Spatial modelling of municipal waste generation: deriving property lot estimates with limited data

Ben Madden[a,*], Nick Florin[a], Steve Mohr[b], Damien Giurco[a]

[a] *Institute for Sustainable Futures, University of Technology Sydney, Ultimo NSW Australia*
[b] *Research and Innovation Division, University of Newcastle, Callaghan NSW Australia*

Abstract

Given recent circular economy policy and waste minimisation targets, there is a significant opportunity to fundamentally change the way waste is managed in Australia, and re-focus waste management to promote resource recovery and efficiency. Detailed data on household waste generation can assist decision makers in targeting waste minimisation incentives, improving resource recovery and circularity, identifying specific technology and infrastructure gaps and informing future development. Unfortunately, high-resolution spatial estimates of waste generation at the property lot scale is typically unavailable. This study presents a novel spatial model developed to estimate waste generation data at the property lot level. Utilising census data at multiple spatial scales and council waste generation data, we apply our model to estimate quantities of residual waste, dry recyclables and garden waste generated for more than 1,200,000 property lots in the Sydney metropolitan area, Australia. Results show the spatial distribution of estimated household waste generation, achieving a high degree of accuracy when compared to validation data. To illustrate the application of our results in the context of identifying ideal areas for waste processing facilities, we analyse the spatial distribution of available garden waste arising from property lots. An area of intense garden waste generation was identified, indicating a supply area of approximately 13km^2 in northern Sydney that can support a facility of approximately 20,000t throughput a year. Our analytical approach presented is novel, and has practical application for locating waste processing facilities; analysing efficient kerbside waste collection services; and in informing data driven urban waste management strategies.

Keywords: Urban waste management, Spatial disaggregation, GIS, Data analytics, Microsimulation

1. Introduction

The management of municipal solid waste presents multiple environmental, human health, logistical and economic challenges for cities and regions (Esin & Cosgun, 2007; Ferrão & Fernández, 2013). In Australia, waste management is in focus with industry and communities committed to reducing waste, avoiding

*Corresponding author
Email address: benjamin.madden@uts.edu.au (Ben Madden)

landfilling, and eliminating the risk of material losses to the environment throughout the resource recovery system. Exporting waste for overseas processing is no longer an option for waste management (Council of Australian Governments, 2020), and new landfill diversion pathways for Australian waste must be prioritised to avoid increases in landfilling in the future. Recently, the government of the Australian state of New South Wales adopted a circular economy framework to guide future waste decision making in the state towards a 'waste-as-a-resource' framework (NSW Government, 2019). Under this updated policy direction, there is a significant opportunity now for the promotion of waste management initiatives focused on improving resource recovery with greater circularity in resource flows.

High resolution data on quantities of waste generated, for example at the property lot or household level, can be very useful for informing decision making around targeted waste management initiatives. For example, such data is useful for: identifying optimal waste treatment or recovery facility locations (Yadav et al., 2017, 2018; Lin et al., 2020); optimising waste collection routing (Hannan et al., 2018; Sarmah et al., 2019; Vu et al., 2019); and in planning for targeted dwelling specific systems, such as insinkerators and basement anaerobic digestion and composting (Lou et al., 2013; Edwards et al., 2016). Waste generation data is often limited at high resolution scales despite its usefulness (Kontokosta et al., 2018), typically due to individual household privacy concerns and the high cost associated with large-scale household bin audits and surveys. There is a need then for an approach to disaggregate available, low spatial-resolution waste generation data down to a finer spatial scale, when such data might be beneficial for informing waste management decisions.

There is a gap in the literature concerning the estimation of waste generation data at high spatial resolutions, where real data is not available. In a recent study, Kontokosta et al. (2018) presents a novel analytical approach for predicting daily and weekly waste generation at the building level in New York City, USA. In Kontokosta et al. (2018), the authors address the issue that understanding patterns of waste generation at building level is limited as data at this scale is generally lacking. For their approach, the authors develop a predictive model utilising detailed census, weather, and waste data to estimate household waste generation at the building level. The resulting model has high accuracy when compared to validation data, and has several important contributions and applications, including in designing more efficient waste collection routing, and for detecting areas of New York City that have higher (or lower) likelihood to recycle for more targeted waste behaviour initiatives.

Kontokosta et al. (2018) use statistical and computational approaches for disaggregating low spatial resolution (e.g., city, region etc.) data to the building level. While studies applying similar, small-area estimation methods are not common in a waste context, spatial disaggregation models are more widespread across other fields including public health (Eberth et al., 2018; Truong & Stein, 2019; Albright et al., 2019), socioeconomic research (Fabrizi & Trivisano, 2016; Buil-Gil et al., 2019), resource assessment (Goerndt et al., 2019), and the management of waste and debris due to natural disasters (Tabata et al., 2019; Hayes

et al., 2021). These small-area estimation methods are widely used for producing estimates of attributes at spatially disaggregated scales where data is limited, by 'borrowing' from data at other scales (Chandra et al., 2012; Buil-Gil et al., 2019). However, small-area estimation models generally require microdata (data available at very fine scale) for calibration and validation. In the absence of such data, microsimulation and other spatial approaches are often used to generate required data for small areas, using appropriate deterministic, probabilistic and/or computational modelling approaches (NATSEM, 2008; Rich, 2018).

The aim of this research was to develop a spatial model for estimating quantities of household waste generated annually at the property lot level. The model described in this paper is motivated by limited availability of data at the property lot level, namely information on household characteristics and quantities and composition of waste generated itself. In this paper, we describe an approach that was used to estimate annual quantities of waste generated for more than 1.2-million property lots in the Sydney metropolitan area (SMA), Australia. The model described first estimates the distribution of dwelling types at the property lot level using council and neighbourhood level census data (Australian Bureau of Statistics, 2016a,b). Average rates of annual waste generation by dwelling type were derived from council reported waste statistics (NSW EPA, 2017) and local kerbside audit data (APC, 2019), and combined with the estimated dwelling type distribution to estimate annual waste generated per property lot in our study area. The model presented in this paper has wide ranging local applications for informing waste management policy, including analyses requiring high spatial resolution waste generation data such as new infrastructure planning. More broadly, the approach developed has modest data requirements and could be readily applied to other jurisdictions.

2. Methodology

For our approach, we first developed a probabilistic model from area-level data to estimate the number of occupied dwellings by type for each property lot within the 31 local government areas (LGA) of the SMA. The estimated number of dwellings by type was then multiplied by estimated dwelling-type waste generation intensities (i.e., tonnes/hh/yr) to estimate annual quantities of household waste generated at property lots for each LGA in the SMA.

Our approach implicitly accounts for variation in waste generation behaviours between different dwelling type occupants amongst the LGAs investigated. We considered two high-level dwelling types in our study: detached dwellings (i.e., a single dwelling located on a property lot); and multi-unit dwellings, including apartments, townhouses, and other residential structures where multiple households are located on a single lot. Focus on these two dwelling types has been made for two key reasons; Firstly, kerbside audits of the municipal waste stream in Southern Sydney, were conducted in 2008 and 2019 for detached and multi-unit dwellings only (APC, 2008, 2019), and found that dwelling type is a significant determinant of waste generation, with detached dwellings generating on average between 1.7 and 2.8 times more total household

waste than multi-unit dwellings. This finding is somewhat expected, given that the number of occupants in detached dwellings is typically greater than multi-unit dwellings (Australian Bureau of Statistics, 2016b), however the composition of waste generated between dwelling types also differs significantly. Secondly, waste generated by multi-unit dwellings presents unique challenges, with increased rates of bin contamination, and often complex bin collection systems making management of waste on site difficult (APC, 2019; Waste Management Review, 2020). As such, high resolution data on the spatial distribution of waste generated from multi-unit dwellings can help inform waste management planning that specifically targets these dwelling types.

Figure 1 provides a high-level overview of the methodological approach. Each component of our approach is described in further detail in the following subsections.

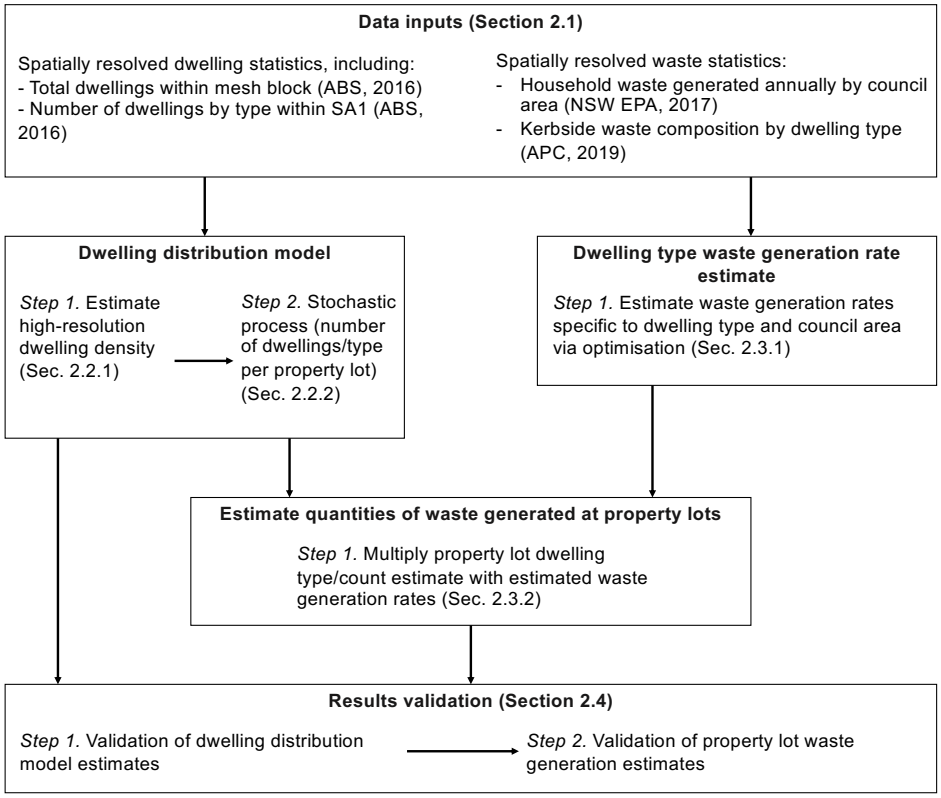


Figure 1: Overview of the methodological approach for this study

2.1. Study area and data

Figure 2 shows the study area located within the Australian state of New South Wales, and for context, quantities of total municipal household waste generated reported by local government areas (LGAs) in 2016 (NSW EPA, 2017).

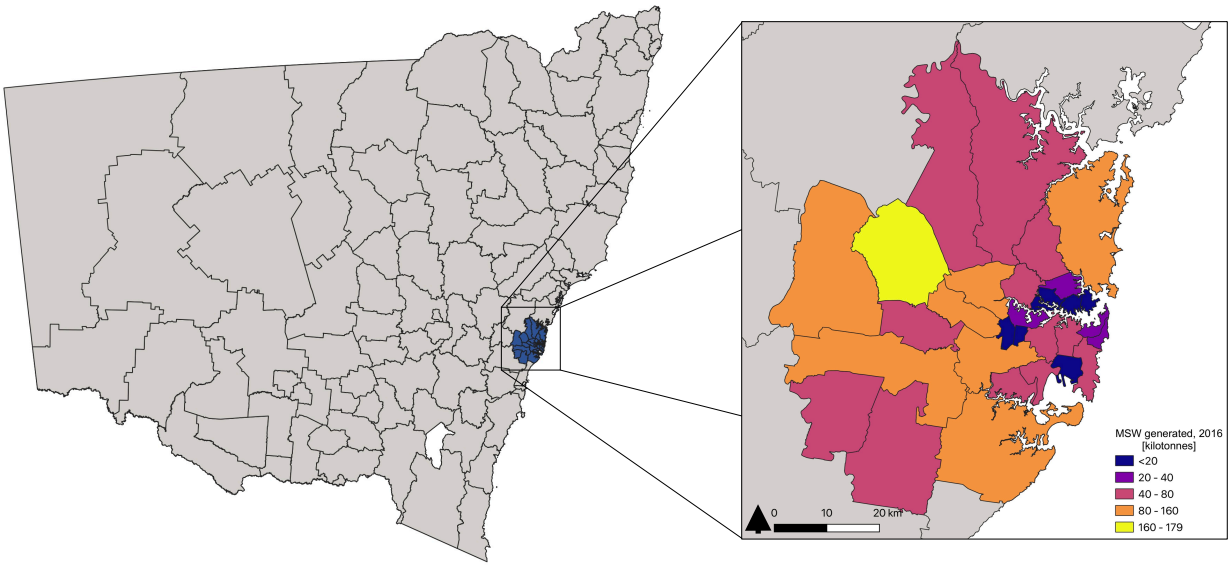


Figure 2: New South Wales and council boundaries. The Sydney Metropolitan Area study area is highlighted, and total municipal household waste generated in 2016 from NSW EPA (2017) is shown

For our study, we are interested in the annual quantity of municipal waste generated at the household level for three key waste streams that are collected by LGAs in the SMA, namely the residual (or non-recyclable) fraction, dry recyclables, and garden organics. The composition of these waste streams are summarised in Table 1. Annual quantities of household waste generated was the variable of interest for this study. Quantities of household waste collected is often used as a proxy for waste generated, given that littering, at home composting, hoarding etc. are not typically measured. Annual waste generation was also chosen, as statistics are typically reported at this interval locally (e.g., NSW EPA (2017); Department of the Environment and Energy (2018)). As a result of this, seasonal variations affecting waste generation behaviours are not directly considered.

Table 1: Composition of the residual, dry recyclable, and garden waste fractions of the municipal waste stream in the Sydney metropolitan area (NSW EPA, 2014)

Material group	Residual waste	Dry recyclables	Garden waste	Total waste
Paper and paper products	20%	55%	<1%	21%
Organics	54%	2%	99%	58%
Glass	4%	30%	<1%	9%
Plastics	11%	8%	<1%	6%
Ferrous material	2%	2%	<1%	1%
Non-ferrous material	1%	1%	<1%	<1%
Other	10%	2%	<1%	5%

Table 2 lists the data used for this study. Waste data on the aforementioned waste streams were derived from LGA reported waste generation data for 2016, which includes total municipal household waste collected (NSW EPA, 2017). This data is available over the 2005 to 2018 period, and the 2016 year was chosen to align with census data. Census data at a number of spatial scales were utilised, which are illustrated in Figure 3. The mesh block scale is the finest spatial resolution for census data available in Australia. At this scale, data describing the total number of occupied dwellings and total residential population is available, however no further information on household characteristics including type is available. Land use characteristics are described at the mesh block level, however only by majority land use within the mesh block. Mesh blocks categories as commercial for example then, still are likely to have occupied residential dwellings in the form of mixed-use zoning. We therefore only consider mesh blocks that contain occupied residential dwellings, as of the 2016 census night. SA1s are built from whole mesh blocks, with target populations of between 200 and 800 persons in urban locals (Australian Bureau of Statistics, 2016a). At the SA1 scale, the total number of dwellings by type is available, and is the highest level of spatial resolution available for this level of detail on dwelling types. Property lot boundaries are taken from the NSW Digital Cadastral Database (Department of Finance, Services and Innovation, 2012). This spatial data does not include characteristics of the property lot, including land use, number of dwellings etc, primarily due to privacy reasons. This spatial scale is the target spatial scale in our study for disaggregation of the area-level (i.e., SA1) dwelling type distribution data.

Table 2: Summary of data sources utilised in this work

Data	Spatial scale	Remarks
2016 census data (Australian Bureau of Statistics, 2016b)	SA1	Total numbers of dwellings by type in the SA1, used in model calibration and validation
	Mesh blocks	Total number of dwellings, used in model calibation and validation
Cadastral data (Department of Finance, Services and Innovation, 2012)	Property lot	Property lot boundaries, used for calibration of model and visualisation
Australian statistical geography standard (Australian Bureau of Statistics, 2016a)	LGA, SA1, mesh block	Spatial boundaries for LGA, SA1 and mesh block scales. Used for calibration and visualisation
Local council waste and resource recovery data (NSW EPA, 2017)	LGA	Annual waste generated by households, by waste fraction for the year 2016. Used for model calibration
Kerbside waste audit data (APC, 2019)	N/A (avg. per-household rates)	Weekly per-household waste generation rates, by dwelling type, from sample of kerbside audit data. Used for model calibration

(a) SA1 scale (b) Mesh block scale (c) Property lot scale

Figure 3: Different geographical scales of analysis, example Burwood local government area in New South Wales

A more detailed description of the approach outlined in Figure 1, including model validation, is provided in the following sections.

2.2. Dwelling distribution model

We developed a probabilistic model to estimate the number of occupied detached and multi-unit dwellings for property lots in our study area. We perform this model for property lots within each SA1, as this is the finest resolution of census data describing the breakdown of dwelling types. We model the number and distribution of dwelling types for a set of property lots within an SA1 as a discrete stochastic process–a family of discrete random variables, indexed over a countable set (Parzen, 2015). In our application, the index set is the set of property lots located within an SA1. Stochastic models can be advantageous, especially for our application where precise approximations of real data are required, despite limited data availability (Fortin et al., 2003; Fuqua & Doty, 2012). With no spatial covariates describing the determinants of dwelling type location at our resolution, we base our model on the observation that SA1s in the study area with higher dwelling densities (number of dwellings per km^2) have a greater number of multi-unit dwellings (Figure 4). This is expected given that multi-unit buildings feature a greater number of dwellings per floor area than detached dwellings, and is also consistent with local and state government urban planning strategies, preferencing high-density dwelling types in urban areas and around transportation corridors (Bunker, 2014; Roberts et al., 2019).

In our approach, numbers of detached and multi-unit dwellings in an SA1 are modelled separately as the stochastic process $X_T = \{X_T(l) : l = 1, \cdots, n\}$; where T is the dwelling type, n is the number of property

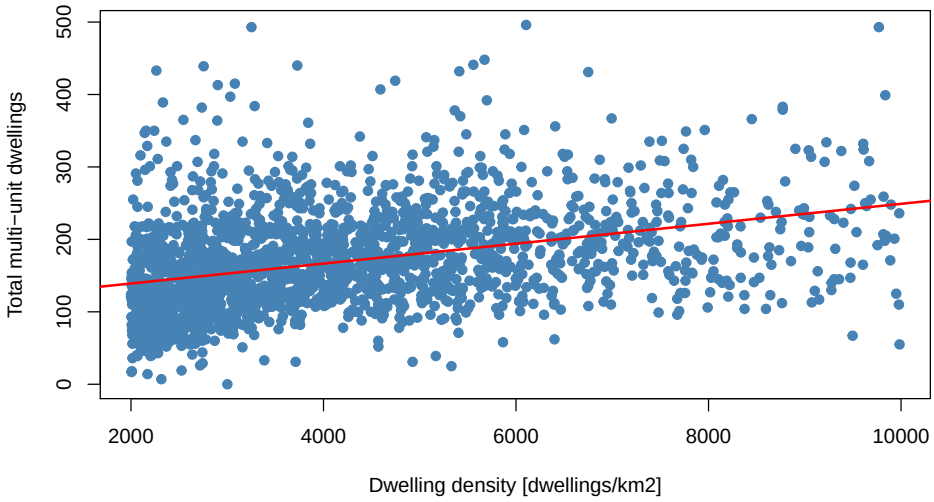


Figure 4: Relationship between dwelling density (number of dwellings/km^2) and number of multi-unit dwellings in an SA1, n = 9,868. The red regression line represents a Pearson's r correlation of 0.37

lots in an SA1, and $X_T(l)$ is a random variable describing the number of dwellings in the property lot l. The random variable $X_T(l)$ is drawn from a probability distribution, which is calibrated on the underlying estimated dwelling density of property lots in the SA1. Subsections 2.2.1 to 2.2.3 describe the elements of the dwelling distribution model in further detail.

2.2.1. Dwelling density surface

The mesh block scale is the highest resolution of dwelling density data available, however for our application, we require estimation of this data at the property lot level to calibrate the stochastic process X_T. Considering the spatial hierarchy in Figure 3, data at the mesh block level can give a coarse indication of the within-SA1 variation in dwelling density. However, the within-mesh block variation in dwelling density is unknown. Without the within-mesh block variation characterised, in our approach property lots within a mesh block would have uniform probability of a particular number and type of dwelling. To estimate dwelling density at the property lot, we generated an interpolated surface of dwelling density values from the mesh block data over the SA1 of interest. This interpolated surface gives a smooth, non-uniform dwelling density estimate over the SA1, with mean dwelling density at areas along the boundary of mesh blocks influenced by dwelling density from adjacent mesh blocks. From this surface, mean dwelling density estimation for each property lot can be achieved through aggregation.

We generated a continuous 2-dimensional dwelling density surface confined to the boundary of the SA1 to calibrate our dwelling density model. We performed a spatial interpolation of calculated mesh block dwelling density using inverse distance weighted (IDW) interpolation. IDW is a deterministic and computationally efficient technique for generating an interpolated surface, which is especially advantageous in our study

given the number of SA1s located within the study area (9,868 SA1s over an approximately 3,600km^2 area). Selection of interpolation approach can be subjective, with no one method necessarily superior than others for a given application (Wu & Hung, 2016). Other interpolation techniques including kriging and spline interpolation were considered for this work. IDW was ultimately chosen as the interpolation approach, as a lack of spatial covariate data at the mesh block scale and below excluded the use of kriging, and spline interpolation has been found to have greater computational demands with little if any improvement in performance over IDW (Mueller et al., 2001; Kravchenko, 2003; Lu & Wong, 2008).

The IDW interpolation approach estimates values at unknown points x on the Cartesian surface S, in our application, a grid of points at 1m^2 resolution, based on known values at N sample interpolation points (Shepard, 1968) (Equations 1 and 2). For our model, sample observation points are the centroids (geometric centre) of the mesh blocks located within the SA1. Some mesh blocks within the study area are irregularly shaped, which can cause estimated centroid locations to fall outside mesh block bounds. In these cases, mesh block centroids are approximated using the *PointOnSurface* method from the *rgeos* library (Bivand & Rundel, 2020) in the *R* statistical computing language.

$$\mathcal{D}(x) = \begin{cases} \frac{\sum_{n=1}^{N} w_n(x)\mathcal{D}(m_n^*)}{\sum_{n=1}^{N} w_n(x)}, & \text{if} \quad d(x, m_n^*) \neq 0 \\ \mathcal{D}(m_n^*) & \text{if} \quad d(x, m_n^*) = 0 \end{cases} \tag{1}$$

$$w_n(x) = \frac{1}{d(x, m_n^*)^\gamma} \tag{2}$$

Where $\mathcal{D}(x)$ is the estimated dwelling density at point $x \in S$, $d(x, m_n^*)$ is the Euclidean distance between points x and mesh block centroids m_n^*, and γ is the power parameter where higher values of γ give greater influence to points closer to sample observation points. $\mathcal{D}(m^*)$ is the estimated dwelling density at mesh block centroid m^* (Equation 3):

$$\mathcal{D}(m^*) = \frac{N_D(m)}{A(m)} \tag{3}$$

Where $N_D(m)$ is the known number of dwellings in mesh block m from Australian Bureau of Statistics (2016a,b), and $A(m)$ is the area of mesh block m in m^2.

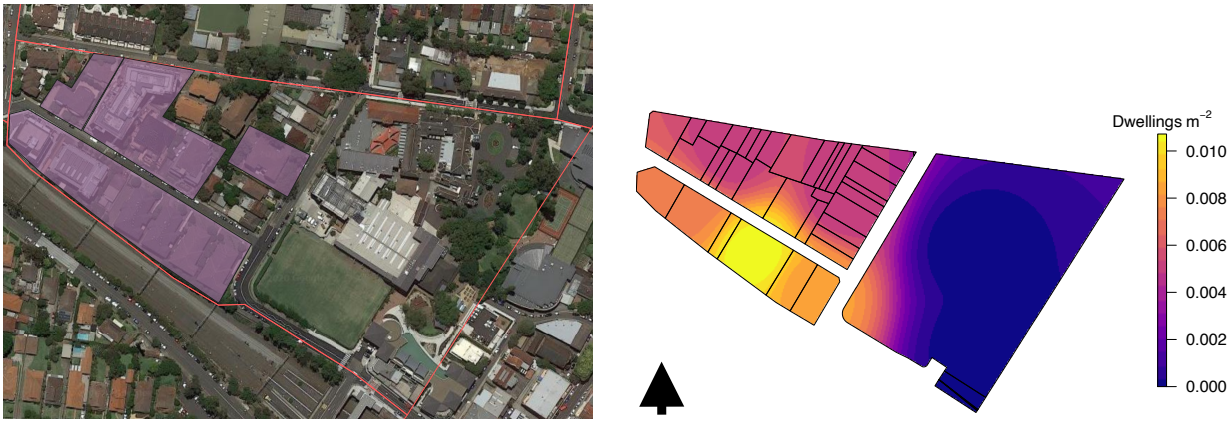
For the surface S, we seek a smooth surface of interpolated dwelling density values across the SA1. In Equation 2, we set $\gamma = 2$, which results in a smooth interpolated surface. As values of γ increase, the resulting IDW interpolation becomes tessellated, with interpolated values approaching that of nearest neighbour interpolation (Bivand et al., 2013). This type of interpolation would result in uniform values where the surface S overlaps with the mesh block, resulting in uniform property lot dwelling densities within, which would be insufficient for our application as previously described.

We finally estimate the mean dwelling density per property lot from S. Let S^* be the set of points in S that lie within the intersection of property lot l and the surface S (i.e., $S^* = \{x : x \in l \cap S\}$). We calculate

mean dwelling density as follows (Equation 4):

$$\mathcal{D}(l) = \frac{\sum_{x \in S^*} \mathcal{D}(x)}{|S^*|} \tag{4}$$

Figure 5 illustrates the described methodology for estimating the dwelling density surface, and aggregated mean dwelling density per property lot for an example SA1 located in the Burwood LGA. The estimated dwelling surface is compared with an aerial image of the SA1, with multi-unit and detached dwellings confirmed visually from Google Streetview imagery (Google, n.d) and highlighted.



(a) Aerial image

(b) Example dwelling density surface estimate

Figure 5: Estimated dwelling density surface for an example SA1, located in the Burwood LGA in the SMA

2.2.2. Estimating dwelling counts for detached dwelling type

We estimated the number of detached dwellings for a property lot l in an SA1 as a random variable drawn from a Bernoulli distribution: $X_{\text{det}}(l) \sim \text{Ber}(p)$, where p is the mean probability that $X_{det}(l) = 1$. The Bernoulli distribution is a discrete probability distribution of a random variable taking a value of 0 or 1, with a value of 1 indicating in our application that a property lot contains a detached dwelling. Values for p can be determined such that the sum of expected values for $X_{\text{det}}(l)$ over all property lots in an SA1 equals the known number of detached dwellings in the SA1.

We determine the value of p indirectly for each property lot from $\mathcal{D}(l)$ in Equation 4. Based on the distribution of estimated mean dwelling density for property lots, we determine a threshold range $\tau = [\tau_{\text{min}}, \tau_{\text{max}}]$ that characterises the likely range of dwelling densities where detached dwellings are most likely. Let L^a be the set of property lots in L that have mean density values within τ (e.g., $L^a = \{l : \mathcal{D}(l) \in \tau\}$). For this paper, we determine τ as the interquartile range of $\mathcal{D}(l)$ values in L^a, assuming that property lots more on the tails of the distribution of average dwelling density values would either contain no residential

properties, or multi-unit dwellings. We then determine p as follows (Equation 5):

$$p = \frac{N_{\text{det}}(\mathcal{S})}{|L^a|} \tag{5}$$

where $N_{\text{det}}(\mathcal{S})$ is the known number of detached dwellings in the SA1 $\mathcal{S}$.

We assumed that property lots with average dwelling density outside of τ to have either no residential properties, or have multi-unit dwellings. Therefore, the number of detached dwellings in a property lot l in an SA1 is then (Equation 6):

$$X_{\text{det}}(l) = \begin{cases} \sim \text{Ber}(p) & \text{if} \quad l \in L^a \\ 0 & \text{if} \quad l \notin L^a \end{cases} \tag{6}$$

To implement the above, we use the R statistical computing language, using the base 'stat' package and spatial methods within the 'sp' (Bivand et al., 2013) and 'raster' (Hijmans, 2020) packages. We estimate $X_{\text{det}}(l)$ using the *rbinom* function in R which simulates a Binomial point process of the form $\mathbf{X} \sim Binom(n, k, p)$ (when $k = 1$, the Binomial distribution is equivalent to the Bernoulli), where each $X \in \mathbf{X}$ are Bernoulli random variables. We set n as the number of property lots in L^a, with values for k and p described in the preceding paragraphs. As described in Equation 6, lots not within L^a are given values of 0.

As a final step, we take the sum of estimated detached dwellings across an SA1 as $\hat{N_{\text{det}}}(\mathcal{S}) = \sum_l X_{\text{det}}(l)$, and repeat the described process iteratively until $\hat{N_{\text{det}}}(\mathcal{S})$ equals the known number in the SA1 ($N_{\text{det}}(\mathcal{S})$), or until an exit condition is reached, in which case the realisation with $\hat{N_{\text{det}}}(\mathcal{S})$ closest to $N_{\text{det}}(\mathcal{S})$ is chosen. The pseudo-code in Algorithm 1 found in Appendix A describes this fitting process.

2.2.3. Estimating dwelling counts for multi-unit dwelling type

We estimated the number of multi-unit dwellings per property lot for an SA1 as a random variable from the Poisson distribution $X_{\text{mul}}(l) \sim \text{Pois}(\lambda(l))$, where $\lambda(l)$ is a non-uniform intensity function (hence an inhomogenous process), giving the average number of dwellings per lot l. The Poisson distribution is a discrete probability distribution, that models the number of occurrences (in our application, multi-unit dwellings) as a discrete random variable taking values $\{0, 1, 2, \cdots\}$ over some fixed interval (Haight, 1967).

We focus on the set of property lots L^b that have mean dwelling density greater than τ_{max}, as these property lots are more likely to contain multi-unit dwellings. Note that L^a and L^b are disjoint sets, and $L^a \cap L^b = \emptyset$.

We estimate $\lambda(l)$ such that the expected (or mean) number of occurrences of multi unit dwellings in L^b (i.e, $\mathbb{E}[N_{\text{mul}}(L^b)]$) equals the known number of multi-unit dwellings in an SA1 (Equation 7):

$$\lambda(l) = \mathcal{D}(l) \times \frac{N_{\text{mul}}(\mathcal{S})}{\sum_{l \in L^b} \mathcal{D}(L)} \tag{7}$$

where $N_{\text{mul}}(\mathcal{S})$ is the number of multi-unit dwellings in an SA1 $\mathcal{S}$.

Similarly with detached dwellings, we let the number of dwellings in property lots not in L^b equal 0. The number of multi-unit dwellings for property lots in an SA1 is then given by (Equation 8):

$$X_{\text{mul}}(l) = \begin{cases} \sim \text{Pois}(\lambda(l)) & \text{if} \quad l \in L^b \\ 0 & \text{if} \quad l \notin L^b \end{cases} \tag{8}$$

$X_{\text{mul}}(l)$ is estimated using the *rpois* function in R which can simulate a Poisson point process of length n, equal to the number of property lots within L^b. Estimates for $X_{\text{mul}}(l)$ go through a similar fitting procedure as detached dwelling estimates. This procedure is described in pseudo-code in Algorithm 2 found in Appendix B.

2.3. Estimating quantities of waste generated at the property lot

We estimated the annual quantities of household waste generated at property lots, by multiplying the dwelling count estimate $X_T(l)$ by an average rate of waste generation specific to dwelling type. Data on per-dwelling generation rates for each LGA in our study area can be calculated from the council-reported waste data described in Table 2. While this rate captures the between-LGA differences in rates of waste generation, it does not account for the differences in waste generation rates between detached and multi-unit dwellings, as reported in APC (2019), and presented in Table 3. Data in APC (2019) includes observed per-dwelling rates of waste generation for each waste fraction and dwelling type of interest for a sample of 13 non-identified LGAs in Southern Sydney. While this data includes direct measurements of waste generation rates, audits were conducted over a small window of time (approximately 3 months) therefore reported rates of waste generation may be influenced by unknown time-of-year effects, and the sample of households surveyed was small.

2.3.1. Estimating dwelling-type specific annual rates of waste generation

We used the ratio between the reported average detached and multi-unit dwelling generation rates from APC (2019), to calibrate adjusted per-dwelling generation rates by dwelling type from the overall council-reported annual waste generation data (NSW EPA, 2017). We did this for each council, such that variations in waste generation behaviours between councils as a result of varied socioeconomic, and other waste drivers, are implicitly accounted for. Impacts of within-LGA variation in socioeconomics on waste generation was not accounted for in our model. This was due to lack of waste generation data at scales finer than the LGA level. Improving the results of our model to account for variation in waste generation behaviours within LGAs could be achieved via sophisticated small-area estimation methods, however is beyond the scope of research.

Table 3: Summary of surveyed per-dwelling generation rates by waste fraction and dwelling type from APC (2019)

Dwelling type	Residual fraction	Dry recyclables	Garden organics
Detached dwellings [kg/hh/wk]	10.6±4	4.6±1	4.1±3
Multi-unit dwellings [kg/hh/wk]	6.3±3	2.5±2	0.7±0.6
Detached:multi-unit ratio (R)	1.6	1.8	6.8

We first define for each LGA an overall per-dwelling generation rate estimate for waste fraction f.This simple estimate is calculated over all dwellings, and does not take into account dwelling types (Equation 9):

$$P_f = \frac{Q_f}{N} \tag{9}$$

Where P_f is the simple per-dwelling waste generation rate, Q is the known quantity of waste generated from the LGA for waste fraction f in 2016 from the data (NSW EPA, 2017), and N is the total number of dwellings in the LGA, from 2016 census data (Australian Bureau of Statistics, 2016b). Given the differences in detached and multi-unit generation in APC (2019), total waste generation in an LGA can be calculated by summing detached dwelling waste, and multi-unit dwelling waste generated annually (Equation 10):

$$\hat{Q}_f = \sum_T (P_{f,T} \times N_T) \tag{10}$$

Where $\hat{Q}_f$ is the estimated waste generated, and N_T are the number of dwellings of type T in the LGA. As $P_{f,T}$ is not known and must be estimated, we introduce a parameter β_T to Equation 10 such that $P_{f,T}$ can be found based on the the overall dwelling generation rate estimate P_f estimate (Equation 11):

$$\hat{Q}_f = \sum_T (\beta_T P_f \times N_T) = (\beta_{\text{det}} P_f \times N_{\text{det}}) + (\beta_{\text{mul}} P_f \times N_{\text{mul}}) \tag{11}$$

Values for β_T in Equation 11 can be found via optimisation. We applied a non-linear constrained least squares approach (Schittkowski, 1988), by fitting β_T parameters that minimise the squared error between the simple per-dwelling rate P_f (Eq.: 9), and the recalculated simple per-dwelling rate $\hat{P}_f$. $\hat{P}_f$ is calculated based on the estimated value $\hat{Q}_f$ in Equation 10 with estimated β_T parameters: $\hat{P}_f = \hat{Q}_f/N$, where N is total number of dwellings in the LGA. We set the constraints, such that the ratio R of the estimated dwelling type generation rates are equal to that in APC (2019), and that β_T take only positive values greater than or equal to 0. We solved the optimisation problem (Equation 12) using the generalised reduced gradient method–a popular method to solve general nonlinear optimisation problems (Lasdon et al., 1978; Maia et al.,

2017), implemented using the *Solver* tool in the Microsoft Excel software:

$$\begin{aligned} \min_{\beta_T} \quad & (P_f - f(\beta_T))^2 \\ \text{where} \quad & f(\beta_T) = \hat{P}_f = \sum_T \beta_T (P_f \times N_T)/N \\ \text{subject to} \quad & \frac{\beta_{\text{det}}}{\beta_{\text{mul}}} = R \\ & \beta_T \geq 0 \end{aligned} \tag{12}$$

Equation 12 was solved heuristically, with initial values for β_{det} and β_{mul} chosen as $\sqrt{R}$ and $1/\sqrt{R}$ respectively, such that the ratio of the estimated dwelling type generation rates equals R in the initial solution. Estimated per-dwelling generation rates estimated for each council from our method can be found in Appendix C.

2.3.2. Estimating quantities of waste generated at the property lot

With values for $\hat{P}_{f,T}$ determined from $\beta_T P_f$, we estimate the total annual household waste generated at a property lot for waste fraction f in LGA c:

$$\hat{Q}_{f,c}(l) = \sum_T [\hat{P}_{c,f,T} \times X_T(l)] \tag{13}$$

We performed sensitivity analysis on selection of initial values, by varying values of R. From data in (APC, 2019), we can determine the likely range of the ratio R, based on the sample of 13 LGAs. We estimate a minimum and a maximum value for R from the 95% confidence range in (APC, 2019). We then calculate a range of values for $\hat{Q}_{f,c}(l)$ in Equation 13, based on $[R_{\min}, R_{\max}]$. Model validation is further described in the following section.

2.4. Model validation

The method for this study was developed owing to a paucity of high spatial resolution data, and model validation is an inherent challenge. Kontokosta et al. (2018) approach this challenge by comparing their building level estimates of waste generation with available data aggregated to larger spatial scales (in that instance, New York City sanitation sub-sections). Goerndt et al. (2019) take a similar approach in their study estimating biomass availability, and so to do Cockx & Canters (2015) in their paper on improved dasymetric population mapping. The approach of comparing aggregated estimates against source data, or data available at other aggregation levels, is typical for quantitatively comparing spatial disaggregation procedures (Li et al., 2007; European Commission, 2019; Monteiro et al., 2019).

Considering the absence of reliable validation data at the property lot level, we verified our model by comparing results aggregated to larger spatial scales where actual data is available. This is done under the

assumption that accurate estimations of waste generation at aggregated spatial units would be associated with reliable estimations at the property lot level.

To verify the dwelling distribution estimates, we first conducted a Monte-Carlo simulation by performing 1,000 iterations of the model as described in the previous sections. Considering estimated dwellings per lot is a random variable, the Monte-Carlo simulation generates a probability distribution of estimated dwelling counts per lot, characterising the sensitivity of dwelling counts to randomness in our model. We aggregate these results to the SA1 level, to compare actual versus estimated dwelling counts by type. We also estimate the mean absolute percentage error (MAPE) of SA1 aggregated dwelling estimates, to give an indication of the variation in accuracy across SA1s in the study area. We also present dwelling count estimates as a range at the 95% confidence level, to compare the actual data with the range of possible estimates from our model.

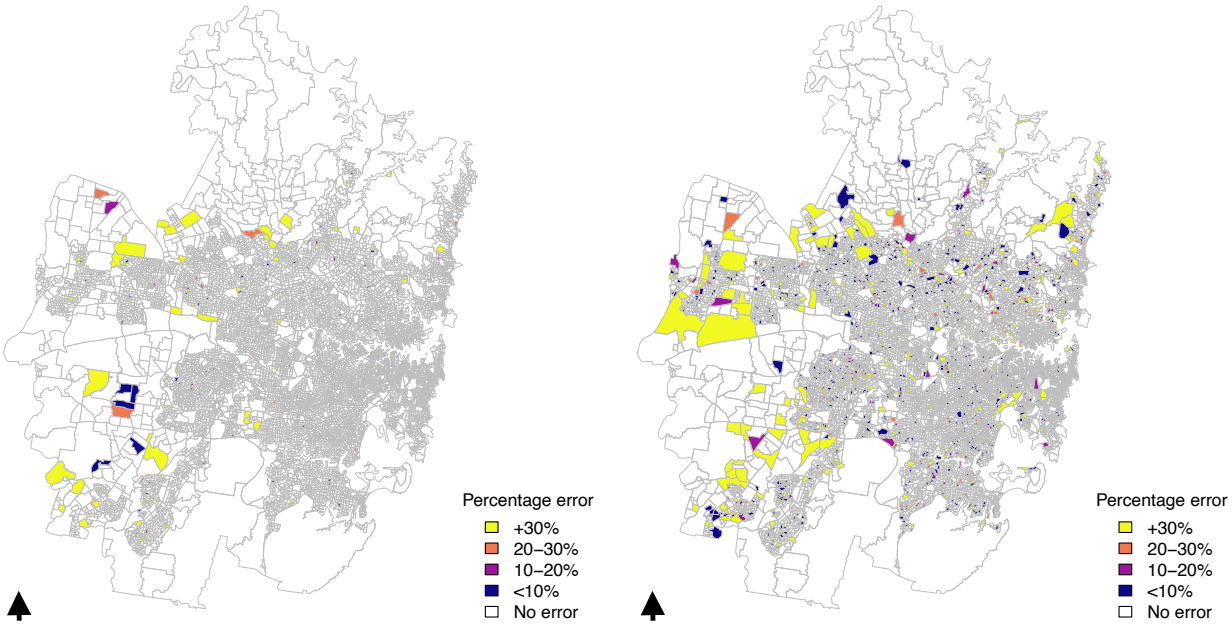
To validate the estimates of waste generated for each property lot, we also aggregate our property lot estimates to the LGA scale, and compare with the LGA reported waste data (NSW EPA, 2017). We first test for sensitivity on the estimation of the dwelling-type waste generation rates, by estimating minimum and maximum (at the 95% confidence interval) estimates for $\hat{P}_{f,c,T}$. Further, we combine the minimum and maximum estimates of $\hat{P}_{f,c,T}$ with the distribution of $X_T(l)$ estimates generated from the Monte-Carlo simulation, to derived a 95% confidence range of $\hat{Q}_{f,c}(l)$ estimates, which are aggregated to the LGA level and compared to actual data. Similar to the dwelling estimates, we calculate MAPE values to characterise the model accuracy when estimates are aggregated to the LGA level.

3. Results and Discussion

3.1. Estimation of dwelling counts at the property lot

Figure 6 shows the distribution of errors when comparing mean dwelling counts with actual data at the SA1 level. This data is also summarised in Table 4. Errors in the dwelling counts at the SA1 level occur where the fitting procedures described (Appendices 1 and 2) fail to converge on the optimal solution (that is, estimated SA1 dwelling count is equal to actual data). Causes for this are generally due to inconsistencies in the data, for example where the number of property lots within an SA1 is less than the number of occupied dwellings. Approximately 98% of all SA1s had prediction errors of 10% or less compared to actual detached dwelling count data. For multi-unit dwelling estimates, 92% of SA1s had prediction errors of 10% or less, indicating that model accuracy is good, and acceptable for the aims of this study.

It was expected that the error for multi-unit dwellings would be greater than detached dwellings, due to the wider range of potential values for the random variable $X_{\text{mul}}(l)$. The Moran's I test for spatial autocorrelation (Bivand et al., 2013) was performed on the absolute percentage error values for each dwelling



(a) Detached dwellings

(b) Multi-unit dwellings

Figure 6: Spatial distribution of estimated dwelling count errors by dwelling type at the SA1 scale, from the dwelling distribution model

Table 4: Summary dwelling distribution estimation errors at the SA1 scale

	No error	<10%	10-20%	20-30%	30%+
Detached dwellings	97.70%	0.46%	0.30%	0.23%	1.31%
Multi-unit dwellings	84.51%	7.10%	1.25%	1.04%	6.10%

type, revealing that errors are randomly distributed, indicating there is no significant systematic spatial bias in model errors (Table 5).

Table 5: Summary of Moran's I analysis of clustering of SA1 absolute percentage values. Values of I approaching 1 indicate clustering of observations, and values approaching -1 indicate perfect dispersion. Values around 0 indicate random dispersion

	I-statistic	p-value	Interpretation
Detached dwellings	0.0045	0.002	Random dispersion
Multi-unit dwellings	0.1148	0.002	Random dispersion

Table 6 shows the total number of estimated dwellings in the study area using our model, compared with actual dwelling counts by type from the 2016 census (Australian Bureau of Statistics, 2016b). The estimated mean dwelling counts and range in Table 6 are derived from the Monte Carlo analysis as described in Section 2.4. Errors are calculated by comparing mean dwelling count estimates with the actual data. The accuracy of the dwelling distribution model is good, with detached dwellings having a mean absolute percentage error (MAPE) value of 1.3%, and multi-unit dwellings a value of 5.2% when aggregated to the SA1 level. The error on multi-unit dwellings being greater compared to detached dwellings is expected, given the additional modelling complexity for estimating multi-unit dwellings. However, there is a consistent under-estimation of dwellings for both dwelling type estimations. This can be attributed to cases where the fitting procedure described in Section 2.2 did not converge on an estimated dwelling count equal to actual data at the SA1 level. Reasons for non-convergence included cases where the underlying data is inconsistent, for example where estimated dwelling density calculated from mesh block data is not consistent with the expected number of multi-unit dwellings from the SA1 level data.

Table 6: Summary dwelling distribution estimation results

	Detached dwellings	Multi-unit dwellings
Actual (Australian Bureau of Statistics, 2016b)	768,709	650,863
Estimated (mean value)	760,906	614,970
Estimated range	760,884 - 760,928	607,640 - 622,299
MAPE	1.32%	5.24%

To examine dwelling distribution model results at a greater level of detail, Figure 7 shows dwelling count estimates for the sample local government area of Burwood. We selected this local government area due to its centralised location within the study area, and the fact that there is a mix of low-, medium- and high-density residential, and commercial land-use types. LGA-wide results presented on the left-hand side of Figure 7 are direct estimations of $\hat{N_D}(l)$ from the first iteration of the model performed. Three test SA1s were selected at random within the council area boundaries to compare estimates with aerial imagery from Google Maps (Google, n.d). Areas with multi-unit developments were confirmed from Google Streetview (Google, n.d)

and highlighted for comparison with model estimates. The task of manually validating estimates for SA1s is labour intensive, so a small number of test SA1s were selected for illustration purposes, however summary results in Figure 6 and Table 4 indicate that results are generally consistent across the study area. Estimates of dwelling count standard deviation on the right side of Figure 7 are presented as heat maps derived from the 1,000 iterations of the model performed, with lighter colours indicating greater deviation from the mean. High standard deviation values indicate which property lots within each SA1 are more likely to contain multi-unit dwellings from our model. Property lots with high standard deviation align well with identified multi-unit dwellings from the aerial imagery. Notable exceptions can be seen in test SA1 3. Differences between the aerial imagery and our estimates for this test SA1 can be explained by differences in the time when the aerial imagery which was captured (2019) compared to census data collection (2016). This was confirmed for test case 3 where it was observed that the actual number of multi-unit dwellings within the SA1 does not align with the number of multi-unit dwellings counted from Google Maps and Streetview, likely owing to additional dwellings being constructed between 2016 and 2019.

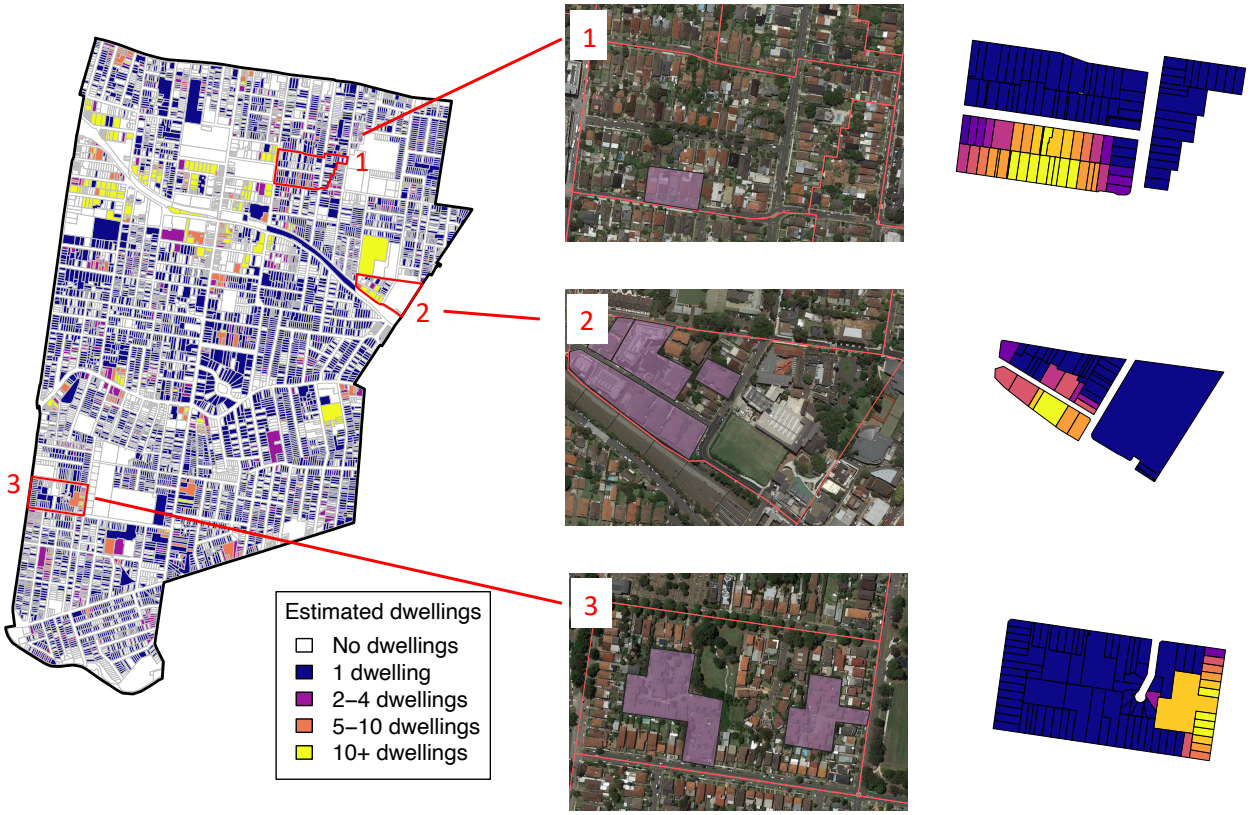


Figure 7: Illustration of property lot-scale results for the example local government area of Burwood. Aerial images correspond to the three sample SA1s in Burwood, with actual multi-unit dwellings highlighted. Estimates of standard deviation on property lot dwelling estimates are shown on the right, with light colours indicating higher standard deviation values

3.2. Estimation of waste generation at the property lot

Property lot waste generation is calculated by multiplying the estimated number of dwellings by type, by the estimated annual waste generation per-dwelling rates as per Equation 13. Table 7 shows a summary of estimated dwelling-specific annual waste generation rate for each fraction across all 31 LGAs, following the optimisation procedure described in Section 2.3. Full estimated per-dwelling generation rates for each LGA in the study area are located in Appendix C

Table 7: Estimated annual per-dwelling waste generation rates by dwelling type for each waste fraction

Dwelling type	Residual fraction	Dry recyclables	Garden waste
Detached dwellings [t/hh/yr]	0.929±0.2	0.371±0.1	0.394±0.02
Multi-unit dwellings [t/hh/yr]	0.551±0.1	0.202±0.04	0.05±0.1
Detached:multi-unit dwelling ratio (R)	1.6	1.8	6.8

Table 8 shows the total quantity of waste generated by fraction in the SMA compared to actual data (NSW EPA, 2017). SMA aggregated estimates closely align with actual data, with the actual reported waste generated totals falling within the range of estimates from our model (at the 95% confidence interval). Mean absolute percentage error is small, and is consistent across the three waste fractions. There is a consistent under-estimation of total waste generated across all LGAs in the SMA from our model. This error can be traced back to errors in dwelling count estimates as shown in Table 6, where a consistent under-estimation was also identified. Improvements to the dwelling distribution model, namely calibration data, and more up-to-date spatial property lot boundary data may improve the estimates in Table 8.

Table 8: Summary of estimated waste generated compared to actual waste generation data

	Residual waste	Dry recyclables	Garden waste
Actual [t/year]	1,108,845	388,040	324,738
Estimated (mean) [t/year]	1,077,900	377,679	319,721
Estimate range [t/year]	522,803 - 1,549,926	181,746 - 537,735	154,987 - 403,810
MAPE	2.98%	2.76%	2.00%

Property lot waste generation estimates for the SMA are shown in Figure 8. The majority of property lots in the SMA generated up to an estimated 2 tonnes of waste from all collected fractions in 2016. Areas with intense waste generation in Figure 8 are associated with property lots with high dwelling count numbers, such as those lots with high-density multi-unit dwelling types (i.e., apartments). These results can inform future waste management development and planning, for example: high resolution estimates as shown in Figure 8 can inform more efficient waste collection routing; quantities of waste at properties can also be combined with other property-level data such as wastewater to estimate and map total household organic flows (e.g., in Turner et al. (2017))–useful in the planning of precinct or district wide organic waste management pathways.

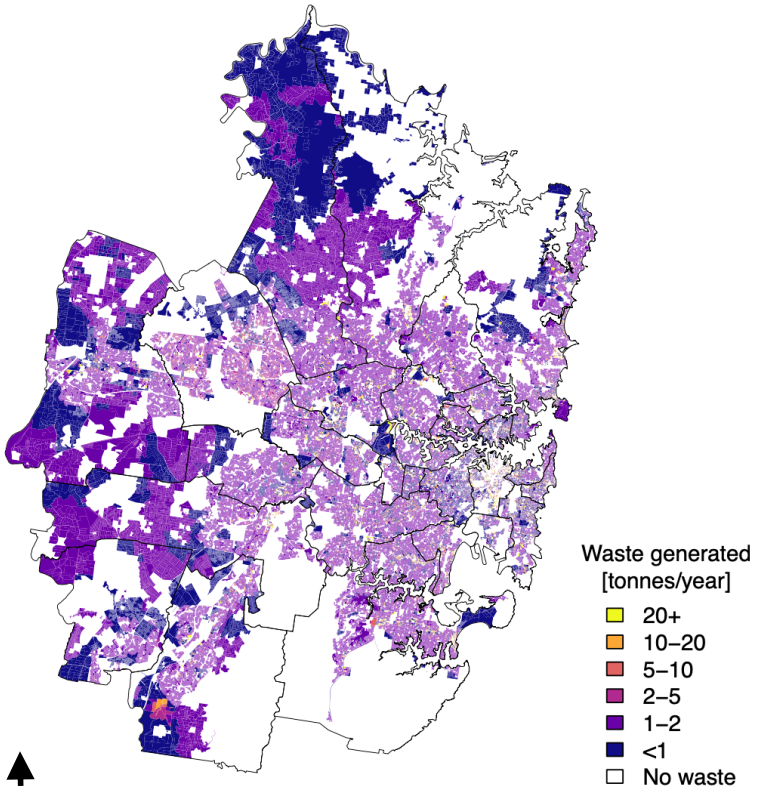


Figure 8: Estimated annual total waste generated at the property lot level for the Sydney metropolitan area

Data in Figure 8 is useful for identifying hot spots of waste generation. For identifying suitable areas for targeted waste interventions, or economically viable locations of recovery facilities, understanding the distribution and availability of supply over an area is important to determine the viability of a new facility (Shi et al., 2008; Sliz-Szkliniarz & Vogt, 2012; Comber et al., 2015; Lozano-García et al., 2020). This application of our results is illustrated in Figure 9, which shows the total garden waste available within an arbitrary 5km collection radius across the study area, without consideration for land use or existing waste management infrastructure. This distance was chosen to represent a small-scale, district-sized recovery process, such as a composting facility. Data in this figure can be used to identify areas with the greatest resource supply to inform decision making around optimal locations for facilities and realistic transportation distances. Locations in the northern areas of the SMA have the greatest available quantities of garden waste in general and are ideal locations for suitable organics recovery. This is expected, as suburbs in this area of the SMA are well vegetated, have large property lot sizes, and have a high proportion of detached dwellings.

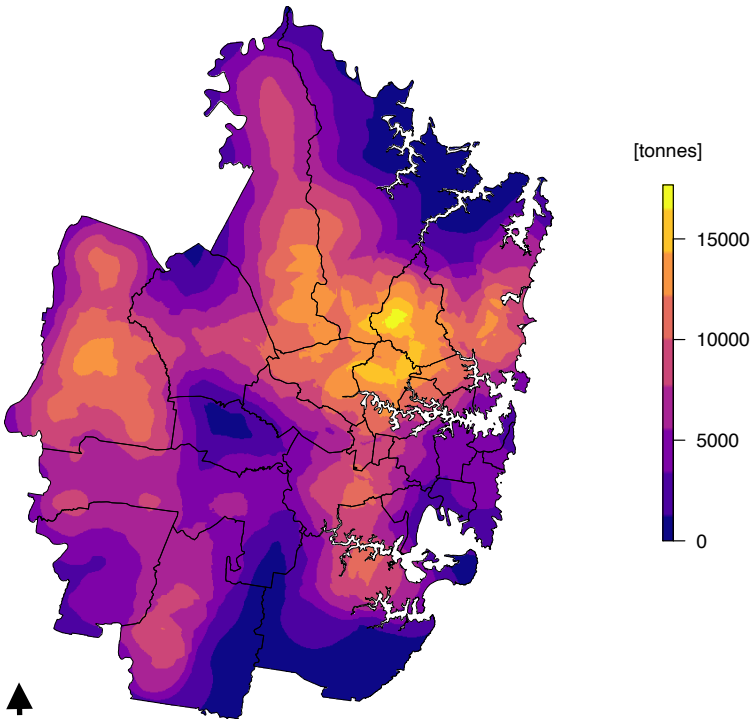


Figure 9: Availability of municipal garden waste within a 5km radius

A clear 'hot-spot' of resource availability can be identified in Figure 9. This garden organics supply area is located within the northern Sydney Ku-ring-gai council area. Figure 10 shows this supply area in greater detail. From this analysis, there is an approximate annual supply of garden waste of 16,800 tonnes of garden waste per year, within a supply area of approximately 13km^2. This feedstock availability would suit medium scale (approx. 20,000t/year throughput), municipal garden organics processing facility, potentially

processing waste from adjacent council areas.

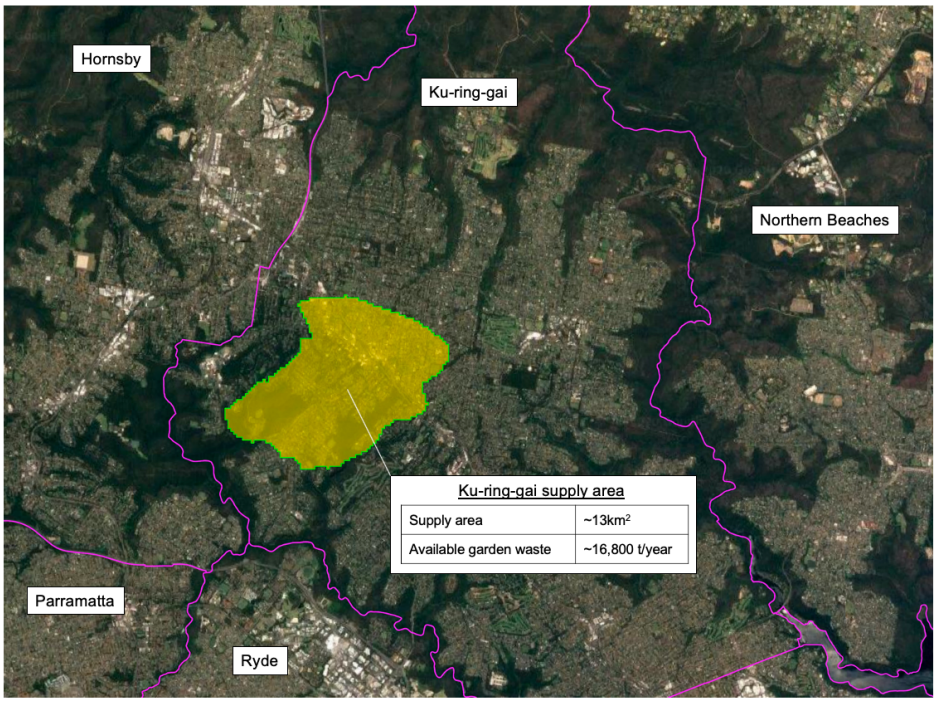


Figure 10: Detail of the identified garden waste supply hot-spot, located in the northern suburbs of Sydney

The data shown in Figures 9 and 10 is illustrative of potential further applications of the model presented in this paper, and a more robust study on the fine scale availability of the waste supply, and optimal locations for economically viable recovery processes is outside the scope of this work. The data in Figure 9 does serve as an example of the capabilities of high spatial resolution waste estimates for informing strategic waste policy and planning.

4. Conclusion

The aim of this research was to develop an approach to estimate high resolution waste generation data at the property lot level, that might be used to inform future waste management planning and infrastructure development. We developed a spatial model to disaggregate council level waste generation down to the property lot scale with a high degree of accuracy. The modelling approach requires modest data inputs, and enables a detailed appraisal of the distribution of household waste generation despite significant gaps in available data. Specific applications for the data generated from this model include waste collection planning and facility location identification, and the data is in a granular form that can be readily combined with other property-level data sets, for example household wastewater flows.

This study adds to the literature on spatial estimation methods for urban waste generation, and on data

driven waste management policy. For the first time, the approach was applied to an important population centre in Australia and the approach could be applied to other jurisdictions. Data generated from our model is accurate, and model performance can be improved if more up-to-date validation data becomes available. Moreover, the approach presented may also have value in estimating commercial and industrial waste at a fine spatial scale, and is worthy of further research.

Acknowledgement

We thank the anonymous reviewers whose thoughtful recommendations helped improve the manuscript. We also thank Dr. Melita Jazbec and Andrea Turner from the Institute for Sustainable Futures for reviewing earlier iterations of this work.

References

Albright, D., Lee, H., McDaniel, J., Kroner, D., Davis, J., Godfrey, K., & Li, Q. (2019). Small area estimation of human papillomavirus vaccination coverage among school-age children in Alabama counties. *Public Health*, *177*, 120–127.

APC (2008). *SSROC Regional Waste Audit Report - 2008*. Technical Report Anne Prince Consulting.

APC (2019). *SSROC Kerbside Waste Audit: Regional Report*. Technical Report Anne Prince Consulting North Sydney.

Australian Bureau of Statistics (2016a). 12701.0.55.001 - Australian Statistical Geography Standard (ASGS), July 2016. URL: `https://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.001`.

Australian Bureau of Statistics (2016b). *Census of Population and Housing*. Technical Report Australian Bureau of Statistics.

Bivand, R., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied Spatial Data Analysis with R*. (Second edition ed.). Springer, New York.

Bivand, R., & Rundel, C. (2020). rgeos: Interface to Geometry Engine - Open Source ('GEOS'). R package version 0.5.3. https://CRAN.R-project.org/package=rgeos.

Buil-Gil, D., Medina, J., & Shlomo, N. (2019). The geographies of perceived neighbourhood disorder. A small area estimation approach. *Applied Geography*, *109*, 102–113.

Bunker, R. (2014). How is the compact city faring in Australia. *Planning Practice and Research*, *29*, 449–460.

Chandra, H., Salvati, N., Chambers, R., & Tzavidis, N. (2012). Small area estimation under spatial nonstationarity. *Computational Statistics and Data Analysis*, *56*, 2875–2888.

Cockx, K., & Canters, F. (2015). Incorporating spatial non-stationarity to improve dasymetric mapping of population. *Applied Geography*, *63*, 220–230.

Comber, A., Dickie, J., Jarvis, C., Phillips, M., & Tansey, K. (2015). Locating bioenergy facilities using a modified GIS-based location-allocation-algorithm: Considering the spatial distribution of resource supply. *Applied Energy*, *154*, 309–316.

Council of Australian Governments (2020). *Phasing out exports of waste plastic, paper, glass and tyes - response strategy to implement the August 2019 agreement of the Council of Australian Governments*. Technical Report Council of Australian Governments.

Department of Finance, Services and Innovation (2012). NSW Digital Cadastre Data Base. URL: `https://sdi.nsw.gov.au/catalog/search/resource/details.page?uuid=921D6586-8594-4473-A923-A446DEBF3336`.

Department of the Environment and Energy (2018). *National Waste Report 2018*. Technical Report Prepared for the Department of the Environment and Energy.

Eberth, J., McLain, A., Hong, Y., Sercy, E., Diedhiou, A., & Kilpatrick, D. (2018). Estimating county-level tobacco use and exposure in South Carolina: a spatial model-based small area estimation approach. *Annals of Epidemiology*, *28*, 481–488.

Edwards, J., Othman, M., Burn, S., & Crossin, E. (2016). Energy and time modelling of kerbside waste collection: changes incurred when adding source separated food waste. *Waste Management*, *56*, 454–465.

Esin, T., & Cosgun, N. (2007). A study conducted to reduce construction waste generation in turkey. *Building and Environment*, *42*, 1667–1674.

European Commission (2019). *Guidelines on small area estimation for city statistics and other functional geographies*. Technical Report European Commission - Manuals and guidelines - General and regional statistics.

Fabrizi, E., & Trivisano, C. (2016). Small area estimation of the Gini concentration coefficient. *Computational Statistics and Data Analysis*, *99*, 223–234.

Ferrão, P., & Fernández, J. (2013). *Sustainable Urban Metabolism*. MIT Press, Cambridge.

Fortin, M., Boots, B., Csillag, F., & Remmel, T. (2003). On the role of spatial stochastic models in understanding landscape indices in ecology. *Oikos*, *102*, 203–212.

Fuqua, M., & Doty, J. (2012). Data-driven stochastic model development for unknown data sources. In *50th AAIA Aerospace Sciences Meeting, Nashville, Tennessee*.

Goerndt, M., Wilson, B., & Aguilar, F. (2019). Comparison of small are estimation methods applied to biopower feedstock supply in the Northern U.S. region. *Biomass and Bioenergy*, *121*, 64–77.

Google (n.d). Google Maps Burwood search area.

Haight, F. (1967). *Handbook of the Poisson Distribution*. John Wiley and Sons.

Hannan, M., Akhtar, M., Begum, R., Basri, H., Hussain, A., & Scavino, E. (2018). Capacitated vehicle-routing problem model for scheduled solid waste collection and route optimization using PSO algorithm. *Waste Management*, *71*, 31–41.

Hayes, J., Wilson, T., Brown, C., Deligne, N., Leonard, G., & Cole, J. (2021). Assessing urban disaster waste management requirements after volcanic eruptions. *International Journal of Disaster Risk Reduction*, *52*.

Hijmans, R. (2020). *raster: Geographic Data Analysis and Modeling. R packaging version 3.0-12*. Technical Report https://CRAN.R-project.org/package=raster.

Kontokosta, C., Hong, B., Johnson, N., & Starobin, D. (2018). Using machine learning and small area estimation to predict building-level municipal solid waste generation in cities. *Computers, Environment and Urban Systems*, *70*, 151–162.

Kravchenko, A. (2003). Influence of spatial structure on accuracy of interpolation methods. *Soil Science Society of America Journal*, *67*, 1564–1571.

Lasdon, L., Warren, A., Jain, A., & Ratner, M. (1978). Design and testing of a generalized reduced gradient code for nonlinear programming. *ACM Transactions on Mathmatical Software*, *4*, 34–50.

Li, T., Pullar, D., Corcoran, J., & Stimson, R. (2007). A comparison of spatial disaggregation techniques as applied to population estimation for South-East Queensland (SEQ), Australia. *Applied GIS*, *3*, 1–16.

Lin, Z., Xie, Q., Feng, Y., Zhang, P., & Yao, P. (2020). Towards a robust facility location model for construction and demolition waste transfer stations under uncertain environment: The case of Chongqing. *Waste Management*, *105*, 73–83.

Lou, X., Nair, J., & Ho, G. (2013). Potential for energy generation from anaerobic digestion of food waste in Australia. *Waste Management and Research*, *31*, 283–294.

Lozano-García, D., Santibañez-Aguilar, J., Lozano, F., & Flores-Tlacuahuac, A. (2020). GIS-based modeling of residual biomass availability for energy and production in Mexico. *Renewable and Sustainable Energy Reviews*, *120*, 109610.

Lu, G., & Wong, D. (2008). An adaptive inverse-distance weighting spatial interpolation technique. *Computers and Geosciences*, *34*, 1044–1055.

Maia, A., Ferreira, E., Oliveira, M., Menezes, L., & Andrade-Campos, A. (2017). Numerical optimization strategies for springback compensation in sheet metal forming. *Computation Methods and Production Engineering*, (pp. 51–82).

Monteiro, J., Martins, B., Murrieta-Flores, P., & Pires, J. (2019). Spatial disaggregation of historical census data leveraging multiple sources of ancillary information. *International Journal of Geo-Information*, *8*, 327.

Mueller, T., Pierce, F., Schabenberger, O., & Warncke, D. (2001). Map quality for site-specific fertility management. *Soil Science Society of America Journal*, *65*, 1547–1558.

NATSEM (2008). *A review of small area estimation problems and methodological developments*. Technical Report National Centre for Social and Economic Modelling, University of Canberra.

NSW EPA (2014). *Domestic kerbside waste and recycling in NSW: results of the 2011 waste audits*. Technical Report State of NSW and Environment Protection Authority.

NSW EPA (2017). *NSW Local Government Waste and Resource Recovery Data Report 2014-15*. Technical Report State of NSW and Environment Protection Authority.

NSW Government (2019). *NSW Circular Economy Policy Statement - Too Good to Waste*. Technical Report NSW Government Sydney, Australia.

Parzen, E. (2015). *Stochastic Processes*. Courier Dover Publications.

Rich, J. (2018). Large-scale spatial population synthesis for Denmark. *European Transport Research Review*, *10*.

Roberts, M., Bruce, A., & MacGill, I. (2019). Opportunities and barriers for photovoltaics on multi-unit residential buildings: Reviewing the australian experience. *Renewable and Sustainable Energy Reviews*, *104*, 95–110.

Sarmah, S., Yadav, R., & Rathore, P. (2019). Development of vehicle routing model in urban solid waste management system under periodic variation: a case study. *IFAC PapersOnLine*, *52*, 1961–1965.

Schittkowski, K. (1988). Solving constrained nonlinear least squares problems by a general purpose SQP-method. In K.-H. Hoffman, J. Zowe, J.-B. Hiriart-Urruty, & C. Lemarechal (Eds.), *Trends in Mathematical Optimization. 4th French-German Conference on Optimization*.

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 ACM National Conference*, (pp. 517–524).

Shi, X., Elmore, A., Li, X., Gorence, N., Jin, H., Zhang, X., & Wang, F. (2008). Using spatial information technologies to select sites for biomass power plants: A case study in Guangdon Province, China. *Biomass and Bioenergy*, *32*, 35–43.

Sliz-Szkliniarz, B., & Vogt, J. (2012). A GIS-based approach for evaluating the potential of biogas production from livestock manure and crops at a regional scale: A case study for the Kujawsko-Pomorskie Voivodeship. *Renewable and Sustainable Energy Reviews*, *16*, 752–763.

Tabata, T., Onishi, A., Saeki, T., & Tsai, P. (2019). Earthquake disaster waste management reviews: Prediction, treatment, recycling, and prevention. *International Journal of Disaster Risk Reduction*, *36*.

Truong, P., & Stein, A. (2019). Model-based small area estimation at two scales using Moran's spatial filtering. *Spatial and Spatio-temporal Epidemiology*, *31*, 100303.

Turner, A., Fam, D., Madden, B., & Liu, A. (2017). *Pyrmont-Ultimo Precinct (PUP) Scale Organics Management Scoping Study*. Technical Report prepared for Sydney Water Corporation and the NSW Environment Protection Authority by the Institute for Sustainable Futures, University of Technology Sydney.

Vu, H., Bolingbroke, D., Ng, K., & Fallah, B. (2019). Assessment of waste characteristics and their impact on GIS vehicle collection route optimization using ANN waste forecasts. *Waste Management*, *88*, 118–130.

Waste Management Review (2020). The MUD dilemma. https://wastemanagementreview.com.au/the-mud-dilemma/.

Wu, Y.-H., & Hung, M.-C. (2016). Applications of Spatial Statistics. chapter Comparison of spatial interpolation techniques using visualization and quantitative assessment. IntechOpen.

Yadav, V., Bhurjee, A., Karmakar, S., & Dikshit, A. (2017). A facility location model for municipal solid waste management system under uncertain environment. *Science of the Total Environment*, *603-604*.

Yadav, V., Karmakar, S., Dikshit, A., & Bhurjee, A. (2018). Interval-valued facility location model: An appraisal of municipal

544 solid waste management system. *Journal of Cleaner Production*, *171*, 250–263.

Appendix A. Algorithm 1: Estimation of detached dwellings at the property lot

Here, $X_{\text{det},\mathcal{S}}$ is the set of all values of $X_{\text{det}}(l)$ for an SA1 $\mathcal{S}$.

Algorithm 1 Estimation of $X_{\text{det},\mathcal{S}}$

Require: k, p, L^a, $n = |L^a|$, $N_{\text{det}}(\mathcal{S})$

Let $\Gamma = \emptyset$ be a container set

Set maximum iterations $max.iter = 1000$

Set counter $i = 0$

while $i \leq max.iter$ **do**

$i = i + 1$

Estimate $\bar{X}_{\text{det},\mathcal{S}}$ for lots in L^a (Eq: 6)

Append $\bar{X}_{\text{det},\mathcal{S}}$ to the container: $\Gamma \cup \bar{X}_{\text{det},\mathcal{S}}$

Calculate $\hat{N}_{\text{det}}(\mathcal{S}) = \sum_l \bar{X}_{\text{det}}(l)$

Calculate $\delta_i = |N_{\text{det}}(\mathcal{S}) - \hat{Y}_{\text{det}}(\mathcal{S})|$

if $\delta_i = 0$ **then**

$X_{\text{det},\mathcal{S}} = \bar{X}_{\text{det},\mathcal{S}}$

return $X_{\text{det},\mathcal{S}}$

else if $i = max.iter$ **then**

$j = \arg\min_i \delta_i$

$X_{\text{det},\mathcal{S}} = \Gamma_j$

return $X_{\text{det},\mathcal{S}}$

end if

end while

Appendix B. Algorithm 2: Estimation of multi-unit dwellings at the property lot

Here $X_{\text{mul},\mathcal{S}}$ is the set of all values of $X_{\text{mul}}(l)$ for an SA1 $\mathcal{S}$.

Algorithm 2 Estimation of $X_{\text{mul},\mathcal{S}}$

Require: λ, L^b, $N_{\text{mul}}(\mathcal{S})$

Let $\Gamma = \emptyset$ be a container set

Set maximum iterations $max.iter = 1000$

Set counter $i = 0$

while $i \leq max.iter$ **do**

$i = i + 1$

Estimate $\bar{X}_{\text{mul},\mathcal{S}}$ for lots in L^b (Eq: 8)

Append $\bar{X}_{\text{mul},\mathcal{S}}$ to the container: $\Gamma \cup \bar{X}_{\text{mul},\mathcal{S}}$

Calculate $\hat{N}_{\text{mul}}(\mathcal{S}) = \sum_l \bar{X}_{\text{mul}}(l)$

Calculate $\delta_i = |N_{\text{mul}}(\mathcal{S}) - \hat{N}_{\text{mul}}(\mathcal{S})|$

if $\delta_i = 0$ **then**

$X_{\text{mul},\mathcal{S}} = \bar{X}_{\text{mul},\mathcal{S}}$

return $X_{\text{mul},\mathcal{S}}$

else if $i = max.iter$ **then**

$j = \arg\min_i \delta_i$

$X_{\text{mul},\mathcal{S}} = \Gamma_j$

return $X_{\text{mul},\mathcal{S}}$

end if

end while

Appendix C. Estimated per-dwelling waste generation rates

Table C.9: Estimated per-dwelling waste generation estimates [kg/household]

	Residual waste		Dry recyclables		Garden waste	
LGA	Detached	Multi-unit	Detached	Multi-unit	Detached	Multi-unit
Blacktown	1371.8	1227.6	231.3	206.9	566.0	391.3
Botany Bay	749.1	592.5	314.1	248.5	86.9	68.7
Burwood	793.6	677.4	256.9	219.3	225.9	192.8
Camden	864.3	536.5	376.0	233.4	449.5	279.0
Campbelltown	836.3	618.6	331.9	245.5	376.6	278.6
Canada Bay	735.3	552.1	330.6	248.2	192.0	144.2
Canterbury-Bankstown	920.8	760.9	287.4	237.5	281.0	232.2
Cumberland	1122.3	945.4	264.8	223.0	161.3	135.9
Fairfield	1265.3	997.3	206.1	162.4	31.4	24.7
Georges River	724.6	593.9	267.5	219.2	276.5	226.6
Hornsby	763.6	558.2	328.8	240.3	393.3	287.5
Hunters Hill	752.7	469.9	296.5	185.1	288.5	180.1
Inner West	739.7	562.3	337.4	256.5	143.2	108.8
Ku-ring-gai	706.6	483.0	360.8	246.7	504.6	345.0
Lane Cove	636.0	422.1	274.0	181.8	537.4	356.7
Liverpool	993.5	716.2	327.1	235.8	304.3	219.4
Mosman	810.7	511.7	321.5	202.9	271.6	171.5
Northern Beaches	587.5	406.9	308.4	213.6	74.7	51.7
North Sydney	768.8	503.5	400.1	262.0	497.6	325.8
Parramatta	796.8	644.7	242.2	196.0	251.4	203.4
Penrith	595.2	411.5	352.6	243.8	566.0	391.3
Randwich	735.1	525.8	336.1	240.4	162.2	116.0
Rockdale	1210.5	933.2	328.0	252.9	2.0	1.5
Ryde	744.0	533.7	292.8	210.0	291.7	209.3
Strathfield	861.4	697.2	254.9	206.3	148.6	120.2
Sutherland Shire	810.1	533.9	364.4	240.1	411.7	271.3
City of Sydney	735.4	543.8	244.9	181.1	28.4	21.0
The Hills Shire	896.8	692.6	318.7	246.1	392.5	303.1
Waverley	843.5	584.3	491.4	340.4	127.0	87.9
Willoughby	709.3	530.9	311.9	233.5	305.2	228.4
Woollahra	838.3	554.3	428.9	283.6	258.1	170.6