Modelling characteristics of the urban form to support water systems planning

Peter M. Bach[1,2,3*], Ana Deletic[1,4], Christian Urich[1] and David T. McCarthy[1,5]

[1]*Monash Infrastructure Research Institute, Civil Engineering Department, Monash University, Clayton VIC 3800, Australia*

[2] *Swiss Federal Institute of Aquatic Science & Technology (Eawag), Überlandstrasse 133, 8600 Dübendorf, Switzerland*

[3] *Institute of Environmental Engineering, ETH Zürich, 8093 Zürich, Switzerland*

[4] *School of Civil and Environmental Engineering, University of New South Wales, Sydney NSW 2052, Australia*

[5] *Environmental and Public Health Microbiology (EPHM) Laboratory, Monash University, Clayton VIC, 380, Australia*

**Corresponding author email:* peterbach@gmail.com

Highlights

- We present a model based on planning regulations to recreate urban form characteristics
- We calibrate the model to three case studies, each using a different source of calibration data
- Model performs reasonably well in reproducing total impervious fractions and roof areas
- Model enables investigation of the links between urban planning and water management

Abstract

A spatial model is presented, based on urban planning concepts for abstracting urban form characteristics in new and existing areas. Requiring input maps of land use, elevation, population and parameters from planning regulations, the model conceptualises (on a spatial grid) attributes including impervious fractions, allotment geometry and roof areas among other relevant characteristics for integrated urban water management. The model is calibrated to three different Melbourne districts, varying in size (10–60km^2) and land use. Performance was evaluated by comparing modelled outputs with observations of total dwelling count, employment and spatial distribution of impervious fractions and residential roof areas. Results not only highlight reasonably good prediction, particularly with spatially variable indicators such as impervious area across all case studies, but also logical contrasts and consistency in the chosen planning parameters across the different case study districts. Discrepancies highlight aspects needing improvement and potential for exploring auto-calibration and model sensitivity.

Keywords

urban planning, integrated urban water management, spatial modelling, land use, impervious fraction, site planning

Highlights

- We present a model based on planning regulations to recreate urban form characteristics
- We calibrate the model to three case studies, each using a different source of calibration data
- Model performs reasonably well in reproducing total impervious fractions and roof areas
- Model enables investigation of the links between urban planning and water management

Software Availability

Name of Software: UrbanBEATSv1.0 – Urban Planning Module (*md_delinblocks.py* and *md_urbplanbb.py* of the **Urban B**iophysical **E**nvironments **a**nd **T**echnologies **S**imulator)

Contact Address: Peter M. Bach, Monash Infrastructure Research Institute, Department of Civil Engineering, 23 College Walk, Monash University, Clayton 3800 VIC, Australia. Tel: +61 4 3217 5283. Email: peterbach@gmail.com

Year first available: 2016

Supported Platform(s): PC (Windows 7, 8, 10), Mac, Linux

Program Language: Python 2.7

Program Size: ~78MB (source files for modules are ~1MB)

Availability: Contract corresponding author to obtain full software, also visit www.urbanbeatsmodel.com for updates.

Cost: Free (GNU General Public License)

1. Introduction

With the emergence of urban ecology in recent decades (Niemelä, 1999, Grimm et al., 2008) and over half of the world's population now living in urban areas (United Nations, 2012), cities have become an important focal point in future sustainable development. Understanding the impact that urban planning can have on environmental outcomes has been of interest in the last two decades (e.g. Pauleit and Duhme, 2000, Alberti et al., 2007). Research has uncovered intricate interactions between urban form and water infrastructure, which include, for example, the effects of land use planning (Lee et al., 2009), impervious cover (Arnold and Gibbons, 1996), density, street layout and residential neighbourhood design (Stone, 2004) on stormwater runoff, water quality, water supply security and other aspects that affect ecosystem services and the overall liveability of cities (Vlachos and Braga, 2001). Despite evolution of urban and water systems planning disciplines over the last few decades (Klosterman, 1997, Brown et al., 2009, Gurran, 2011) towards becoming more complex and 'wicked problems' (Rittel and Webber, 1973, Campbell, 1996, Gauthiez, 2004), considerable advancements have also been concurrently made in the numerical and computational tools to support this process (Geertman and Stillwell, 2004, McIntosh et al., 2007, Bach et al., 2014).

Following the advancements in Geographic Information Systems (GIS), researchers have acquired new and efficient ways of generating, manipulating and communicating spatial information (Harris and Batty, 1993, Chang, 2010, Eggimann et al., 2017). The underlying concepts of spatial data processing have since found their foothold in many existing urban water models and quantitative studies (comprehensive reviews are offered by both Elliott and Trowsdale, 2007 and, Bach et al., 2014). Obtaining and preparing maps of desired urban information as model input (e.g. impervious cover, roof areas, housing demographics, land surface cover) is often a laborious and time-consuming process and fraught with errors and uncertainty that may have originated from the initial digitization or drawing process. Sometimes the information is also non-existent (e.g. a scenario of a future urban development). As such, more systematic and pragmatic methods are often encouraged in the integrated modelling literature (see e.g. Bach et al., 2014, Lerer et al., 2015, Eggimann et al., 2017) that are sufficiently detailed to serve its desired purpose. Of the variety and diversity of studies in the literature, three prominent groups of methods have been identified: (1) empirical relationships, (2) conceptual techniques and (3) procedural methods.

Empirical relationships are used, for example, to estimate impervious surface cover from basic geographic information such as population or land use (Butler and Davis, 2004, Majid, 2006, Chabaeva et al., 2009). Such techniques are also common in assessments of centralised water infrastructure (see e.g. Fu et al., 2009, Sitzenfrei et al., 2013) and urban ecology (e.g. Uuemaa et al., 2005, Alberti et al., 2007) where impact of urbanisation on the natural environment is of interest. In contrast, there are also more complex integrated models that require users to conceptualise the urban landscape in greater detail, either as a subset of demographic input parameters or by selecting suitable templates from a pre-defined database and matching them to available geographic data. Examples of such models include Aquacycle (Mitchell et al., 2001), City Water Balance (Last, 2010) and the ReVisions framework (Ward et al., 2012). Quantitative studies by Bach et al. (2013a) and Stone (2004) also demonstrate how urban form can be conceptualised to assess their interaction with specific urban water system characteristics. A third, but less common methodology (in current urban water modelling research), involves procedural algorithms, i.e. geometric rules (e.g. space syntax, see Hillier and Hanson, 1984) that are used to generate highly detailed geometry of the urban environment, but are also more computationally intensive (e.g. Parish and Müller, 2001, Vanegas et al., 2012). Procedural methods have the potential of generating a much greater level of spatial detail that can support the increasing complexity of integrated urban water models. For example, applications of procedural algorithms by Urich and Rauch (2014) and Mikovits et al. (2014) demonstrate how this richness of spatial information can be used to explore climate and flood adaptation strategies.

Modelling the planning of urban water systems has been increasingly embracing exploratory modelling techniques (Bankes, 1993), evidenced by recent work in both models of the biophysical environment (Sitzenfrei et al., 2010, Urich and Rauch, 2014) and social water system (De Haan et al., 2016). Recent reviews also highlight a progression towards greater participation of affected stakeholders (Voinov and Bousquet, 2010, Bach et al., 2014, Voinov et al., 2016). The success and robustness of these modelling exercises depends not only on an accurate representation of the spatial environment that is being simulated (suited to the planning objective and that stakeholders can relate to), but also on the computational efficiency of these models. Although conceptual methods are more computationally efficient than procedural algorithms, their level of spatial detail is constrained by gross simplification (using highly aggregated parameters and/or limited number of pre-defined templates). As such, their flexibility, transferability and level of realism become questionable. Conversely, procedural algorithms, which are also grounded in architecture and urban planning theory, offer highly detailed representation of urban space, but can require a large amount of input data and powerful hardware or cloud-based solutions when simulating large urban districts.

To cope with the rapidly growing needs for integrated urban water management and the collaborative nature that planning has evolved into (Klosterman, 1997, Voinov and Bousquet, 2010), models should remain pragmatic (Bach et al., 2014), but bridge language, knowledge and communication across disciplines. Designing sustainable urban water technologies or water management policies has embraced the need for better integration with the urban form and demographics and accounting for local context and spatial variability to more effectively harness the multiple benefits that these solutions provide (Kuller et al., 2017). This must not only consider greater and more flexible spatial detail in models, but concurrently make them pragmatic and computationally efficient to support an exploratory process (Bankes, 1993, Urich and Rauch, 2014), facilitate improved dialogue and understanding of interactions and nuances between urban planner, water managers and other stakeholders throughout the process (Tewdwr-Jones and Allmendinger, 1998). Conceptual methods oversimplify the spatial detail with many assumptions and procedural methods are complex and deeply rooted in the architectural and urban planning disciplines. However, we see a necessity in their combination and exploring a new hybrid approach to spatially representing the urban environment. Such a combination leverages the advantages of both conceptual (in terms of simplicity and computational efficiency) and procedural methods (in terms of closer relation to architectural and urban planning language). Although not as prevalent in the urban water literature, the concept of using planning regulations to create abstractions of urban form has been investigated in the energy sector to improve allotment-scale energy calculations for city-scale decision-support models (Yamaguchi et al., 2007, Hargreaves et al., 2017, Salter et al., 2017). Many of these techniques, however, limit the representation of urban form to a pre-defined subset of commonly occurring neighbourhood blocks. Our technique differs in that it does not use pre-defined archetypes, but rather generates the urban form based on geographic input data and planning parameters, which are specified in the form of distributions to account for inherent spatial variability.

Although we previously demonstrate a simpler conceptual approach, which uses planning regulations to conceptualise the urban environment (see Bach et al., 2013a), there are a number of shortcomings: (1) it cannot be adapted directly to real-world data due to its non-spatially explicit nature, (2) it does not cover enough diversity in land use planning both in terms of variety of land uses (e.g. residential, non-residential) and variability within a single land use type (e.g. residential houses or apartments). Furthermore, many of the concepts, whilst they are representative of typical residential urban forms, have neither been validated against real-world data nor been rigorously supported by urban planning theory. In this paper, we build upon this initial concept by developing and testing a more advanced *Urban Planning Module* for characterising the spatial urban environment that, whilst largely a conceptual representation, incorporates more extensive procedural modelling elements. More specifically, this study focuses on:

1.) developing detailed stochastic procedural algorithms for conceptualising urban environments of diverse land uses and varying demographics, which are grounded in urban planning and design;
2.) testing, verifying and validating the algorithms on a range of case studies that represent different regions of a city (e.g. inner-, sub- and peri-urban) and different types of developments (e.g. dense, sprawl, new development);
3.) understanding the differences in model performance when calibrated using different sources of data; and
4.) evaluating the consistency of model structure and input parameters across the range of urban case studies.

The algorithms presented in this paper are part of the *Urban Planning Module* of a much larger planning-support tool known as UrbanBEATS (the **Urban B**iophysical **E**nvironments **a**nd **T**echnologies **S**imulator). We emphasise that the focus of this study is solely on the presentation and validation of the *Urban Planning Module*. One of its possible broader application (i.e. as part of UrbanBEATS) is illustrated in other work (e.g. Bach et al., 2013b, Bach et al., 2015a).

2. Model Description

2.1 General Overview

Our model aims to generate an approximation of urban and catchment characteristics that are sufficiently detailed for the integrated planning and management of urban water systems. This is advantageous in that it avoids large and highly detailed data requirements, especially in cases where data is non-existent (e.g. new urban areas). However, a key difference of our approach in contrast to existing work (e.g. Mitchell et al., 2001, Fu et al., 2009, Last, 2010, Ward et al., 2012), is that the approximation of spatial information is primarily driven by parameters that relate more closely to urban planning theory, architectural standards and existing regulations. As such, they have an added advantage of allowing users to explore how, for example, changing planning policy can affect the urban environment and consequently impact the urban water system.

The Urban Planning Model is subdivided into two parts (see Figure 1):

- **Part 1:** Spatial Delineation, where spatial data are read and collated and catchment characteristics is derived and;

- **Part 2:** Abstraction of Urban Environment: where the abstraction of urban characteristics takes place.

Three spatial input data sets are essential – raster maps at 10m x 10m or finer resolution of: (1) Land use (using a custom water-centric classification detailed in Bach et al., 2015b), (2) Population (measured as density and obtainable from the national census) and (3) Elevation. An optional input map depicting the local waterways can be provided to aid in catchment delineation. In addition to the spatial input data, aggregate or statistical data on employment within the region (if commercial and industrial areas are present), local statutory planning regulations, ordinances and other available demographic information (e.g. household occupancy, car ownership, average floor space per person) are required as input or to aid model calibration (key parameters are summarised in Table 1).

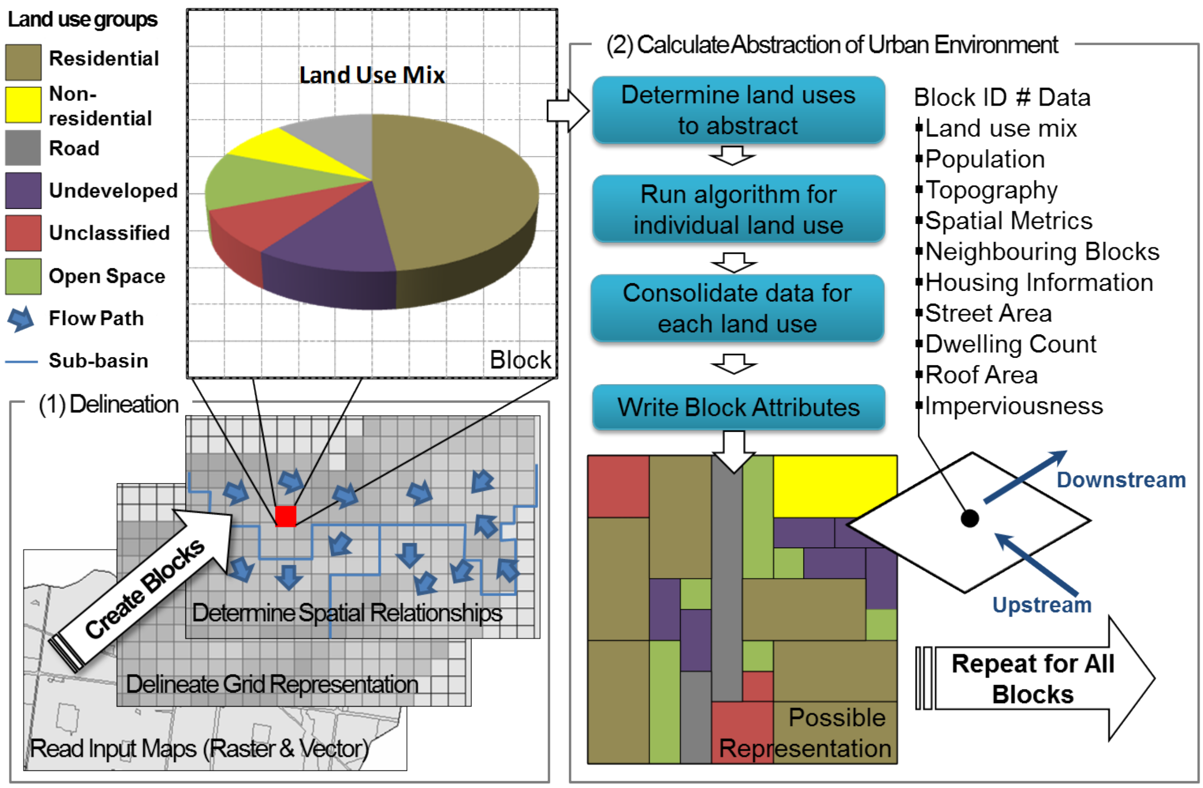

Figure 1. Overview of the developed model; basic steps of its two modules; definition of a "Block", the smallest spatial unit and categorization of different land use groups (each with different sub-categories explained in Bach et al., 2015b) considered

2.2 Spatial Delineation

2.2.1 Block Creation

The spatial representation in the model is based on a coarse square grid, which is common in urban land use and transport models (e.g. Waddell, 2002; Sitzenfrei et al., 2010). Its simple geometry and data management structure enables rapid computation. The finest explicit spatial unit of the model, i.e. a cell of this coarse grid, is referred to as a 'Block' (which has a resolution of 200m x 200m up to 1000m x 1000m defined by the user) and contains the detailed information about land use mix, demographic, biophysical and modelled urban characteristics. During *Block Creation*, the model retrieves, for each Block's spatial extent, the input raster data from each input map (see Figure 1 (1)). The input data is processed into descriptive statistical indicators (e.g. proportions of land uses, total population, maximum and minimum elevation etc.) as well as spatial metrics (e.g. number of land use patches, average size of patches, land use diversity). Unlike conventional grid- or raster-based models in the water management discipline (e.g. Bates and De Roo, 2000, Farolfi et al., 2010, Ziadat et al., 2012) where information is often aggregated, this map is vector-based and each Block is underpinned by a database of these aforementioned statistical and spatial attributes. As such, the procedure avoids aggregation and instead retains the richness of information from the input data in a conceptualised form (see illustration of 'Block' in Figure 1 (2)).

Land use data is summarised as a mix of different land use categories (LUC) and their relative area within the Block's boundary $A_{i\text{-}LUC}$ rather than a single dominant category. Population density is converted based on the available residential area and expressed as a total value for that Block P_i. Although elevation is expressed as a Block average and used solely for delineating gravity-driven water flow directions and catchment delineation, information about slope and variability in topography is also calculated. Finally, if a waterway map is included, the model tracks its flow network and writes, to each Block, a binary number that denotes the presence (value = 1) or absence (value = 0) of a waterway reach.

2.2.2 Flow-path and Catchment Delineation

In the event that model outputs are to be used for management of urban drainage systems, information on sub-catchment boundaries (or 'sub-basins' as they are referred to in the model) and water flow paths is essential. As such, using the coarse grid of Blocks and average elevation, the model runs a D8 algorithm (O'Callaghan and Mark, 1984) to determine the flow paths across the map. Flow direction in the method is based on the single largest elevation drop in the Moore neighbourhood (i.e. all eight adjacent Blocks around the central Block – see e.g. Batty, 2007). Although D8 is normally applied to fine-resolution raster data (preferably of 1m to 10m) and the delineation of catchments in urban environments is more challenging due to the presence of subsurface pipe networks controlling flow directions, previous studies have demonstrated the feasibility of this method at the coarser resolution (using 500m block resolution) (Bach et al., 2013b, Bach et al., 2016) and further testing is conducted in this paper (later in Section 4.1).

Blocks linked by a single connected flow path network are grouped to form 'sub-basins', which represent regional sub-divisions of the input map's spatial extent. A sub-basin is found when a local sink is encountered in the map. The criterion for a sink in the model requires that the elevation of the

Block containing the sink in question is lower than its eight adjacent neighbours and 16 cells adjacent to these neighbours. The assumption is made that at these positions in the urban environment, significant water infrastructure or natural features are usually present. This could, for example be a large conveyance pipe or major water body. For local urban drainage management, these sub-catchments are usually large enough. Using an optional waterway input map will alter the algorithm; delineation of a sub-basin is, instead, terminated once the modelled flow paths reach a Block that contains part of the waterway (i.e. a natural sink). The same criterion for detecting a sink still holds. As the number of sub-catchments will depend on the chosen Block size (since the D8 algorithms is sensitive to the resolution), a general rule-of-thumb is to select the Block size that will yield a sub-catchment layout most appropriate to the modelling objectives (determined by existing information through calibration or by expert judgment).

2.3 Abstraction of Urban Environment

After spatial delineation, the map is scanned by the model Block-by-Block and various abstraction algorithms are called based on the different types of land uses present in each Block. There are, in total, five different abstraction algorithms, one for each of the land use groups (with the exception of 'Undeveloped' land) listed in Figure 1 (top-left). Most of the concepts are based on urban planning literature and local documents. Many of the simplifying assumptions in the method were also validated in discussions with urban and water planners and from a professional consulting firm. We discuss the input and calibration parameters and cover algorithms for each land use in detail in the following sections.

The abstraction algorithms have been developed based on architectural and urban planning standards (e.g. Alexander et al., 1977, Reinhardt and Trudel, 1979, United Kingdom Development Agencies, 1986, Basingstoke & Deane Borough Council, 2008, De Chiara et al., 2009, Austroads, 2010) and planning regulations (e.g. DPCD, 2006, Victorian Building Commission, 2006, City of Santa Maria, 2013, CBSC, 2013). As such, to ensure familiarity of the user with the overarching model structure, parameters were carefully selected. Table 1 lists all key parameters in the model, organised into their respective land use groups and includes their notation, description and key planning regulation or source(s) used for this study. Their use is explained alongside key equations in the proceeding sections. The choice of which parameter is calibrated and which parameter is specified as input is dependent on the case study, input data availability and quality. We show how we distinguish between these in the calibration of the model in Section 3.2.

Table 1. *Summary of Key Input and Calibration Parameters for different land use groups, their notation, description and key planning regulation/source. There are no parameters for 'Undeveloped' Land Use.*

Land use group	**Parameter name [Units] ***	**Symbol / notation**	**Description**	**Key planning regulation / source(s) ^**
Residential (sub-category: Houses)	Occupancy range [persons]	O_{avg} ; O_{max}	Number of people per household, specified as a range based on an average and maximum	ABS (2009)
	Unit floor space [m2/person]	A_{person}	Internal living space per person, used to calculate gross floor area	ABS (2010)
	Max. number of floors [ ]	$N_{floors\text{-}max}$	Maximum number of floors allowable on the allotment	DPCD (2006)
	Front setback range [m] *	$d_{sb\text{-}front}$	Buffer distance of building envelop to allotment frontage	DPCD (2006)
	Side setback range [m]	$d_{sb\text{-}side}$	Buffer distance between building envelopes of adjacent lots	DPCD (2006)
	Max. number of car parks [ ]	N_{cp}	Maximum number of car spaces for a single house	Vic. Building Com. (2006)
	Min. driveway width [m]	$W_{driveway}$	Minimum width of allotment driveway	Vic. Building Com. (2006)
	Max. patio area [m^2]	A_{patio}	Maximum area of an outdoor patio, assumed paved surface	DPCD (2006)
Residential (sub-category: Apartments)	Avg. occupancy [persons]	O_{avg}	Average number of people per apartment unit	ABS (2009)
	Avg. apartment size [m^2]	A_{flat}	Average size of a single apartment unit	ABS (2010)
	Maximum number of floors [ ]	$N_{floors\text{-}max}$	Maximum number of floors allowable for apartment block	Local council regulations
	Outdoor communal space [%]	$p_{exterior}$	Additional external area for the apartment block (taken as % of gross floor area) - variable across developments, specify average	Local council planning schedules & site plans
	Average site setback [m]	$d_{sb\text{-}avg}$	Average buffer distance from apartment building envelope to the property edge	DPCD (2006)
	Parking arrangement	-	Proposed arrangement of parking (Options: On-site, off-site, vary)	Local council regulations
Non-Residential	Employment distribution [employees/ha]	$\rho_{jobs\text{-}LUC}$	Density of employment, expressed as a function of land area and dependent on type of non-residential land use.	ABS (2016)
	Maximum building coverage ratio [ ]	$PR_{max\text{-}LUC}$	Maximum allowable building footprint on site as a ratio to overall site area.	SGS Economics & Planning (2008)
	Non-residential parcel size range [ha]	$A_{par\text{-}min\text{-}LUC}$; $A_{par\text{-}max\text{-}LUC}$	Informs the sub-division of non-residentail land into individual estates, two categories: industrial and commercial/office land	SGS Economics & Planning (2008)
	Maximum number of floors [ ]	$N_{floors\text{-}max}$	Maximum number of floors allowable for non-residential building	Local council regulations
	Minimum site setback [m]	$d_{sb\text{-}min}$	Buffer distance between building envelope and edge of estate	DPCD (2006)
	Carpark requirements [bays per unit area or employees]	$p_{car/person}$; $p_{car/GFA}$	Average car parking space required, taken as a number of bays per employee for industrial and number of bays per floor area for commercial/offices (this varies across different industry types)	DPCD (2006)
	Parking lot dimensions [m]	A_{car}	Dimensions of a single car park	DPCD (2006)
	Loading bay allocation [m2/100m2 GFA]	$A_{bay\text{-}size}$	Required space provisions for loading and unloading of goods and/or services	DPCD (2006)
	% of green landscaping [%]	p_{green}	Proportion of the landscaping area taken as pervious, green space	Consultation of local planning schedules
Roads & Open Spaces	Frontage dimensions for residential [m]	W_{fp} ; W_{ns} ; W_{st}	Range of widths for footpath (FP), nature strip (NS) and street lane (ST) in residential collector streets	DPCD (2006), AUSTROADS (2010)
	Frontage dimensions for non-residential [m]	W_{fp} ; W_{ns} ; W_{st}	Range of widths for footpath (FP), nature strip (NS) and road lane (RD) in non-residential collector streets	DPCD (2006), AUSTROADS (2010)
	Highway reserve dimensions [m] *	W_{lane}; W_{buff}; W_{med}	Range of widths for major arterials and highway road lanes (LANE), median (MED) and side buffers (BUFF)	DPCD (2006), AUSTROADS (2010)
	% of green area of parks [%]	p_{green}	Proportion of park area that is green, pervious space	Defined by land use classification
Unclassified	Imperviousness of 'Unclassified' land use [%]	$p_{imp\text{-}UNC}$	Constant impervious fraction to apply to unclassified land if its area exceeds a given threshold (defined as a > X% of block size)	Consultation of local planning schedules

** A number of parameters in this list can be considered as either input or calibration parameter. The choice is dependent on the type of case study and availability and quality of input or calibration data. Table 2 distinguishes between Input and Calibration parameters for the purpose of this study.*

^ DPCD (2006) refers to the generic planning provisions across the state of Victoria, however, each local council within the metropolitan region prescribes a modified version of these regulations with specific schedules that apply only within their region of jurisdiction, these are useful for determining certain parameter values (e.g. maximum allowable floor areas).

2.3.2 Abstraction Algorithm for Residential Areas

The entire abstraction algorithm for residential areas is illustrated in Figure 2, which depicts the flow and creation of information from initial inputs of Block *i* population P_i and area of residential land use $A_{i\text{-}Res}$ to a range of output residential characteristics (e.g. allotment layout, impervious areas). Literature has placed greater importance on residential urban form as they occupy the majority of a city's urban space (Stone, 2004) and the algorithm for residential land use in this model are more detailed than those for other categories.

For the algorithm to run and development to occur on-site, the total area of residential land use $A_{i\text{-}Res}$ must exceed a minimum development area $A_{i\text{-}Res\text{-}min}$ i.e.:

$$A_{i-Res-min} \geq A_{person} \times O_{avg} \quad \text{Eq. 1}$$

where:

A_{person} – the average floor space per person in residential housing [m^2/person]

O_{avg} – the average occupancy [persons/household].

This threshold ensures that at least one single dwelling can be accommodated on-site. If this criterion is not fulfilled, no development occurs and the algorithm leaves the Block undeveloped. There are, subsequently, two sub-procedures for residential areas: one for houses (blue flowchart in Figure 2 starting at 'Design Houses') and the other for apartment blocks (orange flowchart in Figure 2 starting at 'Design Apartments'). The choice of sub-procedure depends on the areal and demographic site characteristics and is embedded in the concept of *Land Use Intensity* (LUI) (De Chiara et al., 2009). LUI is an index that defines the relationship between floor space, building footprint, open and parking space as well as a range of other characteristics on-site (illustrated for a residential allotment in Figure 3). The value of LUI can be looked up in a reference table (available in De Chiara et al., 2009) upon calculating the Floor-Area-Ratio (FAR), defined as:

$$FAR = \frac{A_{GFA}}{A_{i-Res-min}} \quad \text{Eq. 2}$$

A_{GFA} is the gross floor area (GFA) of a residential site [in m^2] defined as:

$$A_{GFA} = A_{person} \times P_i \quad \text{Eq. 3}$$

FAR determines not only site subdivision, but also building footprint and height. An illustration of this relationship is provided in Figure S1 in the Supplementary Material as a reference for readers, who are unfamiliar with the concept. Based on the FAR and corresponding LUI, a site is likely to comprise of detached or semi-detached dwellings, walk-up apartments and/or high-rise housing. To accurately determine which building typology to use, the algorithm must also satisfy a range of other spatial requirements or *design ratios*. These include ratios for the provision of open (OSR), liveable (LSR) and recreational spaces (RSR) as well as on-site parking (OCR, TCR). These are listed in Figure 2 in the *'Lookup other design ratios'* step and are illustrated in Figure 3 under 'Residential Space Definition'.

Once LUI has been calculated and a residential typology has been chosen (Houses if LUI <= 5.4 and Apartments if LUI > 5.4), the corresponding design ratios for the calculated LUI are used alongside the spatial information to subdivide and characterise the land. At this stage, all stochastic planning parameters are also prepared for Block *i* by sampling a value from their respective distributions (uniform or normal depending on the parameter type). For example, street width W_{st} is sampled from a uniform distribution $U([W_{st\text{-}min}, W_{st\text{-}max}])$ between a minimum and maximum user-defined value. Parameters that are embedded in and modified by the model iteratively through the algorithm are also initialised (e.g. ρ_{dw} – dwellings per allotment in the case of houses).

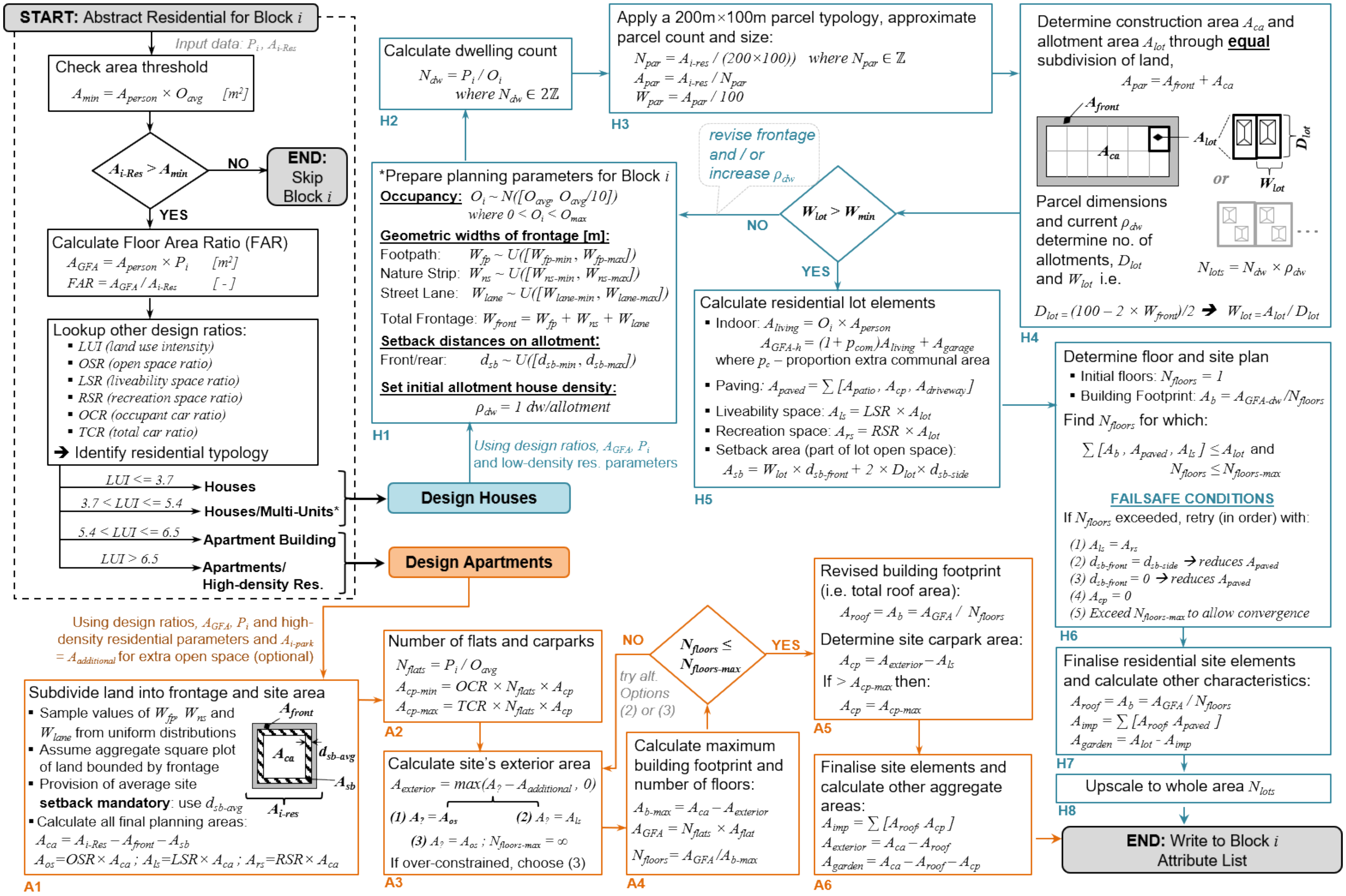

Figure 2. *Flowchart for the Urban Residential Land Use Abstraction Algorithm (refer to Nomenclature list for description of variables and to Table 1 for parameter list)*

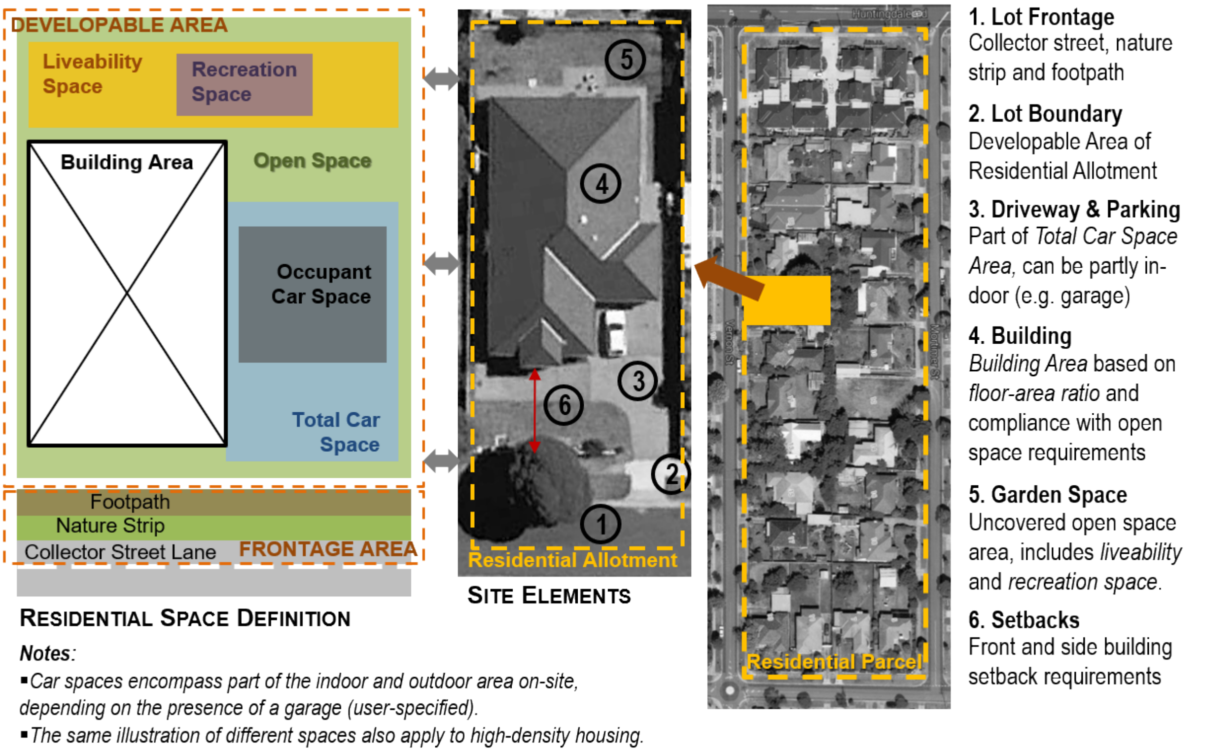

Figure 3. *Definition of different area ratios and site elements on a residential allotment*

2.3.2.1 Abstraction of Houses

Using the initial value of ρ_{dw}, the model calculates the number of dwellings N_{dw} based on the Block population. In the case of houses, a typical residential neighbourhood parcel (approximately 200m x 100m in dimension, such as the one show in Figure 3-right) is used as a template to subdivide $A_{i\text{-}Res}$. The number of possible parcels, N_{par}, is determined from total residential area $A_{i\text{-}Res}$ and rounded down to the nearest integer. We do not constrain the model to the dimensions of 200m x 100m, but rather use it as a guide for the approximate size and street density (the correct dimensions are recalculated in subsequent steps – see Figure 2 Box H4). Residential urban morphology typically adopts different types of spatial structures (Conzen, 2001, Stone, 2004, Siksna, 2006). Our algorithm uses the simplicity of a grid arrangement of allotments as we wanted to avoid the complex geometric computation of highly irregular, ‘unplanned’, curvilinear development patterns, which are commonly referred to as ‘organic’ in form (White et al., 2015), resulting from the spontaneity of urban growth. In the case of Melbourne, for which the model was initially developed, much of the grid-based residential structure is also prevalent in the inner city and suburban areas (Siksna, 2006). Therefore, a pragmatic decision was made to use a flexible grid-based typology that adapts to the different population and dwelling densities encountered in the land use and population data.

Once parcel dimensions have been determined, the model allocates an equal and even number of dwellings N_{dw} to each parcel and subdivides the land into frontage area (i.e. the collector street reserve) and construction area A_{ca}. Collector streets are represented as two road lanes in a single carriageway. Based on how the urban form is conceptualised, we only represent half of the total collector street reserve, which is the sum of footpath, nature strip and area of a single street lane (calculated individually using sampled footpath width W_{fp}, nature strip width W_{ns} and lane widths W_{lane} respectively). Construction area A_{ca} is defined as:

$$A_{ca} = A_{site} - (A_{fp} + A_{ns} + A_{lane}) \quad \text{Eq. 4}$$

where:

A_{site} – total site area of a single parcel of residential land [m^2]

A_{fp} – total footpath area of parcel (based on sampled parameter of footpath width W_{fp}) [m^2]

A_{ns} – total nature strip area of parcel (based on sampled parameter of nature strip width W_{ns}) [m^2]

A_{lane} – total area of street lane of parcel (based on sampled parameter of street width W_{lane}) [m^2].

A_{ca} is equally subdivided into allotments for which the dimensions depth D_{lot} and width W_{lot} can be determined and checked against a minimum acceptable value W_{min} (taken here as 10m based on inspection of aerial photos of Melbourne inner- and suburban houses). If this condition is met, the model proceeds to characterising a single allotment (note that every allotment in a Block has the same characteristics). If the condition is not met, the model increments ρ_{dw}, revises frontage dimensions (W_{fp}, W_{ns}, W_{lane}) and repeats the procedure until conditions have been met.

A single allotment (as illustrated in Figure 3) comprises of building footprint (which can have several floors of indoor floor space), on-site paving (patio, car park and driveway) as well as garden space (recreational space and other spaces due to arrangement of on-site elements). To satisfy urban design requirements for residential housing, liveability space (LSR) and recreational space ratios (RSR) must be met. The respective areas for these (A_{ls} for LSR and A_{rs} for RSR) are calculated from the allotment area using the corresponding ratios based on the FAR of the site (from Eq. 2). Planning provisions also require for buildings to be set back from the front and side allotment boundaries by a prescribed distance (denoted by $d_{sb\text{-}front}$ and $d_{sb\text{-}side}$). Within the setback area A_{sb}, no building construction is permitted (Stone, 2004, DPCD, 2006), but these areas fulfil part of the open space requirement. Parking space and patio area is user-defined for residential houses and are characterised as paved areas abutting the building footprint.

Abstracting a single allotment is an iterative process. In essence, the model searches for the minimum number of floors N_{floors} (not exceeding a maximum $N_{floors\text{-}max}$) such that the sum of building footprint, paved area and liveability space A_{ls} is less than or equal to the total allotment area A_{lot}. Building footprint is calculated by dividing on-site gross floor area $A_{GFA\text{-}h}$ (comprising living space and garage area) by the number of floors N_{floors}. Paved area is the sum of driveway, uncovered parking and patio area. Driveways are assumed to run along the width of the front setback, a similar approach to (Stone, 2004), which is logical as it provides car access from building to front property edge. Its width is taken as that of a car park width. The patio is a resident-defined area and needs to be calibrated as this parameter can vary from house to house depending on the occupants. Car parking can either be open or in a garage. Custom parking dimensions can be obtained from local building and planning regulations (DPCD, 2006, Victorian Building Commission, 2006, CBSC, 2013) or master planning documents for a new development. We use their frequently specified width of 2.6m and depth of 4.6m.

If area requirements cannot be fulfilled upon the first iteration, the model uses a series of failsafe conditions (Figure 2 Box H6) to converge upon realistic values of building footprint, paving and open space. These are applied in order and include: (1) setting the liveability space equal to the recreational space, (2) reducing the front setback to the value of side setback, (3) removing the setbacks entirely (which can occur in some special cases in high-density areas) or (4) removing the on-site parking. If, after applying all four possible conditions, the model still does not meet requirements, then parameters and site characteristics are judged to be over-constrained and a flag is raised to the user (model testing has shown that this occurrence is rare and mainly due to input data error). The model then backtracks to searching for the number of floors N_{floors}, allowing it to exceed the maximum floor limit by one unit (one iteration at a time). The algorithm then resumes to trialling the four alternative conditions for the newly exceeded maximum floor limit until convergence has been achieved. The last calculation step for a single allotment evaluates roof area A_{roof} (which can be determined once the final value of N_{floors} has been determined), on-site impervious area A_{imp} (the sum of roof and paved areas) and the total garden or pervious area A_{garden} (by subtracting the impervious area from total lot area).

2.3.2.2 Abstraction of Apartments

Although the concepts are similar, the algorithm for residential apartments differs slightly from that of houses (see Figure 2 orange flow chart beginning with "Design Apartments" for full algorithm). Rather than individual allotments and a subdivision of the site, the model assumes a single aggregate plot of land bounded by the frontage upon which to place apartment buildings, car parking and open space. The provision of site setback is also mandatory and there are no conditions that would reduce this area. The model is also required to fulfil all five design ratios (OSR, LSR, RSR, OCR and TCR) to ensure that adequate land subdivision has been achieved (De Chiara et al., 2009).

After planning parameters have been sampled and the construction area A_{ca} identified using Eq. 4, the model proceeds to determine the number of apartment units (or 'flats'), N_{flats}, and the minimum and maximum allowable on-site carpark area (Figure 2-Box A2). Similar to houses, N_{flats} is calculated by dividing the total Block population P_i by the prescribed average occupancy for apartments O_{avg}. Then, the required exterior area for the site (to comply with OSR), $A_{exterior}$, is calculated. In the case of high-density residential, users have the option to leverage the presence of nearby parks and open space land use (i.e. only if $A_{i\text{-}park} > 0$) to fulfil these requirements. These provide additional open space area $A_{additional}$ that can ease the occasional difficulty of achieving on-site requirements and, in high-density areas, facilitate the compliance of floor height and building design ratios (Figure 2-Box A3). This step marks the beginning of an iterative loop that, once again, aims to satisfy design ratios and floor restrictions by trying several options.

Maximum building footprint is calculated by subtracting $A_{exterior}$ from A_{ca}. GFA is obtained using Eq. 3, but as the product of A_{flat} (the unit area) and N_{flats} (the number of on-site units). The number of floors N_{floors} is then determined by division of GFA by maximum allowable building footprint $A_{b\text{-}max}$ (Figure 2-Box A4). If N_{floors} are within acceptable range, final building footprint, which is analogous to roof area A_{roof} in the model, is evaluated. On-site carpark area $A_{carpark}$ is taken as the lesser of either maximum required carpark space $A_{cp\text{-}max}$ or ($A_{exterior} – A_{ls}$), which is calculated as LSR × A_{ca} (Figure 2-Box A5). Finally, other site elements are also determined (Figure 2-Box A6) including impervious area (as the sum of roof and carpark area) and garden space (calculated by subtracting A_{roof} and $A_{carpark}$ from A_{ca}).

2.3.3 Abstraction Algorithm for Non-Residential Areas

Non-residential land uses distinguish between industrial and commercial estates and high-density office buildings. These uses are diverse in appearance with a variety of different building types (e.g. process plants, large open storage yards or closed warehousing, factories, mills or high-rise tower). To generalise the urban characteristics, we distilled examples from a broad range of urban planning literature and policy documents from around the world (United Kingdom Development Agencies, 1986, DPCD, 2006, SGS Economics & Planning, 2008, Basingstoke & Deane Borough Council, 2008, Russ, 2009, City of Santa Maria, 2013). The resulting algorithm for non-residential land use is illustrated in Figure 4 (brown-coloured boxes N1 to N9) and explained here.

The smallest urban typology for non-residential land use is the '*estate*', which is a term commonly used in the literature (United Kingdom Development Agencies, 1986). Key site elements for non-residential estates are illustrated in Figure 5 and include: building footprint, on-site car parking, loading bay and landscaping (comprising use of concrete 'grey' features or natural green elements depending on the type of land use). A single estate in the model is assumed to have frontage bounding on two of its four edges (as estates can often be abutting each other, but are also provided enough access to the site to not disrupt business operation).

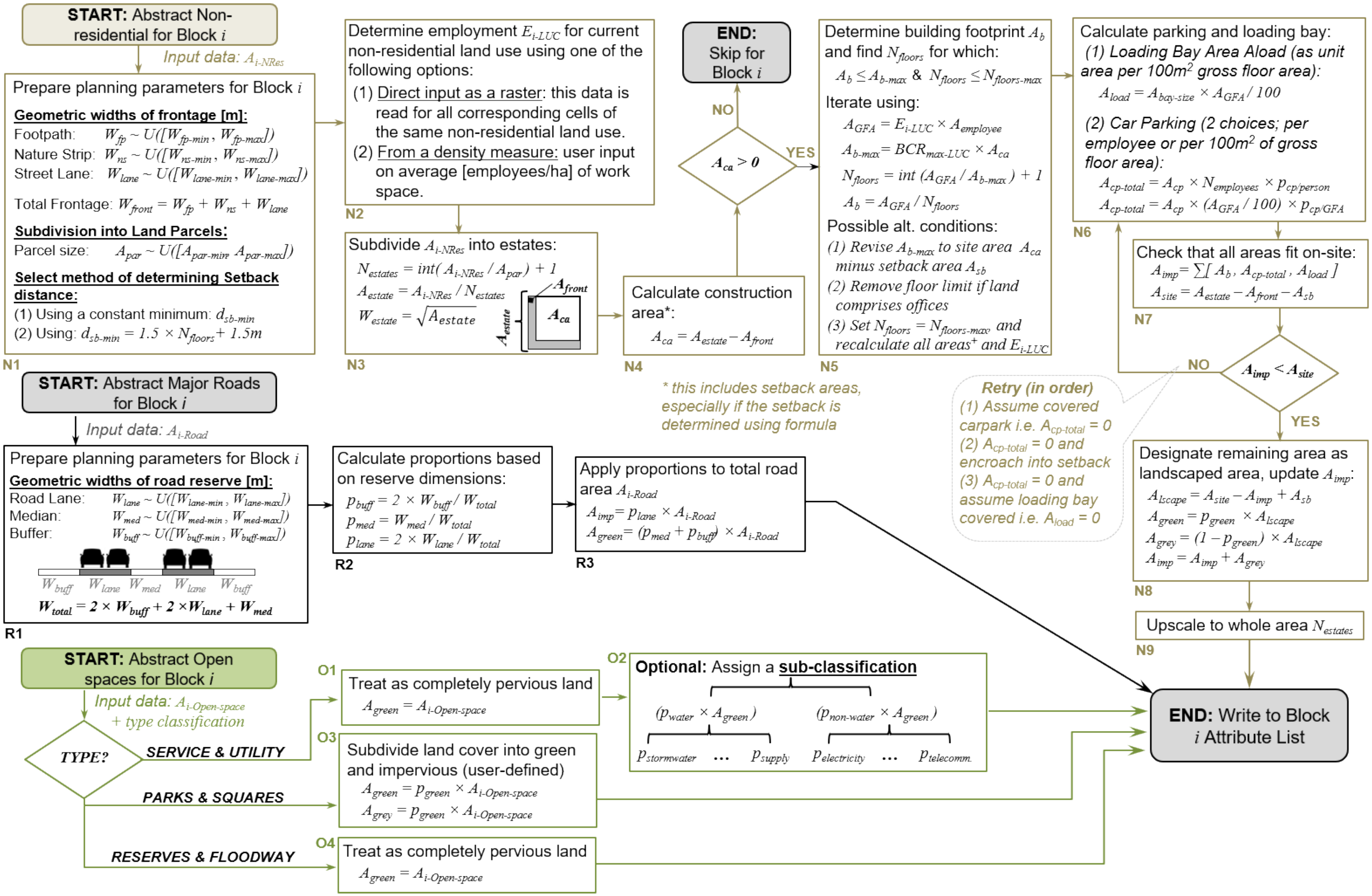

Figure 4. *Flowchart of abstraction algorithm for Non-Residential and other land uses (refer to Nomenclature list for description of variables and to Table 1 for parameter list)*

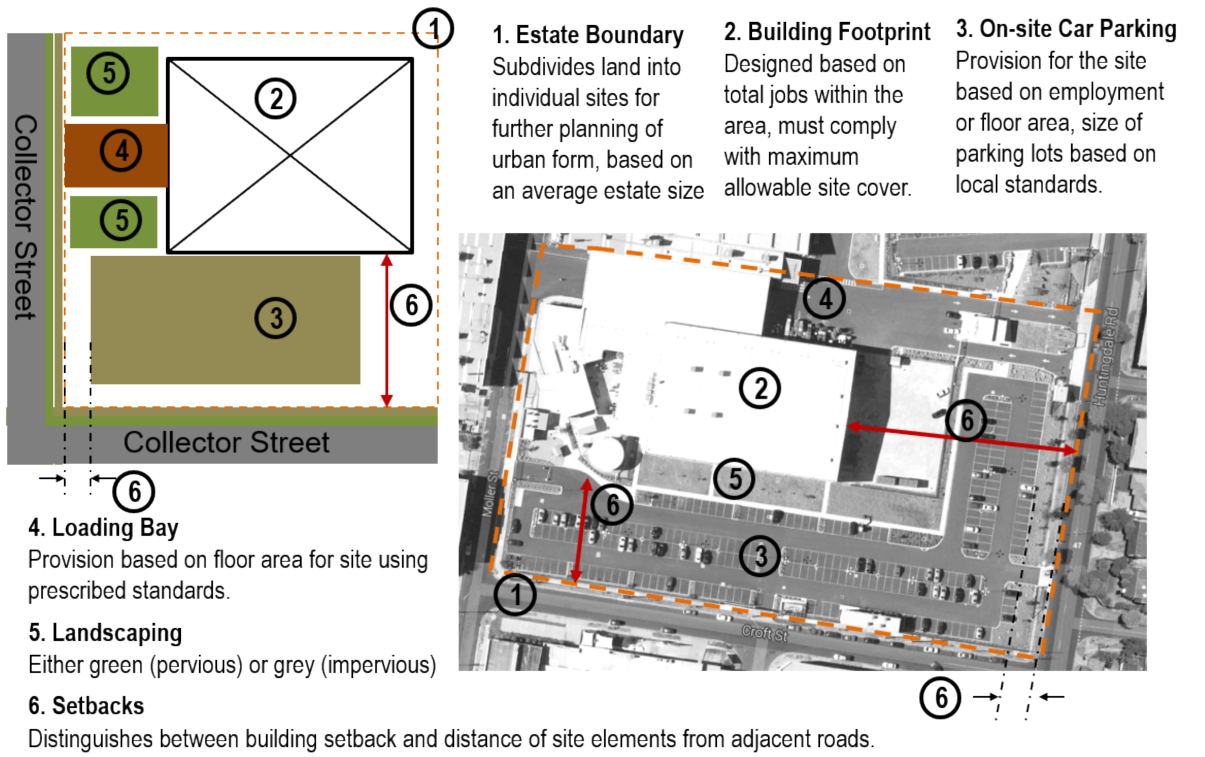

Figure 5. *Definition of different non-residential site elements on a non-residential estate*

The same algorithm is used across all non-residential land use types (which will only differ in terms of their chosen input and calibration parameters – details of which are shown in Table 1). To determine whether a non-residential area is suitable for development is slightly more involved than its residential counterpart (defined earlier by Eq. 1). The model begins by preparing all planning non-residential planning parameters for Block *i* (see Figure 4-Box N1). Widths of frontage dimensions (footpath, nature strip and street) are sampled from their respective distributions. A standard land parcel size (used to subdivide the total land use area into estates) is sampled from the user-defined range. A convention for calculating site setback is chosen by the user as either as a constant minimum value or calculated based on the number of floors.

Input data of non-residential land use area of a given type ($A_{i\text{-}NRes}$) is provided. In contrast to residential land, the algorithm of which is driven by population data P_i, non-residential land works with employment, specified as the number of people working in each of the industrial and commercial land use, $E_{i\text{-}LUC}$. Employment for an industrial or commercial estate is essential information for site planning (United Kingdom Development Agencies, 1986, Phil McDermott Consultants, 2006). It also represents a crucial parameter for linking non-residential land use with its site characteristics (e.g. GFA provision) and, consequently, water management (see e.g. Arbues et al., 2010, Tchobanoglous et al., 2013).

Employment can be determined in the model in two ways (see Figure 4-Box N2): (1) as direct input using a raster map or (2) from a density measure, where employees per hectare is specified for each type of non-residential land use. The second method was used in this study as spatial data on employment is challenging to obtain in Melbourne. Although the Australian census reports on employment, the spatial location of this data is based on households rather than workplace. Furthermore, establishing a ‘journey to work’ relationship is often challenging as there are arguably different trends of relationship between social housing preferences and workplace location depending on location in a metropolitan area (Masuya et al., 2008, Aguilera, 2010). We can, however, obtain an estimate of employment at an aggregate level from available statistical reports (Australian Bureau of Statistics, 2016) and development master plans for newer areas. This allows calibration of employment density, initial estimates for which are available in planning literature and frequently reported by local government as part of summary reports (refer to tables in United Kingdom Development Agencies, 1986, Phil McDermott Consultants, 2006, SGS Economics & Planning, 2008).

Once employment for a non-residential land use $E_{i\text{-}LUC}$ for the site is known, the area of that land use $A_{i\text{-}LUC}$ is subdivided into individual estates using the average parcel size A_{par} (Figure 4-Box N3). The construction area A_{ca} of the estate is determined by subtracting frontage area, calculated in the same way as for residential land use (see Eq. 4 and Figure 4-Box N4), but considered here along two edges of the estate instead (as illustrated in Figure 5). At this point the model assesses whether to proceed with site layout. If A_{ca} is less than zero, the model deems there to be not enough land to develop the site, thereby terminating the algorithm and treating the whole area as undeveloped pervious land. On the contrary, if A_{ca} is greater than zero, the model proceeds to characterise the site of a single estate. Similar to the residential algorithm, every estate of a non-residential land use in a Block is assumed to have the same characteristics.

An iterative process is used to determine the building footprint based on a suitable number of floors (Figure 4-Box N5). GFA for the site is obtained by multiplying $E_{i\text{-}LUC}$ with a unit area $A_{employee}$ specified by the user. Building footprint is then evaluated by dividing GFA by number of floors N_{floors}, which is initially assumed as single storey and incremented if necessary. Two conditions need to be met for the model to proceed: (1) building footprint A_b must be less than a maximum allowable footprint determined by the maximum *building coverage ratio* $BCR_{max\text{-}LUC}$ and (2) the total number of floors N_{floors} must be below the allowable limit $N_{floors\text{-}max}$. BCR is defined as:

$$BCR = \frac{A_b}{A_{site}} \qquad \text{Eq. 5}$$

where:

A_b – on-site building footprint [m^2] and;

A_{site} – net developable site area [m^2].

Regulations usually specify a maximum allowable BCR (DPCD, 2006) as a general guide, but typical limits for specific non-residential uses can also be obtained elsewhere (SGS Economics & Planning, 2008). If a suitable number of floors cannot be found due to the limit imposed by $N_{floors\text{-}max}$, three alternate conditions are tried by the model until convergence has been reached (outlined in Figure 4-Box N5).

With defined values of building footprint and site area, the model proceeds to calculating other site elements including car parking, loading bay and landscaping (Figure 4-Box N6). Car parking is commonly prescribed by local and regional government in their planning regulations. Values for the model can therefore be estimated for typical commercial and industrial uses either from the literature (e.g. United Kingdom Development Agencies, 1986, Russ, 2009) or local/regional regulation (e.g. DPCD, 2006, City of Monash, 2009, CBSC, 2013, City of Santa Maria, 2013). Units for parking requirements are either in 'bays per employee' or 'bays per unit GFA'. These same regulations also prescribe the requirements for loading bays on-site, the dimensions of which are usually scaled based on the size of the estate or its GFA. Incorporating parking, loading bay and building areas within the designated site area A_{site} (which will be less than A_{ca} due to the incorporation of setback area) is, once again, undertaken iteratively (Figure 4-Box N7) and guided by three alternate conditions that alter parking and loading bay area.

Once building footprint, car parking and loading bay have been successfully incorporated, the remaining steps required include the estimation of landscaped area, calculation of pervious and impervious surfaces on-site and up-scaling this information to the Block level across all estates (Figure 4-Boxes N8 and N9). Landscaped area is subdivided into impervious 'grey' landscape and pervious 'green' landscape based on the ratio p_{green} and is applied to the remaining area (i.e. setback area and leftover site area once building, parking and loading bay have been incorporated).

2.3.4 Abstraction Algorithm for Other Land Uses

Open spaces, major roads, undeveloped and unclassified land uses are less complex. The first two are illustrated in the flowchart in Figure 4 boxes R1 to R3 and O1 to O4. Apart from the collector street, nature strip and footpath widths defined for residential and non-residential algorithms, larger arterials such as boulevards or highways are characterised separately in the model. Using parameters of lane, median and buffer widths sampled from user-defined ranges, the model creates a hypothetical bi-directional, two-lane, dual carriageway road reserve (see illustration in Figure 4-Box R1). Unlike the collector streets, where half the reserve is represented (since residential and non-residential parcels are intended to be conceptually 'joined' to create the full reserve), the algorithm for major roads calculates the characteristics of the full reserve. Geometric dimensions can be obtained from national road design guidelines (e.g. Austroads, 2010) or local planning documents if specific provisions have been made for a particular development. Based on the widths, the model calculates proportions of buffer p_{buff}, median p_{med} and road lane p_{lane} (Figure 4-Box R2). These proportions are used to subdivide Block i's total road area, $A_{i\text{-}Road}$, into its sub-components and subsequently total pervious and impervious surfaces (Figure 4-Box R3).

Open spaces encompass three different types of land uses: *Parks & Squares*, *Reserves & Floodways* and *Service & Utility*. Depending on the land use type, different actions are taken in the algorithm. In the case of *Service & Utility* use (Figure 4-Box O1 and O2), the space is treated as completely open space, but a sub-classification of the land can be assigned if the user wishes to perform further assessment with this data. Sub-classifications offered include proportioning the land into area for urban water services (e.g. stormwater, water supply, sewage management) and area for non-water services (e.g. energy, telecommunications). For *Parks & Squares*, a user-defined percentage of green landscaping is used to determine how much of this area is pervious green space and how much is paved open space (Figure 4-Box O3). For *Reserves & Floodways,* all land is treated as completely pervious area (Figure 4-Box O4).

Undeveloped land is treated as completely pervious land, but is accounted for to ensure that the total land area within a Block is consistent with the input data. When the model encounters an area of undeveloped land, it treats it as pervious area and adds it to the Block's total pervious area when consolidation of the information takes place (see Section 2.3.5).

Finally, 'Unclassified' land use comprises land for which there is no concrete information available in the data to classify it into one of the other categories. This can occur for example if there is a patch of open space that has not been officially designated as a recreational or reserve space by the local planning authority. In the context of Melbourne (DPCD, 2006), we define some of the Unclassified land as having the local category 'Special Use Zone', which refers to land with special planning arrangements and unique features (e.g. landmarks or a major harbour). Therefore, in the model, two possible options can be taken:

(1) if the land area is smaller than a certain threshold (which we define based on a percentage of Block area), then the land can be merged into other land uses (e.g. open spaces or roads) that are present in that Block.

(2) if the land area is greater than a certain threshold, then the land can be treated as a special zone for which a constant value of impervious fraction can be defined by the user.

2.3.5 Consolidation of Urban Characteristics Data

The final step of all five algorithms involves the up-scaling and consolidation of data to the Block level. Aggregate or total values are calculated at the Block level before being written as attributes to the Block geodatabase. In residential areas (i.e. houses and apartments), this includes the calculation of total impervious areas (as the sum of all roofed and road, footpath and driveway areas both on and off-site), total roof areas (saved as a separate attribute), total garden space (subdivided into pervious area and its subset, liveable/recreational area) and total area of frontage/road reserve (to distinguish between private and public space). In non-residential areas, total impervious area (made up of all roofs, car park, grey landscaping and road and footpath paving), separate attributes for total roof and carpark area across all non-residential land uses and, once again, the total area of frontage/road reserve are calculated. For other land uses, most of this information has already been derived as a result of their simpler abstraction algorithms.

Once these attributes have been written for each Block, the model evaluates the total impervious area for each Block based on its subset of impervious areas. By dividing it by the Block's total area, we derive the "Block Total Impervious Fraction" (*Block TIF*), which is the primary metric we calibrate against in the later testing of this model. Other metrics that aggregate the information across various land use types calculated in the consolidation step also include: block total road area, block total irrigated green space (determined by the user as areas to include in the irrigation e.g. all residential gardens, parks), block total pervious area, block total roof area and a breakdown of land surface cover information (an additional feature of the model that is, however, beyond the scope of this study). The final output map therefore contains not only detailed information about the characteristics of individual allotments and estates, but also the aggregate information for each specific geographic region delineated by each of the model's Blocks.

3. Model Testing, Setup & Calibration

3.1 Case Study Description and Input Data

We set up and calibrated the model for three distinctly different case studies across Metropolitan Melbourne, Australia (shown in Figure 6). The Yarra Estuary Catchment (60km^2) is located along the lower reaches of the Yarra River, Melbourne's most iconic waterway, contains the central business district, major industrial and trade ports and high density inner-urban residential developments. Much of the development within this region also spans decades of urban planning and major infill. Troups Creek catchment (10km^2), an existing suburban development within Melbourne's sprawl is predominantly residential and features major reserves providing local amenity. Larger undeveloped land masses are in its upstream region. Toolern Precinct (24km^2) is a newly proposed mixed land use development in Melbourne's urban growth zone located on the western fringe, which strives to provide a variety of residential housing types and a commercial zone along an existing train line. Its population density is anticipated to exceed that of Troups Ck catchment.

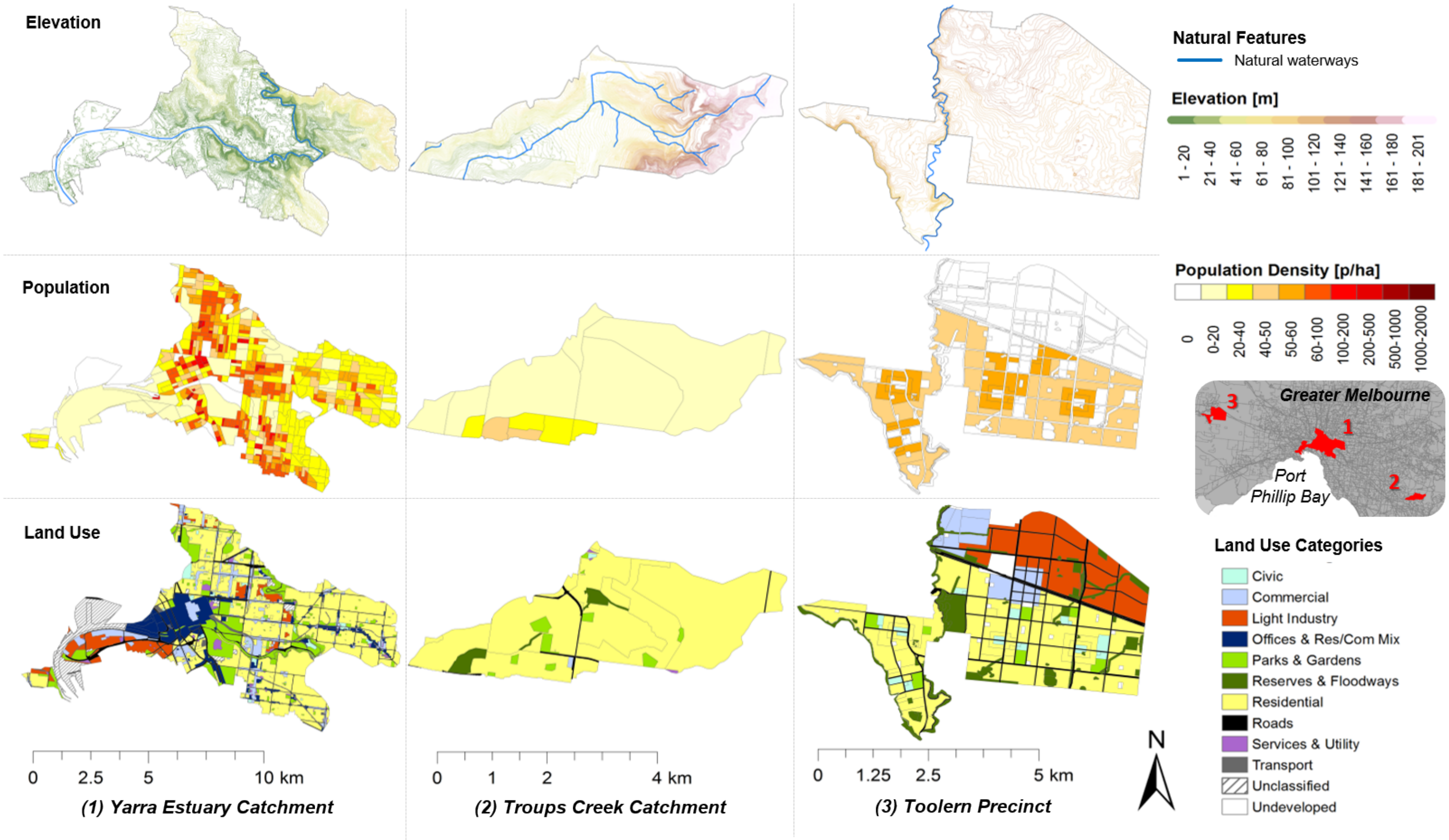

Figure 6. *Overview of the three model case studies, Yarra Estuary Catchment (left), Troups Creek Catchment (centre) and Toolern Precinct (right), spatial input data that includes elevation, population and land use, geographic location and their designation as inner-urban, sub-urban and peri-urban*

We obtained land use, population and elevation data for all three case studies (see Figure 6) from a variety of sources (for further details, refer to Supplementary Section S2). For all case studies, we created 10m x 10m resolution input rasters (in ASCII format). In the case of Yarra Estuary, an additional shapefile of the Yarra River was also provided to the model, this was not used for Troups Ck or Toolern Precinct for several reasons: their sizes are significantly smaller and designated sub-catchments by the local water authority are larger than what the model would produce if river shapefiles were used.

3.2 Model Setup and Calibration

Setting up the Urban Planning Module was undertaken in several steps. We begin by selecting an appropriate Block size to replicate the desired catchment structure. Then, a stepwise calibration approach was adopted. Firstly, we ensured the accurate demographic information (e.g. total population, total dwelling count, total employment) was reproduced by the model. Secondly, we calibrated the urban abstraction algorithm against spatially distributed impervious fractions and spatial distribution of total residential roof area.

3.2.1 Selection of appropriate Block size

To determine the appropriate Block size, we required a map of the existing catchment layout in each of the three case studies. Shapefiles for Yarra Estuary and Troups Ck were obtained from Melbourne's regional water authority. Toolern Precinct's sub-catchments were obtained from the local consultants, who worked on the project. The Urban Planning Module can be operated in stepwise progression. This means that we were able to run only the delineation module and iterate until a suitable Block sizes that produces flow paths and sub-basins that align with the obtained shapefiles was found. We trialled several Block sizes in multiples of 250m (250m, 500m, 750m and 1000m). The aim was not to find the exact number of sub-catchments, but rather, to use a small enough Block size that was computationally efficient and pragmatic to analyse. We judge pragmatic from two possible angles: (1) the shape of the case studies are not rectangular, as such, its approximation using a grid should 'make sense', i.e. there will be Blocks along the edge that only contain minimal data, an appropriate size should minimize this; (2) the borders between sub-catchments should not cross Blocks, it would be preferable to clearly allocate each Block to only a single sub-catchment. For smaller case studies, a smaller Block size was preferred so that it can capture enough variability in the urban environment. Variability is introduced through land use mix, selected parameter ranges as well as the number of Blocks in the map (since sampling of parameters from distributions occurs once per Block). The final decision was to use 500m x 500m Block sizes for Yarra Estuary and Toolern Precinct and a 250m x 250m Block size for Troups Ck.

3.2.2 Calibration of abstraction algorithms – preparing calibration data

We calibrated the abstraction algorithms against 'reference data sets' of spatial distribution of total impervious fraction (TIF) and, in the cases of Yarra Estuary and Troups Ck, also against the spatial distribution of total residential roof area. For TIF, three different sources of calibration data were used, one for each case study, respectively. For Yarra Estuary, we obtained a map of impervious areas from remotely sensed imagery (obtained from the local water authority). In the case of Troups Ck, as this remotely sensed data was non-existent, we used cadastre data of building footprints and road alignment and created a hypothetical impervious map. An illustration of how this data looks like is provided in the Supplementary Material (see Figure S2). For Toolern Precinct, which, at the time of this study, did not have a detailed layout of roads and buildings, we used recommended literature values of impervious fractions suggested by Melbourne's regional water authority modelling guidelines (Melbourne Water, 2010). We were able to calibrate for residential roof areas in existing areas by using a cadastre of building footprints and clipping it to the residential land uses in each case study.

Calibration data sets were created by superimposing the different reference data onto the grid of Blocks derived from the model (i.e. after determining a suitable Block size). For Yarra Estuary and Troups Ck, we clipped the shapefile geometry to each Block's coordinates and calculated TIF and total residential roof area within that Block. For Toolern Precinct, we calculated a total impervious fraction based on the land use mix of each Block.

3.2.3 Calibrating abstraction algorithms – stepwise calibration process

The stepwise calibration process was done manually and iteratively by trial-and-error. The first calibration step involved ensuring that total values (across the entire case study boundary) of population, employment and dwelling count were accurately estimated by the model. A satisfactory outcome meant that the population data was accurately read by the model and that the spatial calibration of residential and non-residential areas in the model would be representative. This calibration step does not require us to alter many of the model parameters. As such, we primarily focussed on the demographic parameters representing occupancies (O_{avg}, O_{max}), employment distribution ($\rho_{jobs\text{-}LUC}$) and some planning parameters, specifically the unit floor areas (A_{person}, A_{flat}) and building coverage ratios ($BCR_{max\text{-}LUC}$) (see Table 1 for descriptions and Table 2 for selected values).

Table 2. *Model default and selected input and calibration parameters for Yarra Estuary Catchment, Troups Creek Catchment and Toolern Precinct case studies*

Land use group	Param. type	Parameter name [units]	INITIAL VALUES	VALUES USED IN CASE STUDIES^		
			Model defaults	**Yarra Estuary**	**Troups Creek**	**Toolern Precinct**
Residential (sub-category: Houses)	INPUT	Side setback range [m]	1.0-2.0	1.0-2.0	1.0-2.0	1.0-2.0
		Max. number of car parks [ ]	2	2	0	2
		Min. driveway width [m]	2.6	2.6	0	2.6
		Max. number of floors []	2	5	1	1
	CALIB.	Occupancy range [persons]*	Avg.: 2.67 ; Max: 5.0	Avg.: 3 ; Max: 4.0	Avg.: 3.1 ; Max: 5.0	Avg.: 3.47 ; Max: 4.0
		Unit floor space [m2/person]	84	40	50	84
		Front setback range [m]*	2.0-9.0	2.0-6.0	2.0-9.0	2.0-4.0
		Max. patio area [m^2]	2.0	25	0	2
Residential (sub-category: Apartments)	INPUT	Average site setback [m]	1.0	5	2	1
		Maximum number of floors	10	20	10	10
		Parking arrangement	On-site	On-site	On-site	On-site
	CALIB.	Avg. occupancy [persons]	1.5	1.2	3.1	1.5
		Avg. apartment Size [m^2]	90.0	55	70	90
		Outdoor communal space [%]	5	15	5	5
Non-Residential	INPUT	Non-residential parcel size range [ha]	Industrial = 4.0-6.0; Com./Offices = 2.0-4.0	Industrial = 4.0-6.0; Com./Offices = 2.0-4.0	Industrial = 4.0 - 6.0; Commercial = 2.0-4.0	Industrial = 4.0 - 6.0; Commercial = 2.0-4.0
		Maximum number of floors []	4	No limit	4	2
		Parking lot dimensions [m]	2.6m x 4.6m	2.6m x 4.6m	2.6m x 4.6m	2.6m x 4.6m
		Loading bay allocation [m2/100m2 GFA]	27.0	27.0	27.0	27.0
		% of green landscaping [%]	100	50	100	50
	CALIB.	Employment distribution [employees/ha]	Industrial = 100; Commercial = 100; Offices =400	Industrial = 100; Commercial = 100; Offices = 400	Industrial = 100; Commercial = 30; Offices = 400	Industrial = 39; Commercial = 20; Offices = 400
		Maximum building coverage ratio [ ]	Industrial = 0.6; Com./Offices = 0.5	Industrial = 0.55; Com./Offices = 0.45	Industrial = 0.6; Com./Offices = 0.3;	Industrial = 0.8; Com./Offices = 0.5;
		Minimum site setback [m]	2.0	1.0	2.0	1.0
		Carpark requirements [bays per unit area or employees]	Industrial = 1.0/employee, Com./Office = 2.0/100m^2	Industrial = 1.0/employee, Com/Office = 1.0/100m^2	Industrial = 1.0/employee, Com./Office = 2.0/100m^2	Industrial = 2.0/employee, Com/Office = 1.0/100m^2
Roads & Open Spaces	INPUT	Frontage dimensions for residential [m] [#]	FP = 1.0 - 3.0; NS = 1.0 - 3.0; ST = 3.0 - 5.0	FP = 1.0 - 3.0; NS = 1.0 - 1.0; ST = 3.0 - 8.0	FP = 1.5 - 3.0; NS = 2.0 - 2.0; ST = 5.0 - 8.0	FP = 1.5 - 1.5; NS = 3.0 - 3.0; ST = 3.5 - 6.5
		Frontage dimensions for non-residential [m] [#]	FP = 1.0 - 3.0; NS = 1.0 - 3.0; ST = 3.0 - 5.0	FP = 1.0 - 3.0; NS = 1.0 - 1.0; ST = 3.0 - 8.0	FP = 1.5 - 3.0; NS = 2.0 - 2.0; ST = 5.0 - 8.0	FP = 1.5 - 1.5; NS = 3.5 - 3.5; ST = 3.5 - 6.5
		% of green area of parks [%]	100	100	100	100
	CALIB.	Highway reserve dimensions [m]*	LANE = 5.0 - 10.0; BUFF = 2.0 - 5.0; MED = 4.0 - 6.0	LANE = 5.0 - 10.0; BUFF = 1.0 - 2.0; MED = 1.0 - 5.0	LANE = 10.0 - 15.0; BUFF = 2.0 - 10.0; MED = 2.0 - 4.0	LANE = 7.0 - 7.0; BUFF = 3.5 - 6.5; MED = 3.0 - 3.0
Unclassified	CALIB.	Imperviousness of 'Unclassified' land use [%]	not considered	57% for large areas (>90% of block size)	not considered	not considered

** The calibration of these parameters is not focussed on the absolute values (these can be obtained from local planning documents), but rather on the range (e.g. min/max, average) that define the sampling distribution of this parameter*

^ The sources listed in Table 1 represent general planning documents within the Melbourne context and were used for Yarra Estuary and Troups Ck case studies. For Toolern Precinct, local master planning documents were also consulted in the selection of parameters.

[#] *FP = footpath, NS = nature strip, ST = road*

The second calibration step, the spatial calibration, was undertaken one land use at a time. We began with residential land uses and addressed each land use in the order presented in Table 2. Changes to all remaining calibration parameters of each land use (see Tables 1 and 2) were made until there were marginal improvements in the modelled Block TIF vs. the observed TIF. We used a visual comparison of modelled vs. observed values as well as the Nash-Sutcliffe coefficient of efficiency (NSE – Nash and Sutcliffe, 1970) and Root Mean Square Error (RMSE) to determine the level of improvement achieved by adjusting calibration parameters. Visual comparison, in particular, allowed us to identify specific Blocks with larger errors. These could be investigated in terms of their individual land uses to determine which parameters could be adjusted next. Calibration was concluded once all calibration parameters were addressed and no further improvement in model fit could be achieved. For clarification, we set carparks and driveway parameters for Troups Ck catchment to zero. This was done deliberately, knowing that the calibration data used for this case study does not contain any of these surfaces. The consequence of using this kind of calibration data is discussed later.

Ideally, the model would have been calibrated automatically. However, this is a complex model with a large number parameters. Due to this complexity as well as the fact that this is the first major piece of work on developing this extensive model, we took a pragmatic approach to do this manually, using the described two-step approach. The authors, however, acknowledge that automatic calibration methods can yield potentially more accurate model fits and allow us to conduct a broader sensitivity analysis. Outcomes of the current work should enable us to apply more automated algorithms and conduct sensitivity analysis in future work.

3.2.5 Evaluation of model performance

The estimation of impervious and roof areas was evaluated in a number of different ways. We statistically compared modelled and observed values of *Block TIF* and *Block Total Residential Roof Area* using the Nash-Sutcliffe coefficient and an RMSE. We also compare the frequency distributions in TIF, which is a technique that has been used to evaluate model performance in previous procedural modelling studies (see e.g. Vanegas et al., 2012). We further evaluated modelled Block TIF by looking at spatial differences in modelled and observed impervious fractions ($TIF_{obs} - TIF_{mod}$) across each Block to better understand where the inaccuracies are located in each case study.

4. Results and Discussion of Model Application

4.1 Results from Block and Catchment Delineation

Part of the challenge in working with coarse spatial grids (250m x 250m and 500m x 500m in this case) is overcoming possible inaccuracies in delineating the sub-catchments for the spatial region. Figure 7 shows the spatial model output for each case study including Blocks, flow paths, selected Block size, number of resulting Blocks and sub-basins. For Yarra Estuary, Blocks that contain the river body have been indicated on the map. These are all outlets for individual sub-basins, which explain the high number of sub-basins in comparison to the other two case studies.

There is consistency in the results for Troups Ck and Toolern Precinct. The former is considered a single catchment that drains into its local waterway, Troups creek. Using a 250m x 250m Block size, the number of sub-basins and flow paths are largely consistent with observed data. Toolern (using a 500m

Block) has four identified sub-catchments. The central catchment drains into Toolern creek, which is the central waterway shown in Figure 7(3). Both the eastern and western parts of the precinct drain correctly in their respective waterways (GAA, 2011). Although there is a minor discrepancy in the west with the model identifying two sub-basins instead of a single catchment, it is worth noting that Melton reservoir, which is the receiving water body in the west, runs along the precinct's western boundary edge and is the reason for why the observed catchment designates this region as a single sub-catchment (GAA, 2011).

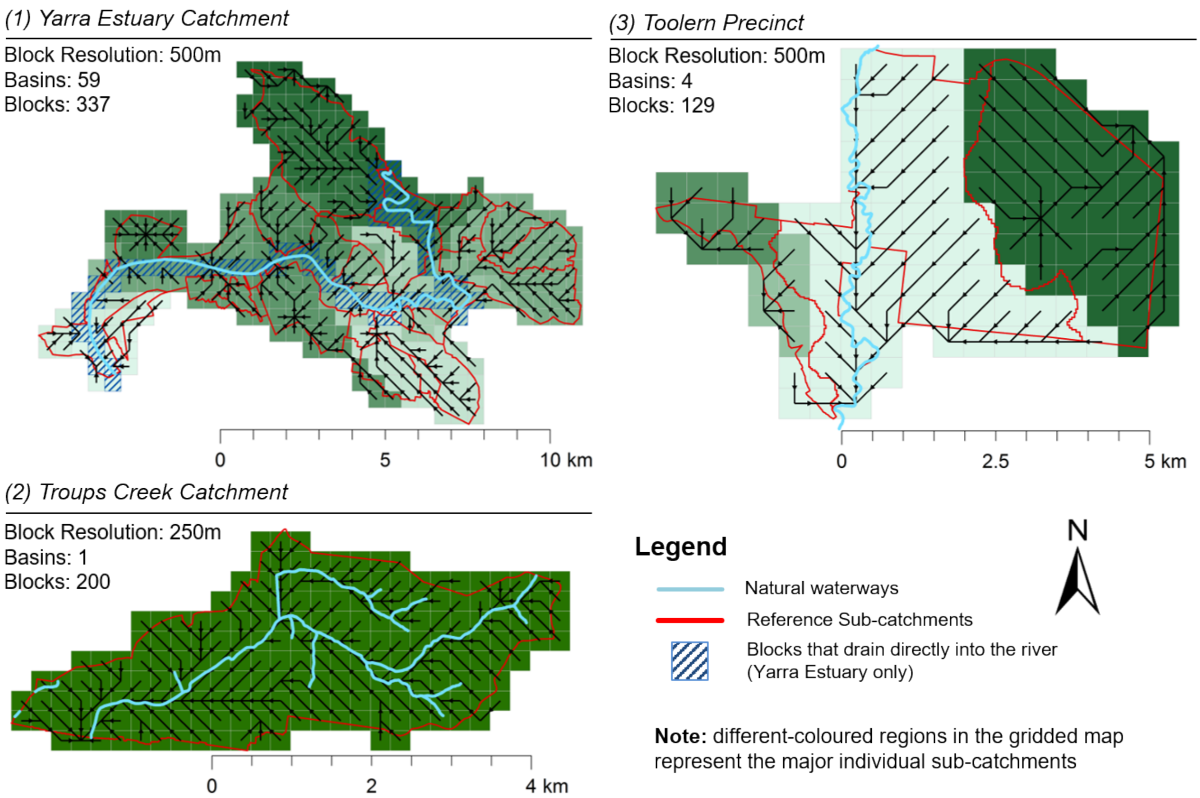

Figure 7. *Delineation of Blocks and Sub-catchments for the three different case studies including their comparison with observed sub-catchments designated by the local water authority*

Yarra Estuary contains 59 modelled sub-basins. Many of these terminate at points along the Yarra River. Based on the actual sub-catchment boundaries, we observe in Figure 7 that the grouping and directions of modelled flow paths are largely consistent and contained within the boundary lines. The spatial discrepancy is likely to be due to the model criterion used for addressing local sinks within the D8 algorithm (O'Callaghan and Mark, 1984) as many large drainage infrastructure is present in the region to cope with large rainfall events in the dense inner-city region and is not accounted for in the model (refer to Section 2.2.2 for key assumptions).

4.2 First Step Calibration – Demographics

Model results from first step calibration are compared to best available information for each of the case studies. As Troups Ck has a negligible amount of non-residential land, no employment data was used. Table 3 shows a summary of the modelled and observed population, dwellings and employment data. The employment estimate for the Yarra Estuary is only known in terms of order of magnitude as this value was based on reports about the inner-city economy (Alford et al., 2009, BITRE, 2013) and had to be transferred (using a density approach and area measurements GIS software) to the case study boundary, which is hydrologically-based.

We observe that the translation of population data in the model is generally consistent and that the calibration of total dwellings and total employment are satisfactory. The case for Toolern Precinct was interesting as the total population of 55,000 was obtained from the master plan's vision statement as an aspirational rather than an exact estimate. Despite having the largest discrepancy in population, we were able to calibrate dwelling count close to the anticipated number and remain internally consistent in selected parameter values with the proposed master plan (GAA, 2011). This suggests a potential inconsistency with the original input population data set, but was not a major concern given the similarities in dwelling and employment characteristics as well as consistency of parameter values with the original master plan. An employment of greater than 300,000 in the Yarra Estuary Region but within a reasonable magnitude was deemed acceptable. This would ensure that the correct densities would be achieved and lead to taller office buildings that represent the central business district, which is spatially delineated as approximately 20 Blocks.

Table 3. *Reference and Modelled (best fit) values of Total Population, Dwelling Count and Employment for Yarra Estuary Catchment, Troups Creek Catchment and Toolern Precinct*

Case Study	**Total Population**		**Total Dwellings**		**Total Employment**	
	Reference	**Modelled**	**Reference**	**Modelled**	**Reference**	**Modelled**
Yarra Estuary Catchment	215,010	215,035	90,164	90,060	>300,000	359,028
Troups Creek Catchment	5,998	5,999	1,892*	1,892	-	-
Toolern Precinct	55,000	52,644	15,860	15,736	22,000	21,936

**calculated by intersecting a GIS data layer of building footprints with the land use map to filter residential roof areas*

4.3 Second Step Calibration – Total Impervious Fraction and Roof Areas

The observed TIF maps that were used in model calibration are shown alongside the best model outputs in Figure 8(a). We also show the spatial difference in modelled and observed TIF in Figure 8(b) to highlight specific regions where the model may have had difficulties. These are accompanied by the performance evaluation criteria (NSE and RMSE) in Table 4. Note that we have normalised the RMSE value by chosen Block size for each case study to allow comparison between case studies that use different Block sizes. Plots of modelled vs. observed points of total impervious area and residential roof area are shown in Figures 9 and 10 (each with the 1:1 line and ±30% error bounds). Additionally, the frequency distribution of TIF for modelled and observed values have also plotted in Figure 9-bottom.

Comparison of modelled and observed TIF in Figure 8 and calibration results in Table 4 show that the model can satisfactorily reproduce spatial TIF across the three case studies (all Nash E values above 0.6). With the exception of Yarra Estuary, most of the errors are minimal, evenly distributed and sit within the ±30% error bounds (see Figure 9). On average, the model appears to over-estimate impervious fractions. With respect to water management, this is more conservative as it will lead to larger, more conservative sizing of drainage systems among other infrastructure.

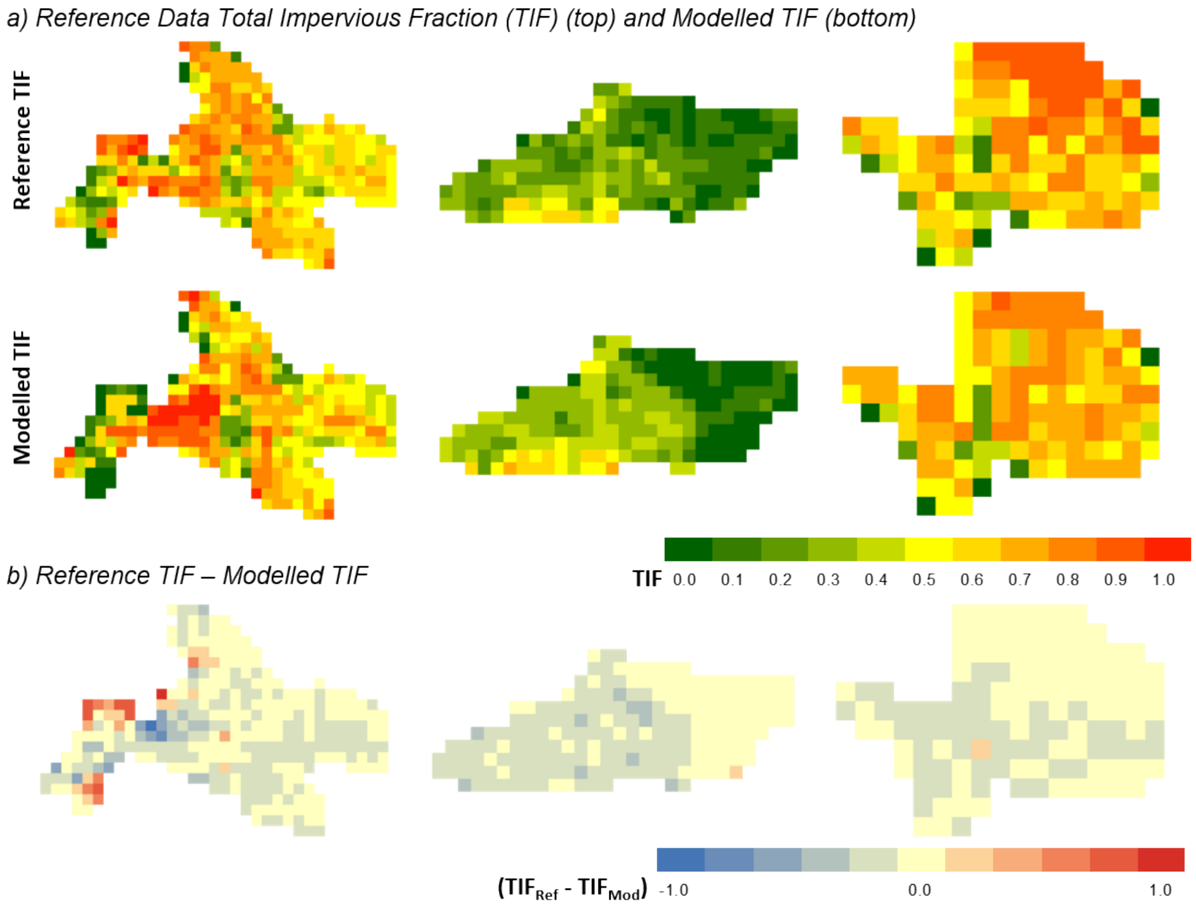

Figure 8. *Comparison of Reference and Modelled Spatial Distribution of Imperviousness in Yarra Estuary Catchment, Troups Creek Catchment and Toolern Precinct*

Differences between modelled and reference TIF for Yarra Estuary in Figure 8 reveal several local areas with higher errors. Based on the original land use map, the western regions are classed as 'Unclassified' land. Melbourne's major industrial port is located here and special planning schedules have been defined for this area (refer to City of Melbourne scheme in DPCD, 2006). The errors in the north are due to an ambiguity in the land use classification. The land is classified as a park, but is, in fact, part of the Melbourne General Cemetery. Its imperviousness is therefore slightly higher than a regular park, but is instead assumed as entirely green area.

We calibrated the model to the residential roof area to ensure that the it is able to distinguish between impervious areas that arise from roofs and pavement. While the estimation of residential roof area still shows good performance (NSE = 0.53 for Yarra Estuary and 0.65 for Troups Ck), certain bias is observable in Figure 10. The model appears to underestimate residential roof area for the suburban low-density case study of Troups Ck and overestimate for the dense inner city Yarra Estuary region.

There are two possible explanations for the bias in roof area estimates. The first is potentially due to the land use intensity and design ratio relationships used in the model structure. These are based on typical urban residential housing types. Many of the inner city houses (which are on relatively small lots, have high lot coverage but nevertheless small roof area) do not necessarily comply with these modern day concepts and design constraints. These are nevertheless present and still inhabited due to their historical significance and heritage value, which has led to their protection and preservation. The very low densities in Troups Ck catchment (due to the presence of some farms) and larger lot size in some areas would have resulted in very low FAR that are outside the minimum design ratios quoted in (De Chiara et al., 2009). Consequently, this would have led to the model to using the minimum available values.

Table 4. *Calibration results for Toolern Precinct, Troups Ck and Yarra Estuary Catchments based on three statistics*

Case Study	**Total Impervious Fraction**		**Total Residential Roof Area**	
	NSE	**RMSE/ A_{Block}**	**NSE**	**RMSE/ A_{Block}**
Yarra Estuary Catchment	0.78	0.129	0.53	0.121
Troups Creek Catchment	0.68	0.083	0.65	0.032
Toolern Precinct	0.91	0.078	-	-

The second explanation applies mainly to the Yarra Estuary, which is one of the oldest areas of Melbourne. Currently, the model is only able to select one of the two predominant residential typologies per block (see Figure 2). For example, if densities in a Block are high enough, the model chooses to design apartments and ignores the possible presence of houses in that Block. Inspecting satellite imagery shows that it is plausible for both residential typologies to be present within a 500m x 500m space. This is, however, a limitation of the current model design and a consideration for further research. This would, however, also need to be accompanied by an algorithm that is able to allocate the residential population to these different housing typologies.

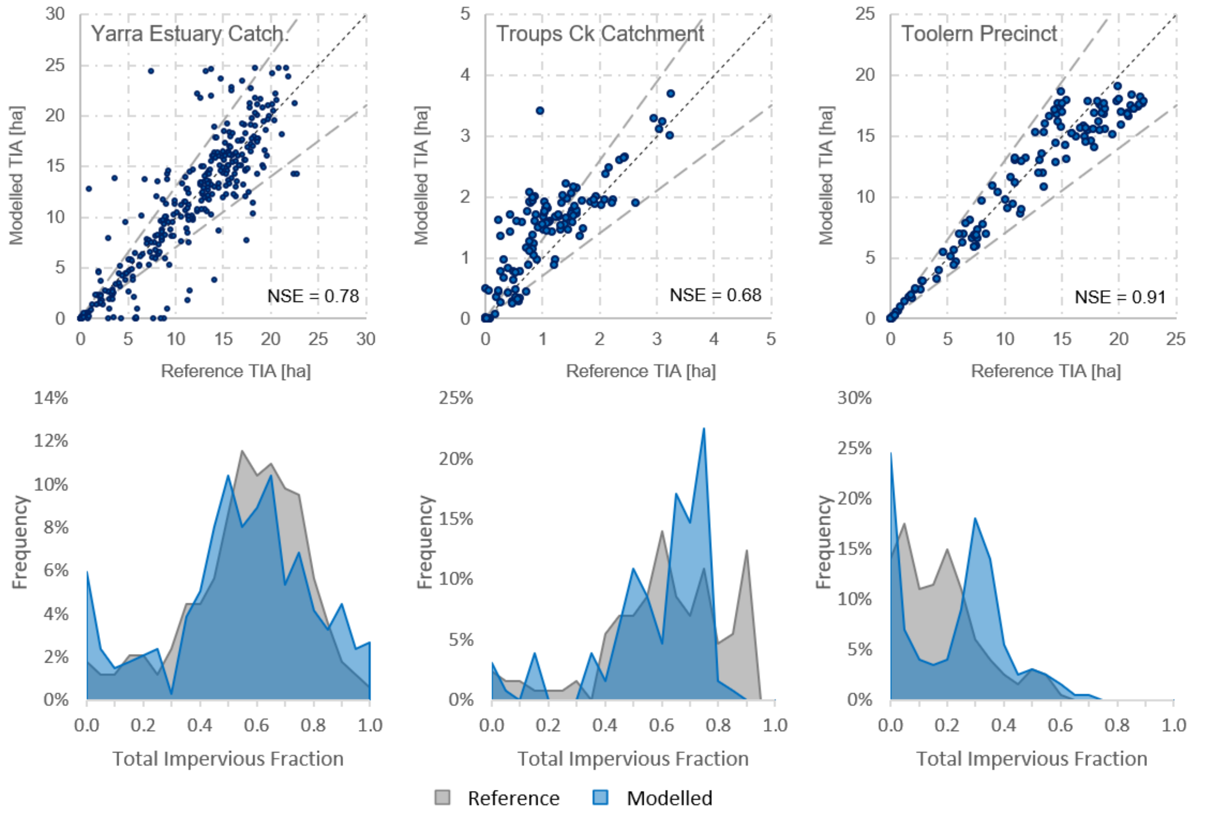

Figure 9. *Calibration Results for Imperviousness for Yarra Estuary Catchment, Troups Creek Catchment, Toolern Precinct, and– Observed vs. Modelled Total Impervious Area (TIA) showing 1:1 line and 30% error bounds (top) and Frequency Distributions of Observed (grey) and Modelled (blue) Total Impervious Fractions (bottom)*

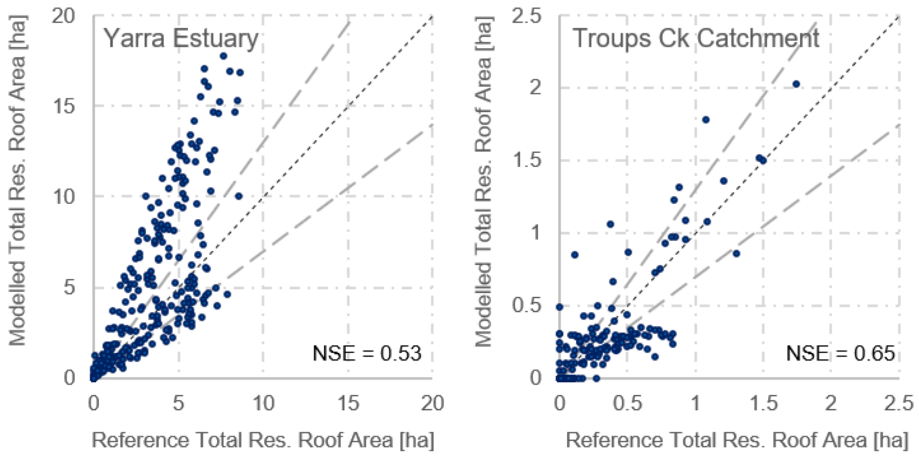

Figure 10. *Calibration of Total Residential Roof Areas for Yarra Estuary (left) and Troups Ck (right) Catchments showing 1:1 line and 30% error bounds*

5. Discussion of the Modelling Approach

The value in our developed hybrid approach (that combines procedural and conceptual techniques) to represent the spatial environment lies not only in the ability to show spatial variations in key characteristics (e.g. impervious fractions), but also the breakdown of different urban characteristics (e.g.

allotment sizes, land subdivision). Our discussion of results shows that reasonably good calibration of typical water management characteristics (impervious fraction and roof areas) can be achieved. To further evaluate the viability of this modelling approach, we focus our discussion on the key points mentioned when we first introduced our objectives: testing under a range of different case studies, quality of calibration data sets used and overall consistency in the selected model parameters against realism.

5.1 Calibration Data Quality

Of the three different calibration data sets used, we can see that the model is easily calibrated against literature values (in the case of Toolern Precinct), which are suggested by local modelling guidelines (Melbourne Water, 2010). Literature values are useful when the layout of the urban environment has not yet been developed (e.g. newly proposed developments in the planning phase). Their simplicity, however, does not truly reflect the variability in urban form unlike cadastre and remotely sensed data sets (shown in Figure S2). Of these two spatial data sets, better calibration with the model was achieved with data from remotely sensed imagery (i.e. Yarra Estuary). This is remarkable as one would expect the uniformity of Troups Ck's density, land use and urban forms to possibly be less challenging for the model to recreate and therefore yield better model results. Part of the problem with Troups Ck's calibration (which is also visible in Figure S2) is that the obtained cadastre data does not provide an entirely accurate representation of impervious and pervious surfaces. Nevertheless, in the absence of remotely sensed data, cadastre data sets can potentially be a more useful calibration data source as it provides the spatial variability that is not only influenced by land use mix (in the case of literature values) but also a number of other factors (e.g. density, urban planning, local population).

5.2 Contrasting Input and Calibrated Parameters

Default values for model parameters in Table 2 were based on local and regional planning information across Melbourne. Setting up and calibrating the model for the specific case studies required these parameters to be altered to suit their local environment. Investigating the calibrated values for all three case studies shows observable and logical contrasts that are consistent with the characteristics of inner, sub- and peri-urban areas. For example, we observe lower dwelling occupancies, floor space and more compact streetscapes in Yarra Estuary compared to Troups Ck and Toolern Precinct. Floor limits are also less strict in the inner city area (20 floors for residential 'no limit' for non-residential) to allow for high-rise development in Yarra Estuary, which contains the Central Business District and several notable areas with very high density residential developments.

Choice of parameters is also dependent on the source of calibration data and input information. We observe, for example, that Toolern Precinct's model setup and calibration is more constrained (e.g. many ranges are reduced to single values e.g. footpath and nature strip widths). This is because our model setup relied more predominantly on the published precinct master plan (GAA, 2011) as a source for many parameter values, whereas the other case studies relied solely on state-wide planning provisions (DPCD, 2006, Victorian Building Commission, 2006), which are broader in their specifications. The influence of calibration data type is most prominent in the case of Troups Ck catchment where any on-site pavement area has been deliberately set to zero since this is not captured by the cadastre data set. With satisfactory model performance, this comparison of parameters shows that there is consistency in the physical meaning of several parameters. This, consequently, can provide users with information to not only compare the structure of two different urban environments, but also the contrasting planning regulations that affect these.

5.3 Application Pathways for the Model

In contrast to existing techniques in environmental studies (e.g. Chabaeva et al., 2009, Schwarz, 2010, Sitzenfrei et al., 2010), our model is a departure from existing spatial representation approaches where many inputs contribute to a single output. This study calibrated the model against impervious area and residential roof area, which are essential indicators for stormwater management (Arnold and Gibbons, 1996, Stone, 2004, Dotto et al., 2010) and rainwater/stormwater harvesting for example (Mitchell et al., 2008, Ghisi, 2010). Our model output is rich in information, in the language of both urban planners and water managers and thus suitable for applications in integrated urban water management. Through this established link, urban water management can be more actively considered throughout the planning process as opposed to simply a requirement at the end. For example, variability in garden space on residential allotments can be assessed for irrigation requirements and the potential of lot-scale decentralised water infrastructure. This could, for example, complement existing tools for sustainable urban drainage systems (a state-of-the-art is offered by various reviews for example e.g. Elliott and Trowsdale, 2007, Bach et al., 2014, Lerer et al., 2015). Household occupancies and commercial floor space can be queried and used to better understand water consumption across the city and manage integrated urban drainage systems (e.g. Fu et al., 2009, Blokker et al., 2010, Benedetti et al., 2013). Policy scenarios for addressing highly impervious regions through special planning regulations can be explored (e.g. to simulate ideas from the literature about stream health restoration Walsh et al., 2005).

5.4 Limitations of the Research and Further Work

There are still several limitations to the current model's design. We acknowledge that the model was developed in an Australian-centric context (using mostly Australian planning regulations and three Australian case studies). Whilst many of the concepts are based on the sprawled city archetype, they also consider urban typologies that are typical of high-density urban environments, therefore making them flexible to overall city structure. The model is adaptable to non-Australian areas and we have previously demonstrated the transferability of the model's concepts in a pilot study of an American urban catchment (see Bach et al., 2016 for more information). Nevertheless, there is still uncertainty about how well current model algorithms can replicate mixed land use developments and the more compact and polycentric cities in Europe and the UK. As the current calibration process is quite rudimentary and lengthy, a logical next step would also be to trial some of the latest developments auto-calibration techniques to not only improve current model performance, but also help us better test the sensitivity of different model parameters. Future work will also investigate alternatives to the Block-based representation, which can more efficiently and accurately handle the translation and processing of spatial input and output data – a natural progression in the model building process that has also been experienced in the land use and transport (Waddell, 2002, Waddell, 2011), architecture (Parish and Müller, 2001, Salter et al., 2017) and energy modelling literature (Monsalvete et al., 2015, Hargreaves et al., 2017), where parcel-based and more spatially explicit representations and urban information management methodologies are being pursued.

6. Conclusion

Growing complexity of integrated urban water management, greater involvement of stakeholders in the planning and modelling process and the computational drawbacks or oversimplification of how geometric information and characteristics of the urban environment are modelled have highlighted the need for new innovative solutions. We develop and test a new model for characterising the urban form in an abstract manner that links back to planning regulations. This model used a grid representation, traditional methods for catchment delineation and five innovative land use abstraction algorithms that are a hybrid between conceptual and procedural approaches and are grounded in urban planning theory. We test the model on three distinctly different case studies using different calibration data sets to identify not only consistency in catchment representation using a coarse grid structure, but also the accurate representation of total impervious area and residential roof area, which are essential information for undertaking key urban water management tasks.

Results from the study produced the following key findings:

- The development of algorithms showed that characteristics of the urban environment (e.g. from the explicit layout of residential allotments and industrial estates to the coverage and spatial variability of impervious and pervious areas) can be procedurally generated from a range of planning documents, guidelines and regulation. This enables a link between urban planning and water management to be established.
- Testing of the model shows that it was able to produce reasonably good spatial estimates of total impervious fractions and roof areas. Inherent biases were present in some case studies and are attributed to constraints in the model structure and the evolutionary nature of urban form
- The use of literature values for impervious fractions is a good alternative if data are not available. Calibration data derived from remotely-sensed imagery have shown to be more spatially accurate and consistent as opposed to aggregating subsets of cadastre maps that may be more inconsistent in distinguishing impervious from pervious land cover.

- Model input and calibration parameters for each case study shows distinct contrasts in urban planning rules across the three case studies that reflect the urban-rural gradient. They are also consistent with the literature and therefore a good indicator of consistency and viability in our developed model.

This study has demonstrated how such a hybrid approach, which combines the 'best of both worlds', namely the conceptual and procedural approaches, can rapidly generate a spatially explicit and viable representation of urban form characteristics that are essential for water infrastructure planning. It is a significant step in allowing planners and urban water stakeholders to engage in more effective dialogue and build links across the multiple disciplines.

Acknowledgements

The authors would like to thank Ross Allen, Peter Hansen and Ben Mahon for their valuable feedback on urban planning and urban design aspects. We would also like to thank the Victorian Planning Authority (formerly Metropolitan Planning Authority and Growth Area Authority) for providing us data on the Toolern Precinct. This project is funded by the Australian Research Council (ARC), Discovery Early Career Researcher Award, project DE17010042, titled "Virtual Reality for Planning of Green Urban Water Infrastructure".

Nomenclature

The following is a list of variable letter, subscripts and abbreviations used throughout the text and should serve as a global reference. Note that some subscripts are not shown as they are self-explanatory.

Abbreviations:	**Variables:**	**Subscripts:**	
BCR – building coverage ratio	*A – area*	*avg – average*	*lot – allotment*
FAR – floor-area ratio	*D – depth*	*b – building*	*ls – liveability space*
GFA – gross floor area	*E – employees*	*buff – road buffer strip*	*lscape – landscaped area*
LUI – land use intensity	*d – distance*	*ca – construction area*	*LUC – land use category*
LSR – liveability space ratio	*N – number*	*com – communal space*	*max – maximum*
NSE – Nash-Sutcliffe Coefficient	*O – occupancy*	*cp – carpark*	*med – road median*
OCR – occupant car ratio	*P – population*	*dw – dwelling*	*min – minimum*
OSR – open space ratio	*p – proportion*	*fp – footpath*	*NRes – non-residential*
RMSE – root mean square error	*W – width*	*front – frontage*	*ns – nature strip*
RSR – recreational space ratio	*ρ – density*	*grey – grey/paved area*	*os – open space*
TCR – total car ratio		*i – referring to a Block*	*par – parcel*
TIA – total impervious area		*imp – impervious*	*Res – residential*
TIF – total impervious fraction		*lane – road lane*	*rs – recreational space*
UNC – unclassified land		*load – loading bay*	*sb – setback (front/side)*

Supplementary Material

S1. Further model concepts

Figure S1 illustrates how floor-area-ratio relates to lot coverage and the number of floors in five different conceptualised examples.

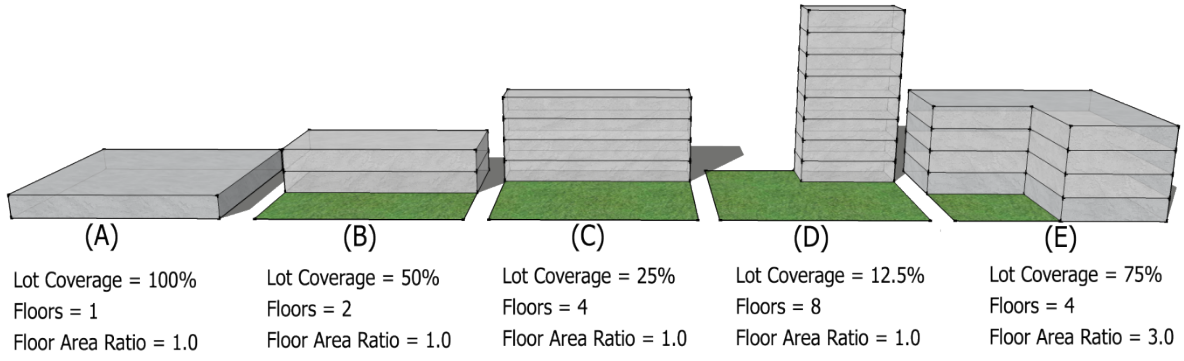

Figure S1. Illustrating the relationship between lot area, number of floors, lot coverage (i.e. building footprint) and floor-area ratio (FAR) to better understand the link in residential planning parameters

S2. Case study data acquisition and pre-processing

Existing land use zoning maps (for Yarra Estuary and Troups Creek case studies) and elevation contours for all three case studies were obtained from the Victorian Government's Open Data repository (Victorian Government, 2015) and reclassified into the model's classification system (guidance for which is provided in Bach et al., 2015b). Population density for Yarra Estuary and Troups Creek were obtained from the latest Australian census (Australian Bureau of Statistics, 2016). For Toolern Precinct, the proposed development plan and projected population densities were obtained from the Metropolitan Planning Authority in the form of a georeferenced drawing and translated into the required raster map and land uses. To aid the translation of this data into the appropriate maps, we also consulted the Precinct Structure Plan (PSP), which prescribes planning ordinances for the precinct (GAA, 2011).

Figure S2 illustrates the two different spatial data used for calibration of impervious areas for Troups Creek Catchment and the Yarra Estuary Catchment.

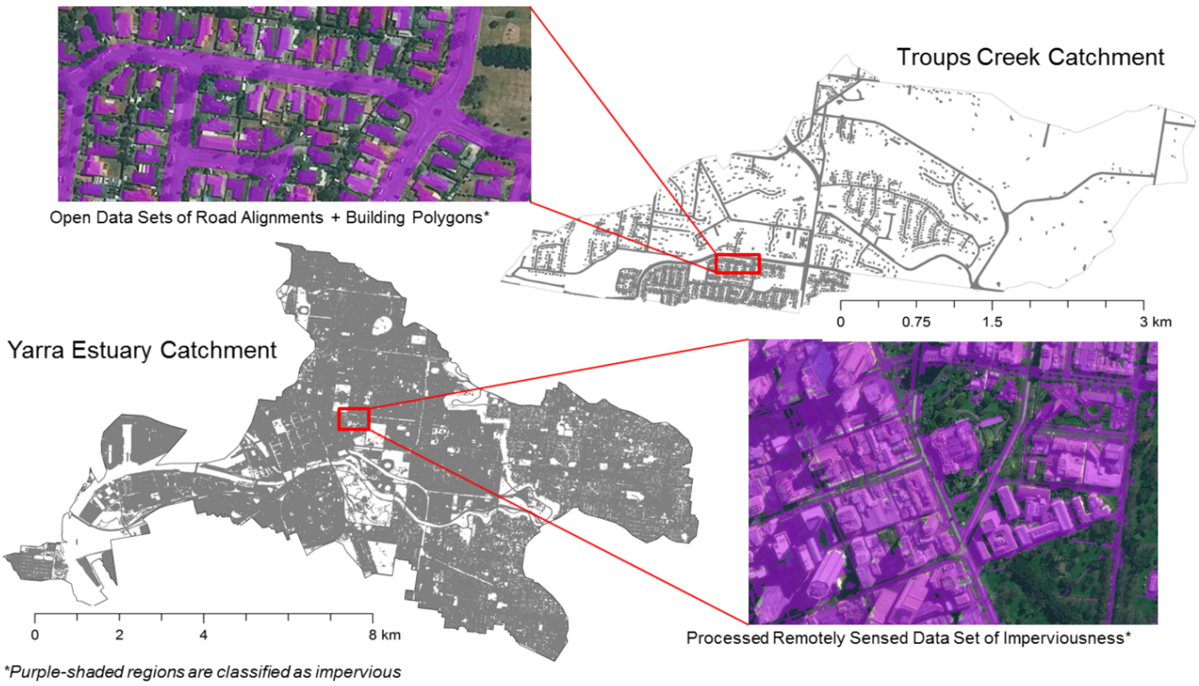

Figure S2. Different spatial reference data sets of impervious areas for model calibration; Troups Ck Catchment uses merged layers of building footprints and road alignments and can have notable errors in the delineation of the impervious road area and exclusion of other paved surfaces; Yarra Estuary Catchment's impervious area was determined from remotely sensed imagery and is more accurate in depicting the majority of areas, but can potentially show errors due to misinterpretation by the automatic image classification algorithm.

References

Aguilera, A. 'Urban Form and Commuting: A critical review of literature', *12th WCTR*, Lisbon, Portugal, 11-15 July, 2010.

Alberti, M., Booth, D., Hill, K., Coburn, B., Avolio, C., Coe, S. and Spirandelli, D. (2007) 'The impact of urban patterns on aquatic ecosystems: An empirical analysis in Puget lowland sub-basins', *Landscape and Urban Planning,* 80, pp. 345-361.

Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I. and Angel, S. (1977) *A Pattern Language.* New York: Oxford University Press.

Alford, G., Whiteman, J. and Transport Policy Analysis and Research Unit (2009) *Macro-Urban Form, Transport Energy Use and Greenhouse Gas Emissions: An Investigation for Melbourne*, Melbourne: Victorian Government Department of Transport.

Arbues, F., Garcia-Valinas, M. A. and Villanua, I. (2010) 'Urban Water Demand for Service and Industrial Use: The Case of Zaragoza', *Water Resources Management,* 24, pp. 4033-4048.

Arnold, C. L. and Gibbons, C. J. (1996) 'Impervious Surface Coverage, The Emergence of a Key Environmental Indicator', *Journal of the American Planning Association,* 62(2), pp. 243-258.

Australian Bureau of Statistics (2016) *Australian Bureau of Statistics*. Available at: http://www.abs.gov.au (Accessed: 30 March 2015).

Austroads (2010) *Guide to Road Design Part 3: Geometric Design.* Sydney, NSW: Austroads Ltd.

Bach, P. M., Deletic, A., Urich, C., Sitzenfrei, R., Kleidorfer, M., Rauch, W. and McCarthy, D. T. (2013a) 'Modelling Interactions Between Lot-Scale Decentralised Water Infrastructure and Urban Form - a Case Study on Infiltration Systems', *Water Resources Management,* 27, pp. 4845-4863.

Bach, P. M., Eisenstein, W., McCarthy, D. T., Hatt, B., Sedlak, D. and Deletic, A. 'Australian water sensitive planning modelling in the San Francisco Bay Area: challenges and implications for model transferability', *2016 International Low Impact Development Conference (LID2016)*, Beijing, China.

Bach, P. M., McCarthy, D. T. and Deletic, A. (2015a) 'Can we model the implementation of Water Sensitive Urban Design in evolving cities?', *Water Science & Technology,* 71(1), pp. 149-156.

Bach, P. M., McCarthy, D. T., Urich, C., Sitzenfrei, R., Kleidorfer, M., Rauch, W. and Deletic, A. (2013b) 'A planning algorithm for quantifying decentralised water management opportunities in urban environments', *Water Science & Technology,* 68(8), pp. 1857-1865.

Bach, P. M., Rauch, W., Mikkelsen, P. S., McCarthy, D. T. and Deletic, A. (2014) 'A critical review of integrated urban water modelling – urban drainage and beyond', *Environmental Modelling & Software,* 54, pp. 88-107.

Bach, P. M., Staalesen, S., McCarthy, D. T. and Deletic, A. (2015b) 'Revisiting Land Use Classification and Spatial Aggregation for Modelling Integrated Urban Water Systems', *Landscape and Urban Planning,* 143, pp. 43-55.

Bankes, S. (1993) 'Exploratory Modeling for Policy Analysis', *Operations Research,* 41(3), pp. 435-449.

Basingstoke & Deane Borough Council (2008) *Appendix 11 - Industrial Development Planning and Design Guidance*, Basingstoke, UK.

Bates, P. D. and De Roo, A. P. J. (2000) 'A simple raster-based model for flood inundation simulation', *Journal of Hydrology,* 236, pp. 54-77.

Batty, M. (2007) *Cities and Complexity Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals.* 1st edn. London: MIT Press.

Benedetti, L., Langeveld, J., Comeau, A., Corominas, L., Daigger, G., Martijn, C., Mikkelsen, P. S., Vezzaro, L., Weijers, S. and Vanrolleghem, P. A. (2013) 'Modelling and monitoring of integrated urban wastewater systems: review on status and perspectives', *Water Science & Technology,* 68(6), pp. 1203-1215.

BITRE (2013) *Population growth, jobs growth and commuting flows - a comparison of Australia's four largest cities, Report 142*, Canberra ACT: Australian Government Bureau of Infrastructure, Transport and Regional Economics.

Blokker, E. J. M., Vreeburg, J. H. G. and van Dijk, J. C. (2010) 'Simulating Residential Water Demand with a Stochastic End-Use Model', *Journal of Water Resources Planning and Management,* 136(1), pp. 19-26.

Brown, R. R., Keath, N. and Wong, T. H. F. (2009) 'Urban water management in cities: historical, current and future regimes', *Water Science & Technology,* 59(5), pp. 847-855.

Butler, D. and Davis, J. W. (2004) *Urban Drainage Second Edition.* London: Spon Press.

Campbell, S. (1996) 'Green Cities, Growing Cities, Just Cities? Urban Planning and the Contradictions of Sustainable Development', *Journal of the American Planning Association,* 62(3), pp. 296-312.

CBSC (2013) *California Building Code*, Sacramento, CA: California Building Standards Commission.

Chabaeva, A., Civco, D. L. and Hurd, J. D. (2009) 'Assessment of Impervious Surface Estimation Techniques', *Journal of Hydrologic Engineering,* 14(4), pp. 377-387.

Chang, K.-T. (2010) *Introduction to Geographic Information Systems (5th Edition).* 5th edn. Singapore: McGraw-Hill International Edition, p. 1-448.

City of Monash (2009) *A Guide to Car Parking in Monash*, Victoria, Australia: Monash City Council.

City of Santa Maria (2013) *Planning Policies and Regulations*. Santa Maria, California, USA: City of Santa Maria. Available at: http://www.cityofsantamaria.org/207-PlanningPoliciesAndRegulations.shtml (Accessed: 20th January 2014 2014).

Conzen, M. P. (2001) 'The study of urban form in the United States', *Urban Morphology,* 5(1), pp. 3-14.

De Chiara, J., Panero, J. and Zelnik, M. (2009) *Time-Saver Standards For Housing and Residential Development, 2nd Edition. Time-Saver Standards*: McGraw Hill, p. 1-1132.

De Haan, F. J., Rogers, B. C., Brown, R. R. and Deletic, A. (2016) 'Many roads to Rome: The emergence of pathways from patterns of change through exploratory modelling of sustainability transitions', *Environmental Modelling & Software,* 85, pp. 279-292.

Dotto, C. B. S., Kleidorfer, M., Deletic, A., Fletcher, T. D., McCarthy, D. T. and Rauch, W. (2010) 'Stormwater quality models: performance and sensitivity analysis', *Water Science & Technology,* 62(4), pp. 837-843.

DPCD (2006) *Planning Scheme of Victoria*, Melbourne, Victoria: Department of Planning and Community Development.

Eggimann, S., Mutzner, L., Wani, O., Schneider, M. Y., Spuhler, D., de Vitry, M. M., Beutler, P. and Maurer, M. (2017) 'The Potential of Knowing More: A Review of Data-Driven Urban Water Management', *Environmental Science & Technology,* 51, pp. 2538-2553.

Elliott, A. H. and Trowsdale, S. A. (2007) 'A review of models for low impact urban stormwater drainage', *Environmental Modelling & Software,* 22, pp. 394-405.

Farolfi, S., Müller, J. and Bonté, B. (2010) 'An iterative construction of multi-agent models to represent water supply and demand dynamics at the catchment level', *Environmental Modelling & Software,* 25, pp. 1130-1148.

Fu, G., Butler, D. and Khu, S. T. (2009) 'The impact of new developments on river water quality from an integrated system modelling perspective', *Science of the Total Environment,* 407, pp. 1257-1267.

GAA (2011) *Toolern - Precinct Structure Plan (Including Toolern Native Vegetation Precinct Plan)*, Victoria, Australia: prepared by Melton Shire Council and Growth Area AuthoritiesSM20110810 3:26).

Gauthiez, B. (2004) 'The history of urban morphology', *Urban Morphology,* 8(2), pp. 71-89.

Geertman, S. and Stillwell, J. (2004) 'Planning support systems: an inventory of current practice', *Computers, Environment and Urban Systems,* 28, pp. 291-310.

Ghisi, E. (2010) 'Parameters Influencing the Sizing of Rainwater Tanks for Use in Houses', *Water Resources Management,* 24, pp. 2381-2403.

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X. and Briggs, J. M. (2008) 'Global Change and the Ecology of Cities', *Science,* 319, pp. 756-760.

Gurran, N. (2011) *Australian urban land use planning: Principles, systems and practice.* 2nd Edition edn. Sydney: Sydney University Press.

Hargreaves, A., Cheng, V., Deshmukh, S., Leach, M. and Steemers, K. (2017) 'Forecasting how residential urban form affects the regional carbon savings and costs of retrofitting and decentralized energy supply', *Applied Energy,* 186, pp. 549-561.

Harris, B. and Batty, M. (1993) 'Locational models, geographic information and planning support systems', *Journal of Planning Education and Research,* 12, pp. 184-198.

Hillier, B. and Hanson, J. (1984) *The Social Logic of Space.* Cambridge, UK: Cambridge University Press.

Klosterman, R. E. (1997) 'Planning-support systems: A new perspective on computer-aided planning', *Journal of Planning Education and Research,* 17, pp. 45-54.

Kuller, M., Bach, P. M., Ramirez-Lovering, D. and Deletic, A. (2017) 'Framing water sensitive urban design as part of the urban form: a critical review of tools for best planning practice', *Environmental Modelling & Software,* 96, pp. 265-282.

Last, E. M. (2010) *City Water Balance, A New Scoping Tool for Integrated Urban Water Management Options.* Doctor of Philosophy, University of Birmingham, UK, Birmingham, UK.

Lee, S.-W., Hwang, S.-J., Lee, S.-B. and Hwang, H.-S. (2009) 'Landscape ecological approach to the relationships of land use patterns in watersheds to water quality characteristics', *Landscape and Urban Planning,* 92, pp. 80-89.

Lerer, S. M., Arnbjerg-Nielsen, K. and Mikkelsen, P. S. (2015) 'A Mapping of Tools for Informing Water Sensitive Urban Design Planning Decisions - Questions, Aspects and Context Sensitivity', *Water Asset Management Journal,* 7, pp. 993-1012.

Majid, M. R. (2006) *Estimating residential imperviousness at subdivision level using remote sensing images and GIS*, Johor, Malaysia: University Teknologi Malaysia.

Masuya, Y., Fujii, M., Tamura, T., Saito, K. and Black, J. 'Commuting Preference Functions from Employment Centres', *Traffic and Transportation Studies Congress 2008*, 197-207.

McIntosh, B. S., Seaton, R. A. F. and Jeffrey, P. (2007) 'Tools to think with? Towards understanding the use of computer-based support tools in policy relevant research', *Environmental Modelling & Software,* 22, pp. 640-648.

Melbourne Water (2010) *MUSIC Guidelines, Recommended input parameters and modelling approaches for MUSIC users*, Melbourne, Australia: Melbourne Water Corporation (ISBN 978-1-921603-87-7.

Mikovits, C., Rauch, W. and Kleidorfer, M. (2014) 'Dynamics in urban development, population growth and their influences on urban water infrastructure', *Procedia Engineering,* 70, pp. 1147-1156.

Mitchell, V. G., McCarthy, D. T., Deletic, A. and Fletcher, T. D. (2008) 'Urban stormwater harvesting - sensitivity of a storage behaviour model', *Environmental Modelling & Software,* 23, pp. 782-793.

Mitchell, V. G., Mein, R. G. and McMahon, T. A. (2001) 'Modelling the urban water cycle', *Environmental Modelling & Software,* 16, pp. 615-629.

Monsalvete, P., Robinson, D. and Eicker, U. (2015) 'Dynamic simulation methodologies for urban energy demand', *Energy Procedia,* 78, pp. 3360-3365.

Nash, J. E. and Sutcliffe, J. V. (1970) 'River flow forecasting through conceptual models Part I - A discussion of principles', *Journal of Hydrology,* 10, pp. 282-290.

Niemelä, J. (1999) 'Is there a need for a theory of urban ecology?', *Urban Ecosystems,* 3, pp. 57-65.

O'Callaghan, J. F. and Mark, D. M. (1984) 'The Extraction of Drainage Networks from Digital Elevation Data', *Computer Vision, Graphics, And Image Processing,* 28, pp. 323-344.

Parish, Y. I. H. and Müller, P. 'Procedural Modelling of Cities', *ACM SIGGRAPH 2001*, Los Angeles, CA, USA, 12-17 August. New York: ACM Press, 301-308.

Pauleit, S. and Duhme, F. (2000) 'Assessing the environmental performance of land cover types for urban planning', *Landscape and Urban Planning,* 52, pp. 1-20.

Phil McDermott Consultants (2006) *Review of Business Land Requirements Review, Report to Smart Growth*, Western Bay of Plenty. Available at: http://www.smartgrowthbop.org.nz/media/1302/i-business_land_requirements_review__-_phil_mcdermott_consultants_-_october_2006.pdf.

Reinhardt, W. and Trudel, H. (1979) *Wohndichte und Bebauungsformern, Praktische Entscheidungshilfen für die kommunale Planung*, Stuttgart.

Rittel, H. W. J. and Webber, M. M. (1973) 'Dilemmas in a General Theory of Planning', *Policy Sciences,* 4, pp. 155-169.
Russ, T. (2009) *Site Planning and Design Handbook (2nd Ed.).* 2nd edn. New York, USA: McGraw-Hill.
Salter, J., Kellett, R., Girling, C. and Inomata, F. 'Modeling Energy for Urban Form Archetypes', *Symposium on Simulation for Architecture and Urban Design (SimAUD 2017)*, Toronto, Canada.
Schwarz, N. (2010) 'Urban form revisited - Selecting indicators for characterising European cities', *Landscape and Urban Planning,* 96, pp. 29-47.
SGS Economics & Planning (2008) *Take up of Industrial Land: Six Regional Victorian Centres*, Melbourne, Victoria: Department of Transport, Planning and Local Infrastructure.
Siksna, A. (2006) 'The study of urban form in Australia', *Urban Morphology,* 10(2), pp. 89-100.
Sitzenfrei, R., Fach, S., Kinzel, H. and Rauch, W. (2010) 'A multi-layer cellular automata approach for algorithmic generation of virtual case studies: VIBe', *Water Science & Technology,* 61(1), pp. 37-45.
Sitzenfrei, R., Moderl, M. and Rauch, W. (2013) 'Assessing the impact of transitions from centralised to decentralised water solutions on existing infrastructures - Integrated city-scale analysis with VIBe', *Water Research,* 47, pp. 7251-7263.
Stone, B. J. (2004) 'Paving over paradise: how land use regulations promote residential imperviousness', *Landscape and Urban Planning,* 69, pp. 101-113.
Tchobanoglous, G., Stensel, H. D., Tsuchihashi, R. and Burton, F. (2013) *Wastewater Engineering: Treatment and Resource Recovery, Fifth Ed.* New York: McGraw Hill.
Tewdwr-Jones, M. and Allmendinger, P. (1998) 'Deconstructing communicative rationality: a critique of Habermasian collaborative planning', *Environment and Planning A,* 30, pp. 1975-1989.
United Kingdom Development Agencies (1986) *Industrial and commercial estates: planning and site development.* London, UK: Telford.
United Nations (2012) *World Urbanization Prospects, The 2011 Revision*, New York: United Nations.
Urich, C. and Rauch, W. (2014) 'Exploring critical pathways for urban water management to identify robust strategies under deep uncertainties', *Water Research,* 66, pp. 374-389.
Uuemaa, E., Roosaare, J. and Mander, U. (2005) 'Scale dependence of landscape metrics and their indicatory value for nutrient and organic matter losses from catchments', *Ecological Indicators,* 5, pp. 350-369.
Vanegas, C. A., Kelly, T., Weber, B., Halatsch, J. and Aliaga, D. G. 'Procedural Generation of Parcels in Urban Modeling', *EUROGRAPHICS 2012*: Computer Graphics Forum, 10.
Victorian Building Commission (2006) *Building Regulations 2006 S.R. No. 68/2006 - Part 4: Siting*, Melbourne, Victoria.
Victorian Government (2015) *Victorian Government Data Directory*. Melbourne, Australia. Available at: https://www.data.vic.gov.au/ (Accessed: 30th March 2015).
Vlachos, E. and Braga, B. (eds.) (2001) *The challenge of urban water management*. Cornwall, UK: IWA Publishing.
Voinov, A. and Bousquet, F. (2010) 'Modelling with stakeholders', *Environmental Modelling & Software,* 25, pp. 1268-1281.
Voinov, A., Kolagani, N., McCall, M. K., Glynn, P. D., Kragt, M. E., Ostermann, F. O., Pierce, S. A. and Ramu, P. (2016) 'Modelling with stakeholders - Next generation', *Environmental Modelling & Software,* 77, pp. 196-220.
Waddell, P. (2002) 'UrbanSim, Modeling Urban Development for Land Use, Transportation, and Environmental Planning', *Journal of the American Planning Association,* 68(3), pp. 297-314.
Waddell, P. (2011) 'Integrated Land Use and Transportation Planning and Modelling: Addressing Challenges in Research and Practice', *Transport Reviews,* 31(2), pp. 209-229.
Walsh, C. J., Fletcher, T. D. and Ladson, A. R. (2005) 'Stream restoration in urban catchments through redesigning stormwater systems: looking to the catchment to save the stream', *Journal of the North American Benthological Society,* 24(3), pp. 690-705.
Ward, S., Farmani, R., Atkinson, S., Butler, D., Hargreaves, A., Cheng, V., Denman, S. and Echenique, M. 'Towards an integrated modelling framework for sustainable urban development', *9th International Conference on Urban Drainage Modelling (9UDM)*, Belgrade, Serbia, 1-12.

White, R., Engelen, G. and Uljee, I. (2015) *Modeling Cities and Regions as Complex Systems: From Theory to Planning Applications.* Cambridge, Massachusetts: The MIT Press.

Yamaguchi, Y., Shimoda, Y. and Mizuno, M. (2007) 'Proposal of a modeling approach considering urban form for evaluation of city level energy management', *Energy and Buildings,* 39(5), pp. 580-592.

Ziadat, F., Bruggeman, A., Oweis, T., Haddad, N., Mazahreh, S., Sartawi, W. and Syuof, M. (2012) 'A Participatory GIS Approach for Assessing Land Suitability for Rainwater Harvesting in an Arid Rangeland Environment', *Arid Land Research and Management,* 26, pp. 297-311.