APPENDIX 3: HUMAN SETTLEMENT AREAS (VERSION 1.2)

SFMC has developed a spatial dataset used to define the extent and name of Human Settlement Areas (HSA) in Tasmania. The term Human Settlement Area is used instead of Community for the purposes of clarity. HSAs are defined as where people live or work.

Dataset Name	HSA_v1_2
Dataset Created	30-01-2014
Data Type	Vector
Data Coverage	Tasmania
Projected Coordinate System	GDA_1994_MGA_Zone_55
Projection	Transverse Mercator

Input Data

- Building point and polygon data from Information and Land Services (ILS), DPIPWE
- Population and Dwelling data from Australian Bureau of Statistics (ABS)
- Cadastral data from ILS
- Localities and Nomenclature data from ILS

Lineage

Version 0 – Beta version made available for comment and feedback to representatives from SFMC, TFS and DPAC (HSA working group).

Version 1 – Minor changes from version 0, based on feedback from HSA working group limited to naming of HSA, no spatial extent changes from version 0.

Version 1.1 – Minor changes from version 1, total HSA count increases from 712 to 715 based on splitting three wrongly dissolved polygons

Version 1.2 (*CURRENT VERSION*) – Minor changes from version 1.1, three polygons had incorrectly identified Suburb and LGA names.

Simplified Model Process

Input Points >>> Weighting >>> Weighted Points >>> Kernel Density >>> Final Output

Methodology

Building polygon data is converted to point (centroid) and merged with building point data to create a continuous building point layer Statewide.

Building points that are deemed not relevant to defining HSAs (as listed below) are removed, percentages are shown for how much of the original (unmodified) merged polygon/point dataset they account for.

- Lighthouses (<0.01%)

ISBN: 978-0-9925970-0-9

- Public toilets (<0.2%)
- Walking huts (<0.05%)
- Sheds* (14.5%)
- Remote Sheds* (<0.2%)
- Ruins (<0.05%)

**Whilst sheds and remote sheds are excluded, "Rural small sheds" and "Rural large sheds" have been included*

Furthermore points falling outside the ABS 2011 Mesh Block Statistical Area are removed (less than 0.03% of the dataset). This is predominantly removing a small number of coastal points sitting outside what the ABS defines as land.

The remaining building points are split into two groups:

a) **Group 1** – Residences (66%), Other (0.4%) and Unknown (8.5%)

b) **Group 2** – Community (3.4%), Commercial (1.9%), Industrial (0.8%), Pumphouse (<0.03%), Rural Large Shed (1%), Rural Small Shed (2.8%), Silo (0.3%) and Hothouse (0.1%)

Group 1 is joined with ABS mesh block data, a population field is then calculated by dividing 'Persons Usually Resident' by 'Dwellings" (both fields from ABS data).

ABS defines these two terms as:

Persons Usually Resident: This is the count of people where they usually live, which may or may not be where they were on Census Night. This data is coded from the address supplied to the question "Where does the person usually live?" (Australian Bureau of Statistics, 2012)

Dwellings: A dwelling is a structure which is intended to have people live in it, and which is habitable on Census Night. Some examples of dwellings are houses, motels, flats, caravans, prisons, tents, humpies and houseboats. All occupied dwellings are counted in the Census. Unoccupied private dwellings are also counted with the exception of those in caravan parks, marinas and manufactured home estates. (Australian Bureau of Statistics, 2012)

Before the population field for Group 1 is applied the population and dwelling fields are modified for special cases. In each mesh block if *Dwellings* are equal to zero and *Persons Usually Resident* does not equal zero, then the *Dwelling* field is set to equal one. This is to account for situations like mesh blocks that only include a school, as there are no defined 'dwellings', but we don't want the population field to be assigned a zero value. Another special case is if the *Persons Usually Resident* field equals zero and *Dwellings* does not equal zero, then *the Persons Usually Resident* is set to equal the *Dwellings* field (thereby giving population of 1). The purpose of this is to give weight to dwellings such as in caravan parks that have not been assigned any persons usually resident.

The cadastral dataset is now used to supplement the population weightings whereby all parcels less than or equal to 2.5 hectares (6.18 acres) are selected. From this selection of parcels any building points that are with in this area are selected. For these selected points the minimum population value is raised to the sum of the mean plus two standard deviations of the entire point population dataset. The main purpose of this is to help better capture the shack communities that have high dwelling counts but low population values. As the census is taken in winter, the *Persons Usually Resident* value assigned to mesh blocks is "where people usually live".

Group 2 is now merged back with Group 1 and the population value for Group 2 is set as the previously calculated mean of Group 1. The purpose of this is to weight industrial, community, and commercial areas that might have otherwise received low weighting depending on the cut up of mesh blocks and population figures.

Using the weighted building point population data a kernel density function is run. Kernel density calculates the magnitude per unit area from the point features using a kernel function to fit a

ISBN: 978-0-9925970-0-9

smoothly tapered surface to each point. In effect the density of point features in a neighbourhood around those features is calculated. Furthermore we use the previously defined population field to further weight those features. The process is run using a cell size of 25m and search radius of 500m.

A cell size of 25m is used as a reasonable compromise between computing processing capability and not over generalising by using too coarse a cell size. Also given a cell size of 25m is a fair approximate of mapping at 1:25,000, this is reasonable for the purposes of this dataset and the input data used. The search radius of 500m was used as it was deemed most appropriate (by repetition of model) for the level of smoothing we required. 500m was also deemed an accurate representative distance when contemplating a radius of defining 'settlement area around a point' i.e. how far to search to build your density layer.

From the resulting output the top 15% of the density surface is taken and converted into a vector polygon dataset. A number of cleaning operations are applied to the dataset, including dropping areas less than 20ha in size. The final dataset is intersected with the existing localities dataset to assign suburb and Local Government area names. Finally a nomenclature name field is added by assigning the closest named nomenclature point to the HSA, from a filtered nomenclature list.

Assumptions, Limitations and Strengths

It is assumed the building point dataset contains a reasonably complete and up to date coverage of the location of buildings across Tasmania. Whilst this is true for some areas of the state it could be considered untrue for others. However one of the strengths of this model to generate the HSA dataset is that if the dataset is missing individual points here and there, it should be reasonably compensated by the complementary methods of weighting. The main weakness is when whole areas of points are missing i.e. an entirely new housing development or another wholesale area is missing from the dataset. For example one such area missing data has already been identified in the North West of the state. ILS has been notified and the building point layer for that area has been updated. Hence, a future version of HSA might include that area depending on the resulting adjusted surface score.

The entire process has been built in a scripted environment; this allows for repetition using different parameters or using updated datasets to generate an updated HSA dataset.

ISBN: 978-0-9925970-0-9