Commercial Buildings Baseline Study

Building Stock Model Overview

30 September 2020

Contents

Figures

Tables

1   Introduction

1.1   Building Stock Model overview

The National Energy Analytics Research (NEAR) Program has implemented an Australia-wide Building Stock Model (BSM), in order to support research into the built environment and the associated energy use. A version has been tailored to provide supporting datasets for the Department of Industry, Science, Energy and Resource's (DISER) Commercial Buildings Baseline Study (CBBS).

1.2   Building Stock Model (BSM) Components

The CBBS BSM has three major components:

- Building Classification
- Floorspace Estimation
- Geospatial Referencing

An overview of these components is provided in the following sections.

1.3   Foundational datasets

The BSM is based upon two geospatial datasets produced by Geoscape – Geoscape Buildings and the Geocoded National Address File (G-NAF). They are linked by an address detail persistent identifier.

Other datasets, such as those used for building classification and geospatial referencing, are linked by joining to Geoscape Buildings or G-NAF using text-based keys (address string, building name) or by spatial joins, such as the intersection of points and polygons.

1.3.1   Geoscape Buildings

Geoscape has endeavoured to catalogue all buildings in Australia with a roof area over 9m^2, which amounts to over 15.5 million buildings. Each building has both a polygon (building footprint outline) and a point (longitude, latitude) spatial representation. The Building Address table includes an 'is residential' flag, which helps filter out residential properties for the CBBS (Geoscape, 2019).

As illustrated in Figure 1, Geoscape building characteristics include:

- Building footprint
- Height – maximum and eaves
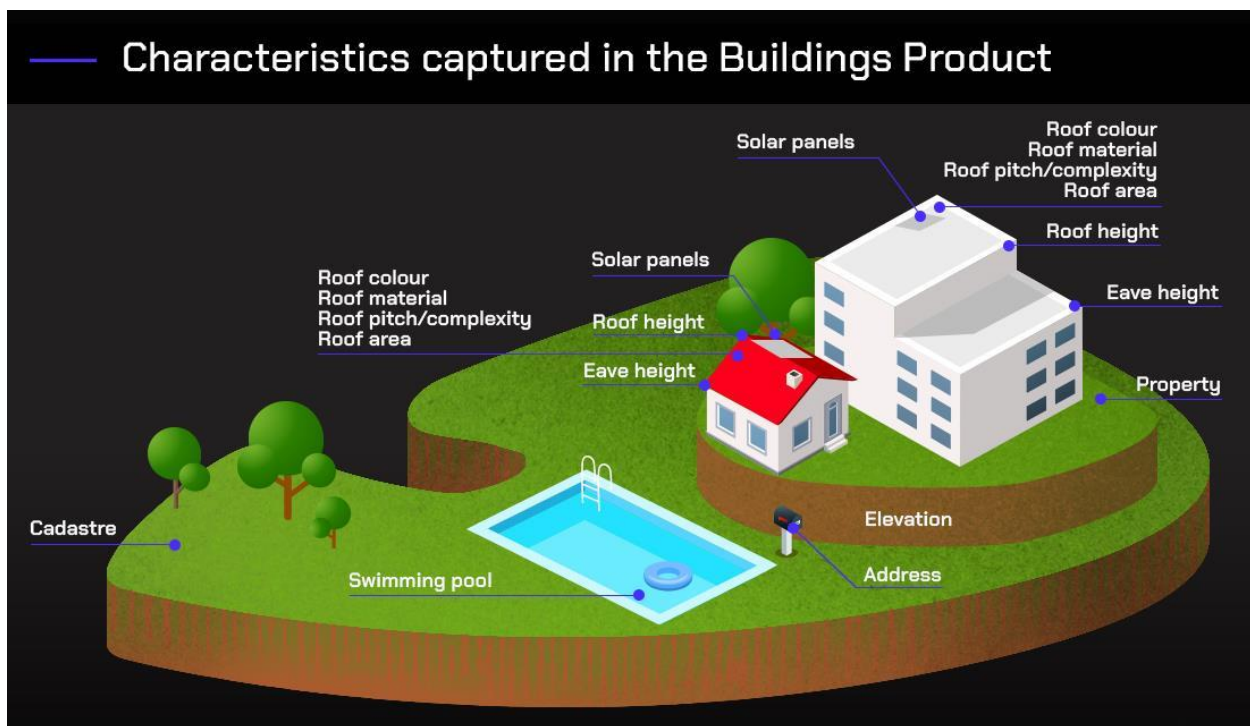- Zero to many G-NAF address references

Figure 1 Geoscape Buildings (Geoscape, 2019)

1.3.2   Geocoded National Address File (G-NAF)

The Geocoded National Address File (G-NAF) is the authoritative source of physical address information in Australia. It aims to capture all addresses, both official and in-use, and currently contains over 14 million entries. Addresses are supplied by over ten entities, including the state and territory land authorities. A Geoscape building may have 0 to many G-NAF addresses. (Geoscape, 2020)

G-NAF contains several variables of interest to the CBBS, for tasks such as classifying buildings, linking with other datasets, and aggregating data by spatial regions. These include:

- Building name
- Address details
    - o Primary address
    - o Unit address(es)
- Legal Parcel ID
- Flat Type Description
- Point location

2 Building Classification

Geoscape Buildings provides physical characteristics of individual buildings, but limited indicators as to their use. G-NAF also provides some limited information, such as through the building name and flat type description.

Enriching datasets need to be linked to Geoscape Buildings to provide taxonomic indicators that can be used to place them into the prescribed CBBS taxonomy. Examples include the NABERS Ratings Register, Commercial Buildings Disclosure (CBD) Downloadable Dataset, and listings of hospitals, schools, aged care facilities and retirement villages.
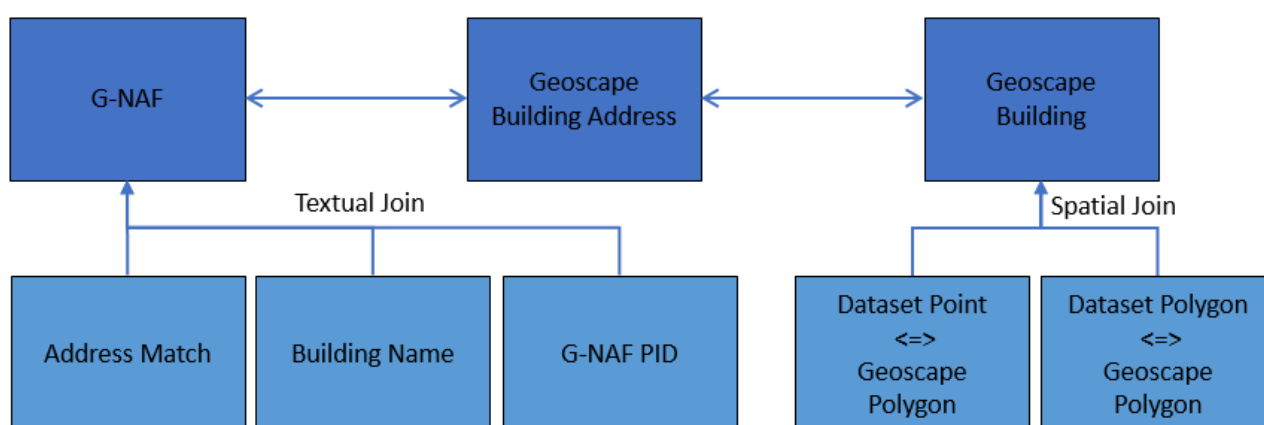
2.1 Methodology

Figure 2 Linking enriching datasets to Geoscape Buildings and G-NAF

1. Link datasets to G-NAF and/or Geoscape Buildings
   As shown in Figure 2, datasets may be linked through text based keys, such as address and/or building name to G-NAF, or by spatial join.

2. Map taxonomic indicators in each enriching dataset to CBBS taxonomy classes
   Datasets may have in excess of 100 feature type indicators to be mapped to a specific CBBS taxon.

3. Apply dataset ranking to determine final class per building, as outlined in Section 2.1.2

2.1.1 CBBS Taxonomy

DISER's industry representative has prescribed a CBBS taxonomy, which approximates the ABS Functional Classification of Buildings (FCB). Some extra classes not in the FCB have been included, such as for Education buildings, where we have more detailed information, and where all that is known is the Mesh Block category. See Table 1 in Appendix A for the CBBS Taxonomy and the building counts per taxon.

2.1.2   Dataset ranking

It is possible that a single Geoscape building could be 'tagged' by multiple datasets. We have implemented a dataset ranking system to give an order of precedence. Consistent, national scale datasets, and those with more specific taxonomic indicators are ranked higher than more general datasets. At the least, a building will be assigned an ABS Mesh Block category.

For example, schools and hospitals appear in numerous datasets, some with a more specific type, such as public hospital versus hospital and public primary school versus school.

The final process is to take the highest ranked dataset for each Geoscape building to ensure that each building is only counted once, such as in the aggregation to the CBBS Data Template.

Table 2 in Appendix B outlines the current dataset ranking hierarchy, and the count of Geoscape buildings 'tagged' by each dataset.

2.1.3   ABS Mesh Block categories

The 2016 ASGS contains 358,122 Mesh Blocks covering the whole of Australia without gaps or overlaps. The ABS has designed Mesh Blocks as an aggregation of land parcels. They are the building blocks for all higher levels of the Australian Statistical Geography Standard (ASGS). Mesh Blocks broadly classify land use into ten categories, including residential, commercial, industrial, primary production and education, and align with major facilities, such as universities, hospitals and retirement villages (Australian Bureau of Statistics, 2020).

Each building will at least be tagged with a MB category, according to the Mesh Block in which the building's centroid lies. Classes have been added to the CBBS taxonomy to capture the cases where this is the only taxonomic indicator for a building.

2.2   Limitations

Ascribing Geoscape buildings to classes requires disaggregated building data with taxonomic indicators. We have found some jurisdictions have better coverage than others. There is a dearth of data on commercial buildings in public domain / non-commercial datasets.

Linking datasets to G-NAF using unsanitised address strings is a challenging process. Problems we have encountered include:

- Differing levels of granularity

- Missing street numbers

- Street number ranges

- Non-standard formats, such as corner of abc st and xyz road

- Inconsistent use of abbreviations, such as st / street, rd /road

These issues may result in a low 'hit rate', where a percentage of the observations will not link to G-NAF and Geoscape. Fuzzy matching techniques can be employed, but this increases the risk of misclassification.

3 Floorspace Estimation

A requirement of the CBBS is to provide an estimate of building floorspace, aggregated by taxonomic class and geospatial regions. This is achieved using the Geoscape building height and footprint (area of building polygon).

3.1 Methodology

Parameters

- Building footprint: the area in square metres of the Geoscape Building polygon
- Building height: the maximum height of the Geoscape Building in metres.
- Floor height: floor height (FH) is currently taken as 3.5 metres. It can be varied as required.

For each building:

If building height is null, or less than floor height parameter:

Floorspace estimate = building footprint

else

Floorspace estimate = floor(building height / FH) * building footprint

We also provide:

- Floors estimate: floor(building height / FH)
- Building shell volume (m^3): building height * building footprint
  for buildings with a non-null height

Examples

- Building height null, building footprint 250sqm
  Floorspace estimate = 250sqm
  Floors estimate = 1
  Building shell volume N/A

- Building height 4.3m, building footprint 250 m^2
  Floors estimate = floor(4.3 / 3.5) = 1
  Floorspace estimate = floor(4.3/3.5) * 250 = 250 m^2
  Building shell volume = 4.3 * 250 = 1,075 m^3

- Building height 10m, building footprint 250sqm
  Floors estimate = floor(10 / 3.5) = 2
  Floorspace estimate = floor(10 / 3.5) * 250 = 500 m^2
  Building shell volume = 10 * 250 = 2,500 m^3

3.2 Limitations

Geoscape Building footprints are a 'bird's eye' view, using data taken from remote sensing, therefore the footprint may encompass verandas, eaves, awnings, pedestal, etc.

Likewise, this methodology is unable to account for the internal complexities of buildings, such as voids, stairwells, multi-level atriums/foyers, or undercover car parks. Underground and basement levels are also not accounted for.

We have noticed the building height seems sensitive to elevation, such as being potentially less accurate in hilly areas or for sloping sites.

It is possible that some Geoscape Buildings may be other structures, such as gazebos, rotundas, or other unenclosed, roofed areas, such as open-air markets and train station platforms.

Geoscape Buildings in many rural areas have no height, and a number have heights less than the floor height parameter, so they are assumed to be single story, and the floorspace estimate is therefore taken as the building footprint.

4   Geospatial Referencing

Geospatial referencing refers to assigning a Geoscape building to a geospatial region of interest. This allows for the aggregation of Geoscape building counts, sum of floorspace estimates and other features of interest.

For the CBBS, we have been asked to aggregate output by:

- ABS Statistical Area 4 (SA4)

- Distribution Network Service Provider (DNSP) supply area

- National Construction Code Climate Zone

Geoscape buildings are assigned to the region of interest using the intersection of the Geoscape building point (longitude, latitude) representation with the spatial region's polygons, that is given a building point – what polygon does it belong to?

For the SA4 referencing, we have used the ABS Mesh Block 2016 spatial data in order that we pick up the Mesh Block category. Given the nested hierarchy of ABS statistical areas, the Mesh Block code allows us to also determine SA1-4, Greater Capital City (GCCSA) and State/Territory, without any further processing.

Note that it is possible for the Geoscape building polygon to straddle the boundaries of more than one region, hence the building centroid (point) is used to avoid double counting. For example, Geoscape Buildings version 1.5 had buildings assigned to more than one Mesh Block and/or planning zone. In version 2, Geoscape has moved to a 'plurality wins' approach.

Output from the Building Stock Model to be aggregated to any regions for which there is geospatial polygon data, such as Postal Areas (POA), Local Government Area (LGA) or NatHERS climate zone.

We have found that geospatial polygon datasets vary in accuracy, such as around the coastline and waterways, meaning some Geoscape buildings may not lie within the boundaries.

5   References

Australian Bureau of Statistics. (2020). *MESH BLOCKS (MB)*. Retrieved April 1, 2020, from
https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.001~July%202016~
Main%20Features~Mesh%20Blocks%20(MB)~10012

Geoscape. (2019, March). Geoscape Data Product Description. Australia: Geoscape.

Geoscape. (2020). *G-NAF*. Retrieved April 1, 2020, from Geoscape: https://geoscape.com.au/data/g-naf/

Appendix A Geoscape Building count by CBBS Taxonomy

Geoscape Building counts by CBBS Taxonomy, includes all structures in Geoscape Buildings product, regardless of footprint.

Table 1 Geoscape Building count by CBBS Taxonomy

CBBS Taxon	Geoscape Building Count
Accommodation - hotel	10,797
Accommodation - other	17,492
Aged Care - nursing home	15,893
Aged Care - other	400
Aged Care - retirement village	12,944
Agriculture - other	182
Agriculture - unresolved	972,796
Carpark commercial	308
Commercial - nec	7,676
Commercial - unresolved	125,994
Data Centre	22
Education - combined	3,484
Education - other	13,697
Education - presecondary	24,416
Education - school	157
Education - secondary	32,549
Education - TAFE	1,112
Education - University	1,895
Education - unresolved	29,449
Entertainment and Recreation - aquatic centre	546
Entertainment and Recreation - cultural	8,285
Entertainment and Recreation - other	1,375
Entertainment and Recreation - sports centre	10,195
Factory - laboratory	12
Factory - other	1,502
Healthcare - other	1,139
Healthcare - unresolved	4,230

Hospital - other	24
Hospital - private	490
Hospital - public	1,000
Hospital - unresolved	2,468
Industrial - agriculture	1,164
Industrial - other	22,547
Industrial - unresolved	190,371
Non-Residential - unresolved	240,816
Office	7,087
Other non-residential - nec	132,035
Religion Building	2,777
Residential	13,758,887
Retail - other	18,865
Shopping Centre	4,540
Supermarket	200
Transport - passenger	1,083
Transport - unresolved	4,198
Warehouse	1,124
Warehouse - refrigerated	9

Appendix B Dataset Ranking

Source datasets, ranked in order of precedence, with the count of Geoscape buildings tagged by each dataset, regardless of size. Note that a building may be tagged by more than one dataset; the highest ranked dataset provides the final CBBS taxonomy assignment.

Table 2 Dataset ranking and Geoscape Building count

Dataset Ranking	Data Source	Building Count
1	Commercial Building Disclosure	4,173
2	NABERS	3,123
3	PCA Shopping Centres Online	2,268
4	Hotels STR	10,194
5	Building Address is_residential	12,685,374
6	G-NAF Building Name	108,523
7	Hospitals Commonwealth	2,864
8	Hospitals Victoria	1,541
9	Hospitals SA	175
10	Hospitals NSW	642
11	Hospitals WA	224
12	AIHW GEN Aged Care	134,291
13	Retirement Village NSW	7,244
14	Retirement Village SA	299
15	Retirement Village Qld	6,790
16	Schools ACARA	2,975
17	Schools Victoria	8,983
18	Schools NSW	8,202
19	Schools SA	5,521
20	Schools WA	8,631
21	CLUE	25,479
22	Victoria FOI Point	13,702
23	Victoria FOI Polygon	125,627
24	Victoria Geographic Names	2,999
25	ACT Building Footprint	54,033
26	NSW POI	9,792

27	Queensland Building Areas	30,840
28	Queensland Building Points	892,091
29	Queensland Landmark Areas	40,132
30	Queensland Community Facilities	2,652
31	Schools Queensland	10,944
32	Tasmania Community Facilities	5,075
33	Tasmania Building Polygons	109,215
34	SA Place Names	1,765
35	WA Geographic Names	1,359
36	G-NAF Flat Type Description	44,677
37	Mesh Block Category	15,666,376

As Australia's national science agency and innovation catalyst, CSIRO is solving the greatest challenges through innovative science and technology.

CSIRO. Unlocking a better future for everyone.

Contact us

1300 363 400
+61 3 9545 2176
csiroenquiries@csiro.au
csiro.au

For further information

Energy
Giles Browne
giles.browne@csiro.au
csiro.au/energy