on GPS data. *ACM conference on Ubiquitous Computing (UbiComp 2008)*, Seoul, Korea. ACM Press: 312–321.
- Zheng, Y., Liu, L., Wang, L., & Xie, X. (2008). Learning transportation modes from raw GPS data for geographic application on the web, *International conference on World Wide Web (WWW 2008)*, Beijing, China. ACM Press: 247-256.

# AUSGeoid09 Performance in Mountainous Terrain: A Case Study in the Blue Mountains

Joseph Allerton  
University of New South Wales  
Sydney NSW 2052, Australia  
joe.allerton@hotmail.com

Volker Janssen  
Land and Property Information  
Bathurst NSW 2795, Australia  
Volker.Janssen@lpi.nsw.gov.au

A.H.W. (Bill) Kearsley  
University of New South Wales  
Sydney NSW 2052, Australia  
w.kearsley@unsw.edu.au

## Abstract

AUSGeoid09 is the latest model used to convert Global Navigation Satellite System (GNSS) derived ellipsoidal heights to heights in the Australian Height Datum (AHD). While previous studies have evaluated the performance of the AUSGeoid09 model across Australia, such studies have not focused on mountainous regions in particular. This paper evaluates AUSGeoid09 in the Blue Mountains region of New South Wales from a practical user's point of view. Along a 90 km stretch of road incorporating flat to mountainous terrain, comparisons were undertaken between AUSGeoid09-derived heights and published AHD heights, using repeated Network Real Time Kinematic (NRTK) GNSS observations. The performance of AUSGeoid09 was also evaluated relative to its predecessor, AUSGeoid98, and the latest gravimetric model AGQG2009. It was found that AUSGeoid09 performs well across the study area and provides a significant improvement over AUSGeoid98. AUSGeoid09 generally allows AHD height determination at the  $\pm 0.03$  m level (1 sigma) in flat terrain and at the  $\pm 0.06$  m level (1 sigma) in mountainous terrain. However, across the entire study area, AUSGeoid09-derived AHD heights are consistently *lower* than the published AHD heights. Comparison of the results obtained with AUSGeoid09 against those using AGQG2009 in flat terrain illustrates the benefit that the introduction of the geometric component of AUSGeoid09 has had on the determination of AHD heights with satellite technology. However, for elevations above 500 m it appears that the geometric component *degrades* the fit to AHD in the study area, indicating that there is room for improvement in regards to future versions of the AUSGeoid model.

**Keywords:** AUSGeoid09, geoid model, height datum, AHD, NRTK, GNSS.

## 1 Introduction

Most countries have adopted Mean Sea Level (MSL) as zero height surface for their national vertical datum (e.g. Featherstone & Kuhn, 2006; Janssen, 2009). Height above MSL is crucial information for a wide range of applications, e.g. flood modelling, emergency management and engineering construction. In Australia, MSL was approximated as the basis of the Australian Height Datum (AHD) by setting to zero the average MSL values of 32 tide gauges around the country for a period of about two years that began in 1966 (Roelse et al., 1971). The national height datum is complemented by the Geocentric Datum of Australia 1994 (GDA94) for horizontal positions (ICSM, 2009).

More than 40 years after its inception, it is well known that shortcomings in the AHD realisation (AHD71 for mainland Australia and AHD83 for Tasmania) resulted in MSL not being coincident with the geoid at the tide gauges involved. These shortcomings included not considering dynamic ocean effects (e.g. winds, currents, atmospheric pressure, temperature and salinity), a lack of long-term tide gauge data, and the omission of observed gravity. This means that the reference surface for AHD is not a truly level (geopotential) surface, although it was intended to be so when created and is generally used as if it was level. In fact, AHD has variations from the level surface of 1.5 m across Australia and is therefore considered a third-order datum (Morgan, 1992; Featherstone & Filmer, 2012). However, for operational convenience and to avoid confusion, the non-level AHD continues to be used as a practical height datum that provides a sufficient approximation of the geoid for many applications. Consequently, in practice, AHD heights are often accepted as being equivalent to orthometric heights.

Over the last two decades, Global Navigation Satellite System (GNSS) technology has become the primary positioning tool due to its accuracy, speed and accessibility. In particular, Network Real Time Kinematic (NRTK) GNSS is being utilised for a wide range of surveying, mapping, agriculture, mining and construction applications, providing users with instant and highly accurate position information over distances of several tens of kilometres. The advantage of NRTK is its ability to provide corrections (accounting for atmospheric and satellite orbit errors) that are based on a Continuously Operating Reference Station (CORS) network rather than a single reference station (e.g. Wang et al., 2010; Janssen & Haasdyk, 2011; Penna et al., 2012).

GNSS-based heights refer to a reference ellipsoid, i.e. a purely mathematical representation of the earth, and therefore have no physical meaning. In most practice, however, heights are required that correctly reflect the flow of water (or at least sufficiently approximate it), e.g. for flood modelling or drainage and pipeline design. Hence, a reliable geoid model is required to derive AHD heights from measured ellipsoidal heights (e.g. Kearsley, 1988). These geoid models provide  $N$  values ( $N$ ), also known as geoid undulations or geoid-ellipsoid separations, that can be used to convert GNSS-derived ellipsoidal heights ( $h$ ) to AHD heights ( $H$ ) and vice versa (provided  $N$  and  $h$  refer to the same ellipsoid):

$$H = h - N \quad (1)$$

For many years, the use of geoid models (or quasigeoid models – see Vaniček et al. (2012) for a discussion of the difference) has helped GNSS users to compute AHD heights from ellipsoidal heights. In the Australian context, AUSGeoid09 is the latest model that best fits AHD (Brown et al., 2011; Featherstone et al., 2011).

While the performance of AUSGeoid09, along with the improvements it provides over its predecessor AUSGeoid98, have been investigated previously (e.g. Janssen & Watson, 2010; Brown et al., 2011), these studies have not focused on mountainous regions. Considering that gravity can change dramatically within a few kilometres on the earth's surface in Australia (Darbeheshti & Featherstone, 2009), especially in mountainous terrain, and that observed gravity data are generally sparse in these areas (usually observed along the roads to allow easy access), it is necessary to investigate mountainous regions in particular.

This paper evaluates AUSGeoid09 performance in the Blue Mountains in New South Wales (NSW), from a practical user's point of view, using NRTK GNSS-derived ellipsoidal heights and published AHD heights. AUSGeoid09 is compared to its gravimetric component (AGQG2009) and its predecessor (AUSGeoid98) in the study area.

## 2 AUSGeoid09

AUSGeoid09 was released in March 2011 by Geoscience Australia, replacing the previous model AUSGeoid98 (Featherstone & Guo, 2001). Both models refer to the GRS80 ellipsoid, which was adopted as the reference ellipsoid for GDA94, and cover the same geographical area between 108°E and 160°E longitude and between 8°S and 46°S latitude. However, AUSGeoid09 is provided as a 1' by 1' grid (approximately 1.8 by 1.8 km), making it four times denser than its predecessor (Featherstone et al., 2011).

Previous versions of AUSGeoid were predominantly gravimetric-only quasigeoids, and it was assumed that these were sufficiently close approximations of AHD – an assumption we now know to be incorrect. In contrast, AUSGeoid09 is a combined gravimetric quasigeoid plus a geometric model, providing a direct connection to AHD and thereby allowing a more reliable determination of AHD heights from GNSS observations (Brown et al., 2011). The empirically derived geometric component accounts for the offset between the gravimetric quasigeoid (AGQG2009 – see Featherstone et al., 2011) and AHD, which is predominantly caused by AHD not taking into account sea surface topography including the differential heating of the oceans. Since the warmer or less dense

water off northern Australia is about 1 metre higher than the cooler or denser water off southern Australia, AHD is about 0.5 m above the quasigeoid in northern Australia and roughly 0.5 m below the quasigeoid in southern Australia (Janssen & Watson, 2010; Brown et al., 2011). The introduction of the geometric component takes care of most of this 1-metre trend across Australia (0.6-metre trend across NSW), thereby providing a better overall fit to AHD.

AUSGeoid09 has been shown to convert ellipsoidal heights to AHD heights with an accuracy of  $\pm 0.03$  m (1 sigma) across most of Australia, with the exception of some pocket areas where the misfit can be larger than  $\pm 0.1$  m due to errors caused by factors such as the ageing levelling network, geoid height variability or data deficiency (Brown et al., 2011). Using a more practical approach, Janssen & Watson (2010) found that AUSGeoid09 generally allows GNSS-based height determination in NSW at the  $\pm 0.05$  m level (1 sigma). In contrast, its predecessor AUSGeoid98 only provides an absolute accuracy of  $\pm 0.4$  m (Featherstone & Guo, 2001; Featherstone et al., 2001).

A recent study by Sussanna et al. (2014) investigated the performance of AUSGeoid09 in two mountainous regions of NSW, based on two sizable GNSS network adjustment datasets. In the Mid Hunter (elevations ranging between 20 m and 1,400 m), AUSGeoid09 was able to deliver AHD heights at the  $\pm 0.04$  m level (1 sigma) and provided a substantial improvement over its predecessor, clearly demonstrating the benefits of its geometric component on GNSS-derived AHD height determination. In the Snowy Mountains (elevations ranging from 200 m to 2,200 m), AHD height determination was achieved at the  $\pm 0.07$  m level (1 sigma) and moderate improvement over AUSGeoid98 was evident. However, a slope was detected for AUSGeoid09 residuals, and it appears that the geometric component may have overcompensated for sea surface topography in this area.

### 3 Testing methodology

Owing to the increased use of GNSS CORS networks, the absolute accuracy of N values is now more important than ever for AHD height determination using satellite positioning techniques (Janssen & Watson, 2010). In this paper, the performance of the AUSGeoid09 model in mountainous terrain is evaluated in the Blue Mountains, based on the comparison of repeated NRTK GNSS observations and published AHD heights.

A number of spirit-levelled benchmarks with known AHD heights of sufficient quality (Class LB Order L2 or better) on public record in the Survey Control Information Management System (SCIMS) were used as test points. SCIMS is the state's database containing about 250,000 survey marks across NSW, including coordinates, heights and other information (Kinlyside, 2013). For a discussion of the terms class and order, the reader is referred to ICSM (2007) and Dickson (2012). It should be noted that the published AHD heights of the benchmarks used as test points are not guaranteed to be error-free. Consequently, they can only be treated as benchmarks defining AHD in the study area (and comparisons are thus made to the national height datum), not as representing true orthometric heights.

In order to replicate a practical scenario, these test points were occupied multiple times using the NRTK GNSS technique to obtain ellipsoidal heights. NRTK observations were based on CORSnet-NSW (e.g. Janssen et al., 2013), an expanding state-wide network of more than 160 GNSS CORS providing fundamental positioning infrastructure for New South Wales (Figure 1). As recommended by Janssen & Haasdyk (2011), the windowing technique was applied to increase reliability of the resulting positions and each test point was re-observed several times (using a tripod for stability) to ensure redundancy and allow for changes in satellite geometry between occupations.

#### 3.1 Windowing technique

The windowing (or averaging) technique is often applied to improve the positioning result for real-time applications if the GNSS rover is allowed to remain stationary for a short period of time. Windowing is achieved by determining the average of several epochs observed at a point, thus increasing the reliability of the resulting position. It is important to note that windowing reduces the effect of extreme, short-lived outlier observations, but can still produce results that are significantly offset from the mean. Janssen & Haasdyk (2011) suggest observing for 1-2 minutes to obtain an averaged position in order to reduce the effects of individual coordinate solution outliers. If high-quality vertical coordinates are required or the user is distant from the nearest CORS, it is advised to use a 2-minute observation window. It was also shown that double occupations, i.e. re-observing each point after waiting about 10-30 minutes or more, increases the reliability of the resulting position.

In a preliminary test, the benefit of using the windowing technique in this study was assessed by occupying several test points multiple times and collecting NRTK-based positions for short (10 epochs) and longer (3 minutes) observation windows. The standard deviations of the ellipsoidal height observations on each test point were then investigated to determine which observation window length should be used in the following analysis.



Figure 1: CORSnet-NSW network map as of February 2015 (LPI, 2015).

### 3.2 Evaluation of AUSGeoid09 in flat and mountainous terrain

Using equation 1, observed NRTK GNSS-derived ellipsoidal heights ( $h$ ) were converted to AHD heights ( $H$ ) using three quasigeoid models (i.e. AUSGeoid09, AUSGeoid98 and AGQG2009) and compared to the published AHD height ( $H_{AHD}$ ) at each test point. The test points were chosen to ensure that an even number of these were located in flat terrain and mountainous terrain in order to allow evaluation of AUSGeoid09 performance in both terrain conditions.

Descriptive statistics were used to quantify the performance of AUSGeoid09 in determining AHD heights from GNSS observations in the study area. Since it is necessary to consider residuals of different signs, the Root Mean Square (RMS) was also utilised. Furthermore, a comparison between AUSGeoid09 and its predecessor AUSGeoid98 as well as AUSGeoid09's gravimetric component (AGQG2009) was performed to quantify the improvement gained in mountainous areas and investigate the effect of introducing the geometric component.

## 4 Study area

The performance of AUSGeoid09 in mountainous terrain was evaluated in the Blue Mountains located west of Sydney, NSW. The study area incorporates 23 test points along Windsor Road and Bells Line of Road between Castle Hill in the east and Lithgow in the west (Figure 2). It exhibits initially flat terrain (blue test points) changing into substantially undulating terrain with large differences in elevation (red test points), thus representing typical mountainous terrain conditions encountered in Australia (Figure 3).

The eastern half of the study area follows Windsor Road, connecting the Sydney suburb of Castle Hill in the east with the townships of Windsor and Richmond towards the west. This line is approximately 40 km long and exhibits relatively flat terrain, gently changing between about 10 m and 125 m in elevation. The western half of the study area is located along Bells Line of Road, a winding mountain road connecting the townships of Richmond and Lithgow. This line is approximately 50 km long and exhibits undulating terrain with elevations ranging from about 185 m to 1,100 m. AUSGeoid09 N values range from 23 m to 26 m in the study area, steadily increasing from east to west.

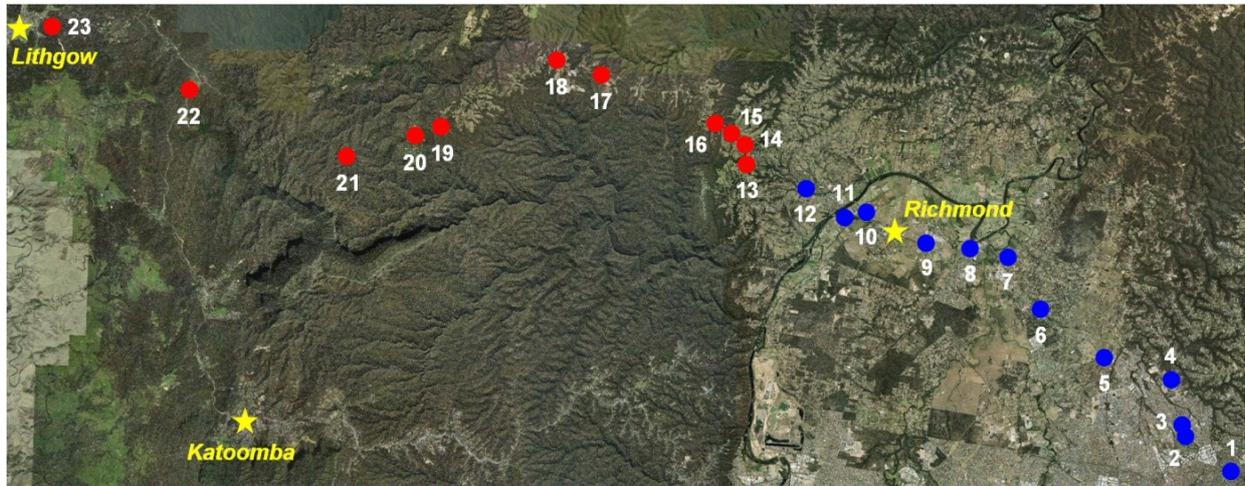


Figure 2: Location of the test points (SCIMS benchmarks) in the study area.

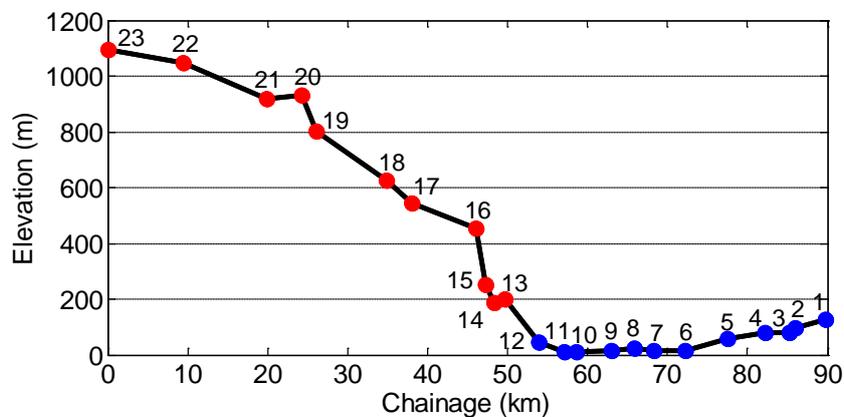


Figure 3: Cross section of the test points, indicating the range in elevation (AHD height).

## 5 Data collection and analysis

### 5.1 Field work issues

Naturally, it is desired to obtain the largest possible sample of high-quality test points and utilise high-precision positioning techniques in order to provide a solid base for analysis. In practice, however, compromises have to be made due to constraints in regards to time and resources. While a lot of first-order levelled benchmarks (LAL1) are available close to Sydney, most of these have an ‘unknown’ classification for the horizontal class and order (UU). This means that the published horizontal coordinates are generally only accurate to several tens of metres, in some cases up to 200 m. Furthermore, the locality sketches for most of these benchmarks were drawn in the early 1960s (when the benchmarks were placed) and did not have many useful references to locate the marks 50 years later. While SCIMS indicates whether a mark has been ‘destroyed’ or ‘not found’ or even ‘found intact’, the marks that have not been classified as such are not guaranteed to still exist. Most of the benchmarks were placed on the side of the road and not located near identifiable features such as road intersections, houses or other physical structures. References mainly consisted of fencing, power poles and mile posts. Understandably, these features have been replaced, removed or have deteriorated over time.

As a result, many of these marks could not be recovered and second-order levelled marks (LBL2) had to be used instead. Along Bells Line of Road, only 8 marks were found out of the initially selected 25 benchmarks, and the dataset was therefore expanded with three LBL2 marks towards the east (Kurrajong) to obtain a larger sample for analysis. While the majority of marks used along Bells Line of Road are classified LAL1, all 12 marks used along Windsor Road are classified LBL2.

Another issue that arose (as expected) was the inability to observe directly over a benchmark due to limited skyview or multipath issues. This problem was overcome by placing an arbitrary (eccentric) mark, i.e. nail in bitumen or hard ground, a short distance away at a location with more favourable observing conditions. The AHD height of the benchmark was then transferred to the arbitrary mark using a total station, measuring four sets of reciprocal zenith angles and slope distances to determine the vertical distance between the two marks. A spirit level and staff were not used because this could not be achieved by only one person in the field. This procedure was required for four test points along Bells Line of Road (SS342, SS341, SS336 and SS319). Finding a larger number

of the initially selected marks would have increased the sample size and may have reduced the percentage of benchmarks observed using arbitrary marks. Fortunately, field work was not hampered by encounters with native wildlife (Janssen, 2012).

## 5.2 Windowing technique and ellipsoidal height quality

A preliminary test was undertaken to assess the benefit of using the windowing technique in this study. A number of test points were occupied several times (generally between 30 minutes and several days apart), collecting NRTK-based positions based on 10 epochs of data as well as 3 minutes of data. Table 1 states the number of occupations and the standard deviation (STD) of the ellipsoidal height observations on each test point. It can be seen that the longer averaging window improves the mean standard deviation slightly and appears to remove two relatively large values of 0.030 m and 0.041 m. Based on these findings, it was decided to occupy all remaining test points for 3-minute observation windows.

Table 1: Assessment of the windowing technique based on the standard deviation of ellipsoidal height observations (STD given in metres).

Test Point	Mark ID	No. Occ. (10 epoch)	No. Occ. (3 min)	STD (h) (10 epoch)	STD (h) (3 min)
17	SS342	4	3	0.030	0.017
18	SS341	8	8	0.009	0.017
19	SS336	6	7	0.041	0.022
20	SS335	6	8	0.017	0.023
21	SS332	8	8	0.016	0.025
22	SS44364	4	4	0.010	0.013
23	SS319	6	6	0.013	0.007
			<b>Mean</b>	<b>0.020</b>	<b>0.018</b>

The average quality of all observed GNSS ellipsoidal heights indicated by the GNSS rover was  $0.014 \text{ m} \pm 0.003 \text{ m}$  along Windsor Rd (flat terrain) and  $0.025 \text{ m} \pm 0.009 \text{ m}$  along Bells Line of Road (mountainous terrain). However, research has shown that these quality indicators can be overly optimistic (e.g. Wang et al., 2010; Janssen & Haasdyk, 2011). In order to get a better appreciation of the quality of the observed ellipsoidal heights, the vertical precision of all 3-minute occupations on each test point was investigated. An average STD of 0.010 m was achieved in flat terrain (ranging from 0.001 m to 0.027 m), while the average STD increased to 0.019 m (ranging from 0.002 m to 0.042 m) in mountainous terrain. These findings, along with a thorough investigation of the collected data, provided confidence that the dataset was free of outliers.

## 5.3 Flat terrain: Windsor Road

Each test point was occupied at least three times with NRTK GNSS using CORSnet-NSW and 3-minute observation windows. The average ellipsoidal height was determined, and AHD heights were calculated by applying the N values from three different quasigeoid models, i.e. AUSGeoid09, AUSGeoid98 and AGQG2009. The AGQG2009 model was kindly provided by Geoscience Australia as it is not readily available to the public in order to avoid confusion between this gravimetric-only model and the published AUSGeoid09. The derived AHD heights for each test point were then compared to the published SCIMS values, which were used as control in this study, to evaluate the performance of AUSGeoid09 and compare the models.

Table 2 shows the differences between NRTK GNSS-derived AHD heights (using the three quasigeoid models) and published AHD heights (AHD<sub>SCIMS</sub>) for the test points situated in flat terrain along Windsor Road. Descriptive statistics show that AUSGeoid09 allows AHD height determination with an accuracy of about  $-0.026 \pm 0.016 \text{ m}$  (1 sigma) in this part of the study area. As all differences are negative, the Root Mean Square (RMS) provides a similar value, i.e. 0.030 m. Using AUSGeoid09 rather than its predecessor AUSGeoid98 resulted in substantially better agreement with published AHD heights, e.g. illustrated by the RMS dropping from 0.211 m to 0.030 m – an improvement by a factor of 7.0.

Comparing the results obtained with AUSGeoid09 against those using AGQG2009 illustrates the benefit that the introduction of the geometric component of AUSGeoid09 has had on the determination of AHD heights with satellite positioning technology. In this case, the RMS improved by a factor of 2.5, dropping from 0.076 m to 0.030 m. For all test points, AUSGeoid09 provided heights that are about 30-50 mm closer to the published AHD values than those obtained using AGQG2009. This improvement is consistent with the geometric component of AUSGeoid09 generally amounting to about -0.05 m or less in this area (Brown et al., 2011), considering that a negative geometric component results in a smaller N value and therefore a larger derived AHD height. While the range of differences for all three models remains at a similar level of 0.06 m, it is obvious that AUSGeoid09 provides heights that are much closer aligned with the published values than the other two models. In other words,

the evolution from AUSGeoid98 to AGQG2009 and AUSGeoid09 has significantly improved the fit between GNSS-derived and published AHD heights.

Table 2: Agreement of NRTK GNSS-derived AHD heights (using different quasigeoid models) to SCIMS in flat terrain (Windsor Road) – all values stated in metres.

Test Point	Mark ID	Class/Order	No. Occ. (3 min)	AHD <sub>SCIMS</sub>	$\Delta$ AHD <sub>AG09</sub>	$\Delta$ AHD <sub>AG98</sub>	$\Delta$ AHD <sub>AGQG09</sub>
1	SS61355	LBL2	3	126.160	-0.037	-0.203	-0.092
2	PM31480	LBL2	4	95.893	-0.026	-0.200	-0.081
3	SS44875	LBL2	3	79.272	-0.015	-0.190	-0.071
4	SS17350	LBL2	4	77.988	-0.020	-0.200	-0.075
5	PM29273	LBL2	4	59.921	-0.031	-0.218	-0.087
6	PM67548	LBL2	6	14.102	-0.028	-0.220	-0.081
7	PM41505	LBL2	6	14.045	-0.035	-0.230	-0.087
8	SS10399	LBL2	6	24.293	-0.004	-0.198	-0.052
9	PM45758	LBL2	6	17.167	-0.015	-0.204	-0.059
10	SS75618	LBL2	4	11.445	-0.037	-0.225	-0.073
11	PM74693	LBL2	3	11.552	-0.062	-0.251	-0.095
12	PM45486	LBL2	3	44.985	-0.002	-0.190	-0.031
				<b>Mean</b>	<b>-0.026</b>	<b>-0.211</b>	<b>-0.074</b>
				<b>STD</b>	<b>0.016</b>	<b>0.018</b>	<b>0.019</b>
				<b>RMS</b>	<b>0.030</b>	<b>0.211</b>	<b>0.076</b>
				<b>Min</b>	<b>-0.062</b>	<b>-0.251</b>	<b>-0.095</b>
				<b>Max</b>	<b>-0.002</b>	<b>-0.190</b>	<b>-0.031</b>
				<b>Range</b>	<b>0.060</b>	<b>0.061</b>	<b>0.064</b>

#### 5.4 Mountainous terrain: Bells Line of Road

The same procedure was applied to the remainder of test points, located in mountainous terrain along Bells Line of Road. The resulting differences between NRTK GNSS-derived AHD heights (using the three quasigeoid models) and published AHD heights (AHD<sub>SCIMS</sub>), along with descriptive statistics, are shown in Table 3.

It can be seen that AUSGeoid09 allows AHD height determination with an accuracy of about  $-0.060 \pm 0.039$  m (1 sigma) in this part of the study area. As all differences are negative, the Root Mean Square (RMS) provides a similar value, i.e. 0.071 m. Again, AUSGeoid09 provides substantially better agreement with published AHD heights than its predecessor AUSGeoid98, e.g. illustrated by the RMS dropping from 0.174 m to 0.071 m – an improvement by a factor of 2.5.

It should be noted that test points 18 & 23 (SS341 & SS319) show much larger discrepancies (at the 120 mm level) to the published AHD heights than all other test points in the study area. This suggests that the published heights of these marks may be incorrect (possibly due to mark movement) or deficiencies may be present in the quasigeoid model used. At SS341, this substantial disagreement is evident across all three models investigated, and neighbouring marks do not exhibit similar behaviour, supporting that subsidence may be the cause. At SS319, the AGQG2009-derived AHD height agrees much better (at the 40 mm level) with the published AHD height, suggesting that in this case the discrepancy may not be due to mark subsidence but rather be caused by deficiencies in the geometric component of AUSGeoid09. In both cases, however, the limited data available precludes a definitive answer.

If these two test points were removed from the analysis, the accuracy of AUSGeoid09-derived AHD heights improves to  $-0.046 \pm 0.025$  m (1 sigma), while the RMS improves to 0.052 m and the range drops from 0.111 m to 0.070 m. In addition, AUSGeoid09 improves the fit to published AHD by a factor of 3.2 over its predecessor AUSGeoid98, with the RMS dropping from 0.164 m to 0.052 m.

Comparison of the results obtained with AUSGeoid09 and AGQG2009 shows that, contrary to the findings in flat terrain, the introduction of AUSGeoid09's geometric component overall has not had a positive effect in this part of the study area. While the range of differences from published AHD heights improved slightly, the RMS actually increased by a factor of 1.4, from 0.050 m to 0.071 m. (If test points 18 & 23 were removed from the analysis, the RMS increases by a factor of 1.3, from 0.039 m to 0.052 m.) Closer inspection reveals that for elevations below 500 m, the geometric component improved the fit to published AHD heights by about 15 mm. However, for elevations above 500 m, the geometric component appears to degrade the fit by about 10-30 mm. For elevations above 1,000 m, this negative effect is even larger, exceeding 70 mm. While it is recognised that the sample size is small, this may indicate possible problems with the geometric component at high elevations – not surprisingly, as it is well known that suitable datasets for the generation of the geometric component are notoriously sparse in mountainous regions. It is also understood that the gravimetric geoid is weaker in mountainous regions because (1) gravity data

are limited, (2) existing gravity data are biased along the ridges (roads) and creeks for ease of access, and (3) the terrain effect is less well modelled in regions of large elevation changes in topography.

Table 3: Agreement of NRTK GNSS-derived AHD heights (using different quasigeoid models) to SCIMS in mountainous terrain (Bells Line of Road) – all values stated in metres.

Test Point	Mark ID	Class/Order	No. Occ. (3 min)	AHD <sub>SCIMS</sub>	$\Delta$ AHD <sub>AG09</sub>	$\Delta$ AHD <sub>AG98</sub>	$\Delta$ AHD <sub>AGQG09</sub>
13	PM45491	LBL2	3	201.224	-0.019	-0.184	-0.037
14	PM45501	LBL2	3	186.888	-0.027	-0.191	-0.045
15	PM1677	LAL1	4	251.366	-0.044	-0.197	-0.060
16	SS347	LAL1	3	454.364	-0.050	-0.184	-0.064
17	SS342	LAL1	3	542.686	-0.014	-0.133	-0.009
18	SS341	LAL1	8	626.173	-0.122	-0.230	-0.109
19	SS336	LAL1	7	801.360	-0.073	-0.165	-0.038
20	SS335	LAL1	8	932.277	-0.071	-0.147	-0.032
21	SS332	LAL1	8	918.139	-0.034	-0.111	0.017
22	SS44364	LBL2	4	1046.127	-0.084	-0.144	-0.010
23	SS319	LAL1	6	1092.527	-0.125	-0.189	-0.043
				<b>Mean</b>	<b>-0.060</b>	<b>-0.170</b>	<b>-0.039</b>
				<b>STD</b>	<b>0.039</b>	<b>0.034</b>	<b>0.033</b>
				<b>RMS</b>	<b>0.071</b>	<b>0.174</b>	<b>0.050</b>
				<b>Min</b>	<b>-0.125</b>	<b>-0.230</b>	<b>-0.109</b>
				<b>Max</b>	<b>-0.014</b>	<b>-0.111</b>	<b>0.017</b>
				<b>Range</b>	<b>0.111</b>	<b>0.119</b>	<b>0.126</b>

Investigating the effect of the geometric component in more detail, it is interesting to note that the improvement of fit to published AHD steadily decreases from east to west in the study area, from about 55 mm in the western outskirts of Sydney to zero near Kurrajong. Heading further west through the Blue Mountains, the geometric component increasingly *degrades* (almost linearly with distance) the GNSS-based determination of AHD heights in the study area, culminating in up to 80 mm at the highest elevations near Lithgow (Figure 4). This can be explained by the decreasing density of datasets available for the empirical determination of the geometric component away from metropolitan areas. It is also interesting to note that across the entire study area, featuring both flat and mountainous terrain, AUSGeoid09-derived AHD heights are always *lower* than the published AHD heights. This indicates that there is room for improvement in regards to future versions of the AUSGeoid model, provided additional datasets are collected in this region.

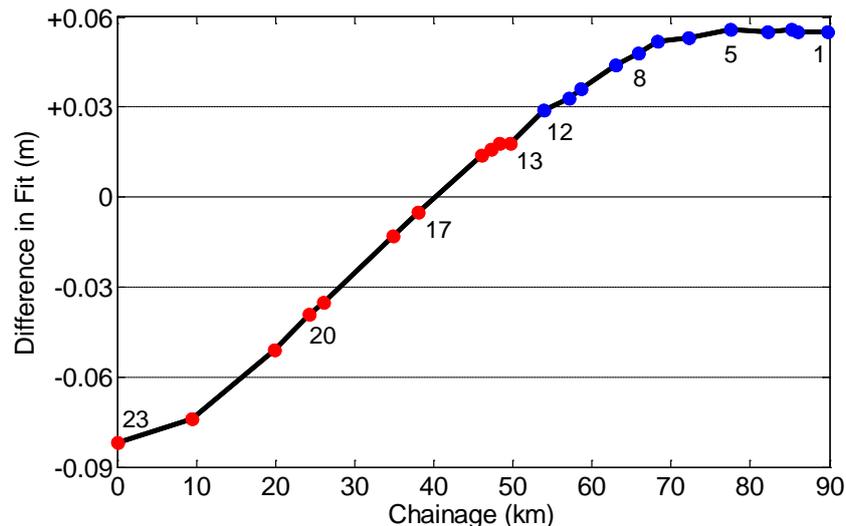


Figure 4: Difference in fit to published AHD heights between AUSGeoid09 and AGQG2009 derived AHD heights in the study area. Positive values indicate improvement of fit due to the geometric component.

## 5.5 Cross sections

In order to provide a visual perspective of the results obtained in the study area, cross sections were generated showing published AHD heights and NRTK GNSS-derived AHD heights based on the three quasigeoid models investigated (Figure 5). The cross sections run from left to right in a west-to-east direction and have been scaled and exaggerated (separately for each part of the study area) to allow a clearer visual comparison.

Across both terrain types, it is clearly evident that AUSGeoid09 provides a far better fit to published AHD values than its predecessor AUSGeoid98. In flat terrain, AUSGeoid09-derived heights are consistently closer to published AHD than AGQG2009-derived heights, showing the benefit of the geometric component. The shape of all quasigeoid-derived cross sections is very similar to the shape of AHD in this part of the study area. In mountainous terrain, the shape of all quasigeoid-derived cross sections is very similar but deviates from the shape of published AHD in several cases. This behaviour is most obvious at test point 18 (SS341), a mark that has already been identified as possibly being affected by subsidence. The cross sections also visualise that the geometric component of AUSGeoid09 appears to increasingly degrade the fit to published AHD for the marks investigated west of test point 17.

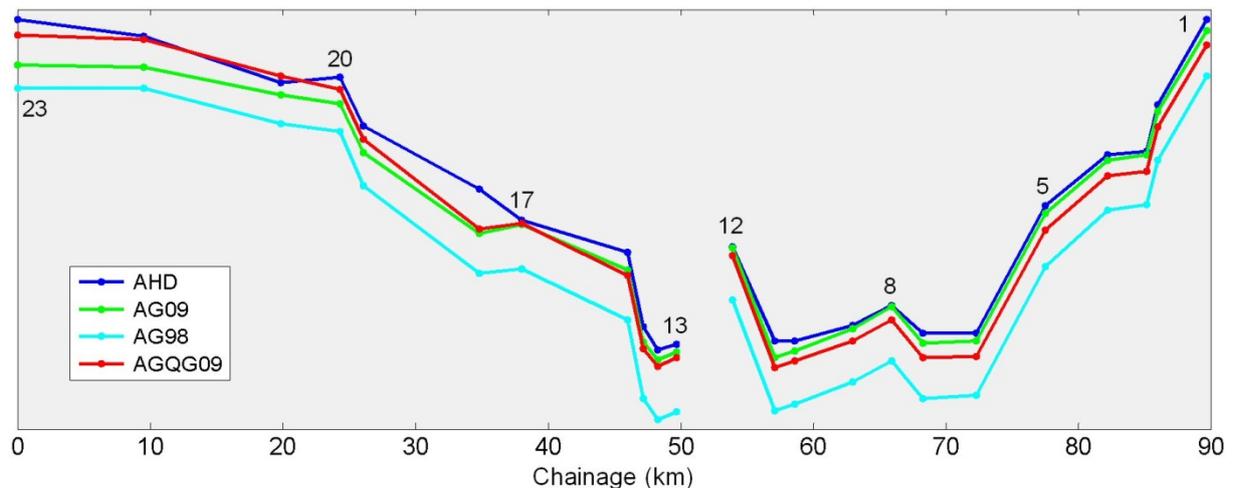


Figure 5: Cross sections showing published AHD heights and NRTK GNSS-derived AHD heights using different quasigeoid models, separately scaled and exaggerated in each part of the study area.

## 6 Conclusion

This paper has evaluated the performance of AUSGeoid09 in the Blue Mountains region of New South Wales, from a practical user's point of view. Along a 90 km stretch of road incorporating flat to mountainous terrain, comparisons were undertaken between AUSGeoid09-derived heights and published AHD heights, using repeated NRTK GNSS observations on 23 test points. The performance of AUSGeoid09 was also evaluated relative to its predecessor, AUSGeoid98, and the latest gravimetric quasigeoid model AGQG2009.

In a preliminary test, it was confirmed that the windowing technique can reduce the effect of extreme, short-lived outlier observations and increase the reliability of NRTK GNSS-derived positions. Consequently, it was decided to occupy the test points for 3-minute observation windows.

It was shown that AUSGeoid09 performs well across the study area and provides a significant improvement over AUSGeoid98. AUSGeoid09 generally allows AHD height determination at the  $\pm 0.03$  m level (1 sigma) in flat terrain and at the  $\pm 0.06$  m level (1 sigma) in mountainous terrain. This level of accuracy agrees well with findings reported in previous studies and is very encouraging, particularly in light of GNSS technology and CORS networks being increasingly used to provide vertical control. However, across the entire study area, AUSGeoid09-derived AHD heights were found to be consistently *lower* than the published AHD heights.

Comparison of the results obtained with AUSGeoid09 against those using AGQG2009 in flat terrain illustrates the benefit that the introduction of the geometric component of AUSGeoid09 has had on the determination of AHD heights with satellite positioning technology. However, it was also found that the geometric component appears to *degrade* the fit to AHD for elevations above 500 m in the study area. This indicates that there is room for improvement in regards to future versions of the AUSGeoid model.

It is recognised that this study is based on a much smaller sample than initially planned. Consequently, it is not suggested that the results presented can be generalised to apply to other mountainous regions. The availability of higher-quality horizontal coordinates for first-order benchmarks in SCIMS would have been very beneficial in regards to locating a larger number of test points in the available time frame. Levelled benchmarks are important for the maintenance of the national height datum, and it should therefore be possible to locate these marks more easily. It is recommended that this well-known issue be addressed by the custodians of SCIMS in the future, particularly in

light of the planned introduction of a next-generation datum for Australia (Haasdyk et al., 2014) and the increasing need to preserve existing survey mark infrastructure (de Belin, 2012; Ward, 2014).

## References

- Brown, N.J., Featherstone, W.E., Hu, G. and Johnston, G.M. (2011). AUSGeoid09: A more direct and more accurate model for converting ellipsoidal heights to AHD heights. *Journal of Spatial Science*, 56(1), 27-37.
- Darbeheshti, N. and Featherstone, W.E. (2009). Non-stationary covariance function modelling in 2D least-squares collocation. *Journal of Geodesy*, 83(6), 495-508.
- de Belin, F. (2012). Preservation? of survey reference marks. *Proceedings of Association of Public Authority Surveyors Conference (APAS2012)*, Wollongong, Australia, 19-21 March, 142-150.
- Dickson, G. (2012). Control surveys: Why things are the way they are and not the way you think they should be! *Proceedings of Association of Public Authority Surveyors Conference (APAS2012)*, Wollongong, Australia, 19-21 March, 12-28.
- Featherstone, W.E. (2001). Absolute and relative testing of gravimetric geoid models using global positioning system and orthometric height data. *Computers and Geosciences*, 27(7), 807-814.
- Featherstone, W.E. and Filmer, M.S. (2012). The north-south tilt in the Australian Height Datum is explained by the ocean's mean dynamic topography. *Journal of Geophysical Research*, 117, C08035, doi:10.1029/2012JC007974.
- Featherstone, W.E. and Guo, W. (2001). Evaluations of the precision of AUSGeoid98 versus AUSGeoid93 using GPS and Australian Height Datum data. *Geomatics Research Australasia*, 74, 75-102.
- Featherstone, W.E., Kirby, J.F., Hirt, C., Filmer, M.S., Claessens, S.J., Brown, N.J., Hu, G. and Johnston, G.M. (2011). The AUSGeoid09 model of the Australian Height Datum. *Journal of Geodesy*, 85(3), 133-150.
- Featherstone, W.E., Kirby, J.F., Kearsley, A.H.W., Gilliland, J.R., Johnston, G.M., Steed, J., Forsberg, R. and Sideris, M.G. (2001). The AUSGeoid98 geoid model of Australia: Data treatment, computations and comparisons with GPS-levelling data. *Journal of Geodesy*, 75(5-6), 313-330.
- Featherstone, W.E. and Kuhn, M. (2006). Height systems and vertical datums: A review in the Australian context. *Journal of Spatial Science*, 51(1), 21-41.
- Haasdyk, J., Donnelly, N., Harrison, C., Rizos, C., Roberts, C. and Stanaway, R. (2014). Options for modernising the Geocentric Datum of Australia. *Proceedings of Research at Locate'14*, Canberra, Australia, 7-9 April, 72-85.
- ICSM (2007). Standards and practices for control surveys (SP1), version 1.7, <http://www.icsm.gov.au/publications/sp1/sp1v1-7.pdf> (accessed Feb 2015).
- ICSM (2009). Geocentric Datum of Australia technical manual, version 2.3(1), <http://www.icsm.gov.au/gda/index.html> (accessed Feb 2015).
- Janssen, V. (2009). Understanding coordinate reference systems, datums and transformations. *International Journal of Geoinformatics*, 5(4), 41-53, 2009.
- Janssen, V. (2012). Indirect tracking of drop bears using GNSS technology. *Australian Geographer*, 43(4), 445-452.
- Janssen, V., Commins, R., Watson, P. and McElroy, S. (2013). Using GNSS CORS to augment long-term tide gauge observations in NSW, *Proceedings of Surveying & Spatial Sciences Conference (SSSC2013)*, Canberra, Australia, 15-19 April, 12pp.
- Janssen, V. and Haasdyk, J. (2011). Assessment of Network RTK performance using CORSnet-NSW, *Proceedings of International Global Navigation Satellite Systems Society Symposium (IGNSS2011)*, Sydney, Australia, 15-17 November, 18pp.
- Janssen, V. and Watson, T. (2010). Improved AHD71 height determination from GNSS using AUSGeoid09 in New South Wales, Australia. *Journal of Global Positioning Systems*, 9(2), 112-121.
- Kearsley, A.H.W. (1988). The determination of the geoid ellipsoid separation for GPS levelling. *Australian Surveyor*, 34(1), 11-18.
- Kinlyside, D. (2013). SCIMS3: The next generation Survey Control Information Management System, *Proceedings of Association of Public Authority Surveyors Conference (APAS2013)*, Canberra, Australia, 12-14 March, 174-186.
- LPI (2015). CORSnet-NSW, <http://www.corsnet.com.au/> (accessed Feb 2015).

- Morgan, P. (1992). An analysis of the Australian Height Datum 1971. *Australian Surveyor*, 37(1), 46-63.
- Penna, N., Clarke, P., Edwards, S. and King, M. (2012). Further testing of commercial Network RTK GNSS services in Great Britain. Report, The Survey Association and Newcastle University, <http://www.tsa-uk.org.uk/for-clients/guidance-notes/> (accessed Feb 2015).
- Roelse, A., Granger, H.W. and Graham, J.W. (1971). *The adjustment of the Australian levelling survey 1970-1971*. Technical Report 12, Division of National Mapping, Canberra, Australia, 81pp.
- Sussanna, V., Janssen, V. and Gibbings, P. (2014). Absolute performance of AUSGeoid09 in mountainous regions, *Journal of Applied Geodesy*, 8(3), 195-203.
- Vaniček, P., Kingdon, R. and Santos, M. (2012). Geoid versus quasigeoid: A case of physics versus geometry. *Contributions to Geophysics and Geodesy*, 42(1), 101-118.
- Wang, C., Feng, Y., Higgins, M. and Cowie, B. (2010). Assessment of commercial Network RTK user positioning performance over long inter-station distances. *Journal of Global Positioning Systems*, 9(2), 78-89.
- Ward, R. (2014). Maintaining the integrity of the cadastre in the Port Macquarie-Hastings Local Government Area. *Proceedings of Association of Public Authority Surveyors Conference (APAS2014)*, Pokolbin, Australia, 31 March – 2 April, 129-144.

# A spatial analysis approach to evacuation management: shelter assignment and routing

Xuefen Liu  
School of Civil & Environmental  
Engineering  
The University of New South Wales  
Sydney, NSW 2032 Australia  
xuefen.liu@student.unsw.edu.au

Samsung Lim  
School of Civil & Environmental  
Engineering  
The University of New South Wales  
Sydney, NSW 2032 Australia  
s.lim@unsw.edu.au

## Abstract

Evacuation planning requires an integrated analysis of heterogeneous spatial datasets including population, road network and facilities. It is a complex and challenging task to delineate evacuation circumstances and make reasonable connections among the datasets which evacuation management of emergency situations will be based on. An evacuation management system requires an easy configuration by evacuation managers who do not necessarily have full knowledge of Geographic Information Systems (GIS) but need to understand the situation promptly and provide decisive instructions that can be fulfilled only when they can manage datasets and develop new workflows in various scenarios. A spatial analysis platform provides toolkits for spatial data acquisition and processing, analysis, and visualisation from/to online open sources. Such toolkits built in typical GIS software are utilised in this paper to show the feasibility of enabling users to manage spatial data and customise their analysis by combining common data analysis tools to meet their requirements in evacuation planning. Case studies are provided to demonstrate the usability of these toolkits in Brisbane flood evacuation management.

## 1 Introduction

Natural hazards are damaging events that may potentially cause casualties, loss or damage of assets, social and economic disruption or environmental degradation. They can be single, sequential or combined in their origins and effects, and different in terms of location, magnitude/intensity, frequency and probability (Alexander, 1991; Thywissen, 2006). Although methods exist for assessing the risks associated with natural hazards and control structures, less effort has been devoted to developing response activities such as evacuation (Johnstone et al., 2009). Evacuation describes the withdraw actions of people from a specific area because of a real or anticipated threat or hazard (Vogt et al., 1992). It is a positive response to disasters and also one of the most beneficial ways to reduce further damage.

Natural hazards threat thousands of lives and a large amount of valuable assets each year. As Australian cities expand due to the increase in population, buildings and infrastructure, disaster events in Australia tend to be more costly (Crompton et al., 2008). Moreover, global climate change has been predicted to be a significant impact on

---

*Copyright © by the paper's authors. Copying permitted only for private and academic purposes.*

In: B. Veenendaal and A. Kealy (Eds.): Research@Locate'15, Brisbane, Australia, 10-12 March 2015, published at <http://ceur-ws.org>

sea level rise and severe storms in Australia (Middelmann, 2007), which increase the possibility of substantial losses of lives and properties. To minimise the negative consequences associated with these potential disasters, policy administrators need to ensure that appropriate emergency plans are in place.

Queensland has a long tradition in dealing with floods. Current work includes academic and government's efforts being conducted in understanding the physical parameters that describe floods' specific characteristics. These parameters determine floods' starts and spreads. Based on the parameters, flood simulation models are built to capture the flood behaviour over the floodplain (Gouldby et al., 2008; Van Der Knijff et al., 2008). More recent researches have been conducted to understand the causes, impacts and lessons learned from severe floods happened in Brisbane (Carter, 2012; van den Honert et al., 2011).

With all the preparation in flood behaviour and impact study, it is time to further investigate into the strategies for evacuation management and address the need for evacuation maps from previous Australian researches. Making the evacuation strategy more convincing and understandable is raised as one of the concerns when less than ten percent of the population living in the flood-prone communities in Grafton responded to the official evacuation warnings (Pfister, 2002). However, more recently, an evaluation of the usefulness of tsunami evacuation maps which specifies inundated zones and potential exits/vertical evacuation buildings was undertaken by interviewing 500 permanent residents, and the results show that it can benefit not only evacuees but also emergency service officers (Dall'Osso et al., 2010).

From the emergency manager's perspective, it is beneficial to preserve an evacuation map showing the arrangement of accommodating potential evacuees. Also, including the official evacuation buildings into the analysis can help evaluate the spatial coverage and effectiveness of existing shelters. To achieve these aims, evacuation managers need to have the skills of acquiring, processing, analysing and visualizing spatial data even though they are not required to understand Geographic Information Systems (GIS) fully. In this case, powerful toolkits for manipulating datasets in GIS software are required.

Spatial analysis is typically defined as a subset of analytic techniques whose results depend on the geographical frame, or will change if the frame changes, or if objects are repositioned within it (Goodchild et al., 1999). These analytic methods for processing data have the objective of solving some scientific or decision-making problem based on the understanding of spatial relationships and patterns in our world. GIS was initially developed as tools for the storage, retrieval and display of geographic information. Capabilities for the geographic analysis of spatial data were either poor or lacking in these early systems (Fotheringham et al., 1994). Revolution in spatial data representation and statistical methodology provide the opportunity for spatial data analysis and GIS coming into contact (Goodchild et al., 2004). Nowadays, at the very heart of GIS technology, spatial analysis is addressing the problems ranging from computational analysis of geographic patterns to finding optimum routes, site selection, and advanced predictive modelling. GIS software is a powerful platform for integrating the databases and hosting the different computational models (Ahola et al., 2007; Church et al., 2000; Taylor et al., 2010). It provides numerous means to process and visualize spatial data. The input for a GIS platform can be a single or combination of data on physical characteristics, demographics and land use, infrastructure and facilities. The quality of output information is largely dependent on the comprehensiveness and accuracy of input datasets.

This paper takes a first step in a spatial analysis approach to assist evacuation management especially in relation to the routing and shelter assignment process. It aims to address three main problems: (1) what is the nearest achievable shelter for a given community in the inundation zone during a flood evacuation? And what are the corresponding detailed instructions for evacuating by car? (2) What is the possible congestion condition given residential locations? (3) How do the existing shelters serve the evacuees? And what is the potential location for new shelters to achieve better supply coverage? To respond to these questions, a case study in 2011 Brisbane River flood scenario is examined. There are 1,186 mesh blocks along the river with a population of around 100,000. This case will be examined through two open-source extensions in ArcGIS, which assist to analyse the situation with real network distances and floodwater impacts on such an inundation emergency.

This paper describes a spatial analysis method for evacuation management in Brisbane River inundation scenario using GIS software. The first section of the paper provides an introduction of floods in Australia, followed by the background of shelter assignment. The next section presents a method for dataset collection and for decision-making in shelter assignment and evacuation routing, including a case study for Brisbane flood. The third section provides an analysis of the results. The paper concludes with a discussion of the applicable scenarios and limitations of this approach.

## **2 Floods in Australia**

Floods occur when water covers land which is normally dry. There are two categories of floods in Australia: localised flash flooding as a result of thunderstorms and more widespread flooding following heavy rain over the catchment areas of river systems. Seasonal flooding occurs in Northern Australia regularly. Since 1990, historic major flood events happened in the states of Queensland, New South Wales, Victoria and Tasmania. Recent floods with high magnitude recurred in Queensland (2011), Victoria (2011) and Queensland and New South Wales (2013). Among these three floods, the 2011 Queensland flooding caused the severest damage; it affected over 200,000 people across the state and led to an estimated reduction in Australia's gross domestic product (GDP) at about \$30 billion.

Town councils and shires have started mapping the 100-year flood areas. However, they still leave the evacuation-related options to residents. People are assumed to know their vulnerability and take the responsibility for their own information acquisition and evacuation preparation (Astill et al., 2014; Bohensky et al., 2014). Therefore, in case residents do not seek evacuation information actively, greater operations from local governments to assist communities exposed to flood threats are required.

Official shelters are places (e.g. showgrounds, churches, clubs...) that are authorized to provide accommodation in an emergency. In a flood scenario, shelters can be located near/in the inundation boundary; yet they have higher elevations to avoid damage. In an inundation emergency, sheltering in vertical evacuation buildings is more feasible and efficient than escaping the whole area during an evacuation (Mas et al., 2013). One main objective of evacuation planning is to match the urgent needs with appropriate resources in the most efficient and timely manner (Alexander, 2005). Shelter assignment in this context refers to giving instructions for evacuees on which shelter to choose, and detailed routes they should take to reach destinations. Assigning evacuees with appropriate shelters is one of the three most discussed topics lying in the approach of sheltering as an evacuation strategy; the other two are evaluating capacity of existing shelters and proposing locations for new shelters (Lämmel, 2011). In Australia, emergency management places reliance on individuals to get alerted and informed before they take self-help approach to protect themselves against risks from natural hazards (Astill et al., 2014; Bohensky et al., 2014; EMA, 2004). However, it might not be efficient to rely on people's initiatives to be aware of their situations, and choose their own shelters and routes in the evacuation process; as this could potentially result in overcrowded shelters and/or severe traffic disruptions. In order to provide residents with information on where and how to escape, shelter assignment strategies need to be fixed early during the evacuation or even before the disasters' occurrences, especially for the scenarios in which the dangerous area can be largely defined according to previous events (e.g. floods).

### 3 Methodology

#### 3.1 Case study in Brisbane

Brisbane is the capital city of Queensland; it locates in south east of the state and has a population of around 2,100,000 inhabitants. Brisbane River, flowing across the city from west to east, is the longest river in Queensland. The river catchment has an area of around 13,570 km<sup>2</sup>, with the Great Dividing Range as the western boundary and various smaller coastal ranges to the north. Most part of the catchment is covered by forestry and grazing land while Brisbane and Ipswich metropolitan areas are also within the range. Flood inundation in this river has been observed since 1823, followed by extensive floods happening during the years before 1990. The second largest flood since the 20th century occurred in 2011 (Figure 1).

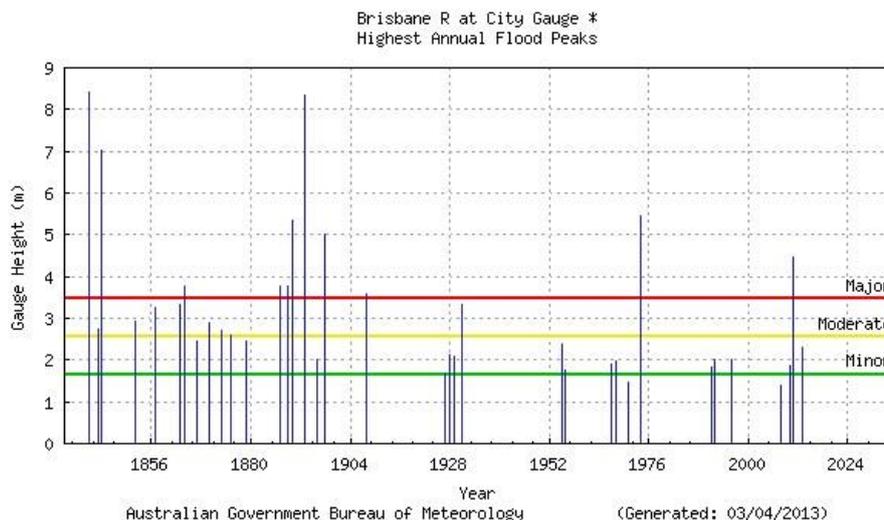


Figure 1: Floods in Brisbane (BoM, 2013)

During the 2011 flood event, Brisbane city experienced a major flood (defined as having a gauge height of 3.5 m or higher) from 10:00 am on 12th January until 6:00 pm on 13th January, accounting for a period of 32 hours. The flood peaked at 5:00 pm on 12th and again at 3:00 am on 13th with a gauge height of 4.25 m and 4.46 m respectively. As a result, in metropolitan Brisbane, over 15,000 properties were inundated and approximately 3,600 households evacuated (van den Honert et al., 2011).

Flood lines which describe the flood extent in years 1974 and 2011 are available from Queensland government. A comparison of these two flood boundaries shows that they are very similar in most areas of Brisbane city (van

den Honert et al., 2011). Therefore, although the Annual Return Interval of flooding is uncertain, the boundary of inundation is largely predictable. In this paper, a case study is conducted. The study area covers the majority of Brisbane (north, west, inner city, south and part of Brisbane east), and the affected area is based on 2011 flood lines (Figure 2).

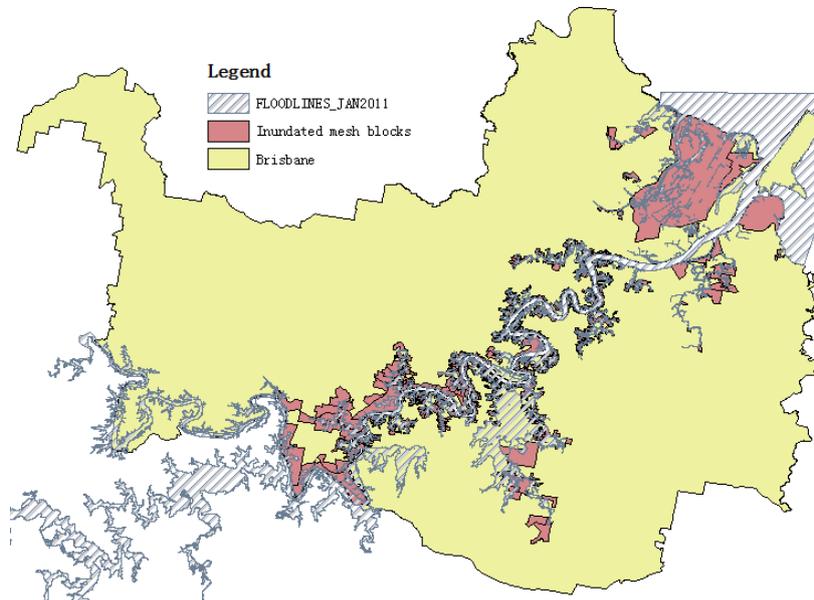


Figure 2: 2011 Brisbane flood lines and affected mesh blocks

### 3.2 Dataset collection and scenario assumption

Datasets (Table 1) have been collected and pre-processed to conduct a simple analysis to generate shelter assignment and routing instructions. Flood extent of 2011 event is downloaded from Queensland Government public database. Mesh block (the smallest census unit) boundaries are acquired from Australian Bureau of Statistics (ABS). Road networks covering the whole Brisbane city can be exported from open data source such as Open Street Map. In particular, five types of roads: trunk, motorway, primary road, secondary road and tertiary are extracted because of people’s preference in choosing familiar, wider streets for evacuation (Tomsen et al., 2014). One evacuation centre can be identified from a news report on 11th January 2011. On 12th January, another news report, with the title “More evacuation centres set up”, delivered the message of setting up five more evacuation centres. Therefore, altogether six shelters are considered in this analysis.

Table 1: Datasets participating in the analysis and their sources

Datasets	Sources
2011 QLD floodlines	Queensland (QLD) Government
Road network for Brisbane city	Open Street Map (OSM)
Official shelters	News report during the event
2011 QLD mesh blocks extent	Australian Bureau of Statistics (ABS)
2011 Australian mesh blocks census counts	Australian Bureau of Statistics (ABS)

Flood lines and mesh block files are integrated to generate locations that are affected by floodwater. There were 100,649 people residing in 1,186 affected mesh blocks, which were confirmed by the news report that “100,000 customers are expected to lose electricity” on 12th January, 2011 (first flood peak). Apparently, six shelters cannot accommodate all those people for evacuation. In reality, most residents experienced low flood inundation and they chose stay-in-place until the water decreased; a good portion of evacuees sought accommodation from family or friends; others turned to official evacuation buildings. Evacuees were approximately 3,600 households in metropolitan Brisbane, which is expected to be more than 8,000 people according to the average household size (namely 2.3 persons per household) in 2011 census data for Brisbane inner city. However, their distribution can be across the 1,186 affected mesh blocks; therefore, shelter assignment analysis within this area is performed.

Prior to developing the shelter assignment strategy for this study, several assumptions had to be made for the emergency scenarios. Firstly, it was assumed that evacuees are distributed across all the inundation areas. Secondly,

the whole routing strategy was car-based, with an evacuation speed of 30 km/h (Mas et al., 2013). Thirdly, each evacuee moves to its closest shelter, using the shortest possible route on the road network (this is a feasible choice in a small community well-known by its inhabitants). Lastly, we chose 30 minutes as the maximum time that evacuees would be willing to travel to shelters by car.

### 3.3 Analysis approach in GIS software

This case study is conducted under two criteria. The first criterion is that the approach is based on real network distances in place of straight-line distances. Therefore, more realistic travel time is captured. The second criterion is that the bridges across Brisbane River are assumed to be unusable or too risky to use during an evacuation; shortest paths passing these bridges are not accounted as solutions.

This analysis is conducted using Network Analyst built in ArcGIS and Urban Network Analysis toolbox developed by the City Form Lab (Sevtsuk et al., 2012; Sevtsuk et al., 2013). Network Analyst extension aims at creating, editing and analysing network datasets. It provides eight solvers to address routing related problems. Four solvers are used in this analysis: route, nearest facility, origin-destination (OD) cost matrix and service area. Service area solver returns polygons which show the service areas generated according to specified search radius or given cost limits. It stores the geometry of the road network into a triangulated irregular network (TIN) data structure. The distance along the road line serves as the height of the locations inside the TIN if the road is usable; otherwise the height is assigned a much larger value. Dijkstra's algorithm (Dijkstra, 1959) is applied to calculate the shortest network distance. The service area polygons are formed by carving out regions covering areas in between the specified break values. Optionally, an origin-destination cost matrix for evacuation from the resident locations to each shelter can be calculated. The results of this matrix can be used to identify residential areas that will be serviced by each shelter within a given drive time.

Urban Network Analysis toolbox contains two sets of tools: centrality tool and redundancy tool. The centrality tool is designed for studying the spatial configurations of cities, and their related social, economic, and environmental processes (Sevtsuk et al., 2013). While the redundancy tool calculates second shortest paths, and the redundant value is to be set by the users. Both of the nearest facility solver in the Network Analyst toolbox and the centrality tool in the Urban Network Analysis toolbox can offer solution for searching the nearest buildings within a given distance. The later one allows weights to be assigned to the destinations, which changes destinations' attractiveness to the origins. However, in this analysis, the weight is difficult to quantify because there is a lack of data describing the capacity and condition of shelters. The centrality tool also offers a "Betweenness" solver. The "Betweenness" of a building is defined as the fraction of shortest paths between pairs of other buildings in the network that pass by building  $i$  (Freeman, 1979); the formula is as follows:

$$Betweenness[i]^r = \sum_{j,k \in G - \{i\}, d[j,k] \leq r} \frac{n_{jk}[i]}{n_{jk}} * w[j]$$

Where

- $r$  is the search radius,  $G$  is the graph,  $d[j, k]$  is the distance from node  $j$  to node  $k$
- $n_{jk}$  is the number of the shortest path from node  $j$  to node  $k$
- $n_{jk}[i]$  is the number of the shortest path from node  $j$  to node  $k$  and pass by node  $i$
- $w[j]$  is the weight of each origin  $j$

If the  $w[j]$  is assigned by demographics for example population, the "Betweenness" can estimate the potential of passersby at different buildings on the network.

## 4 Results

The main target for an evacuation strategy is to specify a certain shelter and detailed route for each mesh block. Figure 3(a) shows one example of how residents from a given mesh block (black dot) can evacuate to the nearest shelter (green dot) following the highlighted path in the graph. Additionally, the criterion of ruling bridges out of the analysis is illustrated as well because the farther shelter was chosen (Site 4) instead of taking the risk of approaching nearer shelters (Sites 2 and 3) across Brisbane River. From Figure 3(b), the evacuees can understand exactly where to turn and which road to choose until they arrive at the shelter.

Based on the assumption that people take the shortest path to evacuate, the possible congested segments can be predicted by calculating "Betweenness" value for each mesh block. The "Betweenness" method estimates how many times each mesh block is passed by evacuees on their way from their locations to shelters along the shortest paths. Figure 4 illustrates the passer-by frequency of each mesh block when residents choose shortest paths for an evacuation with the distance less than 15 km. This can help point out the possible congested road segments. People will need to try avoiding those areas and choose their second shortest routes as alternatives. From the legends in Figure 4, it is clear that the major and medium crowded locations are gathering around the middle part of Brisbane. In particular, these places are either neighbourhoods that are surrounded by Brisbane River on two/three sides or

mesh blocks that have dense population (e.g. central business district). Furthermore, from the busiest points, one can tell that Sites 2, 3 and 4 will be expected for more accommodation than Sites 1, 5 and 6. The congestion condition can be estimated at any location where an observation point is established (in this analysis, the locations of 1,186 mesh blocks).

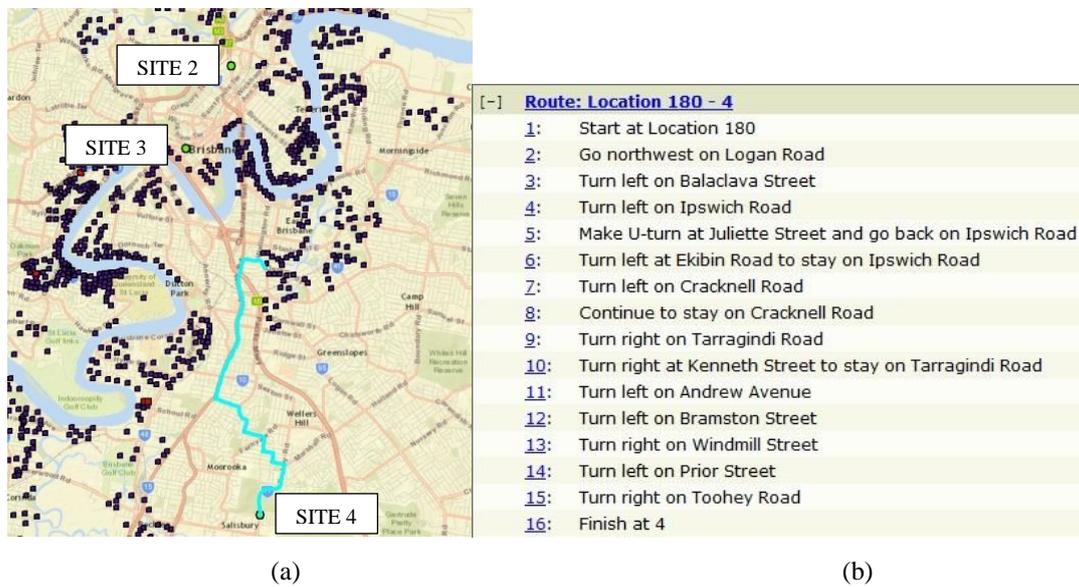


Figure 3: Evacuation route (a) and instructions (b)

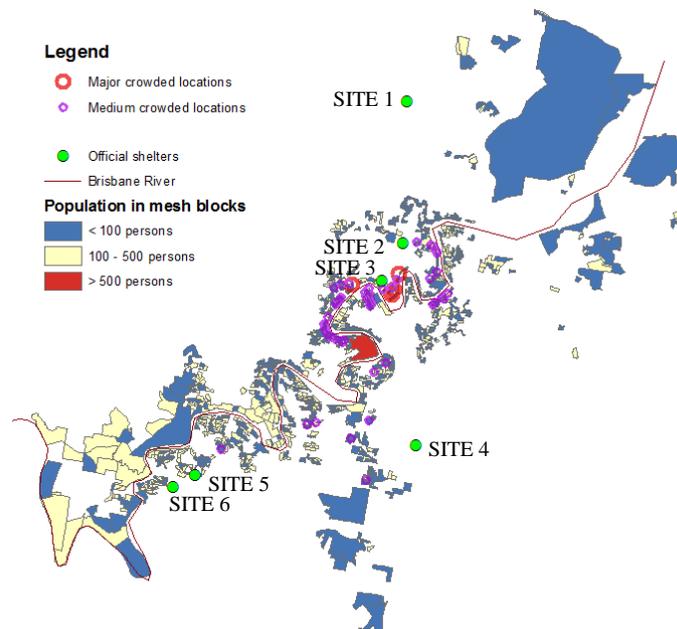


Figure 4: Locations congestion may occur

From the perspective of shelter management, the service areas for each shelter need to be defined and evaluated. In Figure 5, three polygons surrounding each of the six evacuation buildings represent service areas within three different ranges. Mesh blocks located in each of the three polygons from inside out can reach the corresponding shelter during an evacuation in 10 min, 20 min and 30 min respectively. These polygons are calculated based on two criteria: distance along available roads and the exclusion of flooded bridges in the analysis (as they are considered to be risky in an inundation emergency). We notice that some service areas illustrated in the graph do cross a bridge; this is because the algorithm for generating polygons will connect sequential points to form boundary lines, ignoring all the details between adjacent points. As 30 minutes are considered to be the longest time people accept when evacuating, any mesh block that cannot be covered by shelters within 30-minute driving

distance reveal problems in the distribution or/and number of shelters. There are 190 mesh blocks locating in the eastern and western parts of Brisbane which have no shelter accommodation in that sense. Although this number can be exaggerative because of the imprecise presence of road networks (especially in terms of segment connectivity), the analysis results imply that more shelters should be established in the east and west of Brisbane in the future.

Table 2 shows that Sites 2 and 3 provide shelter services to the largest number of mesh blocks. Most of the mesh blocks assigned to these two shelters are located within 10 km. Site 4 will accommodate the second largest number of mesh blocks; the majority of the evacuees escaping to this shelter will travel more than 10 km. Sites 5 and 6 have similar total number of mesh blocks to serve; the reason could be that they are close to each other. The service area of Site 1 only covers 5 mesh blocks, which indicates that it is far for evacuating by car and it may aim to accommodate residents rescued by helicopters.

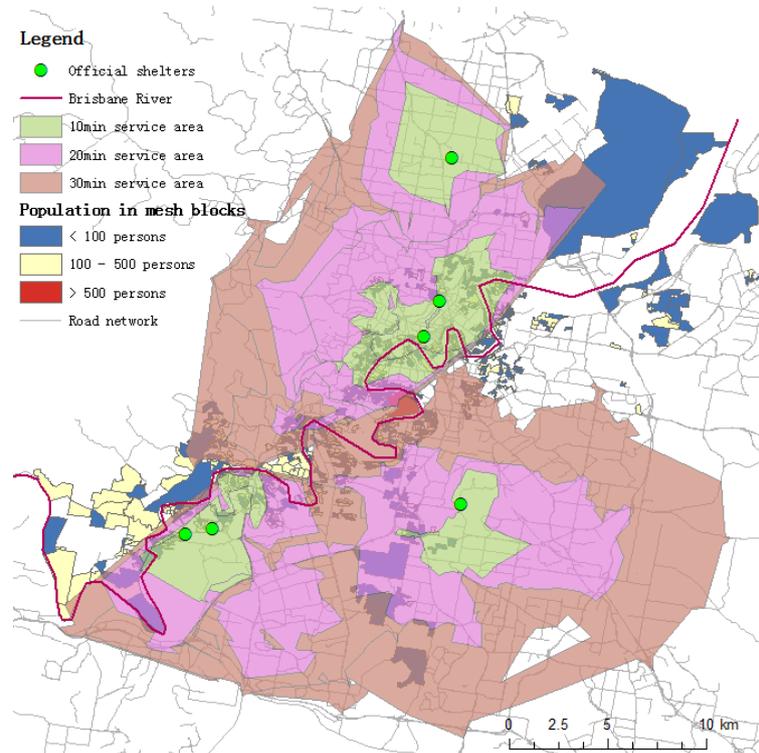


Figure 5: Shelter service areas

Table 2: Number of mesh blocks each shelter covered

	Name	Number of mesh blocks covered			
		Total	Within 0 - 5 km	Within 5 - 10 km	Within 10 - 15 km
Site 1	Kedron Wavell Services Club	5	0	5	0
Site 2	RNA showgrounds at Bowen Hills	314	149	145	20
Site 3	Salvation Army	328	184	120	24
Site 4	Nathan	249	10	58	181
Site 5	Good News Lutheran Church	130	99	1	30
Site 6	St Catherine Anglican Church	106	60	40	6

## 5 Discussion and limitations

The results in the previous section indicate that spatial analysis is an effective tool for evacuation management, especially for shelter assignment and routing. The results show that spatial analysis can provide evacuees with detailed information on shelter selection and possible routes. Congestion conditions at the mesh blocks' locations can also be predicted on the basis of individual choice of shortest paths for evacuation. Furthermore, evaluation on existing shelters and locations is performed in order to establish new shelters. This study has shown the capability

and effectiveness of spatial analysis to address evacuation management problems which need to be accurately resolved by emergency managers.

The applicable scenario mainly involves large-scale evacuation management under flood disasters. Evacuation strategies for the residents exposed to floods are developed. This study focuses on car-based evacuation which requires that the road network is well developed so that the approach applies especially when the study area includes or is close to a central business district. Additionally, a dense population is considered. In a case where only a few people are surrounded by an undeveloped road network, the most suitable solution must be helicopter rescuing. Lastly, the applicability of this study relies on evacuees' familiarity with official shelters.

The results from the Network Analyst toolbox are sensitive to the relative location of residents and nearby road segments. For example, only the mesh block that is reasonably near to the network is considered to be reachable. Furthermore, a service area is illustrated as a polygon which connects adjacent points at the cut-off location along a road segment. This may lead to an unnecessary misunderstanding that some service areas include unreachable segments between adjacent roads. Tools in Urban Network Analysis do not support barriers. Constraints on bridges need to be set up manually in the analysis.

Nevertheless, the analysis tools are beneficial for emergency managers to gain a better understanding of the whole picture of large-scale evacuation and provide a feasible and organised way to manage the evacuation situation. Improvement of the analysis results requires further refinement of the algorithm and models used in the two toolboxes and integration with other useful modules such as the flood planning module and the real time traffic module built in GIS software.

## 6 Concluding Remarks

This paper presents a first step in the flood evacuation management analysis of Brisbane River in a large-scale scenario. The evacuation boundaries are defined based on two major historic flood events. This case study utilised the spatial analysis approach. The results show that the shelters established during 2011 Brisbane flood evacuation are inadequate and new shelter locations are proposed in case of future flood events. The proposed approach is feasible and suitable for emergency managers who do not necessarily have expertise in GIS. Limitations of the conventional GIS toolboxes used in this study are discussed. Further study can be conducted to utilise other analysis modules which manage the floodwater as well as the real traffic condition because the real evacuation situation demands for such temporal information, hence a dynamic analysis is required.

## References

- Ahola, T., Virrantaus, K., Krisp, J. M., & Hunter, G. J. (2007). A spatio-temporal population model to support risk assessment and damage analysis for decision-making. *International Journal of Geographical Information Science*, 21 (8), 935-953.
- Alexander, D. (1991). Natural disasters: a framework for research and teaching. *Disasters*, 15 (3), 209-226.
- Alexander, D. (2005). Towards the development of a standard in emergency planning. *Disaster Prevention and Management*, 14 (2), 158-175.
- Astill, S., & Griggs, P. (2014). Investigating the hazard preparatory information-seeking habits of far north Queensland coastal communities. *Australian Journal of Emergency Management*, 29 (3), 37-45.
- Bohensky, E. L., & Leitch, A. M. (2014). Framing the flood: a media analysis of themes of resilience in the 2011 Brisbane flood. *Regional Environmental Change*, 14 (2), 475-488.
- BoM. (2013). Known Floods in the Brisbane & Bremer River Basin, including the Cities of Brisbane and Ipswich. Retrieved 25th November 2014, from [http://www.bom.gov.au/qld/flood/fld\\_history/brisbane\\_history.shtml](http://www.bom.gov.au/qld/flood/fld_history/brisbane_history.shtml)
- Carter, R. A. (2012). Flood risk, insurance and emergency management in Australia. *The Australian Journal of Emergency Management*, 27 (2), 20-25.
- Church, R. L., & Cova, T. J. (2000). Mapping evacuation risk on transportation networks using a spatial optimization model. *Transportation Research Part C: Emerging Technologies*, 8 (1), 321-336.
- Crompton, R., & McAneney, J. (2008). The cost of natural disasters in Australia: the case for disaster risk reduction. *The Australian Journal of Emergency Management*, 23 (4), 43-46.
- Dall'Osso, F., & Dominey-Howes, D. (2010). Public assessment of the usefulness of "draft" tsunami evacuation maps from Sydney, Australia-implications for the establishment of formal evacuation plans. *Natural Hazards and Earth System Science*, 10 (8), 1739-1750.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische mathematik*, 1 (1), 269-271.

- EMA. (2004). *Emergency Management in Australia: Concept and Principles*. Canberra: Emergency Management Australia.
- Fotheringham, S., & Rogerson, P. (1994). *Spatial analysis and GIS*: CRC Press.
- Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social networks*, 1 (3), 215-239.
- Goodchild, M. F., & Haining, R. P. (2004). GIS and spatial data analysis: Converging perspectives. *Papers in Regional Science*, 83 (1), 363-385.
- Goodchild, M. F., & Longley, P. (1999). The future of GIS and spatial analysis. *Geographical information systems*, 1, 567-580.
- Gouldby, B., Sayers, P., Mulet-Marti, J., Hassan, M., & Benwell, D. (2008). A methodology for regional-scale flood risk assessment. *Proceedings of the ICE-Water Management*, 161 (3), 169-182.
- Johnstone, W., & Lence, B. (2009). Assessing the value of mitigation strategies in reducing the impacts of rapid-onset, catastrophic floods. *Journal of Flood Risk Management*, 2 (3), 209-221.
- Lämmel, G. (2011). *Escaping the Tsunami: Evacuation Strategies for Large Urban Areas Concepts and Implementation of a Multi-Agent Based Approach*. Technischen University Berlin.
- Mas, E., Adriano, B., & Koshimura, S. (2013). An integrated simulation of tsunami hazard and human evacuation in La Punta, Peru. *Journal of Disaster Research*, 8 (2), 285-295.
- Middelmann, M. H. (2007). *NATURAL HAZARDS in AUSTRALIA, identify Risk Analysis Requirements*. Geoscience Australia Retrieved from [http://www.ga.gov.au/corporate\\_data/65444/65444.pdf](http://www.ga.gov.au/corporate_data/65444/65444.pdf).
- Pfister, N. (2002). Community response to flood warnings: the case of an evacuation from Grafton, March 2001. *Australian Journal of Emergency Management*, 17 (2), 19-30.
- Sevtsuk, A., & Mekonnen, M. (2012). Urban network analysis. A new toolbox for ArcGIS. *Revue Internationale de Géomatique*, 22 (2), 287-305.
- Sevtsuk, A., Mekonnen, M., & Kalvo, R. (2013). *Urban Network Analysis toolbox for ArcGIS 10/10.1/10.2*: Singapore University of Technology & Design in collaboration with MIT.
- Taylor, M. A. P., & Freeman, S. K. (2010). A review of planning and operational models used for emergency evacuation situations in Australia. *Procedia Engineering*, 3, 3-14.
- Thywissen, K. (2006). *Core terminology of disaster reduction. Measuring vulnerability to natural hazards: Towards disaster resilient societies*, United Nations University Press, Hong Kong.
- Tomsen, E., Lindsay, J. M., Gahegan, M., Wilson, T. M., & Blake, D. M. (2014). Evacuation planning in the Auckland Volcanic Field, New Zealand: a spatio-temporal approach for emergency management and transportation network decisions. *Journal of Applied Volcanology*, 3 (1), 6-27.
- van den Honert, R. C., & McAneney, J. (2011). The 2011 Brisbane Floods: Causes, Impacts and Implications. *Water*, 3 (4), 1149-1173.
- Van Der Knijff, J. M., Younis, J., & De Roo, A. P. J. (2008). LISFLOOD: a GIS- based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, 24 (2), 189-212.
- Vogt, B. M., & Sorensen, J. H. (1992). *Evacuation research: A reassessment*: Oak Ridge National Lab., TN (United States).

# A Data Model for Integrating GIS and BIM for Assessment and 3D Visualisation of Flood Damage to Building

Sam Amirebrahimi    Abbas Rajabifard    Priyan Mendis    Tuan Ngo  
amis@unimelb.edu.au    abbas.r@unimelb.edu.au    pmendis@unimelb.edu.au    dtngo@unimelb.edu.au

Department of Infrastructure Engineering, The University of Melbourne, VIC 3010

## Abstract

Flood Damage Assessment (FDA) is a key component in modern risk management frameworks providing an effective basis for decision making and the treatment of the risks. Current FDA methods do not consider the distinctiveness of buildings in analysis and therefore, cannot analyse them on a case-by-case basis, which is necessary for a variety of applications like engineering and design evaluation. This is mainly due to the limited input data used in these methods. The information required for such micro-level FDA analysis includes on one hand, complete building information (well-represented in BIM) and on the other hand, flood information that is commonly managed by GIS. While the independent use of BIM and GIS cannot satisfy all the information requirements for detailed FDA, their integration can potentially be used for this purpose. However, existing integration methods are application-specific and their adoption for FDA is challenging. This paper presents a method for BIM-GIS integration to support the requirements of a detailed assessment and 3D visualisation of flood damage to buildings. The data modelling cycle was used to design a new data model as a profile of GML for this purpose. The model was evaluated using a case study and found effective to satisfy the required criteria for micro-level FDA.

## 1 Introduction

It has been recognised in the past decade that a management of flood risks which focuses solely on the hazard is not effective (Birkmann et al., 2013). Therefore, risk-based flood management commonly employs Flood Damage Assessment (FDA) to evaluate the potential consequences of flood in the identification of risks and decisions for their treatment (Thieken et al., 2005).

Efforts in the field of FDA have resulted in development of a variety of FDA methods to serve different applications at different spatial scales (i.e. Macro, Meso and Micro) such as risk mapping, vulnerability assessment or financial appraisal of damage. Within this context, a strong emphasis is made on buildings due to their significant economic importance (Messner et al., 2007). The internationally accepted standard method for FDA is the use of "damage curves" (Merz et al., 2010). Damage in this method is calculated using generalisation of buildings into classes (based on a few of their general characteristics like construction type or age) and applying a predefined curve (or function) to relate the damage level commonly to a single flood parameter, the inundation depth. While

---

*Copyright © by the paper's authors. Copying permitted only for private and academic purposes.*

the effectiveness of the existing FDA approaches is confirmed for large scale applications investigating a large population of buildings (Messner et al., 2007), they are however found unfit for micro level applications (e.g. building design evaluation against flood) where a case-by-case assessment of the flood impacts on individual buildings is required (Pistrika and Jonkman, 2010). This is mainly because of the inability of these methods for capturing and use of a complete representation of the building in the assessment, due to their limited input data (Merz et al., 2010). In this way, uniqueness and distinct behaviour of building against the flood is not included in the analysis of damage. In addition, the outputs provided by these methods are limited to a single number for the overall building damage cost or only an indication of whether the building collapses or not for certain level of water and velocity. Other than these, no further information about the details and the location of damage at the building level is provisioned. Such details are important to reveal the sources of risk for a building for their treatment.

An effective case-by-case analysis of damage to a building at micro level requires the use of two sets of information: the flood parameters (e.g. depth and velocity) causing damage to the building components; and the building components that resist against the flood impacts and are unique to each building (Pistrika and Jonkman, 2010). Geographic Information Systems (GIS) are the key information management tools for the first set with a long history of use in this area. The major strength of GIS is on representation of outdoor and large scale features. However, even with the considerations of the recent efforts for 3D modelling of the details of the buildings in 3D city models (e.g. CityGML), the geospatial domain is still found inadequate for a complete representation of the semantic and geometric aspects of the building and therefore limited for the second set (Zhang et al., 2009). In contrast to GIS, Building Information Models (BIM) from the construction industry provides a comprehensive digital repository of building information. While the strength of BIM lies in the complete representation of every aspect of a building (e.g. geometry, materials and connections), it is restricted to small scale at a building level and currently cannot store geographically extended features like flood.

The independent adoption of BIM or GIS for a micro level FDA would be inadequate to support the information requirements for this application. However, their combination provides the potential to accommodate these requirements under one umbrella to serve the analysis and presentation of the damage to the building (Isikdag and Zlatanova, 2009a). For the overall benefit of integrating BIM with GIS for different applications, numerous efforts were made towards its realisation. These works however, are limited to either (a) addressing the underlying technical integration challenges by one or two-directional geometry or semantic conversion between the two at building level or (b) serving a specific use case (e.g. indoor-outdoor utility management or emergency response) with focus of the integration on the minimum requirements to satisfy the needs of that application. The specific focus of these works limits their adoptability and use for FDA and therefore there is a need for a solution for integration of BIM with GIS for this purpose.

In this paper, a method for combining the BIM and GIS at the data level is proposed to serve the information needs for a micro level assessment and 3D visualisation of potential building damage from floods. It is hypothesised that the use of this method can facilitate an information system to harness the combined strength of the BIM and 3D GIS, allowing for a case-by-case analysis of the building damage by the implementation of certain analysis types at a level of detail that the existing FDA methods are not sensitive towards. The outputs produced by such analyses can facilitate decision making by a range of stakeholders such as engineering and design firms, councils and insurance companies for improving the resilience of the community towards floods and their adverse impacts.

In the remainder of this paper, first, the research background as well as an overview of the related work for the integration of BIM with GIS is provided. Next, the design of the proposed integration method and its details are presented. Further, a demonstration of the integration using a case study is illustrated and the results are discussed. Finally, the paper is concluded and the future research directions are proposed.

## **2 Background and Related Work**

### **2.1 Building Information Model**

BIM is a rich and intelligent digital repository of building information and uses an Object Oriented (OO) approach to describe the characteristics (semantics and geometry) and behaviour of each building element as well as its relationships with others (Eastman et al., 2011). BIM uses Industry Foundation Classes (IFC) as its open standard to establish interoperability in the construction industry. Within the IFC framework, building, its components as well as other relevant construction industry data are described within a single information model enabling the management of such information throughout the lifecycle of a building/facility. Although BIM has a variety of applications and has been previously used for damage assessment for other hazards like fire and earthquake (e.g. Christodoulou et al., 2010), yet, it has not been considered for the case of FDA.

### **2.2 Geographic Information System (GIS)**

GIS, on the other hand, is a platform for managing and presenting spatially referenced information. Within this domain, the exchange of geospatial data and the interoperability between systems are established using the Geographic Markup Language (GML). GML is an Open Geospatial Consortium (OGC) standard data model for defining the data types and constructs for describing the geographic features. With a more specific focus, the

heterogeneous geospatial information about urban data (e.g. buildings, transport, vegetation and water bodies) at different levels of details is integrated within the framework of the 3D virtual city models such as CityGML (Dollner et al., 2006). CityGML is the most comprehensive urban information model within the geospatial domain to digitally represent a city in 3D. The building information in CityGML (and other existing GIS formats) however, is not as complete and mature as BIM and for this reason, multiple extensions (e.g. utility network in Hijazi et al., 2010) have been created separately over the past years to improve the model. CityGML and some of its extensions have considered flood depth (as a water body), however, they were solely used for 3D visualisation of the flood (Schulte and Coors, 2009) and did not include its other parameters (e.g. velocity) and temporal dynamics. These, in addition to the other limitations of CityGML to represent the building details, prevent its effective use for serving the requirements of the micro-level FDA.

### 2.3 BIM-GIS Integration

GIS and BIM originate from different domains and were developed for the specific needs of that field. Their integration creates a seamless and scale-independent view of the world across both domains. It can benefit a variety of applications that meeting their requirements would not be possible by independent use of BIM or GIS (Karimi and Akinci, 2010). This integration however is not simple due to the differences between the two. Such dissimilarities are discussed in terms of spatial scale, level of granularity, geometry representation methods, storage and access methods as well as semantic mismatches between them (Isikdag and Zlatanova, 2009b; Karimi and Akinci, 2010; El-Mekawy and Ostman, 2010).

Various attempts have been made for integrating BIM with GIS that can generally be classified into three groups: at *application*, *process*, and *data* levels.

At the application level, the integration methods use *reconfiguring* or *rebuilding* (Karimi and Akinci, 2010) where an existing GIS or BIM tool is either modified by software patches or is rebuilt from scratch to include the functions of the other. This method is generally costly and inflexible. On the other hand, process level integration methods like OWS-4 project by OGC (2007) use Service Oriented Architecture (SOA) to allow the participation of BIM and GIS systems in those tasks that require the capabilities of both while they simultaneously remain live and distinct. This method provides more flexibility than the first group. However, in this method, the challenges of integration are still to be resolved at the underlying data level to provide interoperability between these systems.

There are a variety of methods developed to integrate BIM with GIS at the data level. *Linking* methods such as ESRI ArcSDE facilitates data transfer between BIM and GIS software by an Application Programming Interface (API) at either side. *Translation/Conversion* methods such as FME (Safe, 2013) and the work by Nagel et al. (2009), on the other hand, were introduced to directly convert between GIS and BIM formats. This method commonly translates the data between IFC and CityGML. Loss of semantics, limitations in geometric conversion and sole focus on the major building elements and neglecting the other aspects (e.g. utilities or connections) are some of the concerns associated with these methods. To resolve the geometry transformation problem, a number of research (e.g. Li et al., 2006; Wu et al., 2010) was conducted which only partially addresses the overall integration.

In the more comprehensive and flexible integration methods, either a new data model as a "meta model" is developed to mediate between BIM and geospatial information at higher level; or a data model at GIS or BIM side (e.g. CityGML or IFC) is extended to incorporate the data from the other. The GeoBIM and utility network extensions by Van Berlo and Laat (2010) and Hijazi et al. (2010) are examples for CityGML. On the other hand, IFC-for-GIS project (IAI, 2005) intended to extend the IFC model to include information from geospatial world. A prominent example of meta models is the "Unified Building Model" (El-Mekawy et al., 2011) that the focus of the integration was to develop an intermediate data model for building for emergency evacuation purposes. In general, these models are application-focused and the integration is made for a particular use case with specific requirements. Therefore, the included concepts and relationships within these models may not suit other applications (with different functional requirements).

In this paper, an integration method for integration of BIM and GIS for micro-level assessment of flood damage on building is presented. It supplies the resistance parameters of the building at high level of details, flood information as well as the other geographically extended features (e.g. elevation) in one unified data model to use for assessment and 3D visualisation of the flood damage to building.

### 3 BIM-GIS Integration for FDA on building

The integration of BIM and GIS in this work is proposed at the data layer and implemented on the GIS side for utilisation of its spatial analysis tools for FDA calculations. The GML standard does not explicitly define the semantics of the geographic features (e.g. buildings or roads), and as discussed previously, there are limitations in representation of building and its components in the CityGML. For this reason, a new data model as a profile of GML is proposed to integrate the BIM information alongside the spatiotemporal dynamics of the flood information for assessing and visualisation of the damage on the building.

### 3.1 Methodology

The proposed data model in this paper was developed based on the "data modelling cycle" (Teorey et al., 2011). This methodology includes five steps for design and implementation of a data model. It commonly starts with the mapping of the real world concepts and their relationships to a conceptual model. The concepts and relationships in the model are identified for a particular or a number of use cases by employing a variety of data requirement gathering methods like survey, interview, review of relevant previous publications, etc. The conceptual data model is further translated to the logical data model defining the structure of the database. The last step in the process involves the development of a physical data model and its implementation (Elmasri and Navathe, 2011).

Next, the details of the above process for designing the proposed data model is explained, the outputs are presented and its benefits for use for a micro-level FDA are discussed.

### 3.2 Use case definition

Use cases are informal scenarios describing the expected behaviour of a system as a response to the needs of the stakeholder(s). Use cases are the basis of the extraction of the required functionalities and the data needs for system design.

In this study, the defined use case is related to the assessment and visualisation of the potential damage to a building from a riverine flood. It requires (i) the assessment of the building safety (structural stability) by assessing the damage to structural and load bearing components, (ii) the estimation of the total cost of repair/replacement of the damaged components (excluding its content), and (iii) the visualisation of the location and mode of damage at component level. This use case was detailed by investigating the land development process, a review of previous publications and liaising with engineering and design firms, councils as well as referral authorities (e.g. Melbourne Water) as potential users of the system outputs. The behaviour of buildings of different types may vary in flood situations and the processes for damage analysis on them may be different from one to another. In this research, the scope was limited to only one type which represents the most common Australian residential construction type: a single-storey slab-on-ground brick veneer house (Geoscience Australia, 2014).

### 3.3 Requirement analysis

Based on the definition of the use cases in Section 3.2, the vulnerable components of the selected house type were identified. This process involved a systematic and extensive review of the literature (e.g. HNFMSC, 2006; CLG, 2007), discussions with engineers and councils, as well as the resources and previous research on the vulnerability of houses in Australia provided by the Geoscience Australia. Our findings indicated that for the construction type in focus, the damage and the incurred costs are mainly from the impacts of floodwater contact and forces on the *floor covering, walls* (including cladding, framing, insulation and lining), *skirting boards* and *cornices, ceiling and its insulation, roof, windows, internal and external doors, eaves lining* and the *interior and exterior utilities* (e.g. *electrical*) depending on their location and materials.

Following the identification of susceptible components, further investigation was undertaken to identify and document (a) the modes of damage for each component from floodwater impacts, (b) engineering methods for their modelling and (c) the information inputs for these calculations. For the first two steps, Australian design and construction standards (e.g. AS3700 - masonry building design) and relevant documents for improving the house resilience to flood (e.g. HNFMSC, 2006; CLG, 2007) were used. Due to the limited space in here, the outputs of these two are not presented in here. Due to the diversity and large number of the extracted data requirements in step (c), only a subset of them is presented in Table 1 according to their importance. Other concepts like details of utility objects and types of different elements like coverings are omitted in this table.

Table 1: Example of the Extracted Data Requirement

Concepts	Details
Spatial Structures	Defining the spatial container for objects. It can have a corresponding element (e.g. building storey or a space in the building) that acts as the container object.
Terrain	Representing the elevation of the area. It is required to be in multiple levels of details. Terrain can be either point-based or surface-based.
Flood	The flood parameters using multiple representations: (a) Spatio-temporal point distribution of depth and velocity vectors (for use in damage calculation); (b) Surface representation of flood (e.g. water level surface).
Buildings	The footprint, address, height and the area.
Building components	Including storeys, walls, stairs, floors, foundation, beams, columns, roof, structural connections (e.g. wall ties), framing members, floorings, ceiling, soffit, skirtings and mouldings, doors, windows and cladding vents (e.g. airbricks).
Utilities	For example, electrical objects like switches, meter boxes and outlets.
Materials	Construction materials of the building elements (single material or multiple)
Cost info.	Including cost of repair/replacement of building and utility components; and the building value.

### 3.4 Data model design

According to the identified requirements, the conceptual data model illustrating the required concepts and their relationships was designed. Throughout the design process, a continuous investigation was undertaken to identify how these concepts are modelled in BIM (IFC) or GIS formats (GML and CityGML). This mapping was used to refine the design to improve interoperability and information translation between IFC or CityGML and the proposed model. The data model consists of seven packages inheriting its high level feature definitions from the GML. These packages are namely: the Core (CoreUrbanFlood), Terrain, Flood, Building, Utility, Valuation and MaterialDomain. In this paper, the UML class diagram was employed for developing and presentation of the data model. Due to the complexity of the model and limitations in space here, the attributes are not illustrated in the class diagrams. In addition, each package and their respective classes are colour-themed for the ease of read.

The "CoreUrbanFlood" package (see Figure 1) includes the necessary high-level objects for a micro-level FDA on a building. "UrbanFloodModel" concept is the highest level entity in the model defining an urban flooding scenario. It is a collection of defined materials, costs, spatial structures (explained in Section 3.3) and urban elements (\_UrbanObject) representing the urban environment in a flooding situation. Urban objects can be the parcels (site), buildings, individual utility elements (\_UtilityObject) or their aggregation as a system (UtilitySystem), flood representation (\_FloodObject), and a simple or complex elevation model (\_TerrainObject). Other urban objects like city furniture, roads, etc were beyond the scope of requirements in this research as the focus is only on damage assessment on buildings.

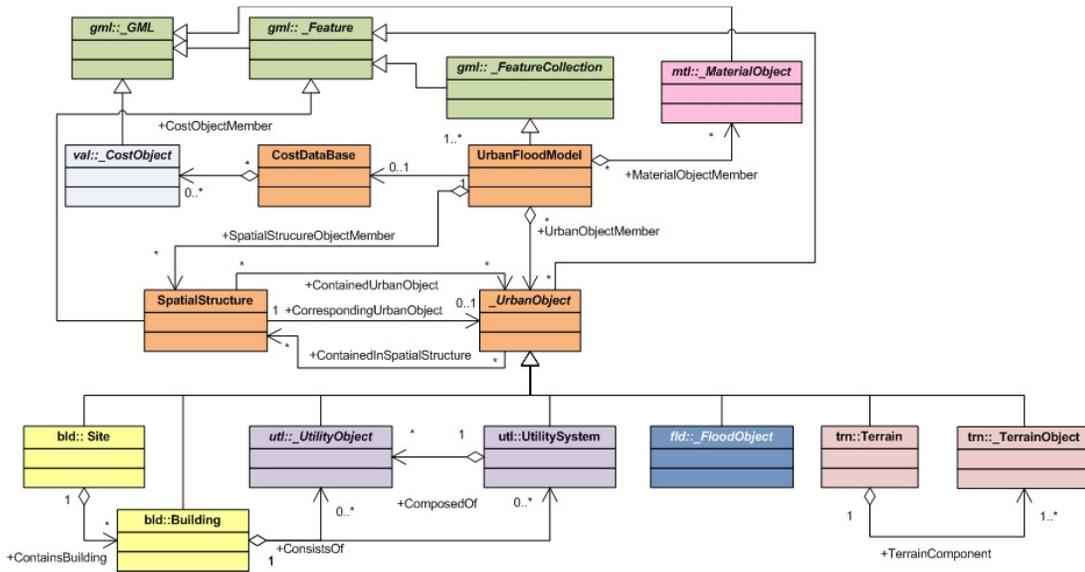


Figure 1: UML Model of the Core Package

The proposed data model adopts a subset of the "Digital Terrain Model" thematic model in CityGML 2.0 for its Terrain package (see Figure 2). An elevation object in here can be stored by an independent subtype of the abstract concept "\_TerrainObject". It can be either a surface-based object (e.g. TIN) represented by "TinTerrain" or in a multi-point form using "MassPointTerrain". On the other hand, a single Terrain can be represented by an aggregation of a number of \_TerrainObjects in different representation forms and levels of details within the "Terrain" object. Each \_TerrainObject has a validity extent - represented by a polygon, to define its effective scope.

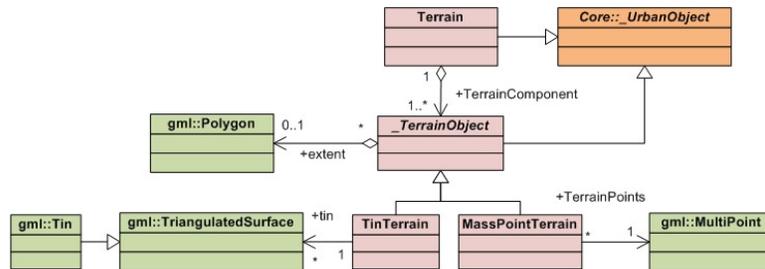


Figure 2: UML Model of the Terrain Package

The Flood package, illustrated in Figure 3, consists of the required classes to represent the flood information. Any flood in the model (`_FloodObject`) is described by its metadata (`FloodMetadata`) that provides information about its exceedance probability, flood duration, number of time steps, duration of each step and the units of measurement for depth and velocity. As a subtype of the abstract `_FloodObject`, the "FloodBody" represents the flood in urban area in either point coverage or GML surface forms. In GML, a coverage class (e.g. `multiPointCoverage`) uses the relationship between `RangeSet` and `DomainSet` to link the geometry with its attributes (see GML 3.2.1 specification for details). The flood extension for CityGML by Schulte and Coors (2009) was adopted in this work and extended via definition of an array of "FloodTimeSeriesElement" classes (containing water depth and velocity components of the flood for a particular time step) to accommodate the temporal aspects of the flood in addition to its spatial components.

Surface representation of the flood was specifically considered for its 3D visualisation. It can either be presented by a single `MultiSurface` object using "RepresentedBySurface" relation for the maximum depths; or using the aggregation hierarchy of `timeSeriesSurfaces` → `TimeStep` → `_floodBoundarySurface` classes, for a surface representation for each time step. While "floodSurface" and "FloodGroundSurface" are used for water level surface and the surface between water and ground, the "FloodClosureSurface" is used to close the enclosure when the flood geometry is not a closed volume.

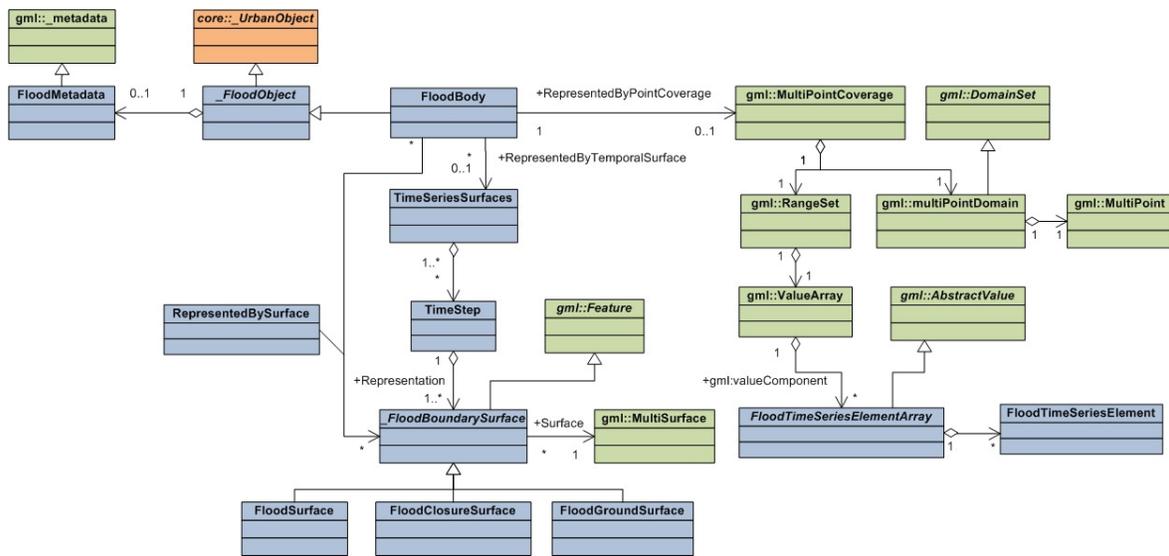


Figure 3: UML Model of the Flood Package

The valuation package (see Figure 4a) contains an abstract concept, the "`_CostObject`", that defines the value of a particular object. The "AssemblyCostObject" and "BuildingValue" realise the `_CostObject` for the repair/replacement value of building components or the construction cost of the building as a whole. Other information describing these classes include issuing institution, date of issue, currency type, etc.

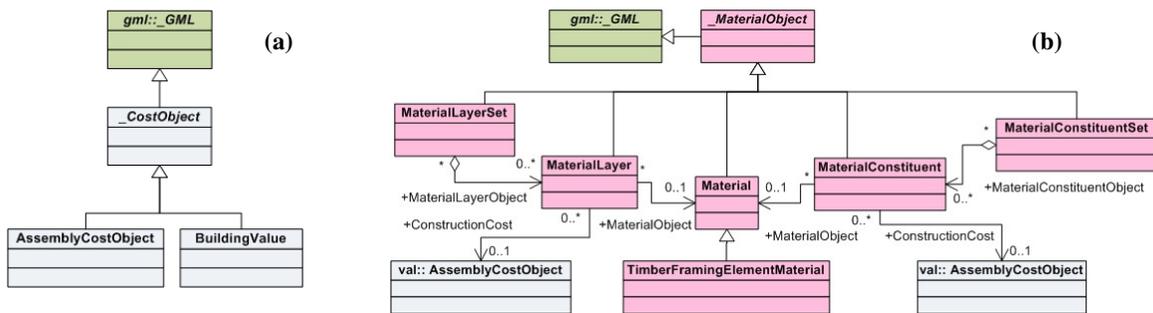


Figure 4: UML Model of the (a) Valuation Package and (b) MaterialDomain Package

The MaterialDomain package contains classes that define the construction materials of the building elements. As illustrated in Figure 4(b), this package consists of three kinds of material definition. "Material" class which defines a single material that can be directly assigned to a building component or be used as a layer in other material classes. "MaterialLayerSet" on the other hand, defines multiple materials as layers. An example here is a wall panel which includes the paint, lining and brick materials. The order of layers defines the position of the material in the object. "MaterialConstituentSet" is another material definition method which consists of one or more "MaterialConstituents" each of which defines the material of a part of component (e.g. "frame" or "glazing" in a window) indicated by its particular name.

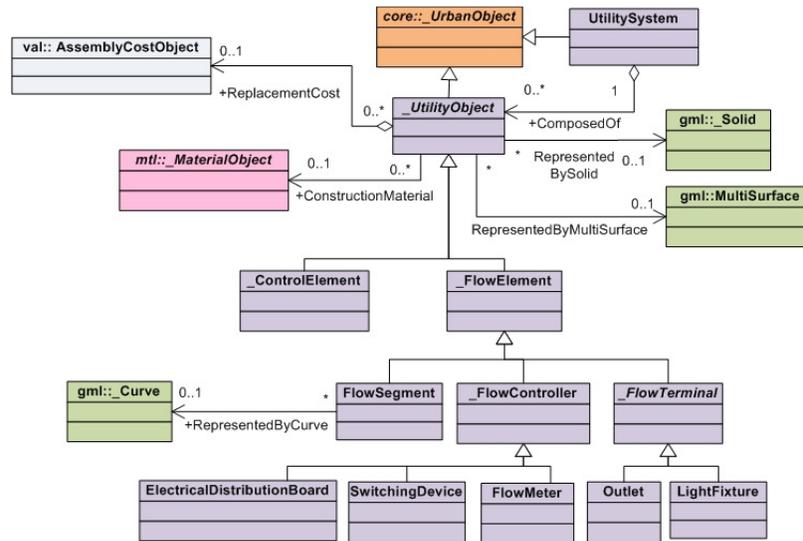


Figure 5: UML Model of the Utility Package

Utility package contains the classes related to the interior or exterior utilities of the house. This package is illustrated in Figure 5. The utility objects can be defined independently or under a particular system (e.g. electrical, water, fuel or HVAC). Each realisation of the abstract "\_UtilityObject" class may have a replacement value and a material object defined. In terms of the representation, utilities are commonly presented by either a GML solid or multiSurface. In this research, for the utilities, the focus was concentrated on the electrical system elements such as lights, outlets, meters, switches and distribution boards which are defined under abstract classes "\_FlowTerminal" and "\_FlowController". In addition, elements such as cables can be defined using the "FlowSegment" class that can be represented by 3D line segments using a GML Curve. The "\_ControlElement" class, on the other hand, is an abstract class reserved for future use to represent utility components that are used to impart control over the other elements in the system. These classes can be directly mapped to the distribution elements in the IFC data model.

The building package (see Figure 6) comprises the classes that represent the building and individual or an aggregation of building components. The "Site" class represents a parcel characterised by an address and a 2D polygon which can contain one or more buildings. Each building has a value associated with it (BuildingValue) and can be represented by either its 2D footprint or the aggregation of 3D geometries of its components (e.g. stories). A building consists of at least one storey defined by "BuildingStorey" class and may contain utility objects or systems, as well as any subtype of the abstract class "\_BuildingElement". Each of utility or building elements has a damage state and replacement cost attributes associated with them that can be used for cost analysis.

The building components defined in the model include slabs (either foundation slab or floor represented by classes of similar names), structural beams and columns, walls (either simple or complex wall represented by its different components using the relationship "consistsOfParts" defined at its supertype), roof (can be defined in the same way as wall), stairs (represented as a single component or by railings and stairflights), framing members (representing the structural framing of the building other than columns and beams), coverings such as ceiling, flooring, soffit, cornices and skirtings), windows (sliding doors are defined similarly), doors, airbricks (vents) or any type of void opening. In addition to these components, "BuildingElementPart" defines a class for a generic part of any other element. Explicit classes for wall parts ("WallComponentElement" such as the cladding) or covering parts (e.g. ceiling insulation and lining) using "CoveringLayerElement" are defined in the model to represent these objects.

In this model, windows and doors are defined either as single object or a combination of a lining (its frame) and a minimum of one panel that may have their own geometry, material and cost. "Space" class in this model defines those elements for representing the internal (e.g. room) or external (e.g. the backyard) spaces for the building.

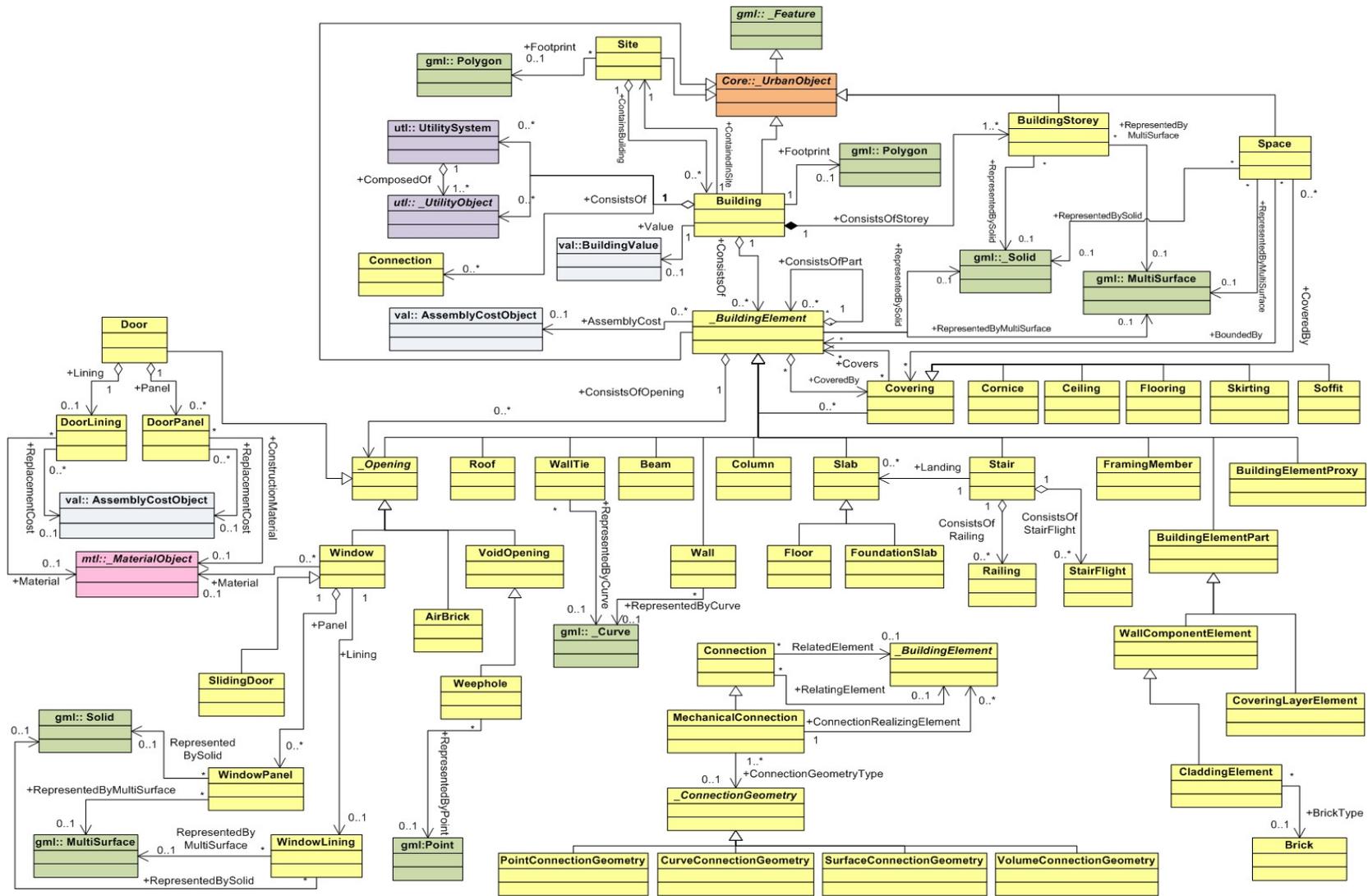


Figure 6: UML Model of the Building Package

On the other hand, "Connection" class here is the supertype for all connections defining a link between a `_BuildingElement` (related element) to another (relating element). The "Mechanical connection", is a specialised connection type that employs an additional linking `_BuildingElement` for the explicit establishment of the connection. An example here is the brick cladding to framing connection using "WallTie". The connection between the relating and related elements can be further detailed by the geometry type (point, curve, surface or volume) that either of the elements or the realising element connects to them. In the wall tie example, the connection is simplified by a `PointConnectionGeometry` class that has a GML point defined on related and relating elements. The connection concept can easily be mapped to its counterpart in IFC model for integration purposes.

To support extensibility, the building model defines a generic class, "BuildingElementProxy" to be used for elements that are not explicitly defined in the current version of the model. A "Slab edge" is an example of this element.

### 3.5 Data Model Implementation

Subsequent to the design of the conceptual model discussed in Section 3.4, the logical and physical models were developed. XML file was selected in this research to implement the integrated information model. Therefore, the physical model was prepared in relation to XML schema specifications. According to the described UML packages in Section 3.4 and the designed physical model, an XML Schema was developed. It defines the structure of the XML file and rules for definition of objects in it. This schema comprises six namespaces, each of which corresponds to and implements one of the UML packages described previously in Section 3.4.

## 4 Case Study: Damage Assessment to a House in Maribyrnong

To verify the application of the designed integrated information model to serve the defined use cases in Section 3.2, and testing the hypothesis of the research, a case study was conducted in collaboration with Maribyrnong Council and Melbourne Water. In this study, damage to a selected house in Maribyrnong was evaluated and visualised.

The required data such as elevation model of the case study area, plans of the building under investigation, and the inputs for the flood simulation (e.g. river discharges) were provided by the council and Melbourne Water. The building value and the component costs, on the other hand, were obtained from the Office of the Valuer General and the Rawlinsons' Australian construction cost handbook 2014. The BIM model of the house was developed based on the provided plans and then exported to an IFC file. On the other hand, a 1-in-100 year flood (commonly used for planning) was simulated using MIKE 21 simulation package and the outputs (spatial distribution of depth and velocity) were exported to 1140 ESRI shape files, each corresponding to a time step. An in-house tool was used to extract flood parameters from these files into a single XML file. This XML file was then mapped to the flood concepts of the proposed model in Section 3.4 and the flood information, including its geometry and attributes, were stored.

In spite of the previous efforts explained in Section 2.3, there was no tool found that could provide a smooth conversion for both geometry and semantics of BIM data to GIS formats. Therefore, a semi-automatic process (illustrated in Figure 7) was designed in this research to import building information from IFC into the implemented database. In this method, the geometry and semantic information are obtained separately and then combined to be stored in the designed database. For geometry extraction, first the IFC elements are converted to ESRI geodatabase feature classes using the ArcGIS Interoperability Extension. Then each feature class is converted to CityGML using the export function in the extension. As the input objects are not known to the converter engine, the created objects in the output are in the form of "GenericCityObject" containing a multiSurface representation of the elements. On the other hand, the attributes and relationships between elements are obtained from an exported XML version of the IFC file and combined with their geometry using a unique identifier of IFC elements, the "TAG". Having the geometry and the semantics of the elements combined, they are then mapped to the proposed data model objects and stored in the database.

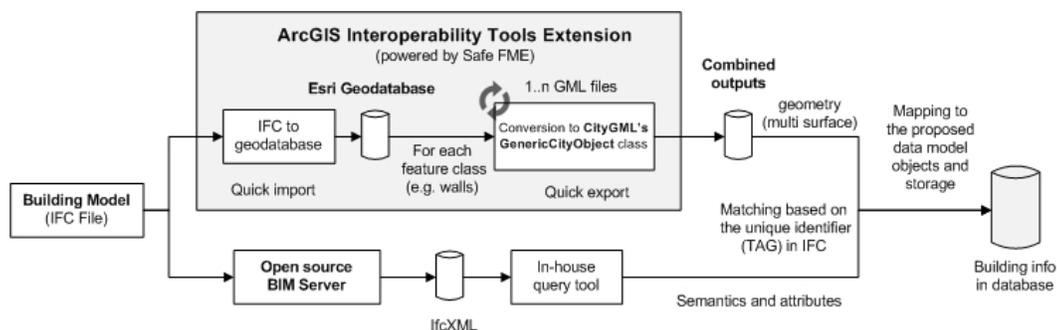


Figure 7: Converting the building information from IFC to the database based on the proposed data model

The flood parameters and building components were then used to evaluate the mode and level of damage to individual components. These calculations are based on the extracted engineering methods for assessing the damage (see Section 3.3) implemented in a prototype system. The cost of damage to each component was calculated based on its damage level and replacement cost. In addition, the geometry of the building components and colour coding were used to visualise the location as well as the level of damage in 3D.

## 5 Results and Discussion

The damage to the building and its components in the case study in Section 4 were assessed by use of the BIM and geospatial information together and the implementation of the functions extracted in Section 3.3. Figure 7 illustrates the study area and the visualisation of flood simulation outputs in 2D and 3D GIS.

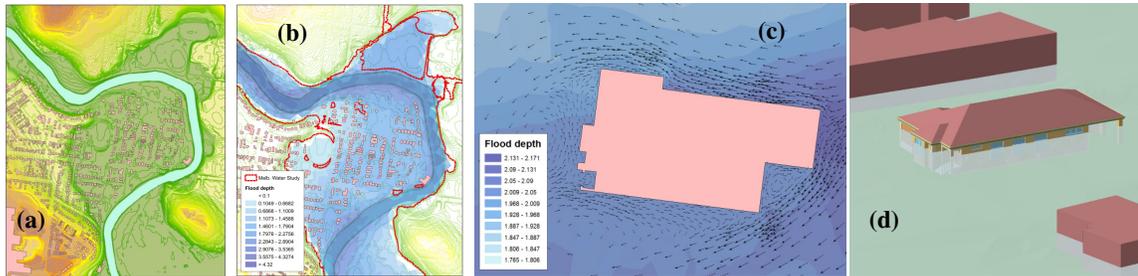


Figure 7: Case study for a house in Maribyrnong: (a) study area, (b) flood simulation output in the area, (c) flood parameters around the house, (d) 3D visualisation of the inundation level for the house

The damage analysis process showed no structural instability as the load bearing elements remained unaffected. However, the building suffered from approximately \$51,000 damage to its other components (e.g. doors and flooring) from water impacts. This number is the sum of damage costs to individual elements which a subset of these costs is presented in Table 2.

Table 2: A subset of the damage to individual building components

Building Component	Count	Total units	Unit of measurement	Unit cost (AUD\$)	Total cost (AUD\$)
Hollow core door (std. 35mm thick)	13	13	each	151.00	1,963.00
Electric meter box	1	1	each	855.00	855.00
Double power point	30	30	each	45.00	1,350.00
Timber skirting	55	137.88	m	15.10	2,081.98
Carpet flooring	6	77.181	sqm	58.50	4,515.08
Timber flooring	1	101.027	sqm	205.00	20,710.53
wall lining (gypsum)	82	418.919	sqm	28.50	11,939.19
Insulation (Rockwool batts)	21	171.56	sqm	13.15	2,256.01

Figure 8 illustrates the 3D visualisation of the location and mode of the damage to building components using a 3D GIS tool (ESRI ArcScene). Elements in these figures correspond to the items presented in Table 2 and can be queried individually. While red represents total damage and replacement is required, green indicates no damage to the component. The grey elements is used for putting the turned on damaged items into the building context and represent those that user has turned their layer off for damage inspection.

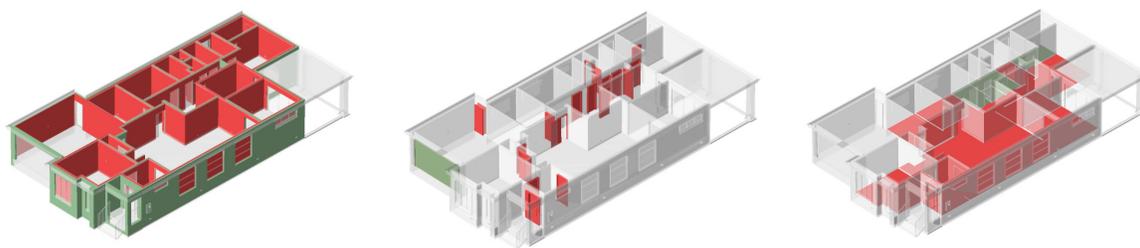


Figure 8: 3D Visualisation of Damaged Walls (left), Doors (middle) and Flooring (right) in ESRI ArcScene

From the above results, all three uses cases defined in Section 3.2 are shown to be successfully satisfied: the building structural safety was assessed and the cost of damage to individual components were estimated and visualised in 3D. The rich database based on the proposed data model in Section 3.4 was profitably used to integrate building information with flood parameters to assess the damage to the selected brick-veneer house at its component level and by taking into consideration its unique characteristics. This method provides a more detailed output illustrating the details of potential damage to a building that cannot be obtained by the current methods for FDA such as damage curves (presenting only a single number for building damage). Therefore, it can potentially overcome the limitations of current methods towards providing a better understanding of vulnerabilities in the building and facilitating an effective decision making for their treatment.

A range of stakeholders (e.g. engineering and design firms, councils, referral authorities as well as the building owners) can benefit from the outputs of the presented method. A majority of these parties are often challenged by similar questions such as "Is a particular development, proposed in an area with risk of flood, resilient to the potential risks and should be permitted for construction?" To answer this question, a tool based on the proposed integration method for BIM with GIS (similar to what was presented in this study) can facilitate engineers to test their designs against the mandated flood performance requirements from Australian Building Code Board. The alternative design options or mitigation measures may be considered for treating the vulnerabilities. In here, a cost-benefit analysis of their deployment using the proposed method can facilitate the selection of the most effective solution. On the other hand, referral authorities and councils can assess the planning and construction requirements of the proposed building more effectively by taking into account the details of the risks. Additionally, a feedback from the council for this study indicated that in case of disputes between owners and council taken to Victorian Civil and Administrative Tribunal (VCAT) in regards to refusal of a particular proposal due to risks, a non-engineering language such as the presented 3D visualisation in here can be beneficial for the communication of risks to owners for their better comprehension of the basis of the decision. The preliminary feedback from some of the aforementioned parties has been positive. However, further systematic investigation is required for understanding the value of the additional detailed information presented to them in this work.

The proposed integration method in this work can complement the use of damage curves for large-scale applications to create a multi-scale framework towards a better and more comprehensive understanding and treatment of the flood risks at different levels. It can help in improving the resilience of the community towards floods and their adverse impacts. The application of the proposed method, however, is limited to one or a few buildings where data and computation resources can be feasibly provided. For larger number of buildings at municipality or city level, the demand for such high level of detail damage evaluation for decision making would be small and other existing methods (e.g. curves) can suitably provide the required decision support.

## 6 Conclusions and Future Work

In this paper, a method for integration of BIM with GIS at the data level using a development of a new data model as a profile of GML was presented. The designed data model allows for a unified and consistent storage of the detailed representation of the building information alongside the flood parameters and other information (e.g. elevation model) in support of the micro-level FDA on buildings. The implementation of the data model and its use for assessing the damage to a selected building in a case study in Maribyrnong evidently supported the proposed hypothesis in this research: the BIM-GIS integration can facilitate a detailed assessment and 3D visualisation of damage costs to a building that is presently not supported by the type of inputs used in the current methods for FDA.

The preliminary feedback from the discussions with engineers and the council has been positive and a number of benefits are highlighted. Further research, however, is envisaged to systematically investigate the value of this extra and detailed information to the stakeholders involved in detailing the use cases in this research.

Furthermore, the data model presented in this paper was designed based on the analysis of the requirements for one particular construction type. The methodology used in this research can also be used for other building types and hazards to extend the data model for all types of buildings towards the development of a comprehensive repository of data for analysis of all types of buildings and events.

## References

- Birkmann, J., Cardona, O. D., Carreno, M. L., Barbat, A. H., and Pelling, M. (2013). "Framing Vulnerability, Risk and Societal Response: The MOVE Framework." *Journal of Natural Hazards* no. 67:193-211.
- Christodoulou, S. E., Vamvatsikos, D., and Georgiou, C. (2010). A BIM-Based Framework for Forecasting and Visualizing Seismic Damage, Cost and Time to Repair. Paper read at 8th European Conference on Product and Process Modelling (ECCPM), at Cork, Ireland.
- CLG (2007). *Improving the flood performance of new buildings - Flood resilient construction*. Department for Communities and Local Government. London.

- Dollner, J., Baumann, K., and Buchholz, H. (2006). *Virtual 3D City Models as Foundation of Complex Urban Information Spaces*. CORP, Vienna. [http://www.realcorp.at/archive/CORP2006\\_DOELLNER.pdf](http://www.realcorp.at/archive/CORP2006_DOELLNER.pdf).
- Eastman, C., Teicholz, P., Sacks, R., and Liston, K. (2011). *BIM Handbook: A guide to building information modeling for owners, managers, designers, engineers and contractors*. Second ed. New Jersey, NY: Wiley.
- El-Mekawy, M., and Ostman, A. (2010). Semantic Mapping: An Ontology Engineering Method for Integrating Building Models in IFC and CityGML. Paper read at 3rd ISDE Digital Earth Summit, 12-14 June, at Nessebar, Bulgaria.
- El-Mekawy, M., Ostman, A., and Shahzad, K. (2011). "Towards Interoperating CityGML and IFC Building Models: A Unified Model Based Approach." In *Advances in 3D Geo-Information Sciences, Lecture Notes in Geoinformation and Cartography*, edited by Kolbe, T. H., 73-93. Berlin: Springer-Verlag.
- Elmasri, R., and Navathe, S. B. (2011). *Database Systems: Models, Languages, Design and Application Programming*. 6th ed. Boston, MA: Pearson.
- Geoscience Australia (2014). *National Exposure Information System (NEXIS)*. Available from <http://www.ga.gov.au/scientific-topics/hazards/risk-impact/nexis> [Last Accessed on 17 November 2014].
- Hijazi, I., Ehlers, M., and Zlatanova, S. (2010). BIM for Geo-Analysis (BIM4GEOA): Set Up of 3D Information System With Open Source Software and Open Specification. Paper read at 5th International 3D GeoInfo Conference, 3-4 November, at Berlin, Germany.
- HNFMSC (2006). *Reducing Vulnerability of Buildings to Flood Damage: Guidance On Building In Flood Prone Areas*. Hawkesbury-Nepean Floodplain Management Steering Committee.
- IAI (2005). *IFC for GIS (IFG)*. Available at [www.iai.no/ifg/Content/ifg\\_use\\_cases.htm](http://www.iai.no/ifg/Content/ifg_use_cases.htm) [Last Accessed in 2012].
- Isikdag, U., and Zlatanova, S. (2009a). A SWOT analysis on the implementation of Building Information Models within the Geospatial Environment. Paper read at proceedings of the Urban Data Management Society symposium 2009, 24-26 June, at Ljubljana, Slovenia.
- Isikdag, U., and Zlatanova, S. (2009b). "Towards Defining a Framework for Automatic Generation of Buildings in CityGML Using Building Information Models." In *3D Geo-Information Sciences: Lecture Notes in Geoinformation and Cartography, 2009, Part II*, 79-96. Springer Berlin Heidelberg.
- Karimi, H. A., and Akinci, B. (2010). *CAD and GIS Integration*. Edited by Karimi, H. A. and Akinci, B. 1st ed: Taylor and Francis Group (CRC Press).
- Li, J., Tor, Y. K., and Zhu, Q. (2006). Research and Implementation of 3D Data Integration Between 3D GIS and 3D CAD. Paper read at XXIII FIG Congress, October 8-13, 2006, at Munich, Germany.
- Merz, B., Kreibich, H., Schwarz, J., and Thielen, A. (2010). "Review Article: "Assessment of Economic Flood Damage"." *Natural Hazards and Earth System Sciences* no. 10:1697-1724.
- Messner, F., Penning-Rowsell, E., Green, C., Meyer, V., Tunstall, S., and Van der Veen, A. (2007). *Evaluating flood damages: guidance and recommendations on principles and methods*. Integrated Flood Risk Analysis and Management Methodologies. Wallingford, UK.
- Nagel, C., Stadler, A., and Kolbe, T. H. (2009). Conceptual Requirements for the Automatic Reconstruction of Building Information Models from Uninterpreted 3D Models. Paper read at GeoWeb 2009 Academic Track - Cityscapes, 27-31 July, at Vancouver, BC, Canada.
- OGC (2007). *OGC Web Services Architecture for CAD, GIS and BIM*. Open Geospatial Consortium.
- Pistrika, A. K., and Jonkman, S. N. (2010). "Damage to residential buildings due to flooding of New Orleans after hurricane Katrina." *Natural Hazards* no. 54:413-434.
- Safe (2013). *FME Software*. Available from <http://www.safe.com/fme/> [Last Accessed on August 2013].
- Schulte, C., and Coors, V. (2009). *Development of a CityGML ADE for Dynamic 3D Flood Information*, Faculty of geomatics, Computer Science and Mathematics, University of Applied Sciences Stuttgart, Germany.
- Teorey, T. J., Lightstone, S. S., Nadeau, T., and Jagadish, H. V. (2011). *Database Modeling and Design: Logical Design*. Fifth ed: Morgan Kaufmann.
- Thielen, A., Muller, M., Kreibich, H., and Merz, B. (2005). "Flood damage and influencing factors: New insights from the August 2002 flood in Germany." *Journal of Water Resources Research* no. 41 (12):W12430.
- Van Berlo, L., and Laats, R. (2010). Integration of BIM and GIS: The Development of the CityGML GeoBIM Extension. Paper read at 5th International 3D GeoInfo Conference, 3-4 November, at Berlin, Germany.
- Wu, H., Zhengwei, H., and Gong, J. (2010). "A virtual globe-based 3D visualization and interactive framework for public participation in urban planning processes." *Computers, Environment and Urban Systems* no. 34:291-298.
- Zhang, X., Arayici, Y., Wu, S., Abbott, C., and Aouad, G. (2009). Integrating BIM and GIS for large-scale facilities asset management: a critical review. Paper read at The Twelfth International Conference on Civil, Structural and Environmental Engineering Computing, 1-4 September 2009, at Funchal, Madeira, Portugal.

# Positioning Eye Fixation and Vehicle Movement: Visual-motor Coordination Assessment in Naturalistic Driving

Qian (Chayn) Sun Dept.  
Spatial Sciences Curtin  
University, WA  
chaynsun@gmail.com

Jianhong (Cecilia) Xia  
Dept. Spatial Sciences  
Curtin University, WA  
c.xia@curtin.edu.au

Nandakumaran Nadarajah  
Dept. Spatial Sciences  
Curtin University, WA  
n.nadarajah@curtin.edu.au

Torbjörn Falkmer  
School of Occupational  
Therapy and Social Work  
Curtin University, WA  
t.falkmer@curtin.edu.au

Jonathan Foster  
School of Psychology and  
Speech Pathology  
Curtin University, WA  
j.foster@curtin.edu.au

Hoe Lee  
School of Occupational  
Therapy and Social Work  
Curtin University, WA  
h.lee@curtin.edu.au

## Abstract

In recent years, many driving studies in the traffic safety literature have undertaken error assessments of driver behaviour. However, few studies have been able to analyse the detailed individual vision and motor behaviours of drivers, due to the lack of reliable data and available technologies. Therefore, little is currently known about drivers' visual-motor coordination involving the use of visual information to regulate their physical movements. This research sets-up a technical framework to investigate on-road drivers' visual-motor coordination via vision tracking and vehicle positioning. The driving behaviour and performance were recorded and analysed using Eye Movement Tracking, Global Navigation Satellite System (GNSS) and Geographic Information Systems (GIS). The eye tracker recorded eye fixations and duration on video images to analyse the visual pattern of individual drivers. Real-time kinematic (RTK) post-processing of multi-GNSS generated vehicle movement trajectory at centimetre-level accuracy horizontally, which encompasses precise lateral positioning, speed and acceleration parameters of driving behaviours. The eye fixation data was then geocoded and synchronised with the vehicle movement trajectory in order to investigate the visual-motor coordination of the drivers. A prototype of implementation of the framework focusing on complex U-turn manoeuvre at a roundabout in five older drivers was presented in this paper. The visualisation of spatial-temporal patterns of visual-motor coordination for individual drivers allows for a greater insight to behaviour assessment. The on-road driving test in this study has also demonstrated a discriminant and ecologically valid approach in driving behaviour assessment, which can be used in studies with other cohort populations.

# 1 Introduction

Driving is an important activity which underpins personal mobility and autonomy in our society. As an activity, driving involves neuropsychological capacities that are mediated by multiple areas of the brain, including visual, attentional, perceptual, cognitive and psychomotor abilities (Anstey, Wood et al. 2005; Molnar, Charlton et al. 2013). Several reviews of driving research have reached the same conclusion: when and where drivers look is of vital importance to driver safety (Underwood, Chapman et al. 2003; Lee 2008). The information that a driver uses is predominantly visual (Sivak 1996). Crundall and Underwood (2011) addressed that the driving task is eminently suited to the application of eye tracking methodologies, and a wide range of specific driving behaviours, from navigation to anticipation of hazardous events, are primarily dependent on the optimum deployment of attention through overt eye movements. Gaze analysis based on eye tracking is a useful tool in understanding the visual behaviours underlying driving, such as fixation of the tangent point when negotiating a curve (MacDougall and Moore 2005). Since fixations are periods of relative stability, during which the eyes focus on something in the visual scene, eye fixations most often reflect the fact that the brain is processing the fixated information (Crundall and Underwood 2011). Therefore, the use of eye tracking measures has greatly increased the understanding of how driving skills develop and what strategies drivers employ to ensure a safe journey (Underwood, Chapman et al. 2003).

The visual scene and the fixated information are tightly linked to driving manoeuvres and result in vehicle positions on the road. Donges (1978) earlier described that the driver's task in steering a vehicle is to extrapolate from the complex information supply of his/her environment, where the vehicle's desired path served as guidance information, and at the same time deduce the vehicle's actual motions related to its desired path which is served as stabilization information. According to both types of information, Donges further stressed that the driver has to intervene in order to keep the vehicle's position continuously in the vicinity of its desired path. Hence, steering a vehicle is a control process with the desired path as forcing function (i.e. the road centreline) and the vehicle's position and attitude relative to the forcing function as output variables. During driving, eye movement and steering are tightly linked (Chattington, Wilson et al. 2007), in other words, the speed control and lane alignment reflect the driver's capability of using visual information to control their physical movement, namely the capability of visual-motor coordination. Since vehicle speed and lane position are two of the many possible factors that lead to crashes on horizontal curves (Fitzsimmons, Nambisan et al. 2012), understanding the visual-motor coordination becomes important to promote safe driving, in particular, for the cohort population of older drivers who have been in the area of high priority related to the increasing elderly population and higher involvement in the car crashes (Lee 2003; Rakotonirainy, Steinhardt et al. 2012), since aging might cause functional decline in vision and motor skill, adversely affecting driving behaviours and performance (McKnight 1999; Fancello, Pinna et al. 2013; Wood, Horswill et al. 2013; Sun, Xia et al. 2014).

Previously, on-road testing, computerized tasks, driving simulation, and clinical measures (physical, visual, cognitive) have all been used to estimate driving competency (Odenheimer, Beaudet et al. 1994), while the on-road driving test is the universal "criterion standard" for licensing new drivers and has been the most widely accepted method for determining driving competency, it generally lacks of standardization and data on reliability or validity (Odenheimer, Beaudet et al. 1994). Many on-road driving assessments have only a pass or fail outcome that was based on driving evaluators' clinical reasoning and not on a quantifiable, numerical test score (Shechtman, Awadzi et al. 2010). As Porter and Whitton (2002) postulated that a standardized on-road driving assessment with a quantifiable score based on Global Positioning Systems (GPS) tracking would allow for greater objectivity in determining whether a driver is fit to drive. Using electronic data collection methods was also recommended by Vlahodimitrakou and Charlton (2013) as future effort upon the DOS (Driving Observation Schedule) approach.

Nevertheless, few driving studies to date have been able to scrutinize detailed individual vision and motor behaviour data (such as speed and acceleration patterns in conjunction with visuospatial skills) due to the lack of reliable data and available technologies. Thus, little is known about drivers' visual-motor coordination in different manoeuvres and the underlying neuropsychological mechanisms. For such reason, we hypothesize in this study that eye tracking technology when coupled with vehicle movement tracking provides an even more detailed assessment of an individual's driving behaviour than standard on-road test. The goal of this paper is to propose a technical framework and to develop methods to study the visual-motor coordination in naturalistic driving, thereby forming a bridge between the literature on visual searching and motor control in driving research.

In this paper we present a framework of driving behaviour study using the combination of eye tracking and GNSS vehicle movement tracking in naturalistic driving. The on-road driving was recorded and analysed using eye tracking synchronised with multiple Global Navigation Satellite System (multi-GNSS) tracking, and Geographic Information Systems (GIS) technologies. The eye tracking equipment recorded eye fixation on video images to analyse the visual patterns of the driver (figure 2), and multi-GNSS tracking and real-time kinematic (RTK) postprocessing technique recorded and processed the precise vehicle movement trajectory, from which we detected lane keeping, speed control parameters of driving behaviours in GIS (figure 3). Previously, in an explorative pedestrian navigation study ETHZ (Kiefer, Straub et al. 2012) combined GPS (Global Positioning System)

positioning with a gaze-overlaid video, however, they came across GPS inaccuracy occurring in urban areas and present in their data. This is the first attempt in the research domain by the time this draft finished, which using surveying technology combining eye tracking to study human behaviours in a naturalistic setting. We investigated the visual perception pattern of the drivers and linked to their speed control and lane keeping, and analysed the discrepancies between individuals and groups. The objectives of this paper are as follows:

- To setup a technical framework which combines eye movement and vehicle movement data to investigate the visual-motor coordination pattern in the drivers.
- To implement above methods in a prototype: the visual-motor coordination of U-turn manoeuvre at a roundabout in older drivers.

## **2 A Framework of Combining Eye Tracking and Vehicle Movement Tracking in Driving Behaviour Assessment**

### **2.1 Tracking Eye Movement in Driving Assessment**

Eye tracking is a technique whereby an individual's eye movements are measured so that the researcher knows both where a person is looking at any given time and the sequence in which their eyes are shifting from one location to another (Poole and Ball 2006). Over the past decade, eye tracking with highly specialized eye wear equipment has been used to record detailed and accurate eye movements and visual direction in many psychology studies (Gilland 2008). Kiefer and his colleagues from ETHZ studied self-localization and human wayfinding using location-aware mobile eye tracking (Kiefer, Straub et al. 2012), which looked at the gazing patterns on the map to determine the participants' critical decision points. Analysis and recording of eye movements has also been an important tool in the investigation of the driver's visual awareness and driver behaviours in dynamic driving situations, particularly a driver's spatial cognition and fixation (Falkmer, Dahlman et al. 2008; Dukic and Broberg 2012). Eye tracking enables the researcher to collect data relating to cognitive processes employed while undertaking any particular task such as turning, these processes may include the order and length of time a viewer directs gaze at any particular object in a visual scene (Falkmer, Dahlman et al. 2008). Fixation is a central aspect of eye tracking analyses, and may be defined as the length of time the eye ceases movement and remains set on any particular focal point (Green, 2002; Zwahlen & Schnell, 1998). Fixation is relatively "stationary" eye behaviour, which allows eyes to focus their gaze on the objects being looked at, and to extract this information (Yang, McDonald et al. 2012). It is during these fractions of a second that the brain is able to receive visual information which has been acquired from the focal point (Gilland 2008). Therefore, this method utilizes eye tracking to gain an accurate picture of a driver's visual pattern and can be used to evaluate driving behaviours.

### **2.2 GPS/GNSS Tracking Vehicle Movement in Driving Assessment**

It can be comparably low-cost and ecologically valid to assess driving behaviours using Global Positioning System (GPS) tracking. GPS provides a feasible way to continuously measure the position, velocity and acceleration of a vehicle under typical driving conditions. In previous work, GPS or the combination of GPS plus video recording can provide a means to assess driving behaviours by tracking vehicle movements (Porter and Whitton 2002; Grengs, Wang et al. 2008; Naito, Miyajima et al. 2009; Cruz, Mac ías et al. 2013; Mudgal, Hallmark et al. 2014). A Multi-GNSS (Global Navigation Satellite System) receiver is the system able to calculate position, velocity and time by receiving the satellite signals broadcasted from multiple satellite navigation systems. Currently, there are a number of GNSS systems under operating including GPS, GLONASS, Galileo, QZSS and Beidou (Kubo, Hou et al. 2014; Noomwongs, Thitipatanapong et al. 2014). Using multiple satellite systems can achieve high position accuracy with increased number of satellites compared to GPS-only positioning, particularly in harsh environments (urban canyon etc.) where the GPS-only positioning becomes difficult (Kubo, Hou et al. 2014; Noomwongs, Thitipatanapong et al. 2014). Due to the nature of satellite signals, GNSS raw data contains noises. The accuracy of GNSS data depends on many factors, the position of the satellites at the time the data was recorded, errors in satellite clocks and orbits, the trips through the layers of the atmosphere, and many other sources contribute inaccuracies to satellite signals by the time they reach a receiver (Sun, Odolinski et al. 2014). For such consideration, the relative GNSS techniques can be introduced to improve the accuracy by minimising the effect of each error source. The most complex relative technique is RTK (real-time kinematic) GNSS, which uses a known position of a base station to computer the moving receiver position. By having the base station over a known position, the errors of atmospheric effects can be estimated and referenced to the observed position of the receiver. This allows for relative positioning accuracies of sub-meter level and even down to decimetre level. From there, vehicle speed, acceleration/deceleration and lateral position data can be generated from the trajectory positions for driving behaviour assessment.

### **2.3 Combining Eye Tracking and Vehicle Movement Tracking in Driving Behaviour Assessment**

Where and how long and how often the driver gazed at during driving are indicative of the visual perception strategy for the drivers to keep the car in an optimal lane position, in fact, cognitive resources involving visuospatial and motor coordination are required for driving manoeuvres. To investigate the visual-motor coordination in driving, the eye movement tracking can be integrated with the vehicle movement tracking by synchronising the two datasets. Figure 1 below illustrates the structural framework of such approach.

The eye tracking data including the fixation objects and duration of visual searching behaviour can be geocoded using vehicle position reference, and linked to the parameters on vehicle movement trajectory, such as speed, acceleration/deceleration, vehicle head angle and vehicle lateral position to the road centreline (Table 1). Such data including spatial and temporal attributes for both visual pattern and vehicle movement, while the latter reflects how the driver controlled the steering wheel and brake and accelerator pedals. The geocoding and synchronising can be accomplished in GIS (Geographic Information Systems) environment; the outcome dataset is presented in table 1. The (x, y) coordinates of the visual-motor behaviour data can be overlaid with other environment and transport information in GIS, by analysing spatial-temporal patterns of the synchronised eye fixation and vehicle movement data, the characteristics of the driver’s visual-motor can be investigated in depth. A prototype implementation of the framework in next section presents the procedure and methods to further investigate the older drivers’ visual-motor coordination of U-turn manoeuvre at a roundabout.

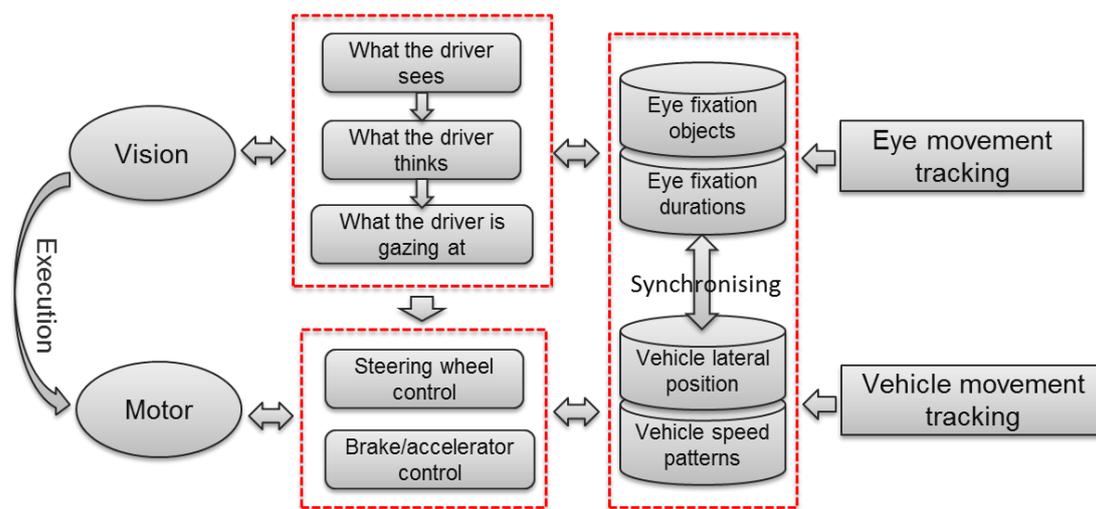


Figure 1: The combination of eye tracking and vehicle movement tracking to assess visual-motor coordination in driving: a technical framework

Table 1: Sample geo-coded eye movement records linked to driving behaviour parameters: the core dataset for visual-motor coordination analysis

FID	Participant_ID	Parameters of Eye movement						Parameters on vehicle movement trajectory				Stages at Roundabout	X_coordinate	Y_coordinate
		Fixation_no	Duration (second)	Background	Object	Gazing direction	Traffic	Speed (km/h)	Acceleration/deceleration (m/s)	Lane Deviation (cm)				
1	009	4082	0.2999	126	2891	Left	1	25.28	7.02	42	Before	394830.708488	6459223.656460	
2	009	4084	0.2001	126	2892	Left	2	24.65	6.85	40	In	394830.626771	6459223.977120	
3	009	4086	1.5992	126	259	Left	1	24.01	6.67	36	In	394830.495008	6459224.387060	
4	009	4088	0.1666	126	243	Middle	2	23.25	6.46	30	In	394830.343245	6459224.904340	
5	009	4090	0.2332	129	2892	Right	2	23.01	6.39	25	In	394830.150961	6459225.503240	
6	009	4092	0.1667	126	2892	Right	2	23.00	6.39	22	In	394829.863753	6459226.237530	

### 3 A Prototype Implementation of the Framework: Older Drivers' Visual-motor Coordination of U-turn Manoeuvre at a Roundabout

#### 3.1 Participants and Recruitment

Previous studies found that older drivers were over-represented in angle collisions, crashes at intersections, turning and changing lanes (McGwin Jr 1999; Clarke, Ward et al. 2010; Marmeleira, Ferreira et al. 2012). Those evidences indicate that the age-related decline in particular function leads to unsafe driving (Dobbs, Heller et al. 1998; McKnight 1999; Fancello, Pinna et al. 2013; Wood, Horswill et al. 2013). Even so, not all older drivers are unsafe, the statistics don't reflect driving abilities of individuals; age itself, isn't the predictor of the fitness of driving (Anstey, Wood et al. 2005). While there is a strong emphasis around the world for older adults to maintain their mobility for as long as possible, the challenge is to develop appropriate evaluation methods to identify those older drivers with hazardous driving behaviours and to provide intervention as early as possible (Lee 2003). Worldwide, the ageing population has brought the issue of older drivers into a sharper focus. In this case study, we set up on-road driving experiment in older drivers in order to investigate their visual-motor coordination behaviour.

Three female and two male older drivers aged from 60 to 79 (mean=67, SD =7.2) were recruited from local community. The eligibility of participation also includes: holding a valid Australian driver license and having an insured vehicle, driving at least 3-4 times a week, and having no mental and physical issues affecting driving. Before the assessment, all subjects provided informed consent for participation in compliance with ethics requirements from the Curtin University HREC (Human Research Ethics Committee), followed by an eye acuity check and the Mini-Mental State Examination (MMSE) to ensure their basic fitness for on-road driving. A mini-questionnaire survey on demographics and driving habits was also conducted prior to the assessment.

#### 3.2 On-road Driving Test and Data Collection

The study area was chosen around the campus of Curtin University in Perth, approximately 1.5km radius distance from the Curtin GNSS (Global Navigation Satellite System) base station. This enables the cost-effective RTK postprocessing techniques for precise vehicle movement data using the base station reference data. The primary purpose of the on-road driving assessment was to track the driver's visual-motor coordination at the complex roundabout manoeuvres. This was achieved by simultaneously recording the driver's eye fixation and the vehicle movement trajectory.

During on-road driving test, participants were asked to drive their own cars with eye tracker mounted on their heads (Figure 2). A 16-point calibration procedure was carried out prior to the on road test. Eye-tracker: A head mounted Arrington Viewpoint™ eye tracker was used to measure the eye movements of participants in 60Hz. Data collected including number of fixations and fixation durations, a proxy to driver's attention and reaction to external stimuli from the environment are captured by the eye-tracker system. Corrected vision was required and the eye tracker can be worn with glasses when necessary. These processes may include the order and length of time a driver directs gaze at any particular object in a visual scene, as well as the visual patterns the driver utilizes while performing any particular driving task (Falkmer, Dahlman et al. 2008). The sequential video clips of drivers view and eye fixation objects (the green dots) in Figure 3 show the visual pattern of the driver during the roundabout: from left to right and top to bottom, the driver looked frequently at the passing cars and edge line of the curves in order to avoid the hazards and coordinate the vehicle positions.

A pair of Trimble R10 GNSS receivers was mounted on each participant's car roof to record the vehicle movement trajectory (Figure 2). The receivers are able to track multi-GNSS systems beyond GPS only approach, this is essential for recording vehicle movement in urban area when the precise vehicle positioning and smoothest trajectory are required. The tracking configuration of Trimble R10 receivers was setup at 10Hz in order to record the car positions at every 0.1 second. Real-time kinematic (RTK) postprocessing technique was used to achieve centimetre to decimetre level horizontal accuracy by minimizing the effect of error sources transmitted between the satellites and GNSS receivers (Sun, Odolinski et al. 2014). The post-processed data was then mapped to calculate the lateral position, speed and acceleration. Figure 4 shows one diver's speed control (left) and lane keeping (right) at the roundabout manoeuvre: slowing down to stop when approaching the roundabout, keeping moderate speed in

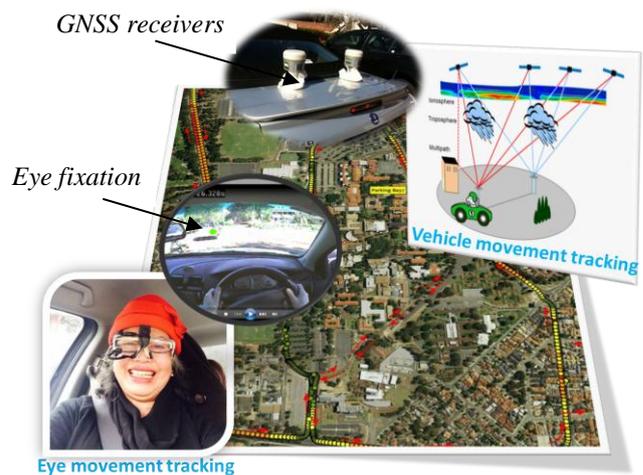


Figure 2: Equipment and on-road driving setting

the roundabout and accelerating to exit the roundabout; the perpendicular red lines between the car position to the road centreline demonstrate the lane deviations of the manoeuvre, in the case of this driver, the exit of the roundabout gave more lane deviation, followed by the stage when entering into the roundabout (west side of the roundabout). Further analysis based on the speed and lane keeping data between the drivers can be seen in figure 6.

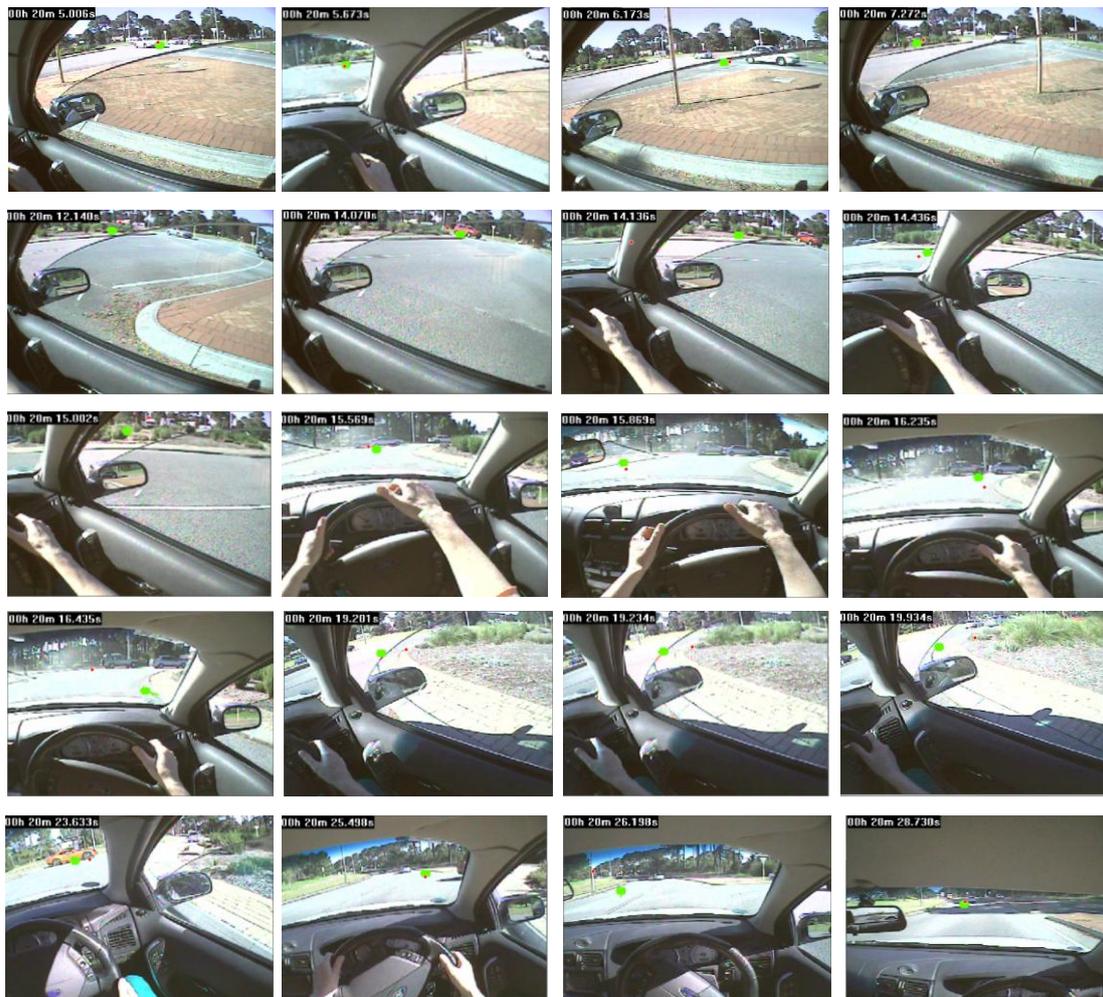


Figure 3: Video clips of driver's view and eye fixation objects recorded by eye tracker (left to right and top to bottom: the sequence of eye movement when taking U-turn through a roundabout, the green dots show the recorded right eye pupil fixation locations).

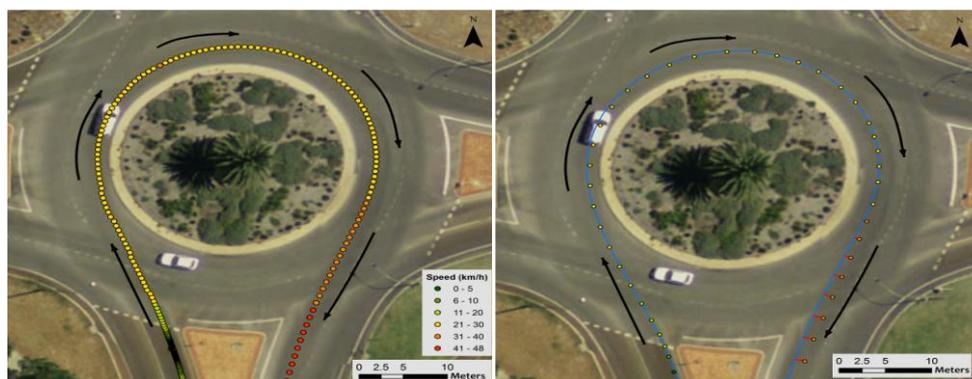


Figure 4: Patterns of speed control (left) and lane keeping (right) at a roundabout based on the multi-GNSS tracking vehicle movement trajectory (the blue line in the right map is the road centreline used as benchmark to calculate the lane deviation. The accuracy of the underneath orthoimagery is +/- 5m horizontally.)

### 3.3 Spatial-temporal Pattern of Visual-motor Coordination in Individual Drivers

All the five older drivers' manoeuvres were analysed in GIS platform, the core dataset combining both the eye movement behaviours and speed control and lane deviation was overlaid with road centrelines and orthoimagery in order to visualise the spatial-temporal patterns of visual-motor coordination in individuals.

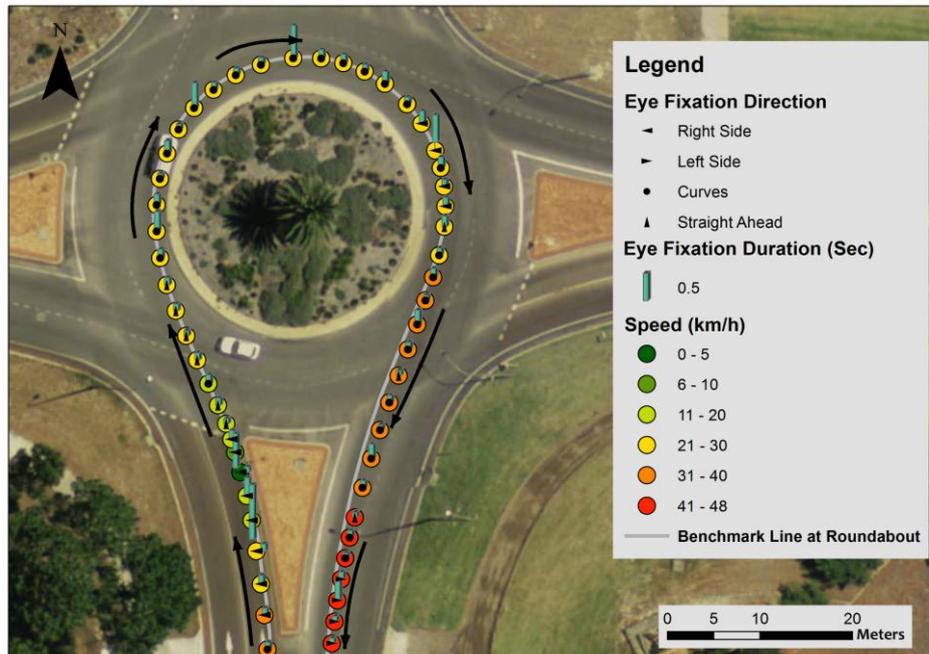


Figure 5: Positioned eye movement fixation behaviour and vehicle movement at a roundabout: A sample

As shown in figure 5 above, the driver smoothly decelerated when approaching the roundabout from 31-40km/h to 21-30km/h then to 11-20km/h to prepare to enter the roundabout and steadily accelerated through the roundabout and while exiting the roundabout to 41-48km/h. Eye tracking data shows that the driver looked right and looked into the curve while preparing to enter the roundabout, specifically at line markings, and other objects. While travelling through the roundabout the participant primarily looked into the curve at the inner rear view mirror, the right rear view mirror, trees/shrubs, and line markings. When exiting the roundabout the participant looked right at the right rear view mirror, looked straight ahead at tree or shrub, looked into the curve at trees/shrubs and line markings, and looked left at trees/shrubs. The participant showed adequate preparation while entering the roundabout and active scanning patterns while travelling through the roundabout, with the majority of fixations being traffic relevant objects. The visualisation of the spatial-temporal patterns of visual-motor coordination would give more insight into the understanding of how the drivers used visual information to control their physical movement in particular manoeuvres.

### 3.4 Statistical Analysis of Visual-motor Coordination in Older Drivers

Figure 6 below shows the variations of visual-motor coordination through roundabout manoeuvres among 5 older drivers. The left graph shows the duration and numbers of fixations within individual drivers; and the right plots the mean and std of speed control and lane deviation. As can be interpreted, Driver4 had the most frequent eye fixation and longest duration at curves, presented the lowest lane deviation, and slight higher mean speed. Driver3 had the least frequent eye fixation overall but average eye fixation and duration at curves, this participant demonstrated the highest lane deviation but the lowest std (standard deviation) value of lane deviation, and the slowest mean speed. Other drivers also used different strategies of visual-motor coordination when negotiating roundabouts. The eye fixation at curves seems associated with lane keeping and speed control, but there might be other conditions leading to the driving performance at roundabouts.

The descriptive statistical analysis in this case study tells that older drivers with better visual searching strategy achieved slightly less lane deviation and higher mean speed at roundabout manoeuvres. The visualisation of spatial-temporal patterns of visual-motor coordination when entering, passing, exiting roundabout gives more insight into the understanding of how the drivers used visual information to control their physical movement in particular manoeuvres. The findings of this study indicated that those individuals at inflated risk of road crashes could be identified using the combination of eye tracking and vehicle movement tracking in order to detect the detailed

behaviours which can be hardly observed by other methods, such as observation by instructors, clinical assessments or driving simulations, etc.

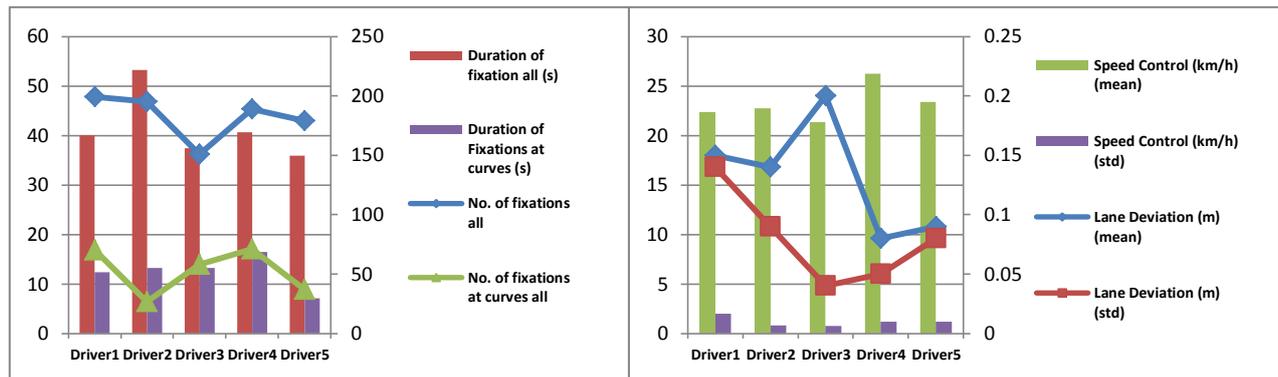


Figure 6: Visual-motor coordination at roundabout manoeuvres among older drivers (left: the duration and numbers of fixations within individual drivers; right: the mean and std of speed control and lane deviation)

In summary, this case study showed that combining eye tracking and precise vehicle movement tracking can detect the variation of the visual-motor coordination in older drivers. This study has also presented a cost-effective and ecologically valid approach in driving behaviour assessment, which can be used in studies with larger sample size in the future. Further advanced statistical analysis method, such as DFA (discriminant function analysis), can be used to predict variables affecting visual-motor coordination in older adults with high discriminant validity.

#### 4 Discussion and Conclusion

In recent years, the behaviour of drivers has become one of the main objectives of safety research; much attention has been focused on the perception and cognition of the drivers. Several studies have shown that mental and physical conditions, driving habits and behaviour when performing certain manoeuvres differ between groups. For example, the mental and physical conditions of older drivers over 65 and their ability to concentrate behind the wheel deteriorate more rapidly; with the result that there is a greater likelihood of older drivers having crashes on the road (Fancello, Pinna et al. 2013). Underwood and Chapman et al (2003) found a different sequences of visual patterns between experienced and novice drivers, suggesting that it may be of benefit to provide suitable intervention/s for novice drivers. Similar interventions may be beneficial, for example, for drivers with autism and post-stroke drivers. Motivated by these considerations, more recent research has primarily focused on behavioural measures to assist driving in particular cohorts. These interventions have been aimed at facilitating certain tasks and improving their performance, so as to promote safe driving.

This paper presented a technical framework and methods for investigating drivers' visual-motor coordination in naturalistic driving using some advanced spatial tracking technologies. We simultaneously recorded drivers' eye movement and precise vehicle movement, and linked both datasets via a GIS platform using sequential time and position information in order to obtain complete attributes of vision and motor behaviours of individual drivers. We attempted to investigate how visual perceptual information is processed with respect to changes in driving patterns due to age or different cognitive conditions, and how these factors altered driving behaviours and vehicle manoeuvres. To address this question we set up a case study to collect detailed individual data and investigated age-related changes in visual exploratory and driving manoeuvre behaviours associated with visual-motor coordination. This prototype implementation of these methods demonstrates how this approach can be used to tackle research questions concerning driving-related spatial problem solving in a novel way.

To conclude, our attempt to investigate the visual-motor coordination behaviour of drivers in a naturalistic (rather than laboratory) setting successfully collected detailed visual and vehicle control data for individuals using eye tracking and vehicle movement positioning. The advanced surveying technology (RTK multi-GNSS) that was used ensures the accuracy of vehicle kinematic positions, which were linked to the visual search behaviour fixation by fixation (or, in other words, the visual behaviour was geo-coded, integrated with the vehicle movements). The GIS platform then provides the analytical and visualising tools to examine the spatial-temporal patterns of the data. This approach offers more insight into how the drivers used visual information to control their physical movement in particular manoeuvres. We are able to analyse not only what and from where a driver is viewing their surroundings, but also how gazing behaviour is associated with vehicle control. The statistical analysis undertaken reveals the relationship between visual searching and driving manoeuvres, and differences between individuals or groups. The preliminary findings obtained suggest that variability in the performance of older drivers may stem from age-related declines in cognitive functioning. It is important that further research effort is directed toward understanding in greater detail behavioural variability in drivers using more samples and cognitive data. Moreover,

the findings obtained in the current study underline the potential value of studies in different populations into particular driving or traffic situations, such as how distraction affects the visual-motor coordination of the driver.

### **Acknowledgements**

The authors would like to thank the GNSS Research Centre, Curtin University for providing base station reference data and the, eye tracking analysis team from the School of Occupational Therapy and Social Work, Curtin University, for analysing eye tracking data.

### **References**

- Anstey, K. J., J. Wood, et al. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical psychology review* **25**(1): 45-65.
- Chattington, M., M. Wilson, et al. (2007). Eye-steering coordination in natural driving. *Experimental brain research* **180**(1): 1-14.
- Clarke, D., P. Ward, et al. (2010). Older drivers' road traffic crashes in the UK. *Accident analysis and prevention* **42**(4): 1018-1024.
- Crundall, D. and G. Underwood (2011). Visual attention while driving: Measures of eye movements used in driving research. *Handbook of Traffic Psychology*: 137-148.
- Cruz, L. C., A. Mac ís, et al. (2013). Risky driving detection through urban mobility traces: A preliminary approach. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. **8276 LNCS**: 382-385.
- Dobbs, A. R., R. B. Heller, et al. (1998). A comparative approach to identify unsafe older drivers. *Accident Analysis & Prevention* **30**(3): 363-370.
- Donges, E. (1978). A two level model of driver steering behavior. *Human Factors* **20**(6): 691-707.
- Dukic, T. and T. Broberg (2012). Older drivers' visual search behaviour at intersections. *Transportation Research Part F: Traffic Psychology and Behaviour* **15**(4): 462-470.
- Falkmer, T., J. Dahlman, et al. (2008). "Fixation identification in centroid versus start-point modes using eye-tracking data." *Perceptual and motor skills* **106**(3): 710-724.
- Fancello, G., C. Pinna, et al. (2013). Visual perception of the roundabout in old age. *WIT Transactions on the Built Environment*.
- Fitzsimmons, E. J., S. S. Nambisan, et al. (2012). Analyses of Vehicle Trajectories and Speed Profiles Along Horizontal Curves. *Journal of Transportation Safety & Security* **5**(3): 187-207.
- Gilland, J. (2008). Driving, eye-tracking and visual entropy: Exploration of age and task effects. Ann Arbor, University of South Dakota. **3340602**: 193.
- Grengs, J., X. Wang, et al. (2008). Using GPS Data to Understand Driving Behavior. *Journal of urban technology* **15**(2): 33-53.
- Kiefer, P., F. Straub, et al. (2012). Towards location-aware mobile eye tracking. *Eye Tracking Research and Applications Symposium (ETRA)*.
- Kubo, N., R. Hou, et al. (2014). Decimeter level vehicle navigation combining multi-GNSS with existing sensors in dense urban areas. *Institute of Navigation International Technical Meeting 2014, ITM 2014*.
- Lee, H. C. (2003). Validation of a driving simulator by measuring the visual attention skill of older adult drivers. *The American Journal of Occupational Therapy* **57**(3): 324-328.
- Lee, J. D. (2008). Fifty years of driving safety research. *Human Factors* **50**(3): 521-528.

- MacDougall, H. G. and S. T. Moore (2005). Functional assessment of head-eye coordination during vehicle operation. *Optometry and Vision Science* **82**(8): 706-715.
- Marmeleira, J., I. Ferreira, et al. (2012). Associations Of Physical Activity With Driving-Related Cognitive Abilities In Older Drivers: An Exploratory Study<sup>1,2</sup>. *Perceptual and motor skills* **115**(2): 521-533.
- McGwin Jr, G. (1999). Characteristics of traffic crashes among young, middle-aged, and older drivers. *Accident analysis and prevention* **31**(3): 181-198.
- McKnight, A. J. (1999). Multivariate analysis of age-related driver ability and performance deficits. *Accident analysis and prevention* **31**(5): 445-454.
- Molnar, L. J., J. L. Charlton, et al. (2013). Self-regulation of driving by older adults: Comparison of self-report and objective driving data. *Transportation Research Part F: Traffic Psychology and Behaviour* **20**(0): 29-38.
- Mudgal, A., S. Hallmark, et al. (2014). Driving behavior at a roundabout: A hierarchical Bayesian regression analysis. *Transportation Research Part D: Transport and Environment* **26**(0): 20-26.
- Naito, A., C. Miyajima, et al. (2009). Driver evaluation based on classification of rapid decelerating patterns. *Vehicular Electronics and Safety (ICVES), 2009 IEEE International Conference on*.
- Noomwongs, N., R. Thitipatanapong, et al. (2014). Driver behavior detection based on Multi-GNSS precise point positioning technology. *Applied Mechanics and Materials*. **619**: 327-331.
- Odenheimer, G. L., M. Beaudet, et al. (1994). Performance-Based Driving Evaluation of the Elderly Driver: Safety, Reliability, and Validity. *Journal of Gerontology* **49**(4): M153-M159.
- Poole, A. and L. J. Ball (2006). Eye Tracking in HCI and Usability Research. *Encyclopedia of Human Computer Interaction*. G. Claude. Hershey, PA, USA, IGI Global: 211-219.
- Porter, M. M. and M. J. Whitton (2002). Assessment of Driving With the Global Positioning System and Video Technology in Young, Middle-Aged, and Older Drivers. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* **57**(9): M578-M582.
- Rakotonirainy, A., D. Steinhardt, et al. (2012). Older drivers' crashes in Queensland, Australia. *Accident Analysis & Prevention* **48**(0): 423-429.
- Shechtman, O., K. D. Awadzi, et al. (2010). Validity and critical driving errors of on-road assessment for older drivers. *American Journal of Occupational Therapy* **64**(2): 242-251.
- Sivak, M. (1996). The information that drivers use: is it indeed 90% visual? *Perception* **25**(9): 1081-1089.
- Sun, Q., J. C. Xia, et al. (2014). Using Vehicle Movement Trajectory to Investigate the Relationship between Spatial Cognitive Ability and Driving Behaviours in Older Persons. Unpublished proposal. Curtin University.
- Underwood, G., P. Chapman, et al. (2003). Visual attention while driving: sequences of eye fixations made by experienced and novice drivers. *Ergonomics* **46**(6): 629-646.
- Vlahodimitrakou, Z., J. L. Charlton, et al. (2013). Development and evaluation of a Driving Observation Schedule (DOS) to study everyday driving performance of older drivers. *Accident Analysis & Prevention* **61**(0): 253-260.
- Wood, J. M., M. S. Horswill, et al. (2013). Evaluation of screening tests for predicting older driver performance and safety assessed by an on-road test. *Accident Analysis & Prevention* **50**(0): 1161-1168.
- Yang, Y., M. McDonald, et al. (2012). Can drivers' eye movements be used to monitor their performance? A case study. *IET Intelligent Transport Systems* **6**(4): 444-452.

# Effectiveness of DOS (Dark-Object Subtraction) method and water index techniques to map wetlands in a rapidly urbanising megacity with Landsat 8 data

**Sean Gilmore**  
Department of Spatial Sciences  
Curtin University, Perth, WA  
sean.gilmore@graduate.curtin.edu.au

**Ashty Saleem**  
Department of Spatial Sciences Curtin  
University, Perth, WA  
Ashty.saleem@postgrad.curtin.edu.au

**Ashraf Dewan**  
Department of Spatial Sciences  
Curtin University, Perth, WA  
A.Dewan@curtin.edu.au

## Abstract

The objectives of this work were to examine the applicability of the Dark-Object Subtraction (DOS) atmospheric correction method and water-based index techniques to map wetlands in Dhaka megacity using Landsat 8 data. With the use of both raw data and DOS-corrected imagery, the analysis revealed that DOS-corrected images performed better in discriminating wetland areas. Furthermore, the Modified Normalised Water Index (MNDWI) was the most superior technique whilst the Normalised Difference Water Index (NDWI) was the least suitable in identifying the spatial locations of wetlands in a rapidly urbanising environment such as Dhaka.

## Introduction

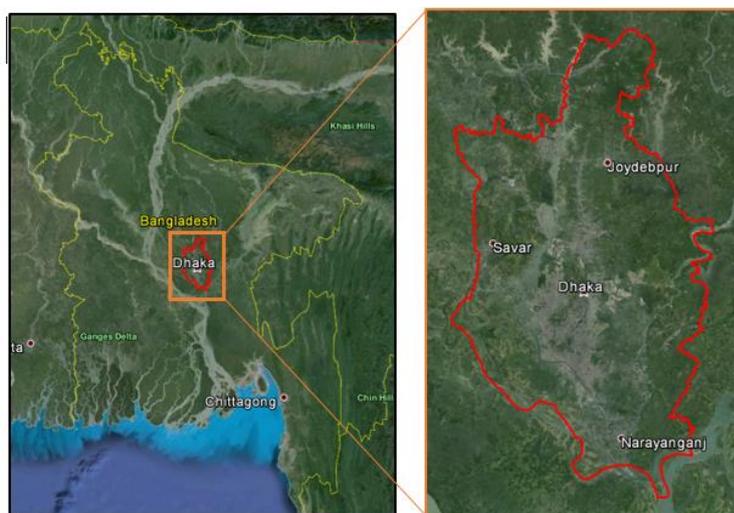
Wetlands comprise roughly 6–9 percent of the Earth's surface (Zedler and Kercher, 2005). The role of wetlands in maintaining environmental quality is well recognised (Ozesmi and Bauer, 2002), and includes the storage of global terrestrial carbon (Mitsch and Gosselink, 2007). In addition, their influence on many aspects of ecology, economy and human welfare has been well documented (Klema, 2011; Ma et al. 2007). Furthermore, wetlands act as an oasis in an urban area which is important in the reduction of surrounding surface air temperature (Sun and Chen, 2012). Changes in the distribution of wetlands either by natural factors or anthropogenic activities could significantly affect the ecosystem services (Barducci et al. 2009) mentioned above. Although they are an important environmental resource, they are heavily abused due to a lack of understanding (Smardon, 2009), particularly in developing countries. Accurate mapping and precise area statistics are therefore of paramount importance in the prevention and management of wetlands and related ecosystem services (Klema, 2011). Satellite remote sensing data have extensively been used to delineate wetlands across the world with a wide range of techniques, including a per-pixel classifier (e.g. supervised classification), semi-automated (e.g. image segmentation) method and spectral water index (e.g. normalised difference water index) (Mwita et al. 2013; Sun et al. 2012; Song et al. 2012; Jiang et al. 2012; Lu et al. 2011; Zhou et al. 2010; Islam et al. 2008; Shanmugam et al. 2006; Lira, 2006; Ouma and Tateishi, 2006). Among these techniques, water-based indices including single-band density slicing (Knight et al. 2009; Frazier and Page, 2000) techniques and band ratios comprising of two reflective bands, are found to perform better in discriminating water features such as wetlands from non-water features (Sun et al. 2012; Xu, 2006). However, deciding the optimal threshold value in isolating wetlands from the surrounding urban and land features remains an inordinate challenge (Zhang et al. 2009). In addition to single-band and band ratio techniques, new automated water-based indices such as the Automated Water Extraction Index (AWEI) has been developed and tested with several sensors in different areas however its applicability to distinguishing wetland areas within a rapidly urbanizing environments has only undergone minor testing.

Various methods have been developed to correct atmospheric influence on remote sensing data. Whilst absolute atmospheric correction methods require in-situ information, image base techniques known as relative scattering correction, on the other hand, are handy and relatively easier to subdue scattering problem in an image. A study by Song et al. (2000) suggests that atmospheric correction of remote sensing data is always not necessary and depends on the nature of the work. However, some researchers strongly favour reducing the effects of atmospheric scattering caused by light scattering (Weng, 2012), particularly in the visible region of the electromagnetic spectrum for all studies involving remotely sensed imagery.

The Dark Object Subtraction (DOS) method is an image-based technique to cancel out the haze component caused by additive scattering from remote sensing data (Chavez Jr, 1988). This method is found to be data dependent and well accepted by the geospatial community to correct light scattering in remote sensing data (Song et al. 2000). However, the DOS method has been developed for early generation Landsat sensors (e.g. TM) and may not work effectively for the new generation data such as Landsat 8 which started delivering data from early 2013. It may be noted that band composition according to electromagnetic energy of Landsat 8 differs from its predecessor Landsat sensors (e.g. TM/ETM+), hence little is known about the effectiveness of DOS method with respect to the scattering correction of the new Landsat 8 sensor. Since this new generation Landsat is expected to deliver data consistently over the next several years, research on the application of the DOS method in analysing Landsat 8 data deserves further examination.

Dhaka megacity, the capital of the people's republic of Bangladesh, has evolved into a rapidly urbanizing mega city as it attempts to accommodate large numbers of people migrating from rural areas since the independence of the country in 1971 (Dewan and Corner, 2014). With a total population of more than 14 million people according to 2011 population and housing census, the city is facing severe environmental degradation, including the rapid decline in natural wetlands due to unplanned urban expansion and related socioeconomic development (Dewan and Yamaguchi, 2009). Studies for example, show that the rapid conversion of wetlands to urban areas aggravated flooding during the monsoon season (Dewan et al. 2012), thus increasing vulnerability of urban dwellers to severe floods. Although a number of studies on the mapping of wetlands in Dhaka have been conducted (Islam, 2009; Sultana et al. 2009), they all are based on a smaller study area. Moreover, none of the studies considered advanced techniques to accurately map wetlands in the megacity. Hence, this paper is expected to contribute significantly to the existing knowledge-base on the spatial locations of wetlands in the Dhaka Metropolitan Development Plan (DMDP) area which is a recently developed planning unit by the policy makers, enforcing local organisations to preserve remaining wetland ecosystems.

Considering the above facts, the objectives of the work are: (i) to understand the effect of DOS correction technique on Landsat 8 in estimating wetlands; and (ii) to analyse the suitability of water-based index in assessing the spatial locations of wetlands in a rapidly urbanising megacity.



**Figure 1** Location map of the DMDP area (source: Google Earth)

## Data and Image Pre-Processing

The imagery was sourced from the USGS Earth Explorer web service, with two images having a path-row of 137-43 and 137-44 required as the study area is split between them. The two images were mosaicked with the resultant image clipped to the study area. As this study examines the impact of the DOS algorithm (Chavez Jr., 1988) on Landsat 8 imagery, a copy of the raw imagery was made and underwent DOS correction. Song et al (2001) describes how the DOS algorithm assumes the existence of 'dark objects', which are pixels having zero to very small reflectance numbers, within a Landsat scene. Therefore the minimum DN (digital number) value in the

histogram is considered to be the effect of atmospheric scattering and is subtracted from all pixels within the scene, thus creating 'dark objects' with a DN value of zero. Elexclis ENVI software was used for performing the DOS correction as it is an automated process which produces a corrected multispectral image. The two corrected images were then mosaicked and co-registered to the raw image was returning a RMSE of < 0.5 pixels which ensured that pixels from both images were positioned almost perfectly on top of each other.

## Methodology

The preceding data preparation produced two images, the 'raw' image and DOS corrected image. Four indices were studied and compared in their ability to accurately classify wetland areas using this imagery. Once calculated, each index underwent dynamic threshold segmentation (Zhang and Wylie, 2009) in order to find the optimal water/non-water threshold value. This was then applied to produce a binary image with water areas having a value of one and all non-water areas a value of zero. A brief outline of each index applied in this study is given below: The Normalized Difference Water Index (NDWI) was calculated using the formula proposed by McFeeters (1996):

$$NDWI = \frac{\text{Green Band} - \text{NIR Band}}{\text{Green Band} + \text{NIR Band}}$$

Where, Green Band (0.52 – 0.60  $\mu\text{m}$ ) represents band 3 and the NIR Band (0.76 – 0.90  $\mu\text{m}$ ) represents band 5 for the OLI (Operational Land Imager) of Landsat 8.

The Modified Normalised Difference Water Index (MNDWI) was calculated using the formula proposed by Xu (2006):

$$MNDWI = \frac{\text{Green Band} - \text{MIR Band}}{\text{Green Band} + \text{MIR Band}}$$

Where, Green Band (0.52 – 0.60  $\mu\text{m}$ ) represents band 3 and the MIR Band (1.55 – 1.75  $\mu\text{m}$ ) represents band 6 for the OLI (Operational Land Imager) of Landsat 8.

The Collective Indices classification was performed using a modified version of the formula developed by Lu et al. (2011). Firstly, the MNDWI was used in place of the NDWI as it performs better in extracting wetland areas in an urban dominated environment which was shown during the MNDWI's development by Xu (2006). Secondly, the formula was reversed by subtracting the NDVI from the MNDWI. This was done because on the first use of the index it gave low negative values to wetland areas and moderate – high positive values to urban and vegetation dominated areas. However, by reversing the formula it was found that the values were reversed. The modified formula is:

$$\text{Collective Indices} = MNDWI - NDVI$$

Where, NDVI is calculated using the formula first developed by Tucker (1979):

$$NDVI = \frac{\text{NIR Band} - \text{Red Band}}{\text{NIR Band} + \text{Red Band}}$$

Where, the NIR Band (0.76 – 0.90  $\mu\text{m}$ ) represents band 5 and the Red Band (0.63 – 0.69  $\mu\text{m}$ ) represents band 4 for the OLI (Operational Land Imager) of Landsat 8.

The Automated Water Extraction Index (AWEI) was calculated using a modified version of the non-shadow formula developed by Feyisa et al. (2014). The original formula had band designations based on Landsat 7 which were updated to the band designations of Landsat 8. The modified formula is:

$$AWEI_{nsh} = 4 \times (\text{BAND 3} - \text{BAND 6}) - (0.25 \times \text{BAND 5} + 2.75 \times \text{BAND 7})$$

Where, Band 3 (Green Band) is 0.53 – 0.59  $\mu\text{m}$ , Band 5 (NIR Band) is 0.85 – 0.88  $\mu\text{m}$ , Band 6 (SWIR1 Band) is 1.57 – 1.65  $\mu\text{m}$  and Band 7 (SWIR2 Band) is 2.11 – 2.29  $\mu\text{m}$ . The non-shadow formula was selected over the shadow formula as the study area is gently sloping to the south-east with no mountainous areas that would produce large shadows. When calculated using the raw Landsat 8 imagery, the indexes DN values ranged between -58555.5 and 60796.5, and was therefore normalized which reduced the range of DN values to between 0 and 53.3, making the thresholding process more manageable. The DOS AWEI was not required to be normalized which is most likely due to the DOS image consisting of reflectance values.

The overall accuracy, producers' accuracy, users' accuracy and kappa were calculated for each index based on the binary map produced using the optimal threshold. As there was no wetland based ground truth data available from the study area, the data was built in the form of 700 random points generated by ArcMap and classifying them as either water or non-water based on the raw imagery using the band combination 7,5,3 (Red, Green, Blue). This combination represented water as blue – black, urban and agricultural areas as pink and vegetation areas as shades of green. In instances, where the points lay within Dhaka's urban centre, historical Geoeye imagery of 2010 was also used as a reference. The 1989 SPOT imagery was also used sparingly when points lay in wetland areas to the north of the study area as these areas have changed significantly throughout the last 24 years. In total 200 points were classified as water and 500 as non-water. Finally, the total wetland area was calculated by multiplying the number of pixels with the value of one from the binary map by 90 m<sup>2</sup>, as each pixel is 30m x 30m.

## Results and Discussion

The MNDWI was the best performing index with raw/DOS overall accuracies of 97.28/98% and kappa values greater than 0.9 (Table 1). It was able to effectively distinguish wetland areas within the urban centre and also in the marshy areas to the north-east in Dhaka megacity. The optimal thresholds were 0.055 for the raw image and 0.15 for the DOS image indicating that the index was correctly giving wetland area pixels positive values.

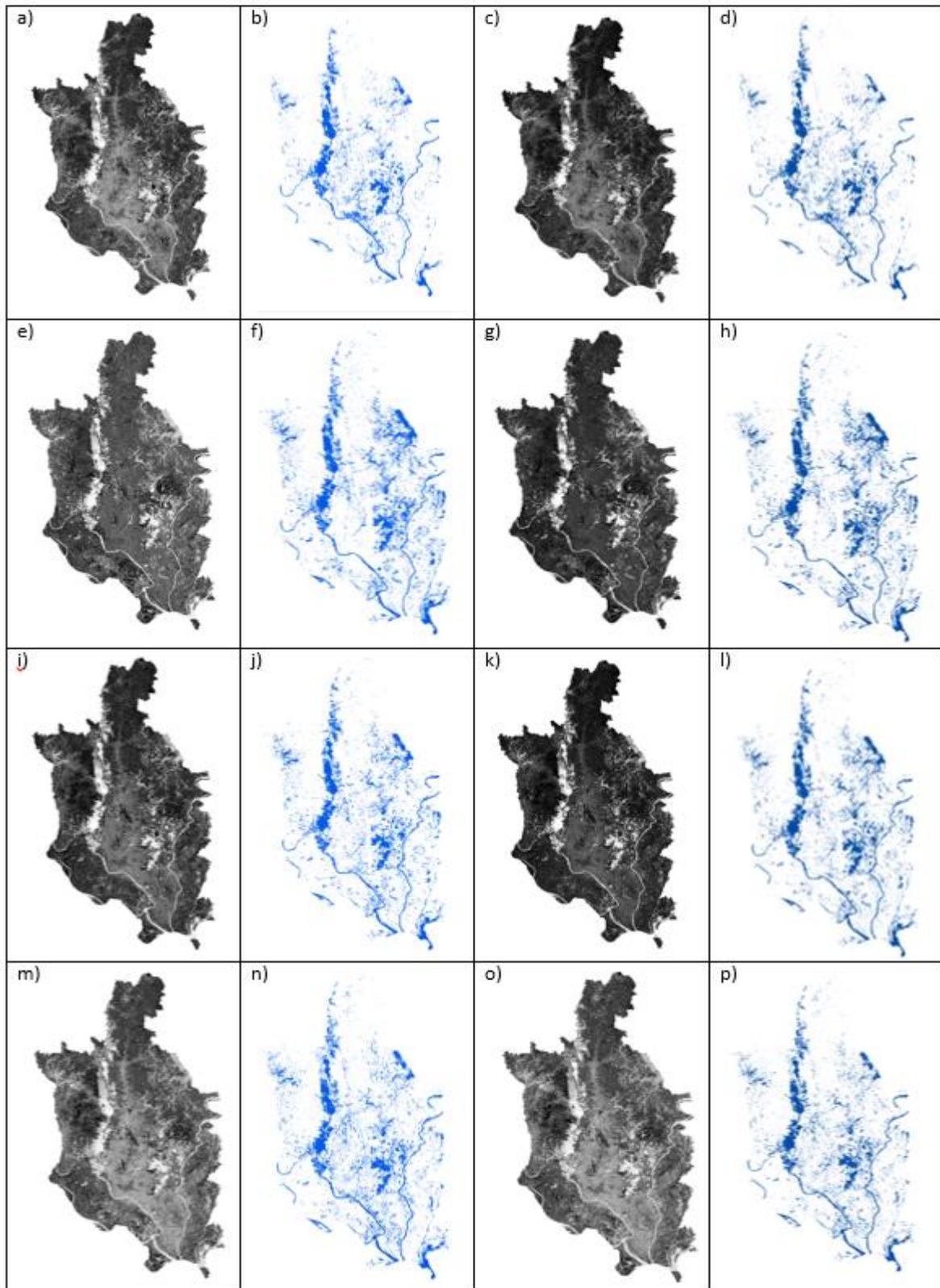
**Table 1** Accuracy assessment on the indices. OA refers to overall accuracy, PA refers to producers' accuracy and UA refers to users' accuracy

		OA	PA-Wetland	PA-Other	UA-Wetland	UA-Other	Kappa
RAW	NDWI	92.7%	98.4%	71.8%	92.8%	92.2%	0.76
	MNDWI	97.3%	98.2%	94.0%	98.4%	93.3%	0.92
	COLLECTIVE	96.6%	98.4%	89.9%	97.3%	93.7%	0.90
	AWEI	94.9%	96.9%	87.3%	96.6%	88.4%	0.85
DOS	NDWI	92.9%	98.2%	73.3%	93.1%	91.7%	0.77
	MNDWI	98.0%	98.9%	94.7%	98.6%	96.0%	0.94
	COLLECTIVE	97.9%	98.7%	94.7%	98.6%	95.3%	0.94
	AWEI	96.9%	98.6%	90.7%	97.5%	94.4%	0.91

The Collective Indices method was the second best performer with raw/DOS overall accuracies of 96.57/97.86% and kappa values greater than 0.89. It was able to correctly extract majority of the marshy wetland areas however misclassified some pixels within the urban area meaning that it may be overestimating the total wetland area. The optimal thresholds were -0.055 for the raw image and -0.1 for the DOS corrected image.

The AWEI was the third best performing index with raw/DOS overall accuracies of 94.85/96.86% and kappa values greater than 0.84. It struggled to accurately extract marshy wetland areas and misclassified several urban areas as wetlands. This is because while the index gives high positive values to clear water, it appears to give marshy wetland and urban areas the same DN values which made it difficult to decide an optimal threshold value. The optimal thresholds were 36.6 for the raw image and -0.725 for the dos corrected image.

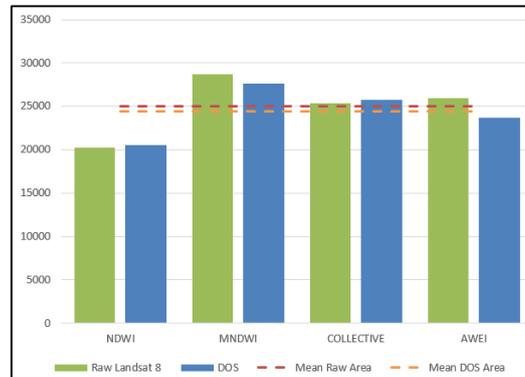
The NDWI was the worst performing index with raw/dos accuracies of 92.7/92.86% and kappa values greater than 0.7. While it effectively extracted clear water, it performed poorly in separating marshy areas of urban areas with large sections of the southern part of Dhaka misclassified. The optimal thresholds were -0.035 for the raw image and -0.075 for the dos indicating that the index is giving some wetland pixels negative values leading to misclassification.



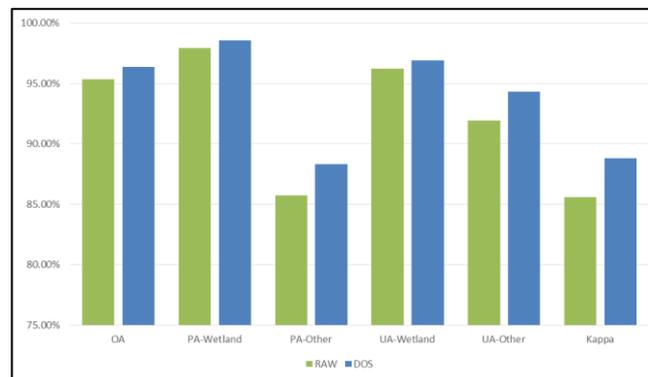
**Figure 2** Spatial locations of wetlands using different band ratio techniques; (a) The NDWI produced from the raw imagery (b) Classified NDWI using the optimal threshold (c) The NDWI produced from the DOS imagery (d) Classified DOS NDWI using the optimal threshold (e) The MNDWI produced from the raw imagery (f) Classified MNDWI using the optimal threshold (g) The MNDWI produced from the DOS imagery (h) Classified DOS MNDWI using the optimal threshold (i) The Collective Indices method produced from the raw imagery (j) Classified Collective Indices method using the optimal threshold (k) The Collective Indices method produced from the DOS imagery (l) Classified DOS Collective Indices method using the optimal threshold (m) The AWEI produced from the raw imagery (n) Classified AWEI using the optimal threshold (o) The AWEI produced from the DOS imagery (p) Classified DOS Collective Indices using the optimal threshold

## Area Results and Comparison Between Raw and DOS Corrected Imagery

The indices produced from the DOS images outperformed the indices run on the raw images in every single accuracy statistic (Figure 3) with the greatest difference being in the overall accuracies and kappa coefficients. The difference in the measured overall accuracy's ranged from a minor 0.16% increase seen in the NDWI, up to a significant 1.3% increase seen in the Collective Index. Despite the accuracy differences, there was only a small amount of variation in the area calculated between the raw and dos indices (Figure 4). On average, there was only a 1.54% difference in the wetland area calculated.



**Figure 3** Calculated wetland area in hectares for each index produced from the raw (green) and DOS-corrected (blue) images. The mean calculated area is shown for the raw (red dashed) and DOS (orange dashed) imagery



**Figure 4** Mean values for area statistics produced from the raw (green) and DOS-corrected (blue) image. OA refers to overall accuracy, PA refers to producers' accuracy and UA refers to users' accuracy

As shown in Figure 2, the DOS-corrected image was consistently sharper in the contrast between wetland and non-wetland areas except for the AWEI. This suggests that maybe a pixel variance based thresholding method such as the Otsu method might be more applicable when classifying wetland areas on DOS-corrected imagery. While the raw imagery had 11 bands, the corrected DOS image produced from ENVI returned only 7, with the Panchromatic, Cirrus, TIRS 1 and TIRS 2 bands missing, however all the major bands required for classifications were included. This makes current atmospheric correction methods still applicable in wetland extraction from remotely sensed images such as from Landsat 8.

## Conclusions

This study assesses the applicability of the widely-used dark-object subtraction method and water-based index techniques in assessing the locations of wetlands in the DMDP area of Bangladesh using new Landsat 8 multispectral data. The study reveals that atmospheric scattering effects should be removed prior to analysis in order to extract land cover information efficiently from Landsat 8 data. As far as water-based index approaches are concerned, the modified normalised difference water index (MNDWI) was found to be the suitable index to accurately determine the spatial locations of wetlands within the DMDP area, followed by the collective indices method. Note that the normalised difference water index (NDWI) was found to be the least suitable method in discriminating wetlands from non-water features which may have stemmed from the noisy characteristics of the urban dominated environment. A multitemporal work is currently underway to study the effectiveness of these methods between early and new generation Landsat sensors which is expected to advance the current knowledge-base.

## References

- Barducci, A., Guzzi, D., Marcoionni, P. and Pippi, I. (2009). Aerospace wetland monitoring by hyperspectral imaging sensors: a case study in the coastal zone of San Rossore Natural Park. *Journal of Environmental Management* 90 (7), 2278-2286.
- Chavez Jr, P.S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction for multispectral data. *Remote Sensing of Environment* 24, 459-479.
- Dewan, A.M. and Yamaguchi, Y. (2009). Land use and land cover change in Greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization. *Applied Geography* 29, 390-401.
- Dewan, A.M., Yamaguchi, Y. and Rahman, M.Z. (2012). Dynamics of land use/cover changes and the analysis of landscape fragmentation in Dhaka Metropolitan. Bangladesh. *GeoJournal* 77(3), 315-330.
- Dewan, A.M. and Corner, R.J. (2014). Dhaka megacity: geospatial perspectives on urbanisation, environment and health. Springer Science + Business Media, Dordrecht, The Netherlands.
- Feyisa, G.L., Meilby, H., Fensholt, R. and Proud, S.R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment* 140, 23-35.
- Frazier, P.S. and Page, K.J. (2000). Water body detection and delineation with Landsat TM data. *Photogrammetric Engineering and Remote Sensing* 66(12), 1461-1467.
- Islam, I. (2009). Wetlands of Dhaka Metro Area: A Study from Social, Economic and Institutional Perspectives. Dhaka, Bangladesh: A H Development Publishing House
- Islam, M.A., Thenkabail, P.S., Kulawardhana, R.W., Alankara, R., Gunasinghe, S., Edussriya, C. and Gunawardana, A. (2008). Semi-automated methods for mapping wetlands using Landsat ETM+ and SRTM data. *International Journal of Remote Sensing* 29(24), 7077-7106.
- Jiang, Z., Qi, J., Su, S., Zhang, Z. and Wu, J. (2012). Water body delineation using index composition and HIS transformation. *International Journal of Remote Sensing* 33(11), 3402-3421.
- Kelmas, V. (2011). Remote sensing of wetlands: case studies comparing practical techniques, *Journal of Coastal Research* 27(3), 418-427.
- Knight, A.W., Tindall, D.R. and Wilson, B.A. (2009). A multitemporal multiple density slice method for wetland mapping across the state of Queensland, Australia. *International Journal of Remote Sensing* 30(13), 3365-3392.
- Lira, J. (2006). Segmentation and morphology of open water bodies from multispectral images. *International Journal of Remote Sensing* 27(18), 4015-4038.
- Lu, S., Wu, B., Yan, N. & Wang, H. (2011). Water body mapping with HJ-1A/B satellite imagery. *International Journal of Applied Earth Observation and Geoinformation* 13, 428-434.

- Ma, M., Wang, X., Veroustraete, F. and Dong, L. (2007). Change in area of Ebinur lake during the 1998-2005 period. *International Journal of Remote Sensing* 28(24), 5523-5533.
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing* 17, 1425-1432.
- Mitsch, W. and Gosselink, J. (2007). *Wetland*, John Wiley and Sons, 4th Edition, Hoboken, NJ.
- Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P. and Macomber, S.A. (2000). Classification and change detection using Landsat data: when and how to correct atmospheric effects? *Remote Sensing of Environment* 75, 230-244.
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing* 17, 1425-1432.
- Mwita, E., Menz, G., Misana, S., Becker, M., Kisanga, D. and Boehme, B. (2013). Mapping small wetlands of Kenya and Tanzania using remote sensing techniques. *International Journal of Applied Earth Observation and Geoinformation* 21, 173-183
- Ouma, Y.O. and Tateishi, R. (2006). A water index for rapid mapping of shoreline changes of five East African Rift Valley lakes: an empirical analysis using Landsat TM and ETM+ data. *International Journal of Remote Sensing* 27(15), 3153-3181.
- Ozesmi, S.L. and Bauer, M.E. (2002). Satellite remote sensing of wetlands. *Wetland Ecology and Management* 10, 391-402.
- Shanmugam, P., Ahn, Y-H. and Sanjeevi, S. (2006). A comparison of the classification of wetland characteristics by linear spectral mixture modelling and traditional hard classifiers on multispectral remotely sensed imagery in southern India. *Ecological Modelling* 194, 379-394.
- Smardon, R.C. (2009). *Sustaining the World's wetlands: setting policy and resolving conflicts*. Springer Science + Business Media, Dordrecht, The Netherlands.
- Song, K., Wang, Z., Li, L., Tedesco, L., Li, F., Jin, C. and Du, J. (2012). Wetlands shrinkage, fragmentation and their links to agriculture in the Muleng – Xingkai Plain, China. *Journal of Environmental Management* 111, 120-132.
- Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P. and Macomber, S.A. (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effects. *Remote Sensing of Environment* 75, 230-244.
- Sultana, M.S., Islam, G.M. and Islam, Z. (2009). Pre- and post-urban wetland area in Dhaka City, Bangladesh: a remote sensing and GIS analysis. *Journal of Water Resources and Protection* 1, 414-421.
- Sun, F., Sun, W., Chen, J. and Gong, P. (2012). Comparison and improvement of methods for identifying waterbodies in remotely sensed imagery. *International Journal of Remote Sensing* 33(21), 6854-6875.
- Sun, R. and Chen, L. (2012). How can urban water bodies be designed for climatic adaptation? *Landscape and Urban Planning* 105, 27-33.
- Tucker C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8, 127 – 150.
- Xu, Hanqiu. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing* 27, 3025-3033.
- Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment* 117, 34-49.
- Zedler, J.B. and Kercher, S. (2005). Wetland resources: status, trends, ecosystem services, and restorability. *Annual Review of Environment and Resources* 30, 39-74.

- Zhang, S., Na, X., Kong, B., Wang, Z., Jiang, H., Yu, H., Zhao, Z., Li, X., Liu, C. and Dale, P. (2009). Identifying wetland change in China's Sanjiang plain using remote sensing. *Wetlands* 29(1), 302-313.
- Zhou, H., Jiang, H., Zhou, G., Song, X., Yu, S., Chang, J., Liu, S., Jiang, Z. and Jiang, B. (2010). Monitoring the change of urban wetland using high spatial resolution remote sensing data. *International Journal of Remote Sensing* 31(7), 1717-1731.
- Zhang, L.J. and Wylie, B. (2009). Analysis of dynamic thresholds of the normalised difference water index. *Photogrammetric Engineering and Remote Sensing* 75(11), 1307-1317.

# Robust methods for feature extraction from mobile laser scanning 3D point clouds

Abdul Nurunnabi<sup>†</sup>, Geoff West<sup>‡</sup>, David Belton<sup>‡</sup>

Department of Spatial Sciences

Curtin University, Perth, Australia

CRC for Spatial Information (CRCSI)

<sup>†</sup>abdul.nurunnabi@postgrad.curtin.edu.au, <sup>‡</sup>{g.west, d.belton}@curtin.edu.au

## Abstract

Three dimensional point cloud data obtained from mobile laser scanning systems commonly contain outliers. In the presence of outliers most of the currently used methods such as principal component analysis for point cloud processing and feature extraction produce inaccurate and unreliable results. This paper investigates the problems of outliers, and explores advantages of recently introduced statistically robust methods for automatic robust feature extraction. The robust algorithms outperform classical methods and show distinct advantages over well-known robust methods such as RANSAC in terms of accuracy and robustness. This paper shows the importance and advantages of several recently introduced robust statistics based algorithms for (i) planar surface fitting, (ii) surface normal estimation, (iii) edge detection, and (iv) segmentation. Experimental results for real mobile laser scanning point cloud data consisting of planar and non-planar complex objects surfaces show the proposed robust methods are more accurate and robust. The robust algorithms have potential for surface reconstruction, 3D modelling, registration, and quality control for point cloud data.

## 1 Introduction and Motivation

Mobile Mapping System (MMS) is an emerging technology for acquiring a three-Dimensional (3D) survey of the environment and objects in the vicinity of the mapping vehicle accurately, quickly and safely. Laser scanning provides explicit and dense 3D measurements that generate sparse point clouds. Due to its cost effectiveness and reasonable data accuracy it has been used in many applications including smart city modelling, road and rail corridor asset and inventory maintenance and management, environmental monitoring, accidental investigation, industrial control, construction management, archaeological studies, marine and coastal surveying, change detection for military and security forces, man-induced and natural disaster management (Tao and Li, 2007; Vosselman and Maas, 2010).

A MMS incorporates various navigation and sensors on a common moving platform. The vehicle (Figure 1) has advanced imaging and ranging devices, such as cameras, laser scanners or Light Detection and Ranging (LiDAR) systems, and navigation/positioning/geo-referencing devices such as a Global Navigation Satellite System (GNSS) for the determination of the position of the moving platform. An Inertial Measurement Unit (IMU) that contains sensors to detect rotation and acceleration used for determining the local orientation of the platform. A Distance Measurement Instrument is often connected to a wheel of the vehicle to provide linear distance in case of GNSS outage. The continuous integration between GNSS and IMU deals with a possible loss of signal sent by the satellite and to constantly maintain the high accuracy of data acquisition. A computer with storage and operational software is used to control the data acquisition task. Laser scanners mounted on the platform (usually at a 45° angle to the vehicle track) swing the laser beam through 360° and time-of-flight is used to determine the distance to

---

*Copyright © by the paper's authors. Copying permitted only for private and academic purposes.*

In: B. Veenendaal and A. Kealy (Eds.): Research@Locate'15, Brisbane, Australia, 10-12March 2015, published at <http://ceur-ws.org>

targets. The unit can rotate up to 200 revolutions per second and the laser records points at frequencies of up to 200 kHz. Performance includes a spatial resolution of up to 1 cm at 50 km/hour, range  $\geq 100$  metres (with 20% reflectivity), and measurement precision  $\pm 7$  mm (Arditi et al., 2010). These configurations and advantages vary for different systems. MMSs are now able to collect more than 600,000 points per second. To extract 3D coordinates of objects and features from the geo-referenced images, modelling and data fusion are required. MMSs produce huge volumes of 3D geospatial point cloud data (Figure 2) defined by their  $x,y,z$  coordinates (or latitude, longitude and elevation). Point cloud data may have colour r,g,b information from co-registered cameras and intensity from the reflected laser beam. The output point cloud data is generally stored in an industry standard format (such as LAS), which encodes the data into a point based binary or text file.



Figure 1: Mobile mapping vehicle with onboard sensors; (courtesy: Department of Spatial Sciences, Curtin University).

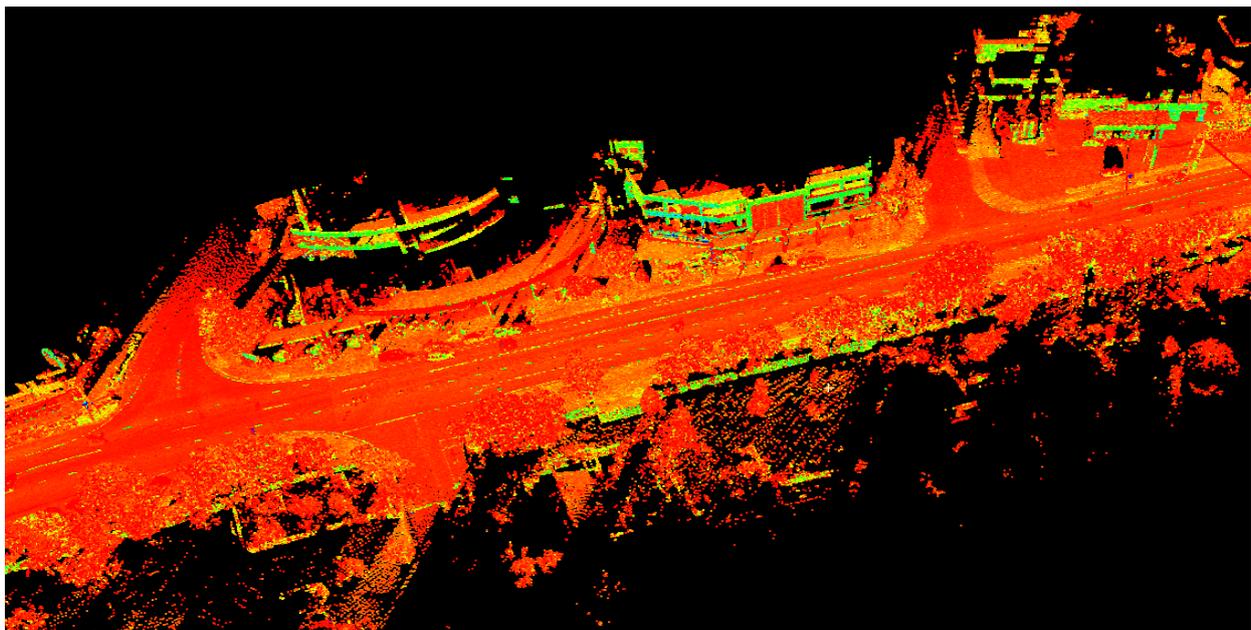


Figure 2: Mobile laser scanning point cloud data with laser intensity (courtesy of AAM<sup>1</sup>).

---

<sup>1</sup> <http://www.aamgroup.com/>

Important tasks in point cloud processing include feature extraction, visualization, analysis, modelling, surface fitting, and reconstruction. Segmentation is required for surface reconstruction, feature extraction, object recognition and modelling (Pfeifer and Briese, 2007). For surface reconstruction, the quality of the results depends on the quality of estimated local surface normals. The data acquired from MMSs typically contains regions of particular surface shape (planes, cylinders), undefined surface shape, of complex topology and geometrical discontinuities, inconsistent point density, sharp features, missing points that can create holes, and features varying in size, density and complexity. Moreover, it is unreasonable to think of point cloud data without outliers. Outliers may occur because of noise, occlusions, multiple reflectance, objects getting in the way including rain, birds and other unimportant features w.r.t. the specific study (Leslar et al, 2010). Inclusion of outliers in point cloud data exacerbates the problems for reliable and robust point cloud processing and feature extraction tasks.

The presence of outliers affects the estimates of normal and curvature, resulting in misleading and inconsistent results. Outliers can be dealt with by robust methods that have been widely used in computer vision, machine learning, pattern recognition, photogrammetry, remote sensing and statistics (Rousseeuw and Leroy, 2003; Meer, 2004). In spite of the recognition of outlier, many applications still use classical non-robust techniques including Least Squares (LS) and Principal Component Analysis (PCA) for point cloud processing tasks (Rabbani et al., 2006). It is well known that most of the classical techniques work well only for high-quality data and fail to perform adequately in the presence of outliers, giving inconsistent and misleading estimates of the model parameters (Nurunnabi et al., 2012a; 2014).

In this paper we investigate outlier problems in classical methods as well as in robust methods such as RANSAC (Fischler and Bolles, 1981) that have been widely used for feature extraction and other point cloud processing tasks. We present some recently introduced robust methods those have been developed based on robust statistical techniques as a solution to outlier influence. This will be demonstrated through four essential point cloud processing tasks: planar surface fitting, robust local saliency features (normals and curvature) estimation, edge detection, and segmentation.

The paper is organised as follows. In Section 2, the relevant literature is briefly discussed. Section 3 gives some general ideas about relevant principles and methods used in feature extraction. Robust feature extraction methods are discussed in Section 4. The limitations of the classical methods and advantages of the robust statistical methods are highlighted using real MMS acquired point cloud datasets in Section 5. Section 6 concludes the paper.

## 2 Background

Accurate plane fitting and the resultant estimates of the plane parameters are essential in point cloud data processing because of the presence of planar objects in the built environment. The LS method is the most well-known classical method for model parameter estimation (Klasing et al., 2009). Hoppe et al. (1992) introduced PCA for plane fitting that has been subsequently used and extended by many authors for point cloud processing (Pauly et al., 2002; Rabbani et al., 2006; Sanchez and Zakhor, 2012). Nurunnabi et al. (2014) showed that PCA based planar surface fitting is better than LS based fitting. Fleishman et al. (2005) proposed a forward-search approach based on a robust moving least squares (Levin, 2003) technique for reconstructing a piecewise smooth surface. The method can deal with multiple outliers, but it requires very dense sampling and a robust initial estimator to initialize the algorithm. The RANSAC algorithm has been used frequently for planar surface fitting and extraction (Schnabel et al., 2007; Masuda et al., 2013). RANSAC is very efficient for detecting large planes in noisy point clouds (Deschaud and Goulette, 2010). The Hough Transform (Duda and Hart, 1972) has been used to detect geometric shapes and for plane detection in point clouds (Borrmann et al., 2011).

Many methods have been developed to improve the quality of estimated normals and curvatures. Combinatorial and numerical approaches (Dey et al., 2005; Castillo et al., 2013) are two methods for normal estimation. Dey et al. (2005) developed combinatorial methods for estimating normals in the presence of noise, but in general, this approach becomes infeasible for large datasets. Hoppe et al. (1992) estimated the normal at each point to the fitted plane of nearest neighbours by applying the ‘total least squares’ method, which can be computed efficiently by PCA. PCA based plane fitting can be shown to be equivalent to the Maximum Likelihood Estimation (MLE) method (Wang et al., 2001). Distance weighting (Alexa et al., 2001), changing neighbourhood size (Mittra et al., 2004) and higher-order fitting (Rabbani et al., 2006) algorithms have been developed to modify PCA for better accuracy near sharp features and to avoid the influence of outliers on the estimates. Oztireli et al. (2009) used local kernel regression to reconstruct sharp features. However Weber et al. (2012) claimed there is a problem with the reconstruction from Oztireli et al. (2009) as it does not have a tangent plane at a discontinuous sharp feature.

Segmentation groups the data points into a number of locally uniform regions. Algorithms proposed can be organised roughly into three types: edge and/or border based, region growing based, and hybrid (Koster and Spann, 2000; Huang and Menq, 2001). In edge/border based methods, points positioned on the edges/borders are detected, and then points are grouped within the identified boundaries and connected edges. In region growing algorithms, generally a seed point is chosen, and then local neighbours of the seed point are combined with the seed point if they have similar surface point properties. Many authors claim that region growing based methods suffer from over

and under segmentation (Liu and Xiang, 2008). This is overcome by hybrid methods that comprise both the boundary/edge and region growing based approaches (Woo et al., 2002). Marshall et al. (2001) used LS fitting and identified surfaces of known geometric features within a segmentation framework. Klasing et al. (2009) identified the limitations of high computational cost for a large number of features. Poppinga et al. (2008) developed an efficient method of plane fitting by mean squared error computation. Castillo et al. (2013) introduced a point cloud segmentation method using surface normals computed by the constrained nonlinear least squares approach.

### 3 Related Classical and Robust Methods used in Feature Extraction

#### 3.1 Principal Component Analysis and Robust Principal Component Analysis

PCA is one of the most popular multidimensional statistical techniques for dimension reduction and data visualization (Johnson and Wichern, 2002). It finds a small number  $d$  of linear combinations of the  $m$  observed variables that can represent most of the variability of the data. Data transformation generates a new set of uncorrelated and orthogonal variables that can explain the underlying covariance structure of the data. The new set of variables (the linear combinations of the mean centered original variables) called Principal Components (PCs) produce the corresponding directions using the eigenvectors of the covariance matrix (Johnson and Wichern, 2002). Many point cloud processing methods analyze the nature of the data for a local neighbourhood of a point of interest  $p_i (x_i, y_i, z_i)$  through the study of the covariance matrix of the neighbourhood. Performing Singular Value Decomposition (SVD; Searle, 2006) on the covariance matrix  $C$  of  $k$  neighbouring points of  $p_i$ ,

$$C_{3 \times 3} = \frac{1}{k} \sum_{i=1}^k (p_i - \bar{p})(p_i - \bar{p})^T, \quad (1)$$

where  $\bar{p}$  is the centre of the data, and solving the eigenvalue equation:

$$\lambda V = CV, \quad (2)$$

produces the diagonal matrix of eigenvalues as its diagonal elements  $\lambda_i$ , and the eigenvector matrix  $V$  that contains eigenvectors or PCs as its columns. Given the required eigenvectors and the corresponding eigenvalues,  $C$  can be rewritten as:

$$C = \sum_{i=0}^2 \lambda_i v_i v_i^T, \quad (3)$$

where  $\lambda_i$  and  $v_i$  are the  $i^{\text{th}}$  eigenvalue and eigenvector, respectively. The eigenvalues denote the variances along the associated eigenvectors (Johnson and Wichern, 2002). The PCs are usually ranked in descending order of explanation of the underlying data variability.

Robust Principal Component Analysis (RPCA) is a robust version of PCA for determining the PCs that are resistant to outliers (Hubert et al., 2005; Feng et al., 2012). For processing 3D point cloud data, where the number of dimensions of a point is very small comparing with the number of points, an efficient method developed by Hubert et al. (2005) has been shown to give good results (Nurunnabi et al., 2014). The method maximizes certain robust estimates of univariate variance to obtain consecutive directions on which the data are projected. Hubert et al. (2005) coupled the ideas of using the robust estimator of the covariance matrix and the well-known Projection Pursuit method (PP) to use advantages from both the approaches. In the RPCA (Hubert et al., 2005), the data are pre-processed to make sure that the transformed data are lying in a subspace whose dimension is less than the number of observations. A useful way for reducing the data space is by using the SVD on the mean-centred data matrix. Then, the  $h$  points, where  $n/2 < h < n$ , i.e. the ‘‘least outlying’’ data points are identified, and a measure of outlyingness is computed by projecting all the data points onto many univariate directions, each of which passes through two individual data points. In order to speed up the computation, the data set is compressed to PCs defining potential directions. Then, every direction for a point  $p_i$  is scored by its corresponding value of outlyingness (Stahel, 1981; Donoho, 1982):

$$w_i = \arg \max_v \frac{|p_i v^T - c_{\text{MCD}}(p_i v^T)|}{\Sigma_{\text{MCD}}(p_i v^T)}, \quad i = 1, 2, \dots, n \quad (4)$$

where  $p_i v^T$  denotes a projection of the  $i^{\text{th}}$  observation onto the  $v$  direction, and  $c_{\text{MCD}}$  and  $\Sigma_{\text{MCD}}$  are the FMCD (Rousseeuw and van Driessen, 1999) based mean and scatter (covariance matrix) on an univariate direction  $v$ . The FMCD based estimators are used as the robust estimators of the mean and scatter in Eq. (4). In the next step, an assumed  $h$  ( $h > n/2$ ) portion of observations with the smallest outlyingness values are used to construct a robust covariance matrix  $\Sigma_h$ . A larger  $h$  can give a more accurate RPCA and a smaller  $h$  is better for more robust results. In this paper for the RPCA we use  $h = \lceil 0.5 \times n \rceil$ . Then, the method projects all the remaining observations onto the  $d$  dimensional subspace spanned by the  $d$  largest eigenvectors of  $\Sigma_h$ , and computes the mean and the covariance matrix by means of the reweighted MCD estimator, with weights based on the robust distance of every point. The

eigenvectors of this covariance matrix from the reweighted observations are the final robust PCs. Using RPCA has the advantages of yielding accurate estimates of outlier free datasets and more robust estimates for contaminated data. In addition, it has the further advantage of classification of points into inliers and outliers (Hubert et al., 2005). Hubert et al. (2012) developed a deterministic algorithm for the MCD (DetMCD) to get the robust mean vector and covariance matrix. It uses the same iteration step but does not draw a random subset. Rather it starts from only a few well-defined initial estimators. The authors claimed that DetMCD is much faster than FMCD and is at least as robust as FMCD. The DetMCD (Hubert et al., 2012) based mean vector and covariance matrix in Eq. (4) can find outlying cases, and use of the DetMCD based mean vector and robust covariance matrix in relevant places in the RPCA algorithm produces the DetMCD based RPCA, which is called Deterministic robust PCA (DetRPCA).

### 3.2 Random Sample and Consensus (RANSAC)

Fischler and Bolles (1981) developed a model-based robust approach: RANdom SAMple Consensus (RANSAC) that has been used in many applications for estimating model parameters from outlier contaminated data such as planar surface fitting, extraction and normal estimation in computer vision, photogrammetry and remote sensing. It can tolerate a large fraction of outliers, depending on the complexity of the model, up to and above 50%. RANSAC classifies data into inliers and outliers by using the LS cost function with maximum support (the number of data points that match with the model). It consists of two steps. First, a subset is randomly sampled and the required model parameters are estimated based on the subset. The size of the subset should be minimal of the random subset that is enough to estimate the model parameters (three points for a plane). In the second step, the model is compared with the data and its support is determined. This two-step iterative process continues until the likelihood of getting a model with better support than the current best model is lower than a given threshold, usually 1% to 5%.

## 4 Robust Algorithms for Feature Extraction

In this section we discuss recently introduced robust statistics based algorithms for both low level and high level feature extraction. We estimate normals and curvature as the low level saliency features that are later used for high level feature extraction. We also detect sharp features such as object edges and corners as the low level features. In this paper we perform robust segmentation (Nurunnabi et al. 2012b) to get individual surfaces (high level features) for the objects acquired by MMS. The methods for feature extraction used in this paper consist of four related tasks (Figure 3): (i) Plane fitting, (ii) edge detection, (iii) normal estimation, and (iv) segmentation.

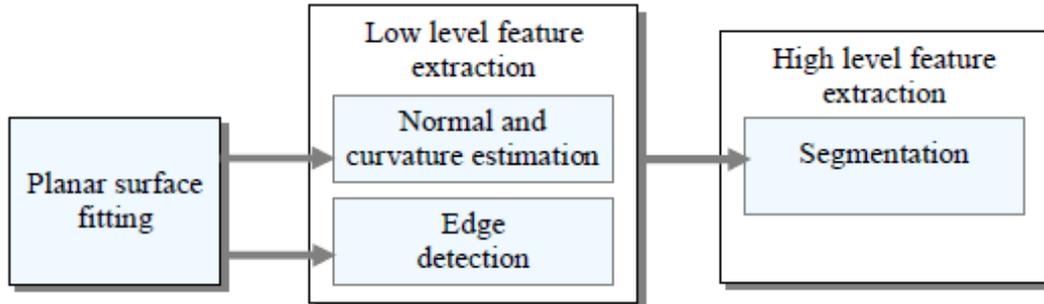


Figure 3: Feature extraction process

Using PCA for 3D point cloud data, the first two PCs form an orthogonal basis for the best-fit-plane, explaining the most variability. The third PC corresponding with the least eigenvalue expresses the least amount of variation, which defines the normal to the fitted plane. The elements of the third PC are used as the estimates of the plane parameters. The PCs (eigenvectors)  $v_2$ ,  $v_1$  and  $v_0$  can be arranged according to the corresponding eigenvalues  $\lambda_2$ ,  $\lambda_1$  and  $\lambda_0$ . Thus  $v_0$  approximates the normal  $\hat{n}$  of the surface, and  $\lambda_0$  indicates the surface variation along the normal. It is known that PCA is extremely sensitive to outliers, so in the following algorithms RPCA is used. Pauly et al. (2002) defined surface variation or curvature as:

$$\sigma_p = \frac{\lambda_0}{\sum_{i=0}^2 \lambda_i}, \quad \lambda_0 \leq \lambda_1 \leq \lambda_2. \quad (5)$$

Nurunnabi et al. (2012a) developed the method for edge detection in point cloud data, where the  $i^{\text{th}}$  point is an edge point if

$$\lambda_0 \geq \text{mean}(\lambda_0) + 1 \times \text{standard deviation}(\lambda_0). \quad (6)$$

In this paper, we use the robust segmentation algorithm (Nurunnabi et al., 2012b) that is based on robust saliency features using DetRPCA. The robust segmentation algorithm uses a region growing approach, where the region growing starts with the least curvature value  $\sigma_p$  in Eq. (5). It uses the  $k$  nearest neighbours defined as the neighbourhood  $Np_i$  for every point in the data. A local planar surface is determined for every point and its neighbourhood in the point cloud using the DetRPCA algorithm. Three measures are used by DetRPCA: Orthogonal Distance  $OD_i$  for the  $i^{\text{th}}$  seed point  $p_i$  to its best-fit-plane, Euclidian Distance  $ED_{ij}$  between the seed point  $p_i$  and one of its neighbours  $p_j$ , and angle  $\theta_{ij}$  between two points  $p_i$  and  $p_j$ . These are defined as:

$$OD_i = (p_i - \bar{p})^T \cdot \hat{n}, \quad (7)$$

where  $\bar{p}$  and  $\hat{n}$  are the centre and the unit normal for the fitted plane for  $p_i$ ,

$$ED_{ij} = \|p_i - p_j\|, \quad (8)$$

and

$$\theta_{ij} = \arccos|\hat{n}_i^T \cdot \hat{n}_j|, \quad (9)$$

where  $\hat{n}_i$  and  $\hat{n}_j$  are the unit normals for the  $i^{\text{th}}$  seed point and one of its neighbours  $p_j$ . The segmentation algorithm grows region by adding more neighbouring points one at a time from the set of data points  $P$  to the current region  $R_c$  and to the seed point list  $S_c$  using the following conditions:

$$(i) OD_i < OD_{th}, (ii) ED_{ij} < ED_{th}, \text{ and } (iii) \theta_{ij} < \theta_{th}, \quad (10)$$

where the OD threshold  $OD_{th} = \text{median}\{OD(Np_i)\} + 2 \times \text{MAD}\{OD(Np_i)\}$ ,  $\{OD(Np_i)\}$  is the set of all OD in the  $Np_i$ , MAD is the Median Absolute Deviation for all the ODs in  $Np_i$ , and ED threshold  $ED_{th} = \text{median}\{ED_{ij}\}$ . If the  $p_j$  fulfil the above conditions in Eq. (10) then the point is removed from the data  $P$  and considered as the next seed point for the region  $R_c$ .  $R_c$  grows until no new point is available in  $S_c$ . After completing the region  $R_c$ , the next seed point is selected from the remaining points in  $P$  in a similar fashion as the first one, and the process of region growing continues until  $P$  is empty.

## 5 Demonstration of the Algorithms and Results Evaluation

In this section we demonstrate the algorithms for planar surface fitting, edge detection, normal estimation, and segmentation. We explore the effects of outliers on the results based on PCA, RANSAC and DetRPCA on real MMS 3D point clouds. The results presented are mainly qualitative as it is difficult, many times, to quantitatively present results compared to, say, ground truth because of the wide variation in the results that depend on the position of outliers and the shape of the surfaces. However the results shown are representative of what would be achieved in reality.

### 5.1 Planar Surface Fitting

We consider the MMS dataset shown in Figure 4a. This figure shows a house and a tree with the planar surface (roof) of interest marked as blue. We label this dataset as the ‘roof’ dataset. The roof plane contains 2293 points, many of which have neighbouring points in the tree. The task is to isolate the roof from the rest of the data requiring the tree points to be detected as outliers. Figure 4b shows the points used to fit a plane using PCA (magenta), which deviate from the orientation of the original points (green) because of the influence of the outlying points in the tree. The black ellipses in Figure 4e shows the fitted and extracted plane contains many outliers projected onto the 2D approximation. That means, the outliers appear as inliers in the PCA determined plane, which clearly shows the masking effect (Hadi and Simonoff, 1993) caused by the presence of multiple outliers. We also use RANSAC algorithm for finding outliers and after removing the outlier points we fit a plane by the classical PCA, which is a diagnostic approach. The RANSAC results are in Figure 4 (c and f). Figure 4f shows many points are missing in the black rectangle and some outliers are still present in the black ellipse. Although RANSAC plane is better because the orientation of the fitted plane (magenta) in Figure 4c is almost coincides to the real points but the fitted plane (magenta) for PCA in Figure 4b is significantly distracted by the outliers. Now we perform DetMCD based RPCA (DetRPCA) to fit the plane. Results in Figure 4d shows that the fitted plane is in right direction without the effect of outliers, and the resultant extracted final plane in Figure 4g is clearly free from the effect of outlier without showing any masking effect. That means, DetRPCA results are robust to outliers.

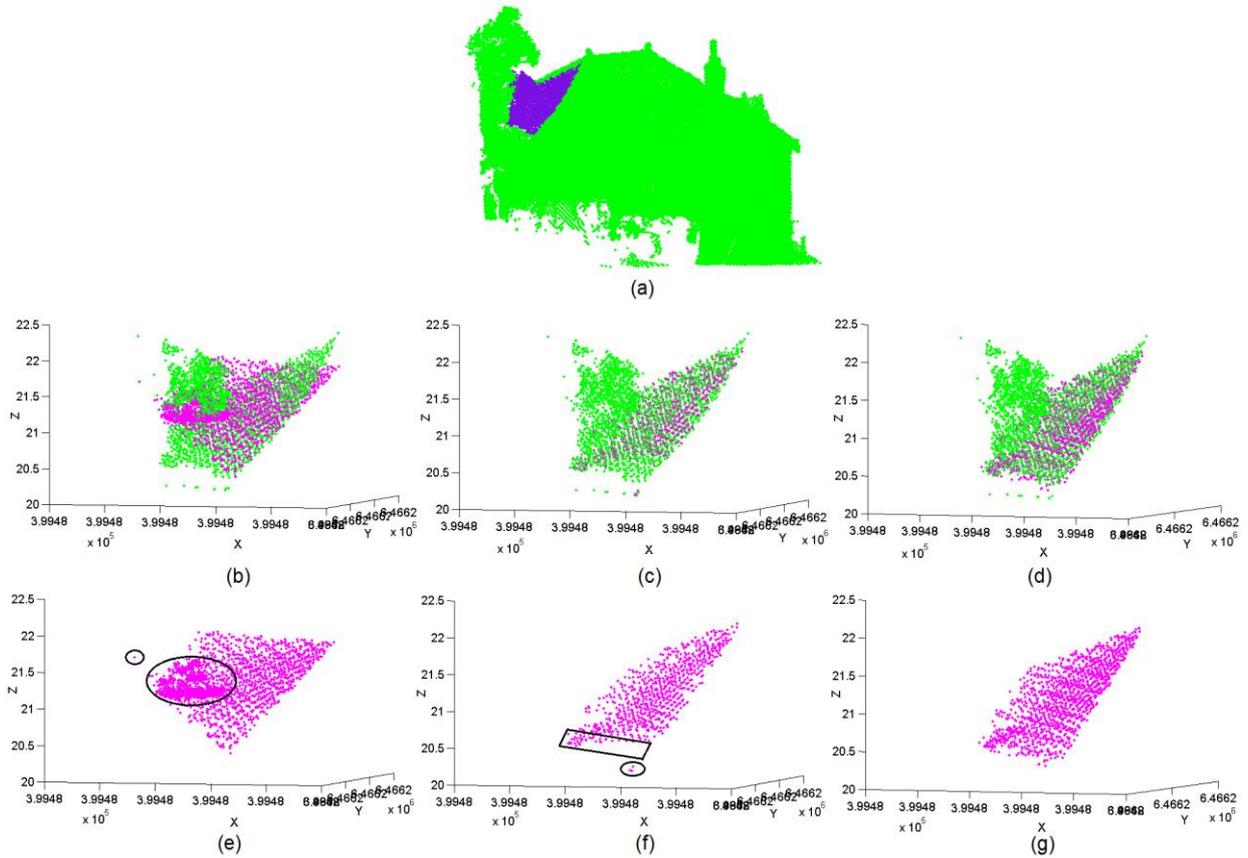


Figure 4: Plane fitting and identification for (a) roof and tree dataset surrounded by a tree, the roof points are in blue. Plane (magenta) orientation by (b) PCA, (c) RANSAC, and (d) DetRPCA. The fitted/extracted plane by (e) PCA, (f) RANSAC, and (g) DetRPCA.

## 5.2 Edge Detection and Normal Estimation

We take a dataset containing part of a building (Figure 5a) consisting of 58,371 points. We label the data as the ‘building and wall’ dataset that has sharp edges and corners. To extract these sharp features we use the classification algorithm (Nurunnabi et al., 2012a). We fit a plane to each point in the data with a local neighbourhood of  $k = 50$ . We choose the value of  $k$  based on similar real data empirically. We calculate the  $\lambda_0$  values and classify the points into inliers (surface points) and outliers (edge or corner points) according to Eq. (6). The classification results are in Figure 5(b, c and d). The results in Figure 5b show that PCA fails to recover the sharp features (edge/corner points), marked by the black ellipse and polygon. Although RANSAC (Figure 5c) is a robust method, it does not correctly classify surface, edges and corners, with many surface points appearing as edge points. Figure 5d shows that the proposed DetRPCA is significantly more accurate than PCA and RANSAC.

Normals on or near sharp features become overly smooth mainly for two reasons: (i) neighbourhood points may be present locally from two or more surfaces, and (ii) the presence of outliers/noise in the local neighbourhood (Nurunnabi et al., 2014). We consider a small part from Figure 5a shown in Figure 6a where three edges and one corner are present. In Figure 6b, the PCA method counts all the points for plane fitting in a local neighbourhood and so misrepresenting the normal at the vertex and smoothing out the sharp features. Figure 6c shows that RANSAC normals are not properly estimated and oriented, with many misclassified points around the edges and corners. This is because RANSAC was unable to get the most homogeneous points within the neighbourhoods. The robust statistical method used in the DetRPCA algorithm groups the majority of points that are homogeneous w.r.t. the neighbouring points. Hence, the estimated normal represents the surface consisting of the majority of homogeneous points. In Figure 6d, we see that robust normals are correctly oriented on the corner and on the edge points and results in normals that preserve the sharp transitions while PCA and RANSAC have smoothly changing normals areas by the black lines.

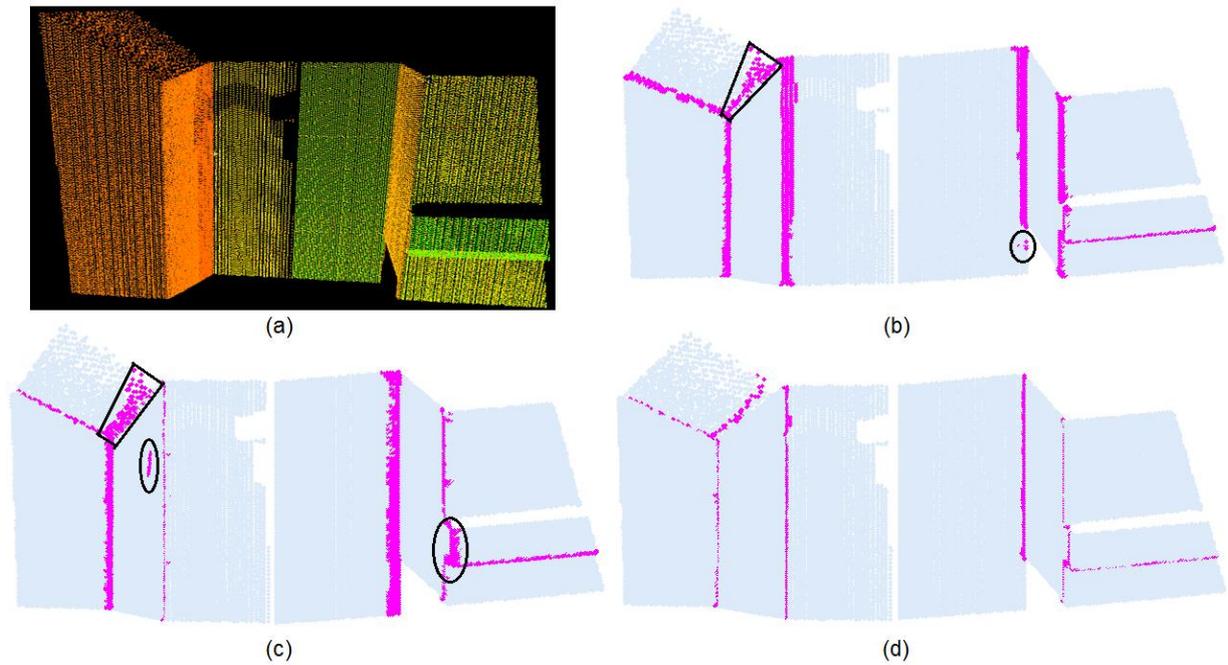


Figure 5: (a) The building and wall data, edge and corner point recovery (in magenta): (b) PCA, (c) RANSAC, and (d) DetRPCA.

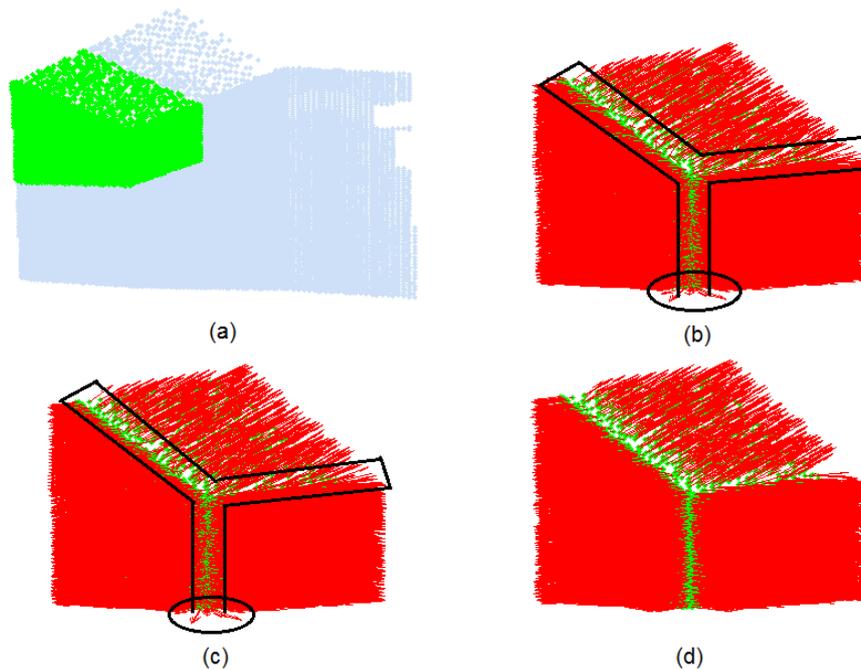


Figure 6: (a) Point cloud data, and normal orientation for: (b) PCA, (c) RANSAC, and (d) DetRPCA.

### 5.3 Segmentation

For segmentation, we consider a MMS dataset shown in Figure 7a of a petrol station, consisting of many planar and non-planar surfaces described by 111,070 points. We name this the ‘petrol station’ dataset. We set the required parameters:  $k = 30$  and angle threshold  $\theta_{th} = 7^\circ$ . Segmentation results are in Figure 7 (b, c and d) in which points that belong to each region are shown in different colours.

Perfect Segmentation (PS) is identified as a true segment from manually determined ground truth i.e. one segment describes a single feature such as the wall of a house that is one planar surface. Over Segmentation (OS) occurs where one true segment is broken into two or more separate segments, and Under Segmentation (US) is where more than one true segment are wrongly grouped together as one segment (Nurunnabi et al., 2014). Figure 7b

shows that PCA produces the worst results and failed to segment most of the different surfaces correctly. Much OS and US occurs in different places, these are highlighted by the black ellipses. Although the results in Figure 7c look slightly better than those in Figure 7b, likewise PCA, RANSAC fails to separate the road, kerb and footpath. In Figure 7c the black ellipse on the roof shows much unwanted OS occurring. DetRPCA based segmentation results in Figure 7d show that most of the surfaces are properly segmented with very little OS and without any US, out of 76 Total Segments (TS), 62 are PS. The road, kerb and footpath are extracted correctly. Performance for the methods is evaluated quantitatively in Table 1.

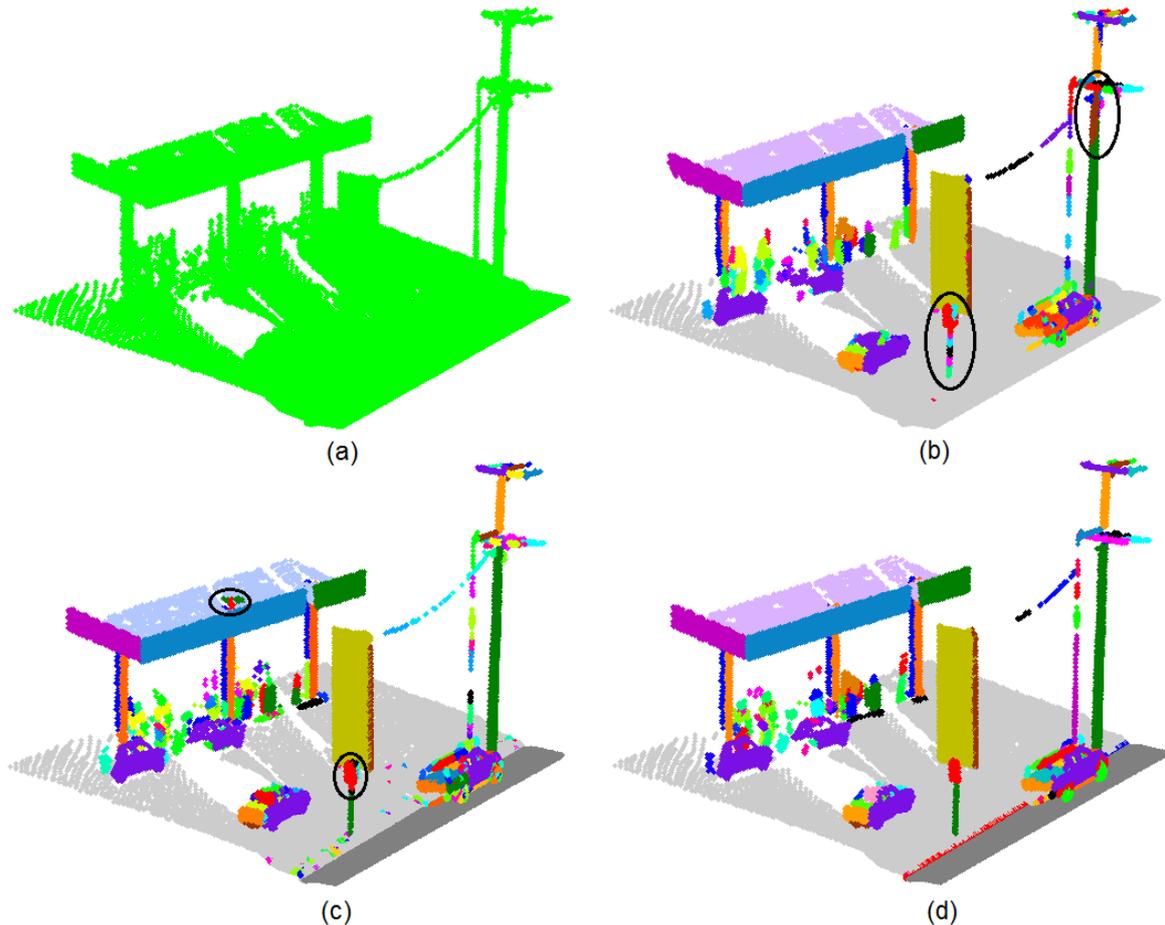


Figure 7: (a) Petrol station dataset, and segmentation results for: (b) PCA, (c) RANSAC, and (d) DetRPCA.

Table 1: Segmentation performance evaluation.

Methods	TS	PS	OS	US
PCA	88	27	34	03
RANSAC	97	32	37	01
DetRPCA	76	62	09	00

## 6 Conclusions

This paper describes some of the recently proposed algorithms for robust feature extraction in MMS 3D point cloud data. It demonstrates that robust statistics based algorithms outperform classical methods and significantly produce much better results than the popular RANSAC method. Results show the planar surfaces were fitted more robustly than the existing methods, edge and corner points were identified more accurately without any misclassification error, normals from the robust method are perfectly oriented, and segmentation results were more accurate with much less over and under segmentation. The algorithms have the limitation that the quality of the results reduces in the presence of more than 50% outliers. However, this is an extreme case but the presented algorithms are much better at dealing with this than others. The robust feature extraction methods are much better than others, and they have potential for surface reconstruction, registration and quality control for the point cloud data.

## Acknowledgements

This study has been carried out as a PhD research supported by a Curtin University International Postgraduate Research Scholarship (IPRS). The work has been supported by the Cooperative Research Centre for Spatial Information (CRCSI), whose activities are funded by the Australian Commonwealth's Cooperative Research Centres Programme. We are also thankful to the AAM Group and McMullen Nolan and Partner Surveyors for the real point cloud datasets.

## References

- Alexa, M., Behr, J., Cohen-Or, D., Fleishman, S., Levin, D., and Silva, C. T. (2001). Point set surfaces. In Proceedings of the 12th IEEE International Conference on Visualization, San Diego, California, USA, 21–28.
- Arditi, R., Garozzo, M., Laddomada, F., Paris, R., Rossi, S., Rotondi, A., and Zampa, F. (2010). New mobile LiDAR and satellite technologies for a better knowledge of roads— application to modern motorways and the case of the ancient appian way. In ASECAP Annual Study and Information Days, Oslo, Norway, 120–128.
- Borrmann, D., Elseberg, J., Lingemann, K., and Nüchter, A. (2011). The 3D Hough Transform for plane detection in point clouds: a review and a new accumulator design. *3D Research*, Springer 2(2), 1–13.
- Castillo, E., Liang, J., and Zhao, H. (2013). Point cloud segmentation and denoising via constrained nonlinear least squares surface normal estimates. In M. Breuß, A. Bruckstein, and P. Maragos (Eds.), *Innovations for Shape Analysis: Models and Algorithms*, Springer, New York, USA, 283–298.
- Deschaud, J.-E. and Goulette, F. (2010). A fast and accurate plane detection algorithm for large noisy point clouds using filtered normals and voxel growing. In Proceedings of the 5th International Symposium 3D Data Processing, Visualization and Transmission, Paris, France.
- Dey, T. K., Gang, L., and Sun, J. (2005). Normal estimation for point cloud: a comparison study for a Voronoi based method. In Proceedings of the Eurographics Symposium on Point-Based Graphics, New York, USA, 39 – 46.
- Donoho, L. (1982). Breakdown properties of multivariate location estimators. PhD Qualifying paper, Harvard University, Boston, USA.
- Duda, R. O. and Hart, P. E. (1972). Use of Hough transformation to detect lines and curves in pictures. *Communications of the ACM* 15(1), 11–15.
- Feng, J., Xu, H., and Yan, S. (2012). Robust PCA in high-dimension: a deterministic approach. In Proceedings of the 29th International Conference on Machine Learning, Edinburgh, Scotland, UK, 249–256.
- Fleishman, S., Cohen-Or, D., and Silva, C. (2005). Robust moving least-squares fitting with sharp features. *ACM Transactions on Graphics* 24(3), 544–552.
- Fischler, M. A. and Bolles, R. C. (1981). Random Sample Consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6), 381–395.
- Hadi, A. S. and Simonoff, J. S. (1993). Procedures for the identification of outliers. *Journal of the American Statistical Association* 88(424), 1264–1272.
- Hoppe, H., Rose, T. D., and Duchamp, T. (1992). Surface reconstruction from unorganized points. In Proceedings of the ACM SIGGRAPH (26), Chicago, USA, 71–78.
- Huang, J. and Menq, C.-H. (2001). Automatic data segmentation for geometric feature extraction from unorganized 3-D coordinate points. *IEEE Transactions on Robotics and Automation* 17(3), 268–279.
- Hubert, M., Rousseeuw, P. J., and Branden, K. V. (2005). ROBPCA: a new approach to robust principal component analysis. *Technometrics* 47(1), 64–79.
- Hubert, M., Rousseeuw, P. J., and Verdonck, T. (2012). A deterministic algorithm for robust scatter and location. *Journal of Computational and Graphical Statistics* 21(3), 618–637.
- Johnson, R. A. and Wichern, D. W. (2002). *Applied Multivariate Statistical Analysis*. Prentice Hall, New Jersey, USA, 5th edition.

- Klasing, L., Altho\_, D., Wollherr, D., and Buss, M. (2009). Comparison of surface normal estimation methods for range sensing applications. In Proceedings of the IEEE International Conference on Robotics and Automation, Kobe, Japan, 3206–3211.
- Koster, K. and Spann, M. (2000). MIR: an approach to robust clustering-application to range image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(5), 430–444.
- Leslar, M., Wang, J. G., and Hu, B. (2010). A comparison of two new methods of outlier detection for mobile terrestrial LiDAR data. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38(Part 1), 78–84.
- Levin, D. (2003). Mesh-independent surface interpolation. In Brunnett, Hamann, and Mueller (Eds.), Geometric Modeling for Scientific Visualization, Springer, New York, USA, 37–49.
- Liu, Y. and Xiang, Y. (2008). Automatic segmentation of unorganized noisy point clouds based on the Gaussian map. Computer-Aided Design 40(5), 576–594.
- Marshall, D., Lukacs, G., and Martin, R. (2001). Robust segmentation of primitives from range data in presence of geometric degeneracy. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(3), 304–314.
- Masuda, H., Tanaka, I., and Enomoto, M. (2013). Reliable surface extraction from point-clouds using scanner-dependent parameters. Computer Aided Design and Applications 10(2), 265–277.
- Meer, P. (2004). Robust techniques for computer vision. In G. Medioni, and S. B. Kang (Eds.), Emerging Topics in Computer Vision, Prentice Hall, New Jersey, USA, 107–190.
- Nurunnabi, A., Belton, D., and West, G. (2012a). Robust segmentation for multiple planar surface extraction in laser scanning 3D point cloud data. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR), Tsukuba Science City, Japan, 1367–1370.
- Nurunnabi, A., Belton, D., and West, G. (2012b). Robust segmentation in laser scanning 3D point cloud data. In Proceedings of the Digital Image Computing: Techniques and Applications (DICTA), Fremantle, Australia, 1–8.
- Nurunnabi, A., Belton, D., and West, G. (2014). Robust statistical approaches for local planar surface fitting in 3D laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing 96, 106–122.
- Öztireli, A. C., Guennebaud, G., and Gross, M. (2009). Feature preserving point set surfaces based on nonlinear kernel regression. Computer Graphics Forum 28(2), 493–501.
- Pauly, M., Gross, M., and Kobbelt, L. P. (2002). Efficient simplification of point sample surface. In Proceedings of the International Conference on Visualization, Washington, D.C., USA, 163–170.
- Pfeifer, N. and Briese, C. (2007). Geometrical aspects of airborne laser scanning and terrestrial laser scanning. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36(Part 3/W52), 311–319.
- Poppinga, J., Vaskevicius, N., Birk, A., and Pathak, K. (2008). Fast plane detection and polygonalization in noisy 3D range image. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS), Nice, France, 3378–3383.
- Rabbani, T., van den Heuvel, F. A., and Vosselman, G. (2006). Segmentation of point clouds using smoothness constraint. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36(5), 248–253.
- Rousseeuw, P. J. and Leroy, A. (2003). Robust Regression and Outlier Detection. John Wiley and Sons, New York, USA.
- Rousseeuw, P. J. and Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. Technometrics, 41(3):212–223.
- Searle, S. R. (2006). Matrix Algebra Useful for Statistics. John Wiley and Sons, New York, USA.
- Stahel, W. A. (1981). Robust Estimation: Infinitesimal Optimality and Covariance Matrix Estimators, PhD thesis, Department of Mathematics, Eidgenössische Technische Hochschule (ETH), Zurich, Switzerland.

- Sanchez, V. and Zakhor, A. (2012). Planar 3D modelling of building interiors from point cloud data. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Florida, USA, 1777–1780.
- Schnabel, R., Wahl, R., and Klein, R. (2007). Efficient RANSAC for point-cloud shape detection. *Computer Graphics Forum* 26(2), 214–226.
- Tao, C. V. and Li, J., Eds. (2007). *Advances in Mobile Mapping Technology*. Taylor and Francis Group, London, UK.
- Vosselman, G. and Maas, H.-G., Eds. (2010). *Airborne and Terrestrial Laser Scanning*. Whittles Publishing and CRC Press, Scotland, UK.
- Wang, C., Tanahashi, H., and Hirayu, H. (2001). Comparison of local plane fitting methods for range data. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 663–669.
- Weber, C., Hahmann, S., Hagen, H., and Bonneau, G.-P. (2012). Sharp feature preserving MLS surface reconstruction based on local feature line approximations. *Graphical Models* 74(6), 335–345.
- Woo, H., Kang, E., Wang, S. Y., and Lee, K. H. (2002). A new segmentation method for point cloud data. *International Journal of Machine Tools and Manufacture* 42(2), 167–178.