Australian Government

Geoscience Australia

GEODATA TOPO 5M 2004

Product User Guide

National Mapping Division,
Geoscience Australia

Published by Geoscience Australia

Published by Geoscience Australia
Department of Industry, Tourism and Resources

1st Edition released: March 2004
2nd Edition released: July 2004 (Minor corrections)

Technical support:
For up to date information on *GEODATA TOPO 5M 2004* refer to the Geoscience Australia website: www.ga.gov.au

Please direct queries to:
Geoscience Australia Sales Centre
GPO Box 378
Canberra ACT 2601
Freecall (within Australia): 1800 800 173
Telephone: +61 2 6249 9966
Facsimile: +61 2 6249 9960
Email: sales@ga.gov.au

Acknowledgments:
Geoscience Australia gratefully acknowledges contributions to map content. Information is supplied by Commonwealth, State, Territory and local government, private sector agencies and individuals. A comprehensive list is available from our web site at www.ga.gov.au/nmd/mapping/acknowledge.htm

About this product user guide
This product user guide sets out the fundamental concepts and characteristics of *GEODATA TOPO 5M 2004*. The guide begins with general information and provides more details in later sections. The overview of data content and structure will allow you to make immediate use of the data.

Contents

1 User information

1.1 User Support/Contact Information

A copy of the licence conditions are supplied at the time of purchase or download, and should be retained for proof of licensing.

Geoscience Australia welcomes feedback on any aspect of its product or services. Please direct your comments or any queries regarding this document or data to:

Geoscience Australia Sales Centre
GPO Box 378
Canberra ACT 2601
Freecall (within Australia): 1800 800 173
Telephone: +61 2 6249 9966
Facsimile: +61 2 6249 9960
Email: sales@ga.gov.au
Internet: www.ga.gov.au

1.2 Geoscience Australia - National Mapping Division

Geoscience Australia is the national agency for geoscience research and spatial information. It serves government and supports the community through its output areas of geoscience for urban centres, oceans and coasts, and regional and rural areas.

The National Mapping Division within Geoscience Australia undertakes national mapping, remote sensing maritime boundary and land information coordination activities in support of Australia's economic and social development.

2 About GEODATA TOPO 5M 2004

2.1 GEODATA TOPO 5M 2004 Components

Your *GEODATA TOPO 5M 2004* data package has four components which combine to give you a complete data product. The components are:

- **Product User Guide**
  This guide describes the structure and content of *GEODATA TOPO 5M 2004*.
- **Data Quality Statement file**
  The Data Quality Statement file carries data quality information relevant to the whole *GEODATA TOPO 5M 2004* product.
- **Data Quality Table**
  This table contains data quality information about each feature instance in the data.
- **Data Files**
  The number of files will vary with the application format of the data.

2.2 The GEODATA TOPO 5M 2004 Product

The *GEODATA TOPO 5M 2004* product is a national seamless data product aimed at regional or national applications. It is a vector representation of the Australian landscape as represented on the Geoscience Australia 1:5M general reference maps and is suitable for GIS applications.

The data was derived from *GEODATA TOPO 250K Series 2* data where available and *GEODATA TOPO 250K Series 1* and has double precision accuracy.

The product consists of eleven data layers, including built-up areas, contours, drainage, framework, localities, offshore, rail transport, road transport, sand ridges, spot heights, and waterbodies.

2.3 The GEODATA Standard

All GEODATA products are:

- **GIS Compatible:**
  Every GEODATA product is designed to be immediately useful within GIS. You save the expense of bringing the data up to standard. For vector products, this means the adoption of a suitable data model and exacting standards for topological integrity;
- **Nationally Consistent:**
  Each GEODATA product adheres to a consistent, national specification. As a consequence, each product offers consistency in the treatment of features and attributes, the criteria for feature selection, the positional and attribute accuracy, and the data point density;
- **Quality Assured:**
  All products undergo independent quality assurance, including tests on vital aspects such as topological integrity, completeness, and positional and attribute accuracy;
- **Comprehensively Documented:**
  Comprehensive documentation accompanies all GEODATA products, allowing you to determine whether a particular product is suitable for your application and to ensure you realise maximum value from the data; and
- **Regularly Maintained:**
  Through timely revisions and upgrades, all GEODATA products remain up-to-date and relevant to changing customer requirements.

2.4 Coordinate System

GEODATA TOPO 5M 2004 data is available in geographic coordinates (latitude and longitude) in decimal degrees using the Geocentric Datum of Australia (GDA94). Elevation data are supplied in metres using the Australian Height Datum (AHD).

3 Data Loading

3.1 Application Formats

The *GEODATA TOPO 5M 2004* data is supplied in three application formats:

- ArcInfo export;
- ArcView shapefile; and
- MapInfo mid/mif.

3.2 Descriptions of Files

The downloaded *GEODATA TOPO 5M 2004* package contains the following files (See Table 1 over page).

Table 1: *GEODATA TOPO 5M 2004* files

Documentation files (documentation folder)		
File name	**Sub-Folder**	**Description**
61114_user_guide.pdf	documentation	This product user guide

Data files / Quality files / Look Up Tables (lut) / Frequency Tables (frq)				
ArcInfo Export (*.e00)	**ArcView Shapefile (*.shp)** ♦	**MapInfo mid/mif (*.mid, *.mif)** ♦	**Sub-Folder**	**Description**
aus5q.dqs	aus5q.dqs	aus5q.dqs	quality	Data Quality Statement (ASCII text file)
aus5dqt.e00	aus5dqt.dbf	aus5dqt.dbf	quality	Data Quality Table
aus5bgd.e00	aus5bgd_obj.shp	aus5bgd_obj.mid/mif	/data/built_up	Built-up areas
aus5cgd.e00	aus5cgd_obj.shp	aus5cgd_obj.mid/mif	/data/contours	Contours
aus5dgd.e00	aus5dgd_obj.shp	aus5dgd_obj.mid/mif	/data/drainage	Drainage
aus5egd.e00	aus5egd_obj.shp	aus5egd_obj.mid/mif	/data/spot_heights	Spot heights
aus5fgd.e00	aus5fgd_obj.shp	aus5fgd_obj.mid/mif	/data/framework	Framework
aus5lgd.e00	aus5lgd_obj.shp	aus5lgd_obj.mid/mif	/data/localities	Localities
aus5ogd.e00	aus5ogd_obj.shp	aus5ogd_obj.mid/mif	/data/offshore	Off-shore
aus5rgd.e00	aus5rgd_obj.shp	aus5rgd_obj.mid/mif	/data/rail	Rail transport
aus5sgd.e00	aus5sgd_obj.shp	aus5sgd_obj.mid/mif	/data/sand_ridges	Sand ridges
aus5vgd.e00	aus5vgd_obj.shp	aus5vgd_obj.mid/mif	/data/road	Road transport
aus5wgd.e00	aus5wgd_obj.shp	aus5wgd_obj.mid/mif	/data/waterbodies	Waterbodies
cla_lut.e00	cla_lut.dbf	cla_lut.dbf	lut	Secondary attribute table for road classification
ele_lut.e00	ele_lut.dbf	ele_lut.dbf	lut	Secondary attribute table for elevation range classification
for_lut.e00	for_lut.dbf	for_lut.dbf	lut	Secondary attribute table for road formation description
gag_lut.e00	gag_lut.dbf	gag_lut.dbf	lut	Secondary attribute table for rail gauge description
loc_lut.e00	loc_lut.dbf	loc_lut.dbf	lut	Secondary attribute table for locality type
per_lut.e00	per_lut.dbf	per_lut.dbf	lut	Secondary attribute table for perennial description
sta_lut.e00	sta_lut.dbf	sta_lut.dbf	lut	Secondary attribute table for state description
stu_lut.e00	stu_lut.dbf	stu_lut.dbf	lut	Secondary attribute table for railway status description
tra_lut.e00	tra_lut.dbf	tra_lut.dbf	lut	Secondary attribute table for rail track type
INFO table ♣	aus5bgd_frq.dbf	aus5bgd_frq.mid/mif	frq	Built-up areas
"	aus5cgd_frq.dbf	aus5cgd_frq.mid/mif	frq	Contours
"	aus5dgd_frq.dbf	aus5dgd_frq.mid/mif	frq	Drainage
"	aus5egd_frq.dbf	aus5egd_frq.mid/mif	frq	Spot heights
"	aus5fgd_frq.dbf	aus5fgd_frq.mid/mif	frq	Framework
"	aus5lgd_frq.dbf	aus5lgd_frq.mid/mif	frq	Localities
"	aus5ogd_frq.dbf	aus5ogd_frq.mid/mif	frq	Off-shore
"	aus5rgd_frq.dbf	aus5rgd_frq.mid/mif	frq	Rail transport
"	aus5sgd_frq.dbf	aus5sgd_frq.mid/mif	frq	Sand ridges
"	aus5vgd_frq.dbf	aus5vgd_frq.mid/mif	frq	Road transport
"	aus5wgd_frq.dbf	aus5wgd_frq.mid/mif	frq	Waterbodies
3.7 Mb	**5.1 Mb**	**3.9 Mb**		Compressed Zip File Size (including documentation).

♦ ArcView and MapInfo files include an additional object identifier (either _p, _a or _r) to indicate object type (point, line, polygon). See *Data File Naming Convention* over page for more information.

♣ Stored as INFO tables that can be viewed after importing the *.e00 data file into an ArcInfo coverage.

Data file naming convention
The data files are named according to a common GEODATA naming convention developed for ArcInfo Export files. The naming convention is slightly different for MapInfo and ArcView formats as there are more data files. This is because the ArcView and MapInfo data is transferred into separate object layers.

The file name of an ArcInfo file is made up of four components:

Aus5<layer_id><coord_id><precision_id> (e.g. aus5bgd.e00)

The file name of an ArcView or MapInfo file is made up of five components:

Aus5<layer_id><coord_id><precision_id>_<object_id> (e.g. aus5bgd_p.shp for a point object)

Where:

Aus5 refers to the GEODATA 5M product

<layer_id> is the single character code identifying the layer (refer to Table 2)

<coord_id> is the single character code identifying the coordinate system (i.e. ‘*g*’ for geographic)

<precision_id> is the single character code identifying the data as double precision (i.e. ‘*d*’)

<object_id> is the single character code identifying the spatial object of the file. The following object codes are used:

‘_***p***’ for Point

‘_***l***’ for Line (or chain or arc)

‘_***r***’ for Region (or polygon)

Table 2: *Codes for GEODATA TOPO 5M 2004 layers*

Layer	<layer_id>
Built-up Areas	b
Contours	c
Drainage	d
Spot heights	e
Framework	f
Localities	l
Offshore	o
Rail transport	r
Sand ridges	s
Road transport	v
Waterbodies	w

4 Data Characteristics and Concepts

4.1 GEODATA TOPO 5M 2004 Essential Characteristics

Truth-in-labelling
GEODATA TOPO 5M 2004 is accompanied by sufficient information to enable you to assess whether the data are fit for use in your application. The information will also assist you to use the data to their maximum potential. Data quality information is provided from three sources:

- **This Product User Guide**
  Chapter 6 of this User Guide covers product quality information such as lineage, accuracy, consistency and completeness.
- **The Data Quality Statement**
  This text file is included in each data transfer and contains information on areas such as transfer format, coordinate systems and point density.
- **The Data Quality Table**
  This table contains data quality information which is linked to each feature instance in the data.

These resources and their content are described in more detail in Chapter 6 - Data Quality Information.

Topological integrity
GEODATA TOPO 5M 2004 data are tested to ensure that they comply with the rules for topological integrity set out below. The data must comply with these rules, and others, to specified levels set out in Chapter 6. Generally the compliance rate is 99.5% or 95% with a 99% confidence. This means that the data may be accepted, even if they contain a small number of errors.

- The data have a node/chain structure. Within a linear network layer or a polygon layer, all linear features are broken by a node at intersections or at the point where an attribute of the feature changes. This is demonstrated in the diagram below. Pseudo will be accepted where:
  - A feature must be broken due to having more than 500 points
  - The specification requires a related feature to be placed on a node, for example, railway stations
  - Symbology or other map attributes change.

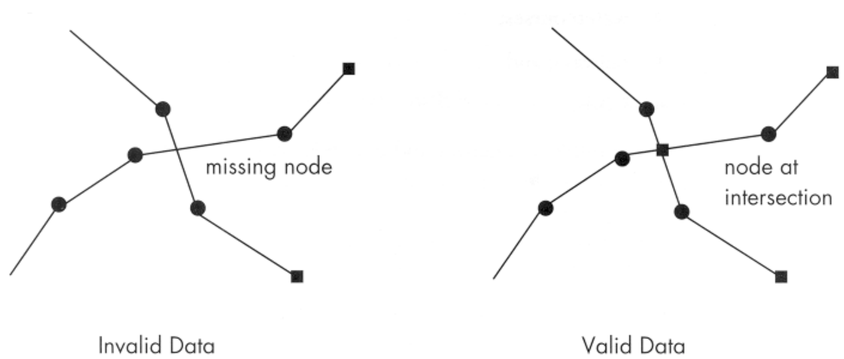

Figure 1: Linear intersections

- Every linear feature instance has a node at each end.
- All polygons are completely closed.
- Every polygon feature contains a polygon label point.

- Polygons in the same layer cannot overlap.
- Within a layer there are no coincident features.
- When two features in separate layers share the same physical position on the source material, they have exactly coincident spatial objects. The same feature instance may occur twice in the data supplied to you. When this occurs the repeated feature instance has exactly the same coordinates.
- There are no undershoots. This possible error is illustrated below.

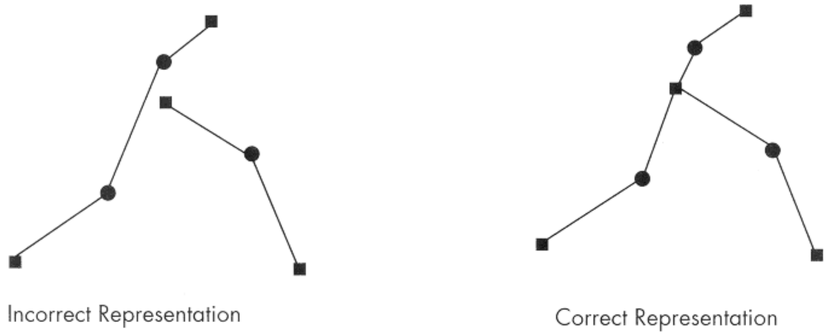

Figure 2: Undershoots

- The spatial data have no overshoots, broken lines or other artefacts of the data capture process. These possible errors in the data are illustrated below.

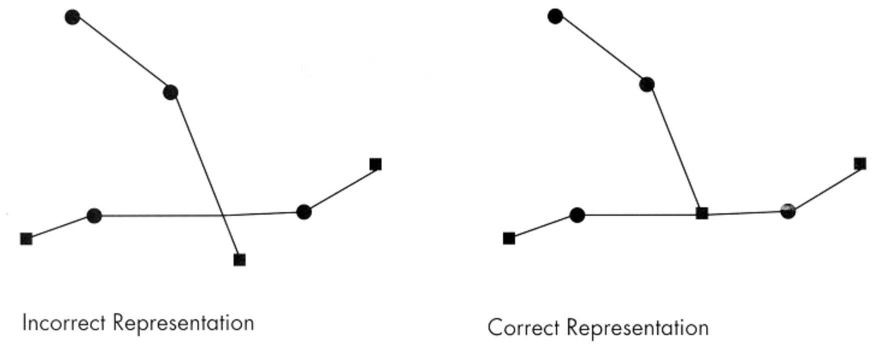

Figure 3: Overshoot

Figure 4: Broken lines

Figure 5: Data spikes

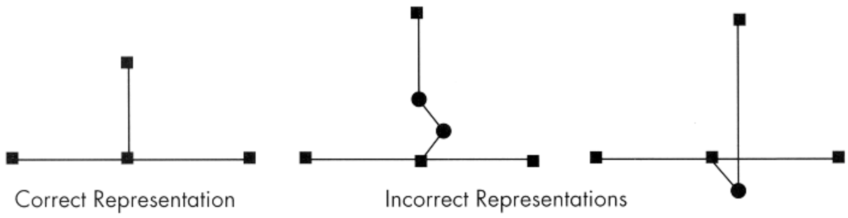

Figure 6: Artefacts in data

Point density reduction

Point density is controlled so that the locational information is conveyed by the minimum number of points while still retaining the smooth shape of the source information.

The following specifications apply for data point reduction for *GEODATA TOPO 5M 2004*:

- The minimum length of a line segment is 0.0044 degrees (or 500 metres); and
- The maximum length of a line segment is 5.07 degrees (or 400 000 metres).

For features other than roads, feature instances of fewer than 20 points will not be filtered (ie: such features are an exception to this rule). Sections of the chains which must be coincident with such features will also be an exception to the rule.

Unique feature identifier

Each entity in *GEODATA TOPO 5M 2004* has an attribute code which is unique to that entity. This attribute, known as the unique feature identifier (UFI), is nationally unique. The UFI is a character string with three component parts as described in the following table.

Table 3: UFI components

Scale identifier		Theme/layer identifier		Sequential identifier
A	1:100 000	A	Hydrography	A sequential number that is unique for each scale/theme combination.
B	1:250 000	B	Infrastructure	
C	1:1 000 000	C	Framework	
D	1:2 500 000	D	Relief	
E	1:5 000 000	E	Coastline and State border (non-topographic) datasets	
F	1:10 000 000	F	1:10 000 000 General Reference	
		G	Vegetation	
		H	Reserved Areas	
		I	1:5 000 000 General Reference	
		J	1:2 500 000 General Reference	

Data quality pointer

Each entity in *GEODATA TOPO 5M 2004* has a data quality attribute attached to it. This attribute named *Q_INFO*, is used to point to a record in the Data Quality Table which is discussed in more detail in Chapter 6.

Table 4: Data quality pointer components

Scale identifier		Source identifier		Sequential identifier
A	1:100 000	F	1:10M General Reference Map Edn 3	A number identifying unique combinations of data quality attributes (see section 4.2).
B	1:250 000	G	1:5M General Reference Map Edn 4	
C	1:1 000 000	H	1:2.5M General Reference Map Edn 6	
D	1:2 500 000	I	250K GEODATA Series 1	
E	1:5 000 000	J	250K GEODATA Series 2	
F	1:10 000 000	K	1:100K Coastline and State Borders Database	
G	Variable	L	Geoscience Australia Mapping Program Railways Database	
H	Not applicable	M	Geoscience Australia Mapping Program Inland Water Features Guide	
		N	WAC Spot Height Database	
		O	State Road Authority Maps	
		P	Satellite Imagery	
		Q	Variable	

4.2 GEODATA TOPO 5M 2004 Data Concepts

Each feature in *GEODATA TOPO 5M 2004* is defined by a spatial object and an attribute object. These features fit into the hierarchy of theme and layer. At the highest level, associated features are grouped into themes. Themes are subdivided into layers according to the spatial objects used to represent the features.

Vector data
Vector data describes spatial data in which the location of a real world phenomenon is defined by points and straight lines (vectors) between these points. The vector data model used for GEODATA also includes polygons - areas bounded by straight lines.

Feature-based data
The GEODATA vector products use a feature-based data model described by the following definitions. These are used to describe data that represent phenomena in the real world:

- **Entity:** A real world phenomenon which cannot be divided into phenomena of the same type.
- **Feature instance:** A single occurrence of a feature which has a unique set of spatial and attribute object values.
- **Attribute:** A descriptive characteristic of a feature. Attributes can be spatial (or locational) and aspatial (or non-locational).
- **Attribute value:** A value assigned to an attribute, either for a feature instance or its attributes.
- **Feature class:** A group of feature instances defined by a set of rules and having common attributes and relationships that are the properties of the corresponding real world phenomena.
- **Entity class:** A group of entities of the same kind, matching the members of a feature class.

The structure of a feature instance in the feature based data model can be summarised as:
feature instance = [*spatial object* + *attribute object*]

Spatial object
Spatial objects are the locational attributes of the feature. In GEODATA, they comprise the special cases of points, chains and polygons. Spatial objects have a spatial address which consists of one or more couplets (x, y) or triplets (x, y, z).

Point
A *point* is a geometric representation defined by a single (x, y) coordinate couplet or a (x, y, z) triplet. Three special points are used.

- **Entity point**
  An *entity point* is used to locate point entities, or area entities represented by a point because of the scale of the source material and/or scale of the final GEODATA product. ●
- **Polygon label point**
  A *polygon label point*, contained within every polygon feature instance, locates information about that polygon. It is linked to the bounding chains of the polygon. In proprietary GIS software packages, this point type is sometimes known as a centroid. +
- **Node**
  A *node* is a junction of two or more feature instances or an end point of a feature instance. Nodes may carry attribute information. ■

Chain
A *chain* is a spatial object composed of a sequence of non-intersecting line segments which is bounded by nodes at each end. Chains may carry topological information such as a reference to the polygons to the left and right (with respect to the direction of digitising) and reference the start and end nodes.

A line segment is a straight line between two consecutive vertices in a chain. Each vertex is defined by a single (x, y) coordinate couplet.

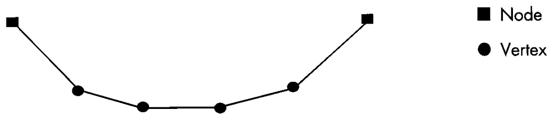

Figure 7: Chain spatial object

Polygon
A *polygon* is a bounded, continuous region consisting of an interior area, and an outer boundary defined by a set of chains. A polygon may also contain one or more non-nested inner boundaries also defined by sets of chains.

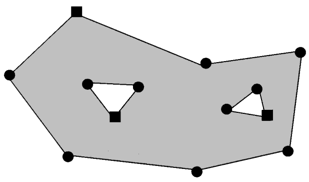

Figure 8: Polygon spatial object

The entity and its spatial object
The spatial object used to depict a feature depends on the size of the entity with respect to the scale of the source material and/or the final GEODATA product. For example, a small lake would be represented as an entity point whereas a large lake would be represented by a polygon. For this reason, a class of feature may be depicted by more than one type of spatial object.

Attribute object
An *attribute object* identifies the class of feature and the non-locational properties of the feature. The following two examples illustrate the possible content of the attribute object:

Attribute	Attribute value
Feature	Watercourse
Name	Darling River
Perennial	2
Data Quality Pointer	BJ000003
Unique Feature Identifier	EI00027096

Attribute	Attribute value
Feature	Island
Name	Whitsunday Island
State	5
Data Quality Pointer	BI000002
Unique Feature Identifier	EI00004515

Figure 9: Examples of attribute objects

It is possible for a feature's attribute object to consist of more than one set of attribute tables. The above attribute tables are known as *primary attribute tables*. Additional descriptive information about a feature instance can be provided by a *secondary attribute table*.

For example, the value for the '*Perennial*' attribute in the following Watercourse example is the link to the relevant row of information in a secondary attribute table called '*per_lut*' which gives the value of '*perennial*' for the Perennial code of 1.

Primary attribute table

Attribute	Attribute value
Feature	Watercourse
Name	Darling River
Perennial	2
Data Quality Pointer	BJ000003
Unique Feature Identifier	EI00027096

Secondary attribute table – per_lut

Code	Description
0	n/a
1	Perennial
2	Non perennial

Figure 10: Primary and secondary attribute table relationship

5 Data Structure and Content

5.1 Data Structure

The spatial object and attribute object as previously defined are the primitive components of GEODATA. When combined, these objects define a feature instance. Features are grouped to form a hierarchy which is used for the capture and transfer of the data.

Theme

The digital spatial data contained in GEODATA are primarily derived from existing map production material. The data on the source material are captured as features and these features may be grouped into themes - each containing logically related geographic information. The theme is the highest level of data grouping in the GEODATA structure. *GEODATA TOPO 5M 2004* is composed of a single theme.

Layer

One or more layers make up a theme. A layer is a grouping of features which have compatible spatial objects. GEODATA may contain four types of layers:

- **Linear network layer**
  Linear layers contain linear features such as watercourses. These layers are composed of nodes and chains.
- **Polygon layer**
  Polygon layers contain area features represented by polygons, such as lakes and reefs.
- **Point layer**
  Point layers contain features that are represented by entity points, such as buildings or aircraft facilities.
- **Point/linear layer**
  Point/linear layers contain a combination of entity point and chain features such as road networks with bridges and river networks with waterfalls and locks.

5.2 Data Layers

GEODATA TOPO 5M 2004 product is divided into eleven layers. The layers in this product are described in the following table :

Table 5: Content of *GEODATA TOPO 5M 2004* layers

Layer	Layer type	Feature	Attributes	Spatial object
Built-up Area	Polygon	bua_void		Polygon
		builtup_a	name	Polygon
		builtup_l		Chain
Contours	Polygon	contour	elevation	Chain
		hypso_a	elev_class	Polygon
Drainage	Chain	canal	name	Chain
		connector	name, perennial	Chain
		watercourse_l	name, perennial	Chain
Framework	Polygon	island	name, state	Polygon
		junction		Chain
		mainland	state	Polygon
		state_border		Chain
		waterline		Chain
Localities	Point	locality	name, locality, population	Point
Offshore	Polygon	offshore_l		Chain
		offshore_void		Polygon
		reef	name	Polygon
Rail transport	Chain	rail_station	name	Point
		railway	name, tracks, status, gauge	Chain
Sand ridges	Chain	sand_ridge		Chain
Spot heights	Point	spot_elevation	elevation	Point
Road transport	Chain	road	name, class, formation, nrn	Chain
Waterbodies	Polygon	junction		Chain
		lake	name, perennial	Polygon
		reservoir	name	Polygon
		sub_to_inund		Polygon
		swamp		Polygon
		watercourse_a	name, perennial	Polygon
		waterline		Chain
		w_body_void		Polygon

In addition to the attributes shown in this table, there are two more attributes attached to every feature instance:

- the unique feature identifier (UFI); and
- the data quality pointer (Q_INFO).

Special features

Connector feature (Drainage layer)
Drainage patterns are made up of both linear (narrow streams) and polygon features (such as lakes and reservoirs) and as such do not constitute a rigorous linear network. To allow linear analysis of drainage networks to be carried out, an artificial feature called a 'Connector' has been added to the data.

This Connector feature is used to bridge the gap in linear Watercourse features where they are separated by Waterbodies such as lakes or reservoirs. The Connector feature is composed of one or more chains in the general location that would be expected if the polygon feature was collapsed to a line. The points that make up this chain cannot be given any value for planimetric accuracy and this is indicated in the Data Quality attribute for the feature by a value of 9999 (not applicable) for the standard deviation of planimetric accuracy.

The Connector is only used if there is flow across a Waterbody polygon feature. Thus if there is only inflow to a Lake and no outflow the Connector feature is not used. The use of the Connector feature is ceased when a Watercourse ends at the Coastline. In cases where the flow is divided (that is, in river deltas or around river islands), the flow is represented by only one of the possible paths which is arbitrarily chosen.

All Connectors contained in Waterbodies that flow into other Waterbodies have been extended to join the Connector on the recipient waterbody.

Tributary Watercourses flowing into a polygon Waterbody are linked to the Waterbody's Connector for the main Watercourse with Connectors. The general rule for the attribution of Connectors is that Connectors carry the attributes of the Watercourse they represent. In the application it has been considered that Watercourses can change their perenniality along their course.

Depiction of the coastal environment
The Framework, Waterbody and Offshore layers contain features depicting the coastal environment. The area of tidal influence is part of the Sea (Universal Polygon) unless it is closed by a Junction feature. The coastline is represented by chains coded as *Waterline* (as are waterbody boundaries) and is the position of mean high water level.

The figure below illustrates how features are used in GEODATA to represent the riverine and maritime environments.

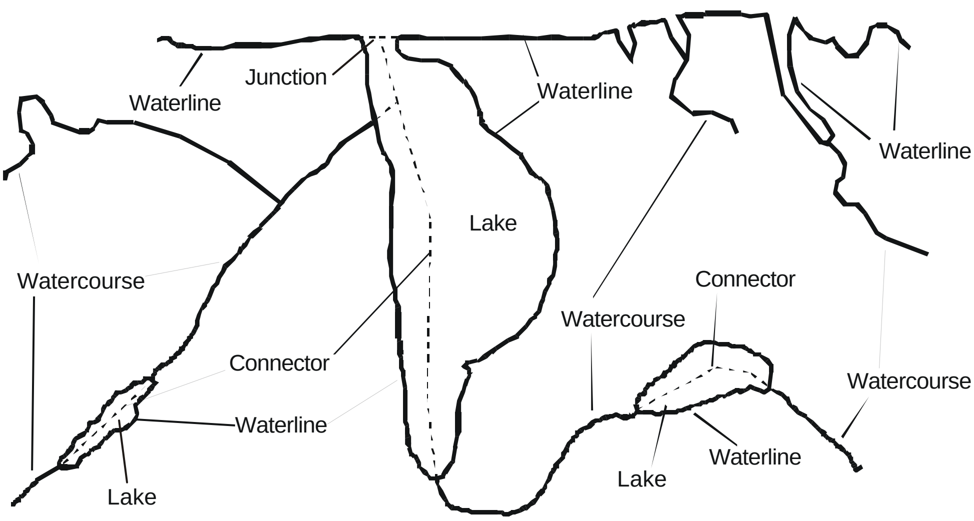

Figure 11: *Coastal feature depiction*

Islands (Framework layer)
Islands are represented as polygons coded 'island' in the Framework layer where they are fully surrounded by sea.

Junction feature (Framework and Waterbodies layers)
The Junction is a linear feature in the Framework and Waterbody polygon layers. It is an artificial line used to separate adjacent polygon areas across which flow can occur. For example, a Junction feature will separate the confluence of two Watercourses where both are depicted as polygons on the source material. A Junction will also close Watercourse polygons where they meet the coastline (e.g. in estuaries).

Junctions usually consist of a single chain segment with two vertices. Three vertex Junction features are permissible where there is a need to 'shape' the junction or control the relationship with the end node on a Connector. Multiple vertex Junction features are permissible in the framework layer. The Junction features in the Framework layer are replicated in the Waterbody layer to allow closure of waterbody polygons.

The Junction feature is arbitrarily placed and cannot be given any value for planimetric accuracy. This is indicated in the Data Quality table by a value of 9999 for the standard deviation of planimetric accuracy.

Junction features have not been placed to:

- Separate two Waterbodies with identical attributes.
- Separate polygons of different feature class, except to separate Watercourse polygons, Lakes, Reservoirs and the sea from one another.

Junction features have been placed to:

- Separate double line watercourses from other water bodies such as lakes and reservoirs.
- Separate waterbody polygons of the same class but with different attributes.
- Fill the coastal gaps in the framework layer.

Localities (Localities layer)
All homesteads are captured as localities in their true planimetric position and have not been displaced in order to position them on top of roads. Localities coincide with the road network only when they do so as entities.

Names
Named features have been attributed with the name in full including the type of feature where it is part of the name (e.g. 'ESK RIVER', 'HUME HIGHWAY'). Usually the type of feature is not part of the name for railway stations and localities. Abbreviations have generally not been used and names are in upper case. Unnamed river anabranches carry the river's name, except for where a river anabranch is named in its own right.

Road and rail intersections (Road transport and Rail transport layers)
A road or rail intersection in digital data contains the same number of nodes as shown on the source material. An intersection node is within 1/6 of the line width of the centre position of the intersection.

The first vertex in each direction from the intersection node are at a distance greater than three times the line width unless there is a bend in the road before this distance.

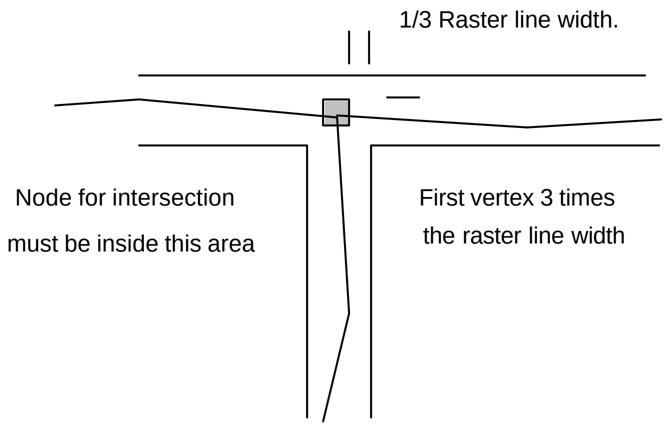

Figure 12: *Intersections*

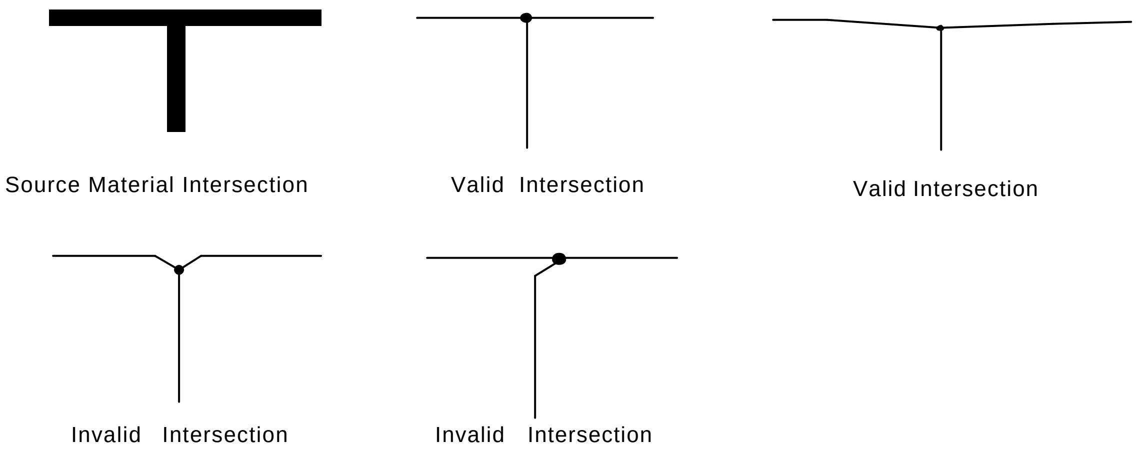

Figure 13: *Valid and invalid intersections*

Void polygons
Polygon features may contain holes or voids which cannot be assigned to any feature class within that layer. For example, a Lake in the Waterbodies layer may have an area of dry land in the middle of it. This appears in the data as a polygon with no centroid. A number of GIS packages cannot operate unless all polygons have centroids, so a centroid has been added. These empty areas within polygons are collectively known as voids.

5.3 Data Dictionary

Characteristics which are common to all features:

- **Unique Feature Identifier (UFI)**
  An attribute code that is unique to that instance.
- **Data Quality Pointer (Q_INFO)**
  This attribute points to a record in the Data Quality Table which holds information on the quality aspects of the feature. The contents of this secondary attribute table are set out in the data quality information within Chapter 6.

Attribute table

Table 6: Attribute table of *GEODATA TOPO 5M 2004*

Feature	Description	Object	Attribute	Attribute values	Selection criteria
BUILT-UP AREA (*builtup_a*)	An area where buildings are close together and have associated road and other infrastructure networks.	Polygon	NAME (*name*)	Name of built-up area	Population is 100,000 of more
BUILT-UP AREA LINE (*builtup_l*)	The bounding line of a built-up area.	Chain			
BUILT-UP AREA VOID (*bua_void*)	An empty or void area in a built-up area.	Polygon			
CANAL (*canal*)	An artificial watercourse conveying water for inland navigation, irrigation or drainage purposes.	Chain	NAME (*name*)	Name of canal	Canals shown on the 1: 5M General Reference Map edition 4.
			PERENNIALITY (*perennial*)	0 = *n/a*	
CONNECTOR (*connector*)	An artificial line used to connect watercourse features across a waterbody feature to allow network analysis of riverine networks.	Chain	NAME	Name of associated watercourse	
			PERENNIALITY (*perennial*)	0 = *n/a* 1 = *Perennial* 2 = *Non Perennial*	
CONTOUR (*contour*)	A line which represents an imaginary line on the ground joining points of equal elevation in relation to the Australian Height Datum (AHD66).	Chain	ELEVATION (*elevation*)	Elevation in metres from the Australian Height Datum	Contours shown on the 1:2.5M General Reference Map edition 6 (note - 1:2.5M contours are

Feature	Description	Object	Attribute	Attribute values	Selection criteria
					used at 5M scale).
HYPSOMETRIC AREA (hypso_a)	An artificial feature representing the height interval between contour lines.	Polygon	ELEVATION RANGE (*elev_class*)	1 = *-200 to 0 metres* 2 = *0 to 200 metres* 3 = *200 to 500 metres* 4 = *500 to 1000 metres* 5 = *1000 to 1500 metres* 6 = *1500 to 2000 metres* 7 = *>2000 metres*	Contours shown on the 1:2.5M General Reference Map edition 6 (note - 1:2.5M contours are used at 5M scale).
ISLAND (island)	An area of land surrounded by the sea.	Polygon	NAME (*name*)	Name of the island	Islands have been included where their size after point density reduction is greater than 0.0002 square degrees.
			STATE (*state*)	0 = *n/a* 1 = *ACT* 2 = *JBT* 3 = *NSW* 4 = *NT* 5 = *QLD* 6 = *SA* 7 = *TAS* 8 = *VIC* 9 = *WA*	
JUNCTION (*junction*)	An artificial line used to separate adjacent hydrographic areas which have differing attributes and across which flow can occur.	Chain			
LAKE (*lake*)	A naturally occurring body of mainly static water surrounded by land.	Polygon	NAME (*name*)	Name of lake	Lakes are included where their size after point density reduction is greater than 0.001 square degrees.
			PERENNIALITY (*perennial*)	1 = *Perennial* 2 = *Non Perennial*	
LAND SUBJECT TO INUNDATION (*sub_to_inund*)	Low lying land usually adjacent to lakes or watercourses, which is regularly covered with flood water for short periods.	Polygon			Land subject to inundation is included where its size after aggregation and generalisation is greater than 0.01 square degrees. Polygons smaller than the

Feature	Description	Object	Attribute	Attribute values	Selection criteria
					minimum size for inclusion and are in groups may be aggregated into one larger representative polygon.
LOCALITY (*locality*)	A named place or area.	Point	NAME (*name*)	Name of locality	Localities shown on the 1: 5M General Reference Map edition 4. All populated places with a population greater than 1000 in the 1996 Census Population and Housing are shown.
			LOCALITY CODE (*locality*)	1 = *Bay-Inlet-Cove* 2 = *Beach* 3 = *Cape-Headland-Point* 4 = *Homestead* 5 = *Road Junction* 6 = *Mountain-Peak -Hill* 7 = *Pass* 8 = *Populated Place* 9 = *Waterbody Island* 10 = *Place Name* 11 = *Gorge*	
			POPULATION (*population*)	Population of the populated place	
MAINLAND (*mainland*)	The area of continental Australia including Tasmania.	Polygon	STATE (*state*)	1 = *ACT* 2 = *JBT* 3 = *NSW* 4 = *NT* 5 = *QLD* 6 = *SA* 7 = *TAS* 8 = *VIC* 9 = *WA*	
OFFSHORE LINE (*offshore_l*)	The line bounding polygons in the offshore layer.	Chain			
OFFSHORE VOID (*offshore_void*)	A void in an offshore polygon.	Polygon			

Feature	Description	Object	Attribute	Attribute values	Selection criteria
RAILWAY (*railway*)	A transportation system using one or more (usually two) rails to carry freight and/or passengers.	Chain	NAME (*name*)	Name of railway line, if a name exists	Primary source is 250K GEODATA Series 2 and Series 1 data. The secondary source is the Railway Reference Database.
			TRACK TYPE (*tracks*)	1 = *One* 2 = *Multiple*	
			STATUS (*status*)	1 = *Operational* 2 = *Abandoned* 3 = *Under Construction*	
			GAUGE (*gauge*)	1 = *Standard: 1435 mm* 2 = *Broad: 1600 mm* 3 = *Narrow: 1067 mm* 4 = *Other* 5 = *Unknown* 6 = *Standard-broad* 7 = *Standard-narrow*	
RAILWAY STATION (*rail_station*)	A recognised stopping place for trains where passengers may board or alight. There may or may not be a platform. The railway station may not be in use.	Point	NAME (*name*)	Name of railway station	Where railway stations fall outside built-up areas, railway stations will be shown: • Where they coincide with a nearby locality with locality codes 8 and 10 and are operation or under construction. • At the end of a branch line. • At the junction of railway lines where the railway line name contains the name of the railway station. Where railway stations fall within built-up areas, only a selection of stations will be shown.

Feature	Description	Object	Attribute	Attribute values	Selection criteria
REEF (*reef*)	An area of rock or coral that is exposed between mean high water and lowest tide, or just below approximate lowest tide, which is visually prominent or a hazard to shipping.	Polygon	NAME (*name*)	Name of reef	Reefs shown on the 1: 5M General Reference Map edition 4.
RESERVOIR (*reservoir*)	A body of water collected and stored behind a constructed barrier for a specific use.	Polygon	NAME (*name*)	Name of reservoir	Reservoirs are included where their size after point density reduction is greater than 0.00028 square degrees.
ROAD (*road*)	A route for the movement of vehicles, people or animals.	Chain	NAME (*name*)	Name of road	
			CLASSIFICATION (*class*)	1 = *Dual Carriageway: divided highway, freeway or tollway with separated carriageways in opposite directions.* 2 = *Principal Road: highways, major through routes and major connecting roads.* 3 = *Secondary Road: connecting roads providing a connection between major through routes and/or major connecting roads, or connections between regional centres.* 4 = *Minor Road: all other roads which form part of the public roads system between Principal roads and Secondary roads.* 5 = *Track: public or private roadways of minimum or no construction which are not necessarily maintained.*	Roads shown on the 1:5M General Reference Map edition 4.
			FORMATION (*formation*)	1 = *Sealed* 2 = *Unsealed* 3 = *Unknown* 4 = *Under Construction*	

Feature	Description	Object	Attribute	Attribute values	Selection criteria
			NATIONAL ROUTE NUMBER (*nrn*)	The national route number assigned to the road. If there are multiple numbers, they are delimited by a minus sign.	
SAND RIDGE (*sand_ridge*)	Sand drifts in long ridges tending parallel to and elongating in the direction of the prevailing winds.	Chain			Sand ridges shown on the 1:5M General Reference Map edition 4.
SPOT ELEVATION (*spot_elevatn*)	A point on the earth's surface, of known elevation, above or below the Australian Height Datum (AHD66)	Point	ELEVATION (*elevation*)	Elevation in metres from the Australian Height Datum.	Spot elevations shown on the 1:5M General Reference Map edition 4.
STATE BORDER (*state_border*)	The boundary defining the division of the Commonwealth of Australia into State and Territory administrations.	Chain			
SWAMP (*swamp*)	Land which is so saturated with water that it is not suitable for agricultural or pastoral use and presents a barrier to free passage and/or the nearly level tract of land between mean high water and the line of the highest astronomical tide.	Polygon			This is an aggregation of saline coastal flat, mangrove and swamp features as defined in 250K GEODATA Series 1. Swamps are included where their size after point density reduction is greater than 0.004 square degrees. Polygons smaller than the minimum size for inclusion and are in groups may be aggregated into one larger representative polygon.
WATER BODY VOID (*w_body_void*)	A void area in a waterbody polygon.	Polygon			Waterbody voids are included where their size after point density reduction is greater than 0.0016 square degrees.
WATERCOURSE (*watercours_l*)	A natural channel along which water may flow from time to time.	Chain	NAME (*name*)	Name of watercourse	Watercourses shown on the 1: 5M General Reference Map edition 4.
			PERENNIALITY (*perennial*)	1 = *Perennial* 2 = *Non Perennial*	
WATERCOURSE AREA	A natural channel along which water may flow from time to time.	Polygon	NAME (*name*)	Name of watercourse	Watercourse areas are included where their size after

Feature	Description	Object	Attribute	Attribute values	Selection criteria
(*watercours_a*)			PERENNIALITY (*perennial*)	1 = *Perennial* 2 = *Non Perennial*	generalisation in greater than 0.0035 square degrees.
WATERLINE (waterline)	A line depicting the boundary of a hydrographic area feature.	Chain			

6 Data Quality Information

6.1 Data Quality Statement

The Data Quality Statement is a text file that accompanies every transfer of *GEODATA TOPO 5M 2004*. It contains information that is specific to the format of the data file. It includes information relating to transfer format, coordinate systems, feature occurrence counts, data capture methods and point density. The content of this information may vary with the format in which you are supplied *GEODATA TOPO 5M 2004*.

6.2 Product Quality Information

Lineage
This section contains information on the lineage of the spatial data in this product. Lineage is the history of the spatial data; the source of the data, how they were captured, prepared, revised etc.

History of the 1:5 Million Scale Mapping Program
Publication of the 1:5 million scale General Reference Map began with the publication of edition 1 in 1966. Revisions of information on the maps have seen several new editions of the maps released, the most recent being edition 4 in 1996.

Methods used to produce source material
The production of the 1:5 million scale General Reference Map was carried out in Geoscience Australia. The content of general reference maps is generally derived from appropriate larger scale mapping and consists of standard base information including road, rail, hydrography, locality and relief themes. The current 5 million scale series has been derived from the 1:1 million scale International Map of the World Series.

Generalised compilations at 5 million scale level of detail were produced from the 1:1 million scale series. These compilations were then joined and reduced to produce the 5 million scale map. A composite positive was produced from which multiple guide images were derived to enable scribing of the various colour separations. These, together with various overlays for names and tints, were used to produce printing plates.

History of the 1:250 000 Scale Mapping Program
The majority of the 1:250 000 topographic mapping program was shared by Geoscience Australia (formerly AUSLIG, NATMAP, the Division of National Mapping) and the Royal Australian Survey Corps (RASvy). A small number of map sheets was also produced by State mapping agencies. The National Topographic Map Series (NTMS) at 1:250 000 scale published by Geoscience Australia and the Joint Operation Graphics (JOG) published by RASvy replaced the R502 Series which consisted largely of uncontoured maps at the same scale.

The 544 sheets in the NTMS/JOG series provided the first nationwide coverage of published, fully contoured, topographic maps. National coverage was completed in 1988.

Because many of the NTMS sheets were converted by RASvy to their specifications and re-published as JOG maps, there are many sheets available today with identical reliability dates but published by both agencies.

The map reproduction material used for the printing of the maps with the latest reliability date was used as the primary source material for *GEODATA TOPO 250K*. However, where both agencies published maps with identical reliability dates, the NTMS material was used.

Most of the 1:250 000 scale NTMS and JOG maps were derived manually from 1:100 000 scale topographic maps and compilations. Some were derived from 1:100 000 and 1:50 000 digital data.

Production methods for the 1:100 000 NTMS

Overall, positional control for the 1:100 000 series was based on the Australian geodetic network using the 1966 adjustment and the Australian National Spheroid. The base data for the map compilation material were obtained through the stereographic observation of aerial photography. The aerial photography was generally flown at a nominal scale of 1:80 000 in blocks which equate to 1:250 000 map sheets. Some photography was at a larger scale.

A variety of specialised aerial photography cameras were used. Propagation of positional control through the blocks of photography was carried out by slotted template adjustment until the mid 1970s when a method using analytical block adjustments was introduced. In some areas radar altimetry was used to extend vertical control.

As with the cameras, a number of models of stereo plotter were used for the plotting of detail and production of contours on the compilation material.

Finally, all map production material was checked against other sources of information. Extensive field checking on the ground and from the air was carried out, local authorities were consulted, and larger scale reference material and supplementary photography were used.

Methods used to produce source material

To produce the 1:250 000 scale map products 1:100 000 scale material was used. Where 1:100 000 maps were published the relevant map repromat was used. In the remote areas where 1:100 000 maps were not published the relevant 1:100 000 compilation material was used.

To ensure accurate registration of all the required drafting materials (clear film, scribe sheets, masks, photographic film etc.), the materials were pre-punched on the same, large format register punch. Reference grids and graticules were computer generated and plotted on a separate layer of stable base drafting film.

Map corner-marks were placed on all separate layers, eg. infrastructure, roads, contours etc. which were then reduced photographically to 1:250 000 scale on a large format Klimsch 'Super Autohorika 101' camera using the same camera setting for all material of the one map area. A sheet of pre-punched clear film was registered to the graticule and all six reduced film positives of each layer registered to the relevant section of the 1:250 000 scale graticule, trimmed, butt joined and fixed to the clear film. A contact negative was produced from each of these assemblies. Each negative was printed down on scribing material as a guide for the cartographer to manually scribe the detail according to map specifications. Map detail was displaced if necessary to avoid overprinting and to ensure a cartographically acceptable product. Masks were cut on 'peelcote' type material to provide infill for area features such as lakes, built-up areas etc.

A colour proof of the line-work was used as a guide to type positioning on a clear film overlay. The proof was thoroughly checked for accuracy, completeness and correct registration of detail. After proof corrections were carried out on the final reproduction material used for platemaking by the printer was produced. Checking forms were used for quality control throughout the production process.

The TOPO 250K Program

In response to significant user demand for high quality digital map data for GIS, Geoscience Australia commenced development and production of a GIS ready topographic data product. This product, TOPO 250K, was designed to meet the needs of users in a broad range of professional activities by providing a nationally consistent data set. Geoscience Australia commenced this program in 1990 and completed full national coverage by mid 1994. This original version of TOPO 250K is now referred to as *TOPO 250K Series 1.*

To ensure national coverage in a relatively short period of time, *TOPO 250K Series 1* included the base topographic features in three themes but did not include all topographic features contained on the paper map product. In total, 43 different feature types were captured. The three themes captured in *TOPO 250K Series 1* were:

- Infrastructure (eg. road, rail, localities);
- Hydrography (eg. streams & lakes); and
- Relief (Spot elevations only).

Production of TOPO 250K Series 1

The key source for the production of *TOPO 250K Series 1* was the repromat from the paper map product. Photographic film positives of the various layers (drainage, roads, railways, etc.) of the original map production material were produced for scanning. These positives were then scanned on Geoscience Australia's Scitex R280 blue argon laser scanner/ plotter at a resolution of 20 pixels per millimetre. The resultant raster image was then checked for content and batch cleaned to remove background noise such as dust spots and film blemishes. The raster files were then converted to vector form by using either vectorising software on the Scitex system or by ProVec software.

Additional editing of the road vector files was required in MicroStation to correct road intersections. This procedure required the vector file to be overlayed on the raster file. All data files were transformed from system coordinates into AMG coordinates through an affine adjustment routine. A residual report was generated which contained the transformation parameters. The average of the residuals had to be less than 50 metres with a maximum individual residual less than 50 metres. If the residuals failed these criteria the control from the scanned file was adjusted and re-transformed.

Verification plots were generated to check content and accuracy. All corrections were attended to before applying a blanket feature code to the vector file (eg. road, river, railway etc). The vector file, original scanned material, residual reports and verification plots were then dispatched to the Geoscience Australia production teams for further editing and tagging with feature codes and attributes.

The Geoscience Australia production teams converted the data into GINA format. The GINA file was then built into a database on the GeoVision GIS system. GeoVision GIS was the production system for the tagging, attributing and structuring of the data. As the database was built, the data were separated into a number of layers which equated to the final product, ie. hydrography, infrastructure and framework.

Initially, the data were checked for positional accuracy (bulk shifts or rotations) by comparing the position of the features close to the edge with a mathematically generated tile edge. Adjustments were made as necessary. The point density of the data was then filtered according to the parameters stated in the specification. Overshoots and undershoots were corrected and gaps in the data were checked by software.

Once all the gaps and undershoots were resolved, the hydrography layer was processed to separate the linear streams network from the waterbody polygons. When the separation was finished, the connector features were positioned. Linear topology for the streams network and polygon topology in the waterbody network were generated using the GeoVision software.

The roads network was given a systematic, visual check on the screen to detect wrongly shaped intersections. Small polygons at intersections were located by custom-written software. Bridges and tunnels were placed and a plot generated to check for completeness, accuracy of tagging and mutual correlation.

The appropriate topology in the remaining layers was generated and a final plot of the entire database was checked for completeness, accuracy of tagging and mutual correlation.

The attributes of the features were generated by two methods:

- Automatically, using in-house software which associates the attributes with the production feature codes. The automatic output was checked to guarantee no abnormalities; and
- Manually, especially for attribute information such as names and road route numbers.

The spelling of names was checked by software against an existing gazetteer. Once the attributing was finalised, the production feature codes were replaced by the TOPO 250K feature classes and the unique feature identifier was generated.

The spot elevations in the relief theme were derived from 1:100 000 scale compilation material, printed maps or digital topographic data which had been captured prior to beginning production of TOPO 250K data.

Geoscience Australia map and data revision program

Following the completion of the *TOPO 250K Series 1* program, it was apparent that the paper map products needed to be updated. However there was also a demand for incorporating all map features into the data product. At this time Geoscience Australia commenced production of a joint map and data specification for topographic data at both 1:100 000 and 1:250 000 scales. At the same time it was recognised that efficiencies could be gained from producing the paper map and digital data product using parallel production methods.

Geoscience Australia specified the production of a central database comprising the information to be included in both products as well as features which are specific to either the data (eg hypsometric areas) or the map (eg. grid and graticule). From this database, the map features and layout can be extracted to produce a postscript file for repromat production, and, the data features can be extracted to form the TOPO 250K tile. This central database is referred to as the 'Working Database' and is held in ArcInfo format. This database will be the key data source for future revisions of both map and data products at 1:250 000 scale.

Another key feature of this program was the introduction of non-standard sheet lines for some map and data areas. Strong customer feedback was received about the availability of maps where the majority of the map was covered with sea and only a very small area of land was shown. Customers were also clearly frustrated by the break up of major cities or features across two map areas. Because of this, Geoscience Australia reviewed all traditional sheet lines and where practical, these were altered to provide optimum representation of the landmass and key map features. As a result, many map sheets were extended beyond the bounds of the traditional map areas and some map areas were rotated to cover an area of 1.5 degrees latitude and 1 degrees longitude.

Production of TOPO 250K Series 2

As an outcome of the Federal Government Budget in 1996, Geoscience Australia was instructed to market test the production of its topographic map and data products. As a result of the market testing, Geoscience Australia outsourced product delivery to the private sector, creating a facilities management contract and a panel of external suppliers.

Under this outsourced arrangement, Geoscience Australia provides the Specification and revision material to the contractors and undertakes independent quality testing of the completed products. Additionally, Geoscience Australia sets the priorities for the program and manages the contractors on behalf of the Commonwealth. Information sources which have been supplied to the Contractors for revision are:

- TOPO 250K Series 1 data;
- TOPO 250K data error reports;
- Horizontal control points data;
- Reproduction material from the previous edition map (contours, cultural features, hydrographic features);
- Satellite imagery (Landsat TM Bands 1,2,3,4,5,7 & SPOT PAN);
- Geographic source information; and
- Map surrounds information.

Satellite imagery has been used to generate the forest areas within the vegetation theme. New features have been identified and placed using satellite imagery, however this information is verified and attributed using other revision intelligence.

A further element of the production of maps and data has been the datum shift from the Australian Geodetic Datum 1966 (AGD66) to the Geocentric Datum of Australia 1994 (GDA94).

To date, all contractors for the production of maps and data under the current revision cycle have used ArcInfo as their base production tool. ArcInfo format is the specified delivery format for the working database and *TOPO 250K Series 2 tile*.

Production of GEODATA TOPO 2.5M 2003

Spatial data for *GEODATA TOPO 2.5M 2003* for the Contours, Offshore and Sand Ridge layers was captured from 1:2.5 million scale mapping by scanning stable base photographic film positives of the various layers of the original map production material. Prior to scanning, each piece of material had twelve control points affixed. The positives were scanned on an Intergraph MapSetter 6000 scanner at a resolution of 20 pixels per millimetre. The resultant raster images were then checked for content and batch cleaned to remove background noise due to dust, film blemishes etc. The Contours, Offshore and Sand Ridge files were then converted to vector form by using ProVec software.

Once vectorised, all files were transformed from system coordinates to geographical coordinates. Verification plots were generated to check for accuracy and content. The vector files in ArcInfo Vector format were then dispatched to the production team.

The Arc Export files resulting from the scanning processes were then built into a database in ArcInfo. ArcInfo was the production system used for the feature coding, structuring and attributing of the data for *GEODATA TOPO 2.5M 2003*. As the database was built, the data were separated into a number of specified layers and fully topologically structured. The point density of the data was filtered according to specified parameters. Overshoots and undershoots were corrected and the presence of gaps in the data checked by software. Feature attributing was carried out using manual and automatic routines.

The key source for production of the Builtup Areas, Drainage, Spot Heights, Framework, Localities, Rail Transport, Road Transport and Waterbodies layers was the *TOPO 250K Series 1* and *TOPO 250K Series 2* products. Due to Australia wide coverage of *TOPO 250K Series 2* being incomplete, *TOPO 250K Series 1* was used where coverage of *TOPO 250K Series 2* did not exist although the Framework and Waterbodies layers have been sourced entirely from *TOPO 250K Series 1*.

The relevant TOPO 250K tiles were appended into one ArcInfo coverage for each layer. In accordance with the selection criteria in the Technical Specifications, all features not required were removed. Some manual generalization of the various layers was then carried out. The point density of the data was filtered according to specified parameters. The various layers were then structured and attributed. The existing attributes were then verified against source material where appropriate. Unique feature identifiers were attached to facilitate 'change-only update of features. Once the appropriate topology for each layer was generated and attributing completed, ArcInfo programs were run to test for invalid values and for inconsistent combinations of attributes. Verification plots of the entire database were produced for checking of completeness, accuracy and mutual correlation.

A further element to the production of *GEODATA TOPO 2.5M 2003* has been the datum shift from the Australian Geodetic Datum 1966 (AGD66) to the Geocentric Datum of Australia 1994 (GDA94).

Production of GEODATA TOPO 5M 2004

Spatial data for *GEODATA TOPO 5M 2004* for the Offshore and Sand Ridge layers was captured from 1:5 million scale mapping by scanning stable base photographic film positives of the various layers of the original map production material. Prior to scanning, each piece of material had twelve control points affixed. The positives were scanned on an Intergraph MapSetter 6000 scanner at a resolution of 20 pixels per millimetre. The resultant raster images were then checked for content and batch cleaned to remove background noise due to dust, film blemishes etc. The Offshore and Sand Ridge files were then converted to vector form by using ProVec software.

Once vectorised, all files were transformed from system coordinates to geographical coordinates. Verification plots were generated to check for accuracy and content. The vector files in ArcInfo Vector format were then dispatched to the production team.

The Arc Export files resulting from the scanning processes were then built into a database in ArcInfo. ArcInfo was the production system used for the feature coding, structuring and attributing of the data for *GEODATA TOPO 5M 2004.* As the database was built, the data were separated into a number of specified layers and fully topologically structured. The point density of the data was filtered according to specified parameters. Overshoots and undershoots were corrected and the presence of gaps in the data checked by software. Feature attributing was carried out using manual and automatic routines.

The Builtup Areas, Contours, Drainage, Spot Heights, Framework, Localities, Rail Transport, Road Transport and Waterbodies layers were sourced from the GEODATA TOPO 2.5M 2003 product.

In accordance with the selection criteria in the Technical Specifications, all features not required were removed. Some manual generalization of the various layers was then carried out. The point density of the data was filtered according to specified parameters. The various layers were then structured and attributed. The existing attributes were then verified against source material where appropriate. Unique feature identifiers were attached to facilitate 'change-only update of features. Once the appropriate topology for each layer was generated and attributing completed, ArcInfo programs were run to test for invalid values and for inconsistent combinations of attributes. Verification plots of the entire database were produced for checking of completeness, accuracy and mutual correlation.

Total Quality Management

Procedures throughout the entire production process are designed to ensure that the data are produced right-first time, every time. The Total Quality Management system guarantees quality from the product planning stage, right through to product delivery to the client.

To ensure that the product meets the customers' needs, Geoscience Australia uses market research in the product development stage and incorporates customer feedback. Geoscience Australia's commitment to truth-in-labelling for GEODATA products (complete information on data quality) helps customers to use the data appropriately.

A detailed technical specification is developed for each GEODATA product. The technical specification is a controlled document. This means that only correct and current copies of the specification are available within Geoscience Australia. The technical specification is the basis for the development of production procedures and other documentation such as the Product User Guide.

Post-production validation and testing

After the production team completes work the data are dispatched for post-production validation and testing. A separate cell has been set up totally independent of the production areas to carry out the validation and testing. The validation and testing cell tests the correctness of the data using ArcInfo. The data are statistically sampled and tested to ensure compliance with all aspects of the technical specification. Aspects tested include topological integrity, completeness, positional accuracy, attribute accuracy and filtering.

Statistical sampling procedures are used to ensure that the data pass pre-determined conformance tests. For some tests, such as polygon closure, a zero failure tolerance is set. For other tests sufficient samples are tested to ensure that an error rate of up to 5% is permitted in the data. If any data do not satisfy the conformance criteria they are sent back to the production area for reprocessing.

Positional accuracy

The positional accuracy of spatial objects is an estimate of the degree to which the planimetric coordinates and elevations of a feature instance agree with the true values or values accepted as being true. The measure of accuracy given for *GEODATA TOPO 5M* is the standard deviation.

Planimetric accuracy

Well-defined points: a well-defined point is one which can be accurately identified on the source material and in the digital data. Most commonly the points used in tests are nodes at intersections.

Geoscience Australia has carried out both error budget analysis and independent testing to verify the positional accuracy of the data. *GEODATA TOPO 5M* data comply with the following statement of planimetric accuracy: The summation of errors from all sources represents an error of 1700 metres on the ground for 1:5 000 000 data. This will be rounded to 2000 metres for the *GEODATA TOPO 5M* product. An alternative and equal way of expressing this error is '*Not more than 10% of well-defined points are in error by more than 2800 metres*'.

An estimate of the standard deviation of planimetric error of each feature instance is given in the Data Quality Table. The value for point and linear features is generally 2000 metres. A value of 9999 is used when the positional accuracy of the feature is not definable or not applicable. For example, the coordinates of a connector feature do not carry any meaning with respect to positional accuracy and so the value given is 9999.

Cartographic generalisation: Most *GEODATA TOPO 5M* data were digitised from existing map production material at 1:5 000 000, 1:2 500 000 and 1:250 000 scale and some features may be subject to cartographic generalisation. Cartographic generalisation can have a major effect on the selection and positioning of features on small-scale cartographic products. It may involve the processes of selection, displacement, simplification, exaggeration or aggregation.

Selection is necessary to reduce the clutter on the map. Features are selected because of their relative importance. For example, in densely settled areas only the largest cities and towns are shown and some cities and towns may be excluded. In sparsely settled areas very small settlements may be shown.

Displacement is the movement of one feature with respect to another to allow them to be clearly portrayed at map scale. Displacement has not been carried out on layers sourced from *TOPO 250K Series 1* and *TOPO 250K Series 2* other than that which already existed at that scale.

Simplification is a smoothing of the detail of a line to reduce clutter on the map and to clarify portrayal of a feature. For example, at small scale the course of a meandering river is simplified. Exaggeration allows small but significant features to be shown. For example, a small island may be exaggerated because it is a well-known landmark. Aggregation allows a number of small features to be symbolised by a single feature of the same type. For example, a number of small reefs may be aggregated and portrayed as a single reef.

Methods used to verify planimetric accuracy

Geoscience Australia has carried out error budget analysis to verify the planimetric accuracy of the data. The planimetric accuracy attainable in *GEODATA TOPO 5M* is the sum of errors from two sources:

- The positional accuracy of the source material; and
- Errors due to the digitising process.

There is an expectation that the source data comply with the following statement: Not more than 10% of well-defined points are in error by more than 0.5 mm measured on the source material. Statistically, this relates to a standard deviation on the map of 0.31 mm.

The errors caused by the digitising process depend on the accuracy of the set-up, systematic errors in the equipment, errors caused by software, and errors specific to the operator. An accepted standard for digitising is that the line accuracy should be within half a line width. The majority of features in *GEODATA TOPO 5M* have a line width of 0.2 mm or greater. The half line width is taken as 0.1 mm and this is interpreted as one standard deviation for the distribution of errors.

The total statistical error from the source material and digitising process discussed above is given by taking the square root of the sum of the squares of the component standard deviations. This gives a total standard deviation of 0.34 mm. This represents an error of 1700 metres on the ground for *GEODATA TOPO 5M*. A more conservative estimate of 2000 metres for the standard deviation is used in any data quality information on this product. An alternative and equal way of expressing this error is '*Not more than 10% of well defined points are in error by more than 2800 metres'*.

Vertical accuracy

The table below gives the estimates of the planimetric and vertical accuracy for the spot elevation and linear relief features in terms of their standard deviation.

Table 7: Elevation and planimetric accuracy estimates

Feature	Elevation accuracy (Std Dev)	Planimetric accuracy (Std Dev)
Spot Elevation	5	2000
Contour, Isobath	100	2000

Attribute accuracy

Attribute accuracy is a measure of the degree to which the features and their attributes are correct. For this product, attribute accuracy is a measure of the degree to which the attribute values of a feature agree with the information on the source material. The allowable error in attribute accuracy ranges between 0% and 5%.

Description of testing procedure used

Where less than 1% of attribute errors are permissible the entire population is tested. Where a less stringent limit is set for allowable errors a random subset of the relevant features in the tile may be tested. The sample size is determined from statistical tables using the known population size of the relevant feature. The following table sets out the checks on the data and gives the test sample size and the allowable error.

Table 8: Attribute testing

Attributes tested	Test procedures	Test sample size	Allowable error
Unique feature identifiers (UFI) are valid and within the allocated range.	ArcInfo Program	Full population	No errors
All features have a valid data quality pointer.	ArcInfo Program	Full population	2%
Attribute values other than for UFI and Data Quality Pointer are within the valid ranges as specified in the Data Dictionary.	ArcInfo Program	Full population	2%
Features have the correct feature code	ArcInfo Program	Statistical subset	1%
The 'state' attribute of mainland and islands is correct.	ArcInfo Program	Full population	No errors
Islands have the correct name.	ArcInfo Program	Statistical subset	5%
Watercourses have the correct name.	On screen	Statistical subset	5%
The perenniality of lakes is correct.	On screen	Statistical subset	5%
The perenniality of watercourses is correct.	On screen	Statistical subset	5%
Railway track status is correct	On screen	Statistical subset	5%
Railway track number is correct	On screen	Statistical subset	5%
Railway track gauge is correct	On screen	Statistical subset	5%
Railway names (ie. Railway Lines and Railway Stations) are correct.	On screen	Statistical subset	5%
Road classification is correct.	On screen	Statistical subset	5%
Road formation is correct.	On screen	Statistical subset	5%
Roads have the correct name.	On screen	Statistical subset	5%
National Route Numbers on roads are correct and compete.	On screen	Statistical subset	5%
Localities have the correct name.	On screen	Statistical subset	2%
Localities have the correct locality code.	On screen	Statistical subset	5%
Names of the watercourses are carried on the entire length where intended.	ArcInfo Program Onscreen	Statistical subset	2%
Names of the roads are carried on the entire length where intended.	ArcInfo Program On screen	Statistical subset	2%
National Route Numbers of roads are carried on the entire length where intended.	ArcInfo Program On screen	Statistical subset	2%
Status of the railways is carried the entire	ArcInfo Program	Statistical subset	2%

Attributes tested	Test procedures	Test sample size	Allowable error
length of the feature where intended.	On screen		
Names of railways are carried the entire length where intended.	ArcInfo Program On screen	Statistical subset	2%
Gauge of railways is carried the entire length of the feature where intended.	ArcInfo Program On screen	Statistical subset	2%
Built-up area names are correct.	On screen	Full population	5%
Waterbody names are correct.	On screen	Statistical subset	5%
Population ranges for Localities is correct	ArcInfo program	Statistical subset	5 %

Logical consistency
Logical consistency is a measure of the degree to which data complies with the technical specification. Validation of logical consistency may involve tests to check table and file names are set out as in the data dictionary. Also included are graphical tests, which check such things as intersections, polygon closure, minimum size of polygons and topological relationships. The allowable error in logical consistency ranges from 0% to 5%.

Description of testing procedure used
The logical consistency of the data is tested using a mixture of ArcInfo commands and UNIX programs. Where less than 1% of attribute errors are permissible the entire population is tested. Where a less stringent limit is set for allowable errors a random subset of the relevant features in the tile may be tested. The sample size is determined from statistical tables using the known population size of the relevant feature.

The following table sets out the checks on the data and gives the test sample size and the allowable error.

Table 9: Logical consistency checking

Logical consistency check	Test procedure	Test sample size	Allowable error
Names of export files and data quality table are correct.	ArcInfo Program	Full population	No errors
Table names are valid.	ArcInfo Program	Full population	No errors
Item names in coverages are valid.	ArcInfo Program	Full population	No errors
Item names are present in coverage attribute files.	ArcInfo Program	Full population	No errors
Linear features have more than one coordinate pair.	ArcInfo Program	Full population	No errors
Line segments are not greater than 5.07 degrees long (approximately 400 000m)	ArcInfo and UNIX Program	Full population	0.5%
Line segments are not less than 0.0044 degrees long (approximately 500m)	ArcInfo and UNIX Program	Full population	5%
The ArcInfo coverages can be generated, have attributes and can be built.	ArcInfo Program	Full population	No errors
There are no coincident line segments within a single coverage of intersecting arcs without a node or double digitised points.	ArcInfo Program	Full population	0.5%
In polygon coverages, there are no label errors, i.e. every polygon has one and only one polygon label point.	ArcInfo Program	Full population	No errors
There are no pseudo nodes present, i.e. nodes separating arcs with the same attributes excepting the cases outlines in Chapter 4.	ArcInfo Program	Full population	2%
There are no overshoots, i.e. arc overhangs at intersections.	ArcInfo and UNIX Program	Full population	1%
There are no undershoots, i.e. arcs failing to meet at intersections.	ArcInfo Program	Statistical subset	0.5%

Logical consistency check	Test procedure	Test sample size	Allowable error
There are no polygons smaller than minimum specified area.	ArcInfo Program	Full population	5%
There are no new linear features shorter than the minimum length.	ArcInfo Program	Full population	5%
No arcs separate polygons with identical attributes i.e. abutting polygons do not have the same attributes.	ArcInfo Program	Full population	1%
There are no invalid artefacts such as spikes or deviations visible at 1: 2 500 000.	ArcInfo Program On screen	Statistical subset	5%
Junction line segment numbers correct.	ArcInfo Program	Full population	5%
Tile quality and frequency tables are correct.	ArcInfo Program	Full population	No errors
Separate covers have exactly coincident lines where intended.	ArcInfo Program] On screen	Statistical subset	5%
The separation of the digital data linework from the source material due to filtering is not greater than 1000 metres.	On screen	Statistical subset	5%
In the Data Quality Table the data quality pointer values are unique, complete and correct.	ArcInfo Program	Full population	No errors
Data format, tolerances, projection and data type are correct.	ArcInfo Program	Full population	No errors
GEODATA resolution is correct.	ArcInfo Program	Full population	No errors
Rail stations are coincident with nodes in the rail network.	ArcInfo Program	Full population	1%
Well-defined point features are located within 1700 metres of their position on the source material.	ArcInfo Program	Full population	1%
There are no Spot elevations in relevant waterbodies.	ArcInfo Program	Full population	1%
Features labelled as junction features occur only between valid features.	On screen	Statistical subset	1%
Connector features occur only within a mainland waterbody	ArcInfo Program	Full population	1%
Datum shifting has occurred and is correct.	ArcInfo Program On screen	Full population	No errors
Features labelled as waterbodies i.e. late etc., occur only within the mainland or on an island.	ArcInfo Program	Full population	1%
Spot heights agree with contours and hypsometric areas.	ArcInfo Program	Full population	1%
Roads and railways do not fall within reservoirs or perennial waterbodies.	ArcInfo Program	Full population	1%
There are no road formation changes <0.03 degrees in length.	ArcInfo Program	Full population	1%
Railway lines do not cross unspecified watercourses.	ArcInfo Program	Full population	1%
Localities positioning valid with respect to waterbodies.	ArcInfo Program	Full population	1%
Populated places are coincident with roads where specified.	ArcInfo Program	Full population	1%
Coastline is cloned to zero in Contours coverage.	ArcInfo Program	Full population	1%
State border sourced from COAST 100K and vertice removal and cloning requirements met.	ArcInfo Program On screen	Full population	No errors

Completeness

Completeness is the measure of the degree to which all features listed in the technical specifications have been captured in accordance with the selection criteria, definitions and other rules specified. All instances of a feature and its attribute values that appear on the source material are captured unless otherwise indicated in the selection criteria.

Description of testing procedure used

The completeness of the data is tested by assessing the data against the supplied and quoted source material, carrying out visual comparisons and data comparisons as required. Where feature populations are small the entire population will be tested, while a statistical subset may be tested where the numbers are larger.

Table 10: Completeness checking

Completeness check	Test procedure	Test sample size	Allowable error
All features on the source material have been captured as per the selection criteria.	On screen	Statistical subset	2%
Revision has been applied correctly.	On screen	Sample area	2%
Localities with population >100 000 included in Built-up Areas coverage as Built-up Area polygons.	ArcInfo Program	Full population	1%
Built-up Area polygon aggregation and generalisation from source material valid.	On screen	Statistical subset	2%
Swamp polygon aggregation and generalisation from source material valid.	On screen	Statistical subset	2%

6.3 Data Quality Table

The Data Quality Table (*aus5dqt.**) is a look-up table which contains data quality information about each feature instance in the dataset. There is a Data Quality Table per data file. The link between the table and each feature instance is the Data Quality Pointer and is shown in the figure below.

Attribute	Attribute value
Feature	Locality
Name	Derby
Locality	8
Population	3236
Data Quality Pointer	BJ000004
Unique Feature Identifier	EI00006935

Data quality pointer	Feature reliability	Attribute reliability	Planimetric accuracy	Elevation accuracy
BI000002	01/04/2003	01/04/2003	9999	9999
BJ000003	01/04/2003	01/04/2003	2000	9999
BJ000004	01/04/2003	01/04/2003	9999	9999
DH000005	01/04/2003	01/04/2003	2000	100
BI000006	01/04/2003	01/04/2003	2000	100

Figure 14: Attribute and data quality pointer relationship

The attribute content of this table is described in the following table.

Table 11: Description of attributes within the Data Quality Table

Attribute	Description
DATA QUALITY POINTER (*q_info*)	An index value held in the data quality pointer attribute. This is the link field to the primary attribute table.
FEATURE RELIABILITY (*feat_rel*)	Date of photography, field verification or other event which verified the existence of the feature. Only month and year information are significant. The default will be the first day of the respective month. If the month is not known then the default is 1 January of that year. If unknown, a date of 01/01/1901 is recorded.
ATTRIBUTE RELIABILITY (*att_rel*)	Date on which attribute information of the feature was last verified. If one attribute of the feature is amended it is assumed that all attributes have been verified. The default will be the first day of the respective month. If the month is not known then the default is 1 January of that year. If unknown, a date of 01/01/1901 is recorded.
PLANIMETRIC ACCURACY (*plan_acc*)	The accuracy of the horizontal position in metres of a feature on the map sheet used as source material. If a planimetric accuracy for the feature is not applicable, relevant or cannot be reliably quoted then this field shall contain 9999.
ELEVATION ACCURACY (*elev_acc*)	The accuracy of the elevation in metres of a feature on the map sheet used as source material.

Appendix A: Metadata

Note: This dataset description is metadata (data about data) which describes the actual dataset in accordance with the ANZLIC (Australia New Zealand Land Information Council) Core Metadata Guidelines Version 2.

Dataset citation

ANZLIC unique identifier: ANZCW0703006341

Title: GEODATA TOPO 5M 2004

Custodian

Custodian: Geoscience Australia

Jurisdiction: Australia

Description

Abstract:

The GEODATA TOPO 5M 2004 product is a national seamless data product aimed at regional or national applications. It is a vector representation of the Australian landscape as represented on the Geoscience Australia 1:5M general reference maps and is suitable for GIS applications.

The data has been derived from GEODATA TOPO 250K Series 2 data where it was available (else GEODATA TOPO 250K Series 1) and has double precision accuracy.

The product consists of eleven data layers, including: built-up areas; contours; drainage; framework; localities; offshore; rail transport; road transport; sand ridges; spot heights; and waterbodies.

ANZLIC search words:

- BOUNDARIES Administrative Mapping
- LAND Geography Mapping
- LAND Topography Mapping
- MARINE Coast Mapping

Geographic extent name:

AUSTRALIA EXCLUDING EXTERNAL TERRITORIES - AUS - Australia - Australia

Note: The format for each Geographic extent name is: `Name - Identifier - Category - Jurisdiction` (as appropriate) See GEN Register

Geographic bounding box:

North bounding latitude: -9
South bounding latitude: -44
East bounding longitude: 154
West bounding longitude: 112

Data currency

Beginning date: Not Known

Ending date: 01 November 2003

Dataset status

Progress: Complete

Maintenance and update frequency: Irregular

Access

Stored data format:
Digital: ArcInfo

Available format type:
Digital: ArcInfo Export
Digital: ArcView Shapefile
Digital: MapInfo mid/mif

Access constraints:
The data are subject to Copyright. Data files may be downloaded from Geoscience Australia's website at www.ga.gov.au/download/. A licence agreement is required.

Data quality

Lineage:
Data for the Offshore and Sand ridge layers was captured from 1:5 million scale mapping by scanning stable base photographic film positives of the original map production material. The key source material for Built-up areas, Contours, Drainage, Spot heights, Framework, Localities, Rail transport, Road transport and Waterbodies layers was sourced from the GEODATA TOPO 2.5M 2003 product.

Positional accuracy:
GEODATA 5M 2004 data comply with the following statement of planimetric accuracy: 'Not more than 10% of well-defined points are in error by more than 2800 metres'.

Attribute accuracy:
For this product, attribute accuracy is a measure of the degree to which the attribute values of a feature agree with the information on the source material. The allowable error in attribute accuracy ranges between 0% and 5%. Where less than 1% of errors are permissible, the entire population is tested. A full description of the checks on the data, the test sample size and the allowable error are provided in the product user guide of GEODATA TOPO 5M 2004.

Logical Consistency:
Logical consistency is a measure of the degree to which the data complies with the technical specification. The data was tested using a mixture of ArcInfo and UNIX programs. This included checking of the naming of tables and files and graphical tests to check intersections, polygon closures, and minimum size of polygons and topological relationships.

Completeness:
All instances of a feature and its attribute values that appear on the source materia are captured unless otherwise indicated in the selection criteria. The completeness was tested by assessing the data against the supplied and quoted source material, carrying out visual comparisons and data comparisons as required.

Contact information

Contact organisation: Geoscience Australia

Contact position: Geoscience Australia Sales Centre

Mail address: PO Box 378

Locality: CANBERRA

State: ACT

Country: Australia

Postcode: 2601

Telephone: Australia Freecall 1800 800 173

Facsimile: +61 2 6249 9960

Electronic mail address: sales@ga.gov.au

Metadata information

Metadata date: 2004-03-08

Additional metadata

Metadata reference XHTML: http://www.ga.gov.au/meta/ANZCW0703006341.html

Metadata reference XML: http://www.ga.gov.au/meta/ANZCW0703006341.xml

Size of dataset: 3.7 to 5.1 Mb depending on the format

Scale/resolution: 1:5 million

Projection/datum: Geographical coordinates using the Geocentric Datum of Australia 1994 (GDA94). Elevation data are supplied in metres using the Australia Height Datum (AHD).

Appendix B: Product maintenance form

You can help us keep our products up-to-date and error free by letting us know about those features which are missing or incorrectly shown in digital form (or on a map), or any other errors that you have identified. When we receive your information, it will be reviewed before being incorporated into the next version of the product. Amendments to maps and data could take some time to appear in new editions because revision cycles can span several years.

To add a suggestion or identify a problem, please complete this form and return it to:

Geoscience Australia Sales Centre
Geoscience Australia
GPO Box 378
CANBERRA ACT 2601
Please feel free to send us a simple map of the correction by facsimile on +61 2 6249 9960.

Name:..

Phone:...

Date:...

Organisation:...

Product: GEODATA TOPO 5M 2004

Data Format: ArcInfo
ArcView
MapInfo

Coordinate System: Geographical Coordinates

Layer:	Name:	Version:
....	....	....
....	....	....
....	....	....

Feature code:..

Error Description:...
....
....

UFI(s): Coordinates:
List multiple UFIs if the same error applies equally to a number of features. The approximate coordinates of the feature if the UFI is not known or is insufficient to identify the feature

1...	....	....
2...	....	....
3...	....	....
4...	....	....
5...	....	....

Thank you for your interest and assistance!

Glossary

Attribute
The descriptive characteristic of a feature. An attribute has a defined set of attribute values.

Attribute object
The attribute object holds the non-locational or semantic information about the feature instance.

Australian Geodetic Datum 1966 (AGD66)
This datum was adopted in 1966 and is defined by the parameters of the Australian National Spheroid and the coordinates of the Johnston Geodetic Station. This datum is used for the determination of coordinates for some Geoscience Australia products. Superseded by the Geocentric Datum of Australia (GDA94).

Australian Height Datum (AHD)
The datum used for the determination of elevations in Australia. The determination used a national network of benchmarks and tide gauges and set mean sea level as zero elevation.

Australian Map Grid (AMG)
A Cartesian coordinate system based on the Universal Transverse Mercator Projection and the Australian Geodetic Datum. The unit of measure is the metre.

Chain
A line composed of a sequence of non-intersecting line segments bounded by nodes. Chains reference the polygon to the left and right of the chain.

Connector feature
An artificial linear feature used to connect a linear network across an area feature.

Data Quality Statement (DQS)
A text file which carries information about the quality of the spatial data contained in each GEODATA data transfer.

Datum
A mathematical surface from which heights or positions are referenced.

Entity
A real world phenomenon which cannot be divided into phenomena of the same type.

Entity class
A group of entities of the same kind, matching the members of a feature class.

Entity point
An entity point is used to locate point entities represented by a point because of the scale of the source material.

Feature
A feature is the cartographic or digital representation of a class of entity.

Feature class
A feature class is a group of feature instances defined by a set of rules and having common attributes and relationships that are the properties of the corresponding real world phenomena.

Feature instance
A single occurrence of a feature which has a unique set of spatial and attribute object values.

Generalisation
A process which may involve the selection, displacement, simplification, exaggeration or aggregation of features from their true position for the sake of cartographic clarity.

Geocentric Datum of Australia 1994 (GDA94)
The set of geographic coordinates based on the Geocentric Datum of Australia. It is compatible with Global Positioning Systems (GPS). Adopted in 1994 and implemented in the year 2000. Used in production of new editions of 1:100 000 and 1:250 000 NATMAPs.

GEODATA
Geoscience Australia's brand of high quality digital data products for use in Geographic Information Systems (GIS).

Geodetic datum
A datum defines the basis of a coordinate system. A local or regional geodetic datum is normally referred to an origin whose coordinates are defined. The datum is associated with a specific reference ellipsoid which best fits the surface (geoid) of the area of interest. A global geodetic datum is now related to the centre of the earth's mass, and its associated spheroid is a best fit to the known size and shape of the whole earth. The position of a point common to two different surveys executed on different geodetic datums will be assigned two different sets of geographical coordinates.

Geographical coordinates
A position given in spherical coordinates commonly known as latitude and longitude.

Geographic Information System (GIS)
A spatial database which is manipulated via a set of spatial operators or commands.

Latitude
The latitude of a feature is its angular distance on a Meridian, measured northwards or southwards from the terrestrial Equator.

Layer
The features in a theme are subdivided into one or more layers on the basis of the spatial objects used to represent the features. Linear networks, polygons and point features are placed in separate layers.

Linear Network
A layer consisting of linear features which are connected and which form a pathway along which movement is possible.

Longitude
An angular distance measured east or west from a reference meridian (usually Greenwich) on the earth's surface.

National Topographic Map Series (NTMS)
A civilian map series comprising a set of consistent topographic maps nationwide, at scales of 1:100 000 and 1:250 000.

NATMAP
Geoscience Australia's brand for its popular topographic map range.

Node
A point that is a junction of two or more chains or which is the end point of a chain. Connectivity of chains is indicated by the sharing of nodes at their intersections.

Node/chain structure
The structuring of linear features in a layer so that they consist of chains broken by nodes at intersections or at the point where an attribute of the feature changes.

Point
A geometric representation defined by a single (x,y) coordinate pair or an (x,y,z) triplet.

Polygon
A continuous area defined by a set of bounding chains. There is only one external polygon and there may be one or more internal, non-nested inner boundaries.

Polygon label point
A point within a polygon feature instance used to locate labels or information about that polygon. This point is sometimes known as a centroid.

Positional accuracy
Statistical estimate of the degree to which planimetric coordinates and elevations of features agree with their real world values.

Primary Attribute Table
A data table which contains information directly related to the feature instance.

Projection
Any systematic way of representing the meridians and parallels of the earth upon a plane surface or map.

Repromat
Colour-separated reproduction material on a stable base used for the printing of maps.

Secondary Attribute Table
Additional descriptive information about a feature instance which is related to the feature instance by way of information in the primary attribute table.

Segment
A direct line between a pair of points or a point and a node.

Spatial object
The spatial object holds the locational information of a feature instance. It is composed of either a point, chain or polygon.

Theme
The information contained in map production material can be divided into themes which contain logically related geographic information. Each theme is capable of being used as a dataset in its own right.

Tile
The area of a spatial database included in a data transfer.

Tile edge
An artificial linear feature which indicates the boundaries of the tile. The tile edge closes off polygon features which are situated in more than one tile.

Topological integrity
The measure of how well spatial data conform to the sophisticated data structure required for GIS, especially with respect to connectivity and adjacency.

Unique Feature Identifier (UFI)
An attribute code which is unique to each feature entity and is attached to every feature instance. It is primarily used to facilitate 'change only' updates.

Vector Data
Vector data uses points and straight lines (vectors) to describe features on, or characteristics of, the earth’s surface. Vector data can also include polygons, which are areas enclosed by a number of vectors. To record additional information, data attributes can be attached to individual vector features.

Vertex
The connecting point of two line segments.