منظمة الأغذية والزراعة للأمم المتحدة
联合国粮食及农业组织
Food and Agriculture Organization of the United Nations
Organisation des Nations Unies pour l'alimentation et l'agriculture
Organización de las Naciones Unidas para la Agricultura y la Alimentación

APCAS/10/22
April 2010

Agenda Item 8

ASIA AND PACIFIC COMMISSION ON AGRICULTURAL STATISTICS

TWENTY-THIRD SESSION

Siem Reap, Cambodia, 26-30 April 2010

GEO-CODING AND SPATIAL DISPLAY OF DATA FROM THE AUSTRALIAN AGRICULTURAL CENSUS

Introduction

Up until 1996-97, Australia conducted an Agricultural Census (complete enumeration of farms) every year. Since then there have been Agricultural Censuses in 2000-01 and 2005-06, and the intention is to conduct an Agricultural Census every five years in the same year as the Census of Population and Housing (Population Census). The next Agricultural Census will be conducted in 2010-11. In the intervening years large sample surveys are conducted.

The main objective of the Agricultural Census is to provide benchmark information on the agriculture sector for small geographic areas. Prior to the 2005-06 Agricultural Census, data has been produced at Statistical Local Area (SLA) level. SLAs are generally equivalent to local government areas, where local governments are the third tier of government in Australia after National and State governments. For many years this was a suitable level of output for most users.

However, in recent years the role of agriculture in natural resource management (NRM), water policy, and in rural and regional communities, has become a focus across a range of government agencies. When considering these issues, the geographic areas of interest often relate to 'natural' boundaries rather than administrative or government boundaries. Thus areas such as water catchments, soil

types, riparian zones etc are important for policy considerations. As a result, the various governments involved have agreed on a range of Natural Resource Management (NRM) Regions which have become the focus of investigation, funding, activity, monitoring and evaluation. Demand for Agricultural Census data to be able to be produced for a range of different regional geographies (including these NRM regions) has become very strong.

This paper describes the approach the ABS took in relation to the provision of small area data from the 2005-06 Agricultural Census.

2005-06 Agricultural Census

The 2005-06 Agricultural Census was a complete enumeration of all farms in Australia with an Estimated Value of Agricultural Operations (EVAO) of $5,000 or more. This is the same scope as in previous agricultural censuses and surveys, and includes relatively small farms.

The 2005-06 Agricultural Census was conducted using a mail-out / mail-back collection methodology with a questionnaire of 20 pages. Questionnaires were dispatched to 190,000 agricultural businesses in June 2006 and over 93% were returned.

Given the mail-out / mail-back methodology, the only contact details maintained for each farm were mailing address details, while the only information about the farm location was the Local Government Area as reported by the farmer. Thus there was no information which could accurately locate the farm in a spatial sense.

Mesh Blocks

Mesh Blocks are the smallest area geographic unit of the new Australian Statistical Geography Standard (ASGS). They have been designed primarily for use in the Australian Census of Population and Housing, and for that reason they usually contain between 30 and 60 dwellings. In urban areas, mesh block boundaries have been designed to align with suburbs, while in rural areas they have been designed to align with water catchment boundaries.

The small size of mesh blocks means that they will aggregate and/or align more closely to most administrative, social, economic and environmental boundaries.

Coding farms to Mesh Blocks

In the 2005-06 Agricultural Census, the ABS asked farmers to provide location address details for their main agricultural property. This address was then used to code each farm to a latitude and longitude coordinate which could then be aggregated to 'mesh block' level.

7 Where is this business's main agricultural property located?

Note
- Please provide the actual location address, not a post-box or road-side mail box.
- Please report in every applicable line.
- If more than one property operated, show location of main property.

Property name … …

Road/Street number

Road/Street name …

Suburb, rural locality or town … … … …

Shire … … … … … …

State … … … … … … Postcode

Respondents were also asked whether they operated properties in more than one Shire (or Local Government Area). In some instances where land holdings in other localities were large in area (greater than 2,000 hectares), their additional property location addresses was recorded, the units were then 'split' and treated as separate units for estimation purposes.

The address provided by the agricultural business operator was matched to the Geocoded National Address File (G-NAF). This is a listing of all valid addresses in Australia, with the associated latitude and longitude coordinates.

Where farmers provided a full and accurate address, a match to G-NAF was found and the farm accurately coded. However, approximately 36% of respondents did not report a road number, which was an important input into the address coding. Where insufficient address information was provided, a polygon overlay approach used other available information including property name, town, postcode and shire to determine the coordinate.

Also, addresses in rural areas tended to be less precise than those in urban areas and around 72,000 addresses required some manual intervention to code to a latitude / longitude coordinate. The ABS used commercially available software such as Google Earth and developed a range of other tools to help code remaining farms as accurately as possible to mesh block. In some cases farmers were contacted to obtain better address information.

A quality assurance review at the completion of the address coding project indicated that approximately 94% of latitude / longitude coordinates of all units coded were within 5 kms of the farm gate or centroid of the land parcel.

It should be noted that all activity of the farm was allocated to the coordinate point (and mesh block) to which the farm was coded, even though some agricultural activity may be on land located quite some distance away and in different mesh blocks.

Once farms were coded to a latitude / longitude coordinate, it was possible to produce a range of outputs from frequency tables or cross tabulated datasets to maps and animated outputs for a range of different geographic levels.

OUTPUTS – some constraints

The main constraint to release of data is the need to protect the confidentiality of information of individual agricultural businesses or providers. The ABS obtains very high response rates to its surveys and censuses and its record for protecting confidentiality is one of the main reasons for such good response. The legislation under which the ABS operates also has provisions to protect the confidentiality of data, and provides for severe penalties for any breaches.

Another constraint to the release of data is the quality of data at very detailed levels. In terms of geographic areas, it is simply not feasible to have quality control checks on the data at mesh block level (there are at least 26,000 mesh blocks across Australia with agricultural activity), therefore output editing is done at a higher level of geography. Thus while it is possible to produce data for mesh block and other small area regions, ABS cannot be confident of the quality of detailed data at that level.

Thus, in terms of outputs,

- data was <u>not</u> released at individual farm or business level
- detailed data was <u>not</u> released at mesh block level
- data was <u>not</u> released for aggregations of mesh blocks which equate to very small geographic areas

Some data has been released at mesh block level, but this was generally in ranges (eg crop yields between 1.5 and 2.0 tonnes per hectare, density of dairy cattle etc) rather than absolute numbers (eg wheat yield of 1.86 tonnes per hectare). The maps presented later give some examples of mesh block level releases. As a way of protecting confidentiality, the number of agricultural businesses in a mesh block has also not been released, nor has detailed commodity production data for each mesh block.

OUTPUTS – what has been possible

The objective of using mesh blocks (built up from individual unit latitude / longitude coordinates) was to increase as much as possible the flexibility to produce a range of geographic data, including custom regions identified by clients, both internal and external.

Standard Outputs

Standard outputs are those that are available free of charge on the ABS website (www.abs.gov.au). These are generally available as data cubes in spreadsheet format, as data suitable for use in GIS systems and possibly as maps.

Standard outputs were produced for the following regional structures based on the coded latitude / longitude coordinates:

- NRM regions (a set of Natural Resource Management regions defined by the Federal and State Governments for focus for activity in relation to natural resource management)
- Drainage Divisions and River Basins

- Murray-Darling Basin region [an important region in terms of contribution to Australia's value of agricultural production (40%) and utilises 65% of Australia's irrigated land]

Data presented in standard outputs included:

- A full range of commodity production items
- Selected data on water use and source
- Selected data on land use

Detailed water use and source and irrigation activity were also released for the following defined water regions, based on latitude / longitude coordinates:

- Surface Water Management Areas
- Ground Water Provinces
- Interim Biogeographic Regions of Australia (IBRA)
- Capital City Water Supply Areas

Customised data

The ABS has also met client demand by providing a wide range of customised reports for defined regions of interest. Such requests have included unusual geographic areas, detailed data items, and cross-tabulations of several data items. This work has been undertaken on a fee for service basis.

Maps and animated flyovers

In addition to the frequency tables and tabulated outputs, it has also been possible to produce a range of outputs in the form of maps and animated formats for a range of different geographic levels, based on mesh block level data.

Various maps have been produced to display the presentation power of mesh block level data compared with the previous lowest level of data released, which was at Statistical Local Area (SLA) level. The maps presented in the attachment to this paper provide some examples of the increased detail provided by mesh block level releases.

Taking the presentation of estimates a step further than tables or static maps, the ABS has manipulated various 2005-06 Agricultural Census mesh block level data to produce animated aerial 'fly-overs' as another and exciting way of displaying data. This process uses Feature Manipulation Engine (FME) software with data displayed using Google Earth.

One such example was the amalgamation of mesh blocks for regions adjoining the Murray River in South-Eastern Australia. The data were manipulated to illustrate the gross value of irrigated and non-irrigated production and to also display the gross value of production in areas of dryland agricultural production. This animated presentation can be viewed on the ABS website, in Ag Mag – The Agriculture Newsletter, December 2009 (cat. no. 7101.0

Other potential uses of fly-over presentations

ABS is currently investigating the suitability and wider application of data displayed using flyover and Google Earth methodologies, including during editing and data

quality phases of its collections. The potential to display crop production and yield data for given geographies in such a format could assist in more quickly identifying unusually high or low estimates which should be examined. Not only is it possible to display the aggregate estimate, but also to use Google Earth functionality to display the location and details of those units contributing to the estimate and to view their individual contributions as part of a significance editing approach.

Improvements in data availability

One way to illustrate the improvement in usefulness of data based on mesh blocks is through the use of maps. Attachment A provides a series of maps which show the difference in data precision for selected areas when using data based on mesh blocks compared with data which was previously available based on administrative Statistical Local Areas (SLAs).

Map 1 shows the presence of dairy cattle at SLA level for the State of Tasmania, Australia's southern island state. The SLA regions represent quite large physical areas, including parts which may have no or minimal agricultural activity. The use of SLA boundaries therefore does not accurately reflect, in a spatial sense, where the dairy farming activity is actually being undertaken.

Map 2 also displays dairy cattle information, but this time at mesh block level. As can be seen, the level of spatial detail is significantly improved.

Map 3 shows potato production in Tasmania, again presented at mesh block level.

Whilst the use of mesh block level estimates offer significant improvements over previous standard ABS geographies, particularly for mapping purposes, there are instances where the use of farm latitude / longitude coordinates to produce estimates may be more appropriate.

Maps 4, 5 and 6 display how the use of farm business coordinates compare favourably with mesh block boundaries, particularly when seeking to align with user-defined or non-ABS boundaries. Mesh blocks, which can cover quite small geographic areas, do not always align well with regions of interest specified by clients. If using a mesh block approach, it is necessary to determine which mesh blocks fall within the specified custom region. This can be done in several ways, including using a mesh block centroid approach or setting a minimum percentage of mesh block area to determine whether or not the mesh block should be included.

In such instances, it may be preferable to use the latitude / longitude coordinates of agricultural businesses to determine the in-scope population and to then produce weighted estimates for this population. However, an important factor to consider in the context of the accuracy of mesh block estimates, is that they are based on single coordinate or location which is then attributed to a single mesh block. It is possible, particularly for larger properties and holdings (several Australian beef cattle properties have an area of holding of more than 1 million hectares with the largest almost 2.4 million hectares), that the reported agricultural activity is undertaken across several mesh blocks. As a result, it is possible a particular activity may not necessarily occur in the mesh block specified and the Area of Holding and other estimates of agricultural activity may exceed or not account for all activities within

that mesh block. For these reasons, ABS includes a caveat for users indicating that the quality of estimates may be lower for some small area geographies.

Future Surveys

The mesh block estimation approach is really only applicable in census years, as a full enumeration is needed to be confident of data quality at small area levels. In survey years geographic flexibility is more limited, with standard outputs released at the larger Natural Resource Management (NRM) region level.

However, now that good spatial information is available for each farm business, sample design possibilities (in terms of geographic stratification) in survey years are increased. In addition, ABS has continued to record property location information for both new agricultural businesses selected in the annual agricultural surveys and also for businesses where the main location may have changed as a result of land acquisitions or disposals.

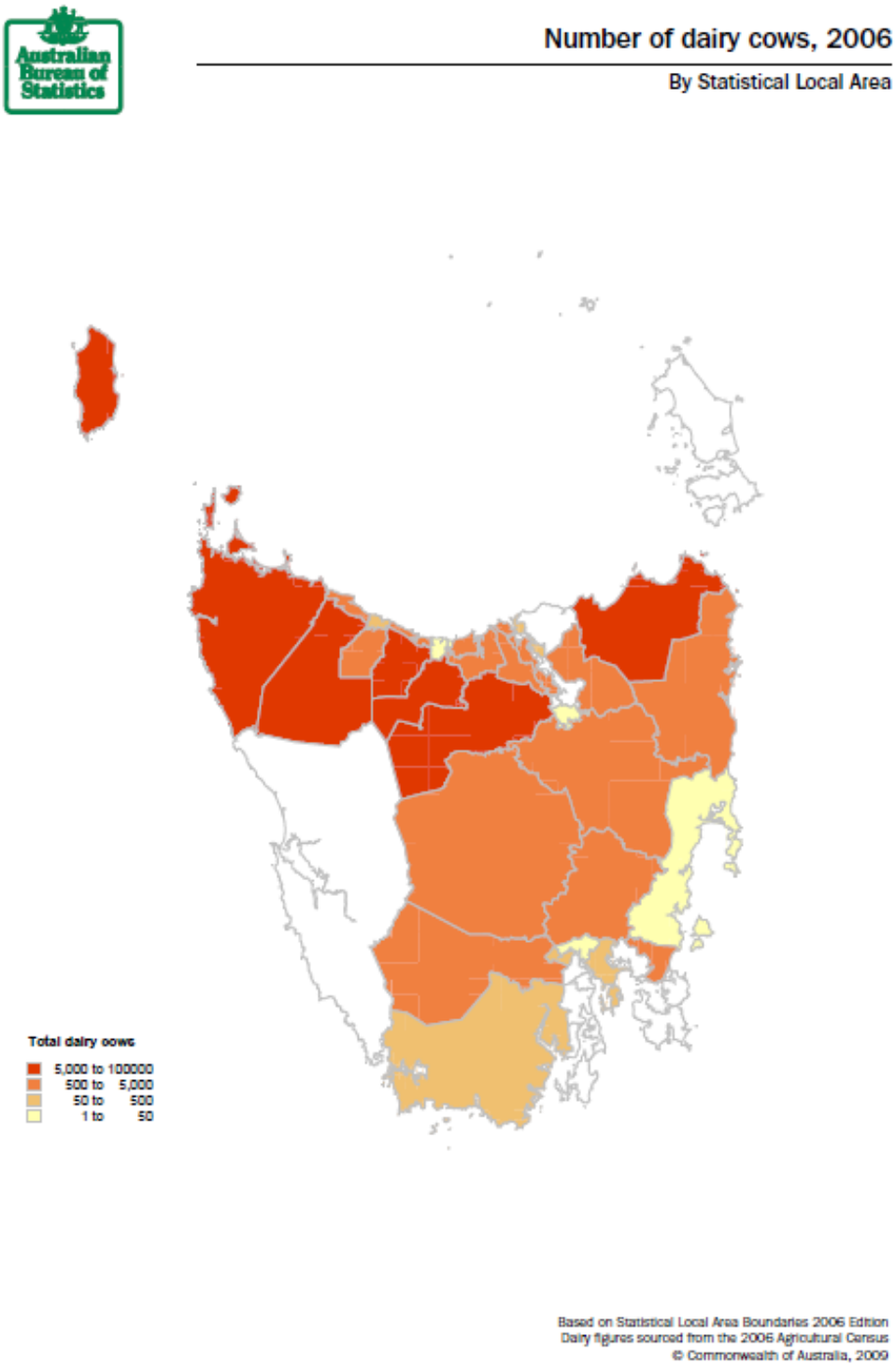
The geography strategy for the 2010-11 Agricultural Census will be similar to the strategy adopted for the 2005-06 Census and will support the production of flexible regional level outputs. This will include the collection of detailed property location address data on the Census form (both paper and e-form versions) and the assignment of a latitude and longitude coordinate to the main agricultural property holding.

ABS output dissemination systems have also been identified for particular attention in the 2010-11 Agricultural Census. The aim is to improve the user experience in locating and importing data into their own systems as well as providing for the integration of other ABS and non-ABS data at the small geographic level.

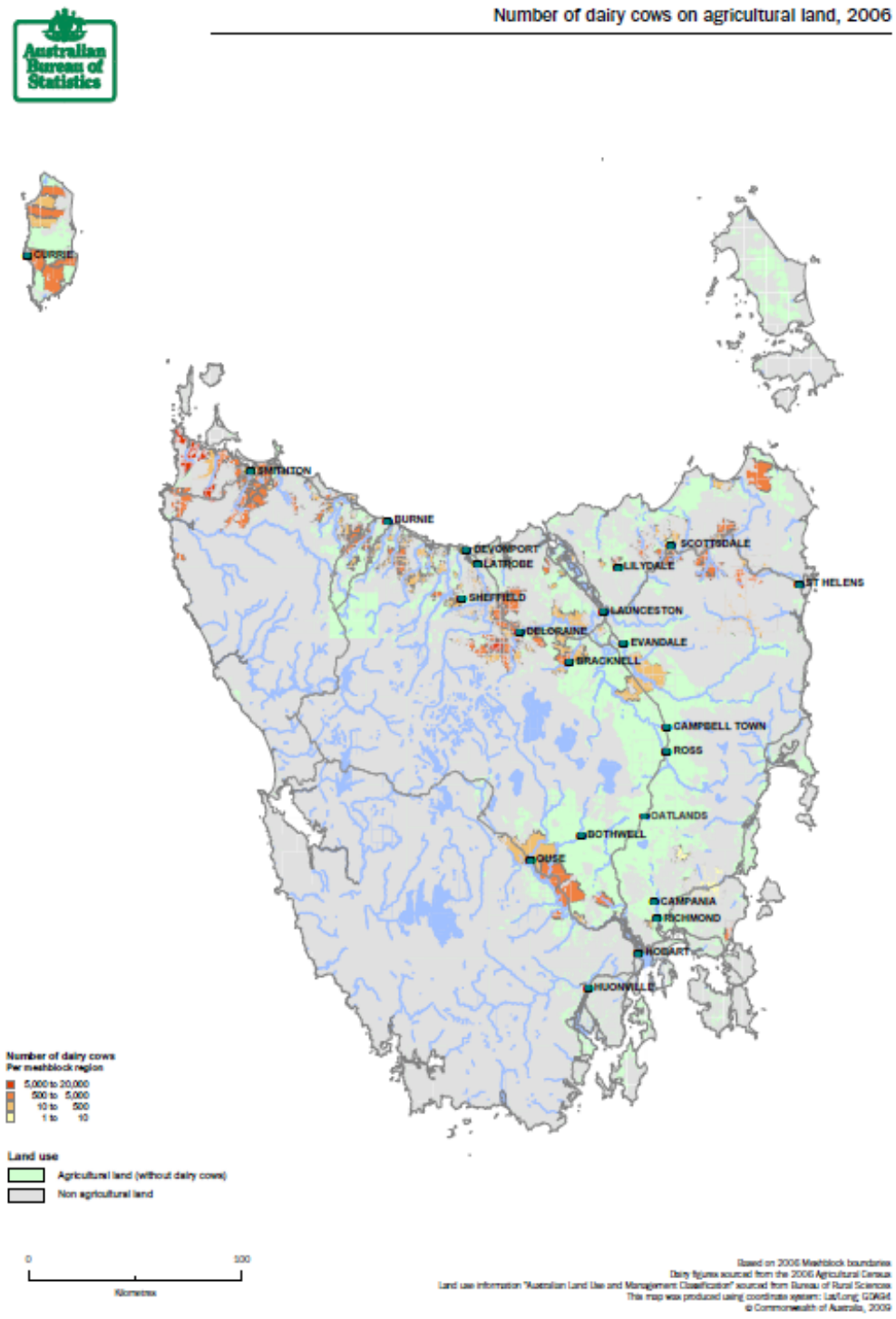
Environment and Agriculture Statistics
Australian Bureau of Statistics
April 2010

Attachment

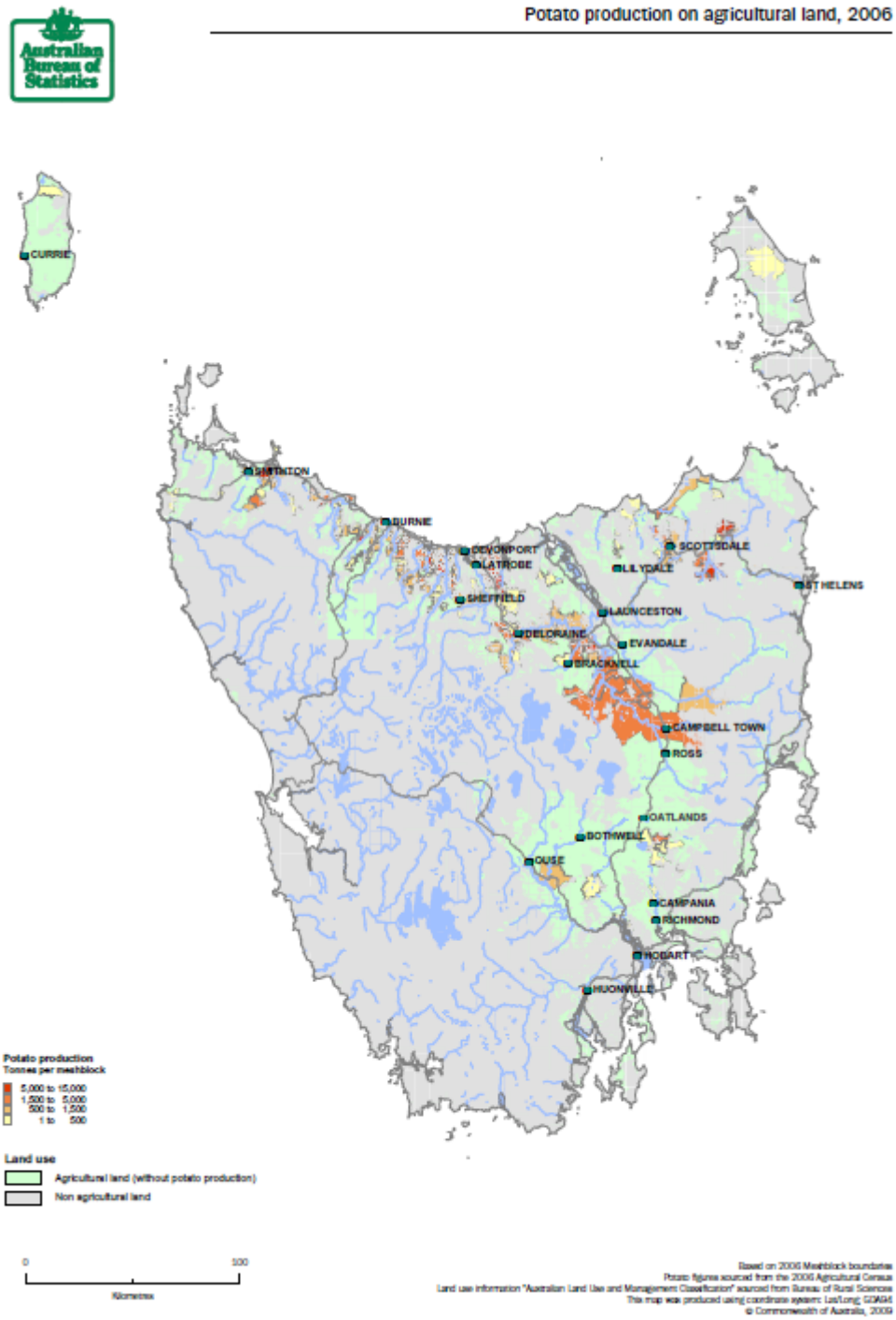
Map 1 (SLA boundaries) – Dairy cattle in State of Tasmania



Map 2 (Mesh Blocks) – Dairy cattle in State of Tasmania

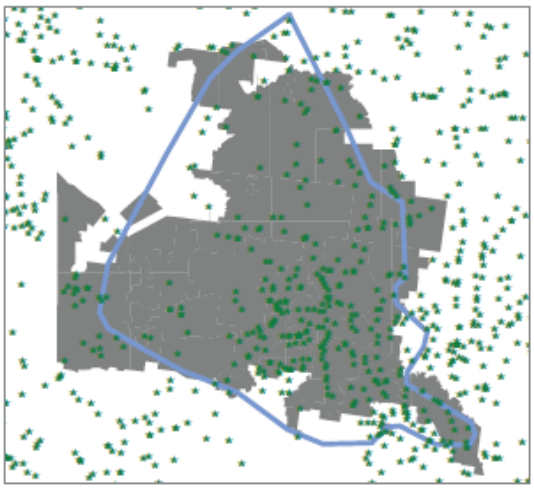


Map 3 (Mesh Blocks) – Potato production in State of Tasmania

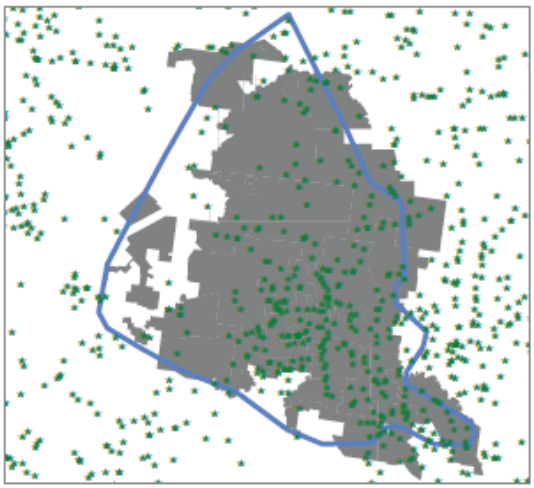


Maps 4, 5 and 6 Examples of mesh blocks approximating non-ABS boundaries

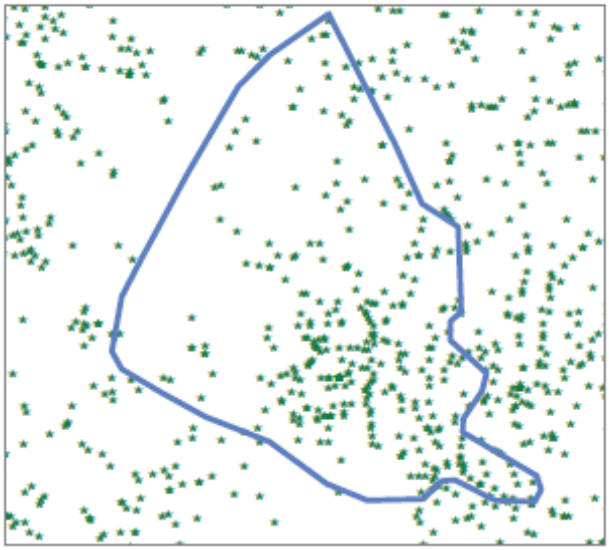
Map4
Meshblock centroids within Casterton region



Map5
Meshblocks with more than 50 percent area within Casterton region



Map6
Farm business coordinates from the 2006 Agricultural Census



User defined "Casterton" boundary

Meshblock approximation

Geocoded farm business location

0 20

Kilometres