Australian Government

Geoscience Australia

GEODATA TOPO 10M 2002 (Double precision)

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: July 2002
2nd Edition released: April 2003
3rd Edition released: October 2003
4th Edition released: July 2004 (Minor corrections)

Technical support:
For up to date information on *GEODATA TOPO 10M 2002* 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
Internet: www.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 http://www.ga.gov.au/nmd/mapping/acknowledge.htm

About the product user guide
This product user guide sets out the fundamental concepts and characteristics of *GEODATA TOPO 10M 2002*. 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.

The information in this product user guide was correct at the time of publication and is subject to change. Geoscience Australia assumes no liability resulting from any statements, errors or omissions in the publication or from the use of information contained in this product user guide.

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 10M 2002

2.1 GEODATA TOPO 10M 2002 components

Your *GEODATA TOPO 10M 2002* 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 10M 2002*.
- **Data Quality Statement file**
  The Data Quality Statement file carries data quality information relevant to the whole *GEODATA TOPO 10M 2002* 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 10M 2002 product

The 2002 version of the product has been completely revised, including the source information. In October 2003, the data was released in double precision coordinates. Unlike the previous version which was sourced from 1:10M scale General Reference Maps, the majority of *GEODATA TOPO 10M 2002* spatial data was derived from *GEODATA TOPO 250K Series 1* data that have been produced from 1:250K topographic maps. The product meets Geoscience Australia's GEODATA quality standard and is suitable for GIS applications.

GEODATA TOPO 10M 2002 provides a fundamental base layer of geographic information on which you can build a wide range of applications. *GEODATA TOPO 10M 2002* is particularly suited to State-wide and national applications, and can be combined with resource and environmental datasets, satellite imagery and demographic data. Typical applications for *GEODATA TOPO 10M 2002* include:

- Management of natural resources including forests, agriculture, water and minerals;
- Conservation of the environment including wildlife management, pollution control and environmental monitoring;
- Transportation including infrastructure planning;
- Demonstrating the concepts underpinning other GEODATA products such as *GEODATA TOPO 250K*;
- Demonstrating the advantages of well structured, quality assured spatial data;
- Demonstrating the versatility of digital data when combined with GIS technology; and
- Cultivating the imagination of potential users of GIS.

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 10M 2002 is available in geographical coordinates (latitude and longitude) in decimal degrees using the Geocentric Datum of Australia 1994 (GDA94). Elevation data are supplied in metres using the Australian Height Datum (AHD).

3 Data loading

3.1 Application formats

The *GEODATA TOPO 10M 2002* data is supplied in three application formats:

- ArcInfo Export;
- ArcView Shapefile; and
- MapInfo mid/mif.

3.2 Description of files

The downloaded *GEODATA TOPO 10M 2002* package contains the following files.

Table 1: *GEODATA TOPO 10M 2002* files

Documentation files				
File name	**File content**			
60803_user_guide.pdf	This product user guide			
Data files				
File name	**ArcInfo Export (*.e00) File size (KB)**	**ArcView Shapefile (*.dbf, *.prj, *.shp, *.shx) File size (KB)**	**MapInfo mid/mif (*.mid, *.mif) File size (KB)**	**File content**
aus10bgd	75	56	52	Built-up areas
aus10cgd	1 570	1 670	1 639	Contours
aus10dgd	951	416	370	Drainage
aus10dgt	2	2	2	Data Quality Table
aus10egd	13	7	3	Spot heights
aus10fgd	673	727	717	Framework
aus10g.dqs	3	4	3	Data Quality Statement
aus10lgd	124	82	31	Localities
aus10ogd	378	320	216	Offshore
aus10rgd	635	321	236	Rail transport
aus10vgd	1 933	986	749	Road transport
aus10wgd	467	360	331	Waterbodies
cla_lut	1	1	1	Secondary attribute table for road classification
ele_lut	1	1	1	Secondary attribute table for elevation range classification
for_lut	1	1	1	Secondary attribute table for road formation description
gag_lut	1	1	1	Secondary attribute table for rail gauge description
per_lut	1	1	1	Secondary attribute table for perennial description
pop_lut	1	1	1	Secondary attribute table for population range
sta_lut	1	1	1	Secondary attribute table for state description
stu_lut	1	1	1	Secondary attribute table for railway status description
Total	6 832 KB/ 6.7 MB	4 959 KB/ 4.9 MB	4 357 KB/ 4.3 MB	

Please note:

- ArcView and MapInfo files are arranged into folders by layer to assist users with the management of the files.

- There are attribute frequency tables (*.frq) for the ArcView and MapInfo files.

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 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:

Aus10<layer_id><coord_id><precision_id> (e.g. aus10bgd.e00)

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

Aus10<layer_id><coord_id><precision_id>_<object_id> (e.g. aus10gbd_p.shp)

Where:

Aus10 refers to the GEODATA TOPO 10M 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 10M 2002 layers*

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

4 Data characteristics and concepts

4.1 GEODATA TOPO 10M 2002 essential characteristics

Truth-in-labelling
GEODATA TOPO 10M 2002 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.

Resolution of coordinates
Coordinates of all spatial objects are quoted to the nearest 0.001 degrees (approximately 100 metres) for the 1:10 million scale product. However, in some formats the rounding of coordinates may not be able to be maintained.

Area of tile coverage
The geographic area covered by each file of spatial data is described as a tile. *GEODATA TOPO 10M 2002* comprises a single tile covering the whole of Australia. The extents of the tile and the land included in the data are shown in the diagram below.

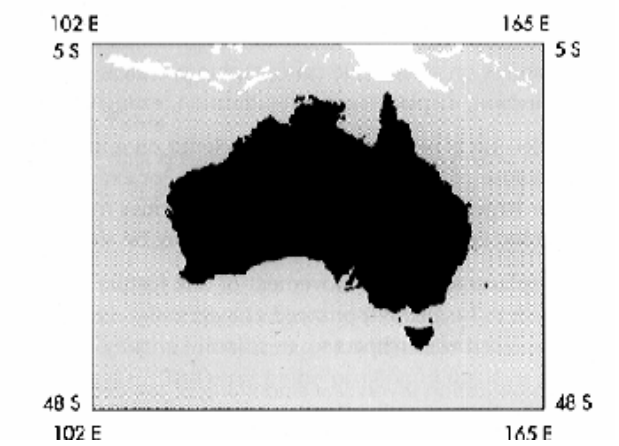

Figure 1: *GEODATA TOPO 10M 2002* tile edge extents

Topological integrity
GEODATA TOPO 10M 2002 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 the Data Quality Statement in Chapter 6. Generally the compliance rate is 99.5% or 95%. 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 following diagram.

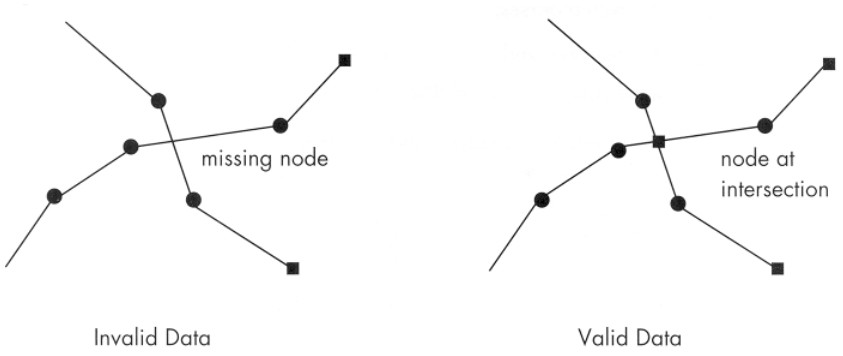

Figure 2: Linear intersections

- 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.

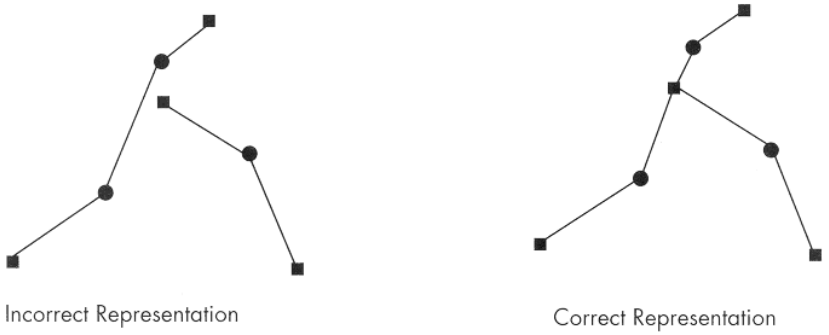

Figure 3: Undershoots

- 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.

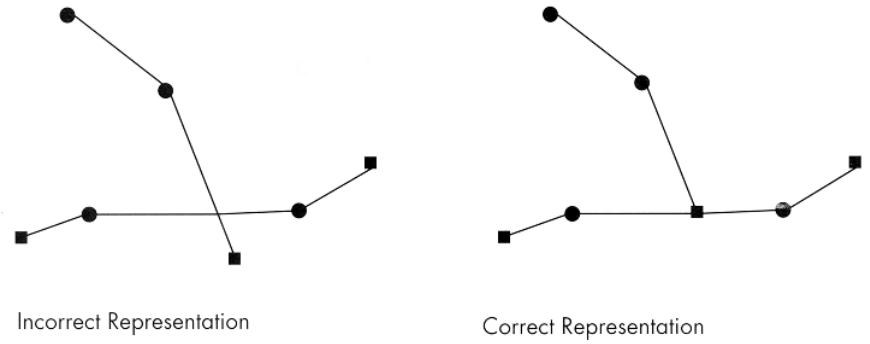

Figure 4: Overshoot

Figure 5: Broken lines

Figure 6: Data spikes

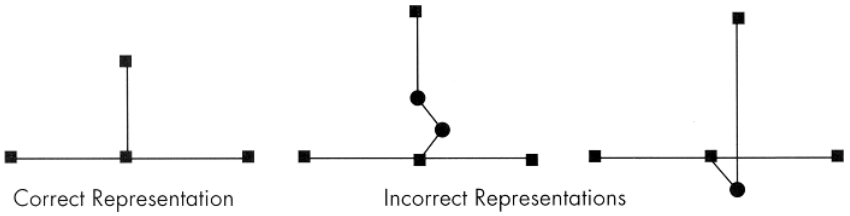

Figure 7: Artefacts in data

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 10M 2002*:

- The minimum length of a line segment is 0.0044 degrees (approximately 500 metres); and
- The maximum length of a line segment is 3.52 degrees (approximately 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 10M 2002* 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. The *GEODATA TOPO 10M 2002* product numbers have been allocated from FF00010001. UFI's from the 10M Series 1 data are carried over. These start from FF00000001
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 (not used at 1:100K and 1:250K)	
	G Vegetation	
	H Reserved Areas	

Data quality pointer
Each entity in *GEODATA TOPO 2.5M 2003* 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 *1:10 General Reference Map	A sequential number identifying unique combination of data quality attributes.
B 1:250 000	G Master Names File	
C 1:1 000 000	H Thematic overlay	
D 1:2 500 000	I *WAC Spot Height Database	
E 1:5 000 000	J Geoscience Australia Railways Map	
F 1:10 000 000	K SYMBAS Guide to Classification of Inland Water	
G Variable	L Review of Australia's Water Resources AGPS 1975	
H Not applicable	M State Railway Annual Reports	
	N Royal Automotive Association maps	
	O 'The National Highway' – Dept Transport and Communication map	
	P State Road Authority maps	
	Q *250K GEODATA Series 1	
	R *Geoscience Australia Mapping Program – Revision Information Railways Database	
	S *Satellite imagery	
	T *1:2.5M General Reference Map Edition 6	
	U *100K Coastline and State Borders Database	
	(* Indicates sources used for TOPO 10M Version 2)	

4.2 GEODATA TOPO 10M 2002 data concepts

Each feature in *GEODATA TOPO 10M 2002* 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 8: 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.

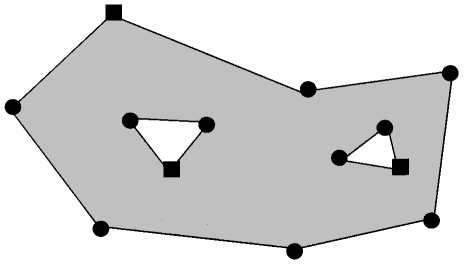

Figure 9: Polygon spatial object

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	Lake
Name	LAKE ARGYLE
Perennial	1
Data Quality Pointer	BQ000019
Unique Feature Identifier	FF00011233

Attribute	Attribute value
Feature	Watercourse
Name	TWENTY MILE CREEK
Perennial	2
Data Quality Pointer	BQ000018
Unique Feature Identifier	FF00011351

Figure 10: 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 *'Non Perennial*' for the Perennial code of 2.

Primary attribute table

Attribute	Attribute value
Feature	Watercourse
Name	TWENTY MILE CREEK
Perennial	2
Data Quality Pointer	BQ000018
Unique Feature Identifier	FF00011351

Secondary attribute table – per_lut

Code	Description
0	n/a
1	Perennial
2	Non Perennial

Figure 11: 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 10M 2002* is composed of only 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.

GEODATA TOPO 10M 2002 product is divided into ten layers:

5.2 Data layers

Table 5: Content of *GEODATA TOPO 10M 2002* layers

Layer	Layer type	Feature	Attributes	Spatial object	Attribute source
Built-up areas	Polygon	builtup_a	name	Polygon	*name*: 1996 census
		builtup_l		Chain	
Contours	Polygon	contour	elevation	Chain	
		hypso_a	elev_class	Polygon	
		isobath	depth	Chain	
		tile_edge		Chain	
		waterline		Chain	
Drainage	Linear	connector	name, perennial	Chain	*name*: 250K GEODATA Series 1 *perennial*: inland water features guide Sept 1966
		watercours_l	name, perennial	Chain	
Spot heights	Point	spot_elevatn	name, elevation	Point	*name, elevation*: GA Spot Height Database 1:1M
Framework	Polygon	island	name, state	Polygon	*name*: 250K

Layer	Layer type	Feature	Attributes	Spatial object	Attribute source
		junction		Chain	GEODATA Series 1
		mainland	state	Polygon	
		sea		Polygon	
		state_border		Chain	
		tile_edge		Chain	
		waterline		Chain	
Localities	Point	pop_place	name, pop range	Point	*name*: 250K GEODATA Series 1 *pop_range*: 1996 Census
Offshore	Polygon	offshore_l		Chain	*name*: National Topographic Map series, 100K & 250K
		reef	name	Polygon	
Rail transport	Point/linear	rail_station	name	Point	*status, gauge, name*: GA Mapping Program – Revision Information Railways Database 250K scale.
		railway	status, gauge	Chain	
Road transport	linear	road	name, class, formation, nrn	Chain	*name class, formation, nrn*: 250K GEODATA Series 1, AAA road maps
waterbodies	polygon	junction		Chain	*name*: 250K GEODATA Series 1 *perennial*: inland water features guide Sept 1996
		lake	name, perennial	Polygon	
		reservoir	name	Polygon	
		waterline		Chain	

Special features

Connector feature (Drainage layer)
Drainage patterns are made up of both linear (eg. narrow streams) and polygon features (eg. 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 water bodies 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 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. All Connectors contained in waterbodies that flow into other waterbodies have been extended to join the Connector on the recipient waterbody.

The use of the Connector feature ceases when a watercourse runs into the sea. 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.

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 river they represent. In the application of the rule it has been considered that rivers can change their perenniality along their course.

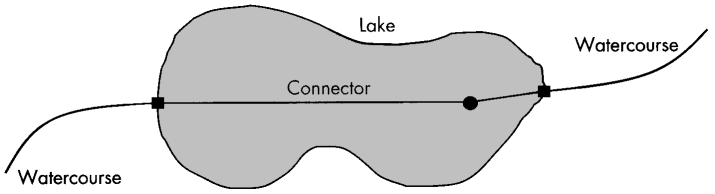

Figure 12: Connector feature for linear watercourses

Depiction of the coastal environment
The Framework, Waterbodies and Offshore layers contain features depicting the coastal environment. The area of tidal influence is part of the Sea feature unless it is closed by a Junction feature.

The coastline is represented by chains coded as *Waterline* (as is waterbody boundaries) and is the position of mean high water level.

The diagram below illustrates how features are used in *GEODATA TOPO 10M 2002* to represent the riverine and maritime environments.

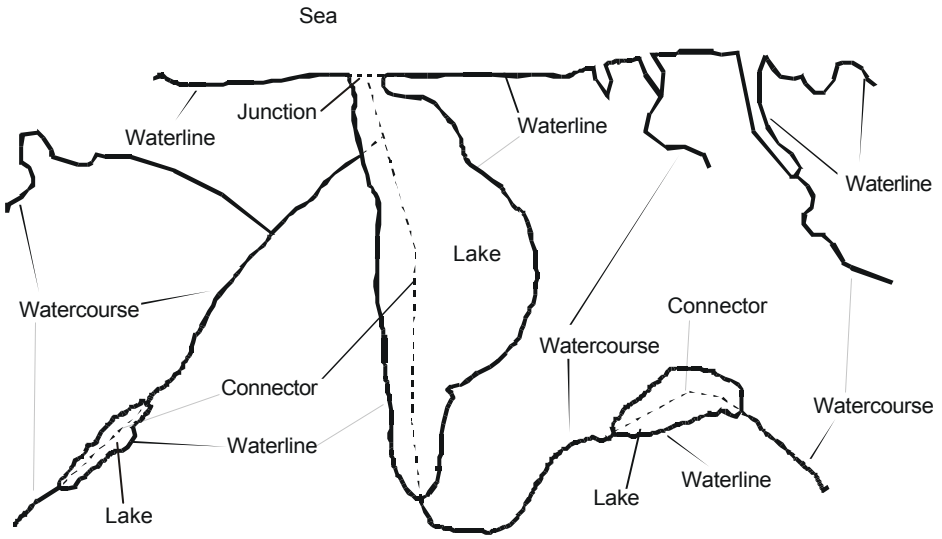

Figure 13: Riverine and maritime feature representations

Framework polygon boundaries (Framework layer)
Where the bounding chain of the sea feature is coincident with the tile edge, it comprises the tile edge feature. This means the framework layer is bounded by one continuous series of tile edge feature instances.

Hypsometric area feature (Contour layer)
The hypsometric area feature is a polygon bounded by the linear elevation features held in this layer. It is included to facilitate GIS analyses involving relief data. For example, the hypsometric area polygon could be used to quickly identify all localities within the elevation class 0 to 200 metres.

The bounding chains of a hypsometric area polygon may be composed of one or more of the following classes of feature:

- waterline;
- contour;
- isobath; and
- tile edge.

Islands (Framework layer)
Islands are represented as polygons coded ‘island’ when they are fully surrounded by sea.

Junction feature (Framework and Waterbodies layers)
The Junction is a linear feature in the Framework and Waterbodies 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 also separates watercourse polygons from the Sea. 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 vertice junction features are permissible in the framework layer. The Junction features in the Framework layer are replicated in the Waterbodies 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; and
- 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.; and
- Fill the coastal gaps in the framework layer.

Locality feature (Locality layer)
Only localities that are significant populated places are included in *GEODATA TOPO 10M 2002*. The population of each locality feature is provided as a population range attribute. Populations for each town and city have been determined according to four classes as defined in the data dictionary in Section 5.3.

Road and rail intersections (Road transport and Rail transport layers)
A road or rail intersection in digital 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 is at a distance greater than three times the line width unless there is a bend in the road before this distance.

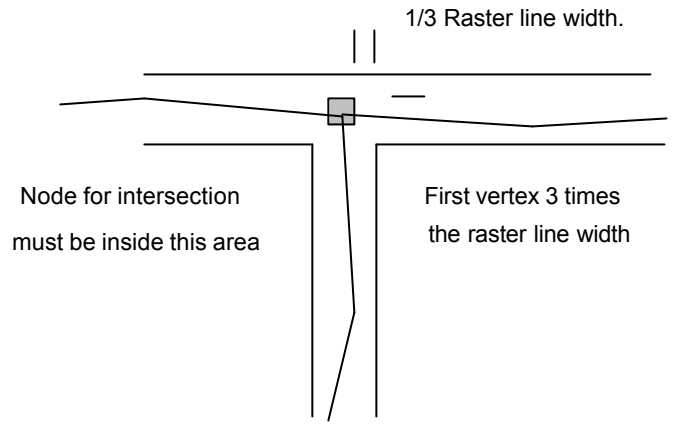

Figure 14: Intersections

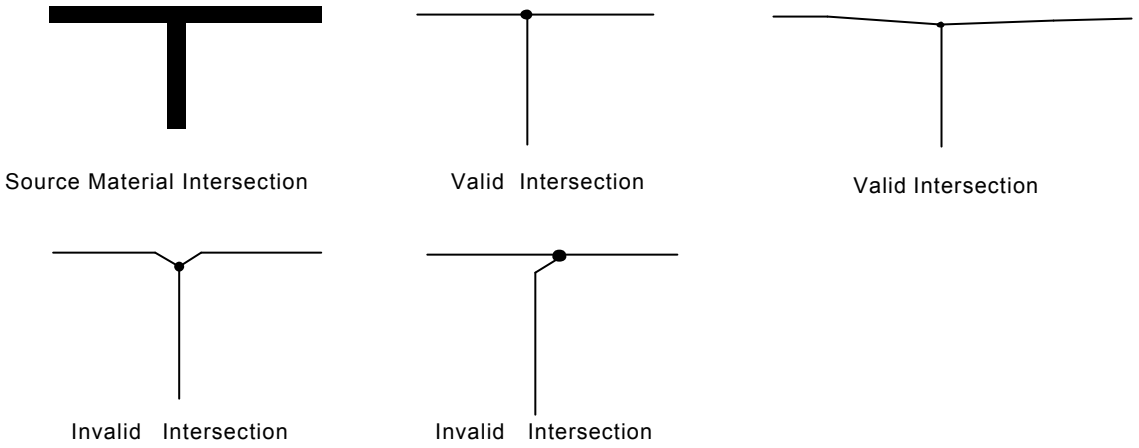

Figure 15: Valid and invalid intersections

Waterbody polygon boundaries (Waterbodies layer)
The boundaries of all lake and reservoir features are represented by the Waterline feature.

Waterline feature (Contour layer)
The waterline feature in this layer represents the coastline and is positionally coincident with the waterline feature instances held in the Framework layer. Used in the context of the Relief layer, the waterline represents a continuous line of zero elevation. This has been done to avoid the topologically illegal duplication that would occur if a zero elevation contour and a zero depth isobath were included in the data.

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 10M 2002*

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	Cities with a population of 100 000 or greater at the 1996 Census of Population and Housing.
BUILT-UP AREA LINE (*builtup_l*)	The bounding line of a built-up area.	Chain			
CONNECTOR (*connector*)	An artificial line used to connect watercourse features across a waterbody feature to allow network analysis of riverine systems.	Chain	NAME (*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 (AHD)	Contours at the intervals shown on the 1:10M General Reference Map edition 3.
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*	

Feature	Description	Object	Attribute	Attribute values	Selection criteria
				6 = *1500 to 2000 metres* 7 = *>2000 metres* 8 = *–200 to 0 metres (seabed)* 9 = *<–200 metres (seabed)*	
ISLAND (*island*)	An are of land fully surrounded by the sea.	Polygon	NAME (*name*)	Name of the island	All islands contained in GEODATA TOPO 250K Series 1 where the area is greater than 0.002 square degrees after generalisation.
			STATE (*state*)	0 = *n/a* 1 = *ACT*: Australian Capital Territory 2 = *JBT*: Jervis Bay Territory 3 = *NSW*: New South Wales 4 = *NT*: Northern Territory 5 = *QLD*: Queensland 6 = *SA*: South Australia 7 = *TAS*: Tasmania 8 = *VIC*: Victoria 9 = *WA*: Western Australia	
ISOBATH (*isobath*)	A line on a map that represents and imaginary line of points of equal depth in relation to a common datum.	Chain	DEPTH (*depth*)		Isobars at the intervals shown on the 1:10M General Reference Map edition 3.
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 the lake	Lakes contained in GEODATA 250K TOPO Series 1 where the area is greater than 0.01 square degrees after generalisation.
			PERENNIALITY (*perennial*)	0 = *n/a* 1 = *Perennial* 2 = *Non Perennial*	
LOCALITY (*pop_place*)	A populated place.	Point	NAME (*name*)	Name of locality	All populated places shown in GEODATA TOPO 10M Series 1.

Feature	Description	Object	Attribute	Attribute values	Selection criteria
			POPULATION RANGE (*pop_range*)	1 = *> 99999* 2 = *25000 – 99999* 3 = *5000 – 24999* 4 = *< 5000*	
MAINLAND (*mainland*)	The area of Continental Australia including Tasmania	Polygon	STATE (*state*)	0 = *n/a* 1 = *ACT*: Australian Capital Territory 2 = *JBT*: Jervis Bay Territory 3 = *NSW*: New South Wales 4 = *NT*: Northern Territory 5 = *QLD*: Queensland 6 = *SA*: South Australia 7 = *TAS*: Tasmania 8 = *VIC*: Victoria 9 = *WA*: Western Australia	
OFFSHORE LINE (*offshore_l*)	The line bounding polygons in the offshore layer.	Chain			
RAILWAY (*railway*)	A transportation system using one or more rails to carry freight or passengers.	Chain	STATUS (*status*)	1 = *Operational* 2 = *Abandoned* 3 = *Under construction*	All operational, or under construction non-tourist/historical standard, broad and narrow gauge lines. Where scale limits the depiction of all tracks especially in Builtup Areas, only those that depict the general representative pattern will be shown. Short branch lines which are insignificant at 10M scale are not shown.
			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 or freight be loaded or unloaded. There may or may not be a platform. The railway station may not be in use.	Point	NAME (*name*)	Name of the railway station	Railways are shown: • where they coincide with a locality that has been included in the locality coverage.

Feature	Description	Object	Attribute	Attribute values	Selection criteria
					• At the end of a branch line.
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 the reef	Reefs shown on the 1:10M General Reference Map edition 3.
RESERVOIR (*reservoir*)	A body of water collected and stored behind a constructed barrier for a specific use.	Polygon	NAME (*name*)	Name of the reservoir	All reservoirs contained in GEODATA 250K TOPO Series 1 where the area is greater than 0.01 square degrees after generalisation.
ROAD (*road*)	A route for the movement of vehicles, people or animals.	Chain	NAME (*name*)	Name of the road	All roads contained in TOPO 10M Version 1 are shown.
			FORMATION (*formation*)	1 = *Sealed* 2 = *Unsealed* 3 = *Unknown* 4 = *Under construction*	
			NATIONAL ROUTE NUMBER (*nrn*)	The national route number assigned to the road. If there are multiple numbers, they are delimited by a minus sign.	
			CLASSIFICATION (*class*)	2 = *Principal road*: highway and regional road 3 = *Secondary road*: arterial and connecting road 4 = *Minor road/track*: minimal use road	
SEA (*sea*)	The water area surrounding the Australian continent (including Tasmania) and its offshore islands.	Polygon			
SPOT ELEVATION (*spot_elevatn*)	A point on the earth's surface, of known elevation, above or below the Australian Height Datum (AHD66)	Entity point	NAME (*name*)	Name of the mountain or hill.	Spot elevations shown on the 1:10 million General Reference Map edition 3.
			ELEVATION (*elevation*)	Elevation in metres from the Australian Height Datum.	
STATE BORDER (*state_border*)	The boundary defining the division of the Commonwealth of Australia into State and Territory administrations.	Chain			

Feature	Description	Object	Attribute	Attribute values	Selection criteria
TILE EDGE (*tile_edge*)	The line defining the limit of the data.	Chain			
WATERCOURSE (*watercours_l*)	A natural channel along which water may flow from time to time. Included terms are natural channels such as river and creek.	Chain	NAME (*name*)	Name of the watercourse.	All watercourses contained in TOPO 10M GEODATA Series 1 are shown.
			PERENNIALITY (*perennial*)	0 = *n/a* 1 = *Perennial* 2 = *Non Perennial*	
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 a *GEODATA TOPO 10M 2002* tile. It contains information that is specific to the tile covered by the data file. It includes information relating to tile identification, 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 10M 2002*.

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 1:10 Million Scale Mapping Program
The current series of the 1:10 million scale General Reference Map began with the publication of edition 1 in 1975. It was produced as a general purpose small-scale map and was distinguished from its predecessors by continental relief shading. It was frequently distributed free by the Commonwealth Government in the interest of national development and international relations.

Later, it became an integral part of the Atlas of Australian Resources, and was the foundation of the small-scale thematic map series which presented individually some twenty different geographical themes. Revisions of the map saw the second and third editions produced in 1984 and 1990 respectively. Edition 3 was markedly different to the earlier editions in both design and content level.

Methods used to produce source material
The production of the 1:10 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, population and relief themes. The current 1:10 million scale series has been derived from the 1:5 million scale General Reference Map, which has been produced from the 1:2.5 million scale General Reference Map, which in turn originated from the 1:1 million scale International Map of the World Series.

Generalised compilations at 1:2.5 million scale level of detail were produced from the 1:1 million scale series. These compilations were then joined and reduced to produce the four-sheet 1:2.5 million scale map. After joining the reduced composite positives of the four 1:2.5 million scale sheets, further cartographic generalisation techniques were used to produce the final reproduction material for the 1:5 million scale map.

Production of the 1:10 million map involved reducing the 1:5 million scale material. A composite positive was produced from which multiple guide images were derived to enable cartographic generalisation and scribing of the various colour separations. These, together with various overlays for names and tints, were used to produce printing plates.

Technical description of digitisation process for the Spot heights, Contours and Offshore layers
Spatial data for *GEODATA TOPO 10M Version 1* has been used for the Spot Heights, Contours and Offshore Layers in *GEODATA TOPO 10M Version 2* and was captured from 1:10 million scale mapping by scan digitisation. Stable base photographic film positives of the various layers of the original map production material were produced for scanning. Prior to scanning, each piece of material had twelve control points affixed. The positives were scanned on a Scitex R280 blue argon laser scanner/plotter. The resultant raster image was then checked for content and batch cleaned to remove background noise due to dust, film blemishes etc.

The contour layer was separately processed through purpose written batch programs on the Scitex system to produce a vector file. An interactive program was used to tag all contours with their correct height. Once vectorised, all files were transformed from system coordinates to geographical coordinates. Verification plots were generated on a Precision Image electrostatic plotter to check for accuracy and content. The vector file in GeoVision GINA format, verification plots, residuals reports and feature code count listings were then dispatched to the production team.

Feature coding and structuring
The GINA files resulting from the digitisation processes were then built into a database on the GeoVision GIS. GeoVision was the production system for the feature coding, structuring and attributing of the data. 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, and unique feature identifiers were attached to facilitate 'change-only' update of features. Once the appropriate topology for each layer was generated and attributing completed, Oracle RDBMS routines 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.

History of the 1:250 000 Scale Mapping Program
The majority of the 1:250 000 topographic mapping program was shared by Geoscience Australia's National Mapping Division (formerly AUSLIG) 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 the former AUSLIG 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, *GEODATA TOPO 250K,* was designed to meet the needs of users in a broad range of professional activities by providing a nationally consistent data set. AUSLIG commenced this program in 1990 and completed full national coverage by mid 1994. This original version of TOPO 250K is now referred to as *GEODATA TOPO 250K Series 1*.

To ensure national coverage in a relatively short period of time, *GEODATA 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 *GEODATA TOPO 250K Series 1* were:

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

The key source for the production of *GEODATA 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 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 AUSLIG production teams for further editing and tagging with feature codes and attributes.

The 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/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 *GEODATA TOPO 250K* feature classes and the unique feature identifier was generated.

Production of *GEODATA TOPO 10M Version 2*

The key source for production of the Builtup Areas, Drainage, Framework, Localities, Rail Transport, Road Transport and Waterbodies layers was the *GEODATA TOPO 250K Series 1* product. All *GEODATA TOPO 250K Series 1* 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 various layers were then structured and attributed. The existing attributes were then verified against source material where appropriate.

Some revision of the Builtup Areas, Road Transport, Rail Transport and Waterbodies layers was carried out using the latest available satellite imagery. The point density of the data was then filtered according to the parameters stated in the technical specifications. The primary source for the Spot Heights, Contours and Offshore layers was the *GEODATA TOPO 10M Version 1* product. The *GEODATA TOPO 10M Version 1* files in GINA format were converted to ArcInfo coverages, the Waterline (coastline) in the Contours layer was replaced with Waterline (coastline) from the Framework layer.

A further element to the production of *GEODATA TOPO 10M 2002* has been the datum shift from the Australian Geodetic Datum 1966 (AGD66) to the Geocentric Datum of Australia 1994 (GDA94).

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 the ArcInfo GIS. 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 10M 2002* 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 10M 2002* data comply with the following statement of planimetric accuracy: The summation of errors from all sources represents an error of 3400 metres on the ground for 1:10 000 000 data. This will be rounded to 4 000 metres for the *GEODATA TOPO 10M 2002*. Alternative and equal ways of expressing this error are:

- Not more than 10% of well-defined points are in error by more than 5600 metres.
- In the worst case, a well-defined point is out of position by 12 000 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 4 000 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: *GEODATA TOPO 10M 2002* data were digitised from existing map production material at 1:10 000 000 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 are shown and some cities 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 *GEODATA TOPO 250K Series 1* other than that which already existed in that product. 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 10M 2002* 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 10M 2002* 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 0.33 mm. This represents an error of 3 300 metres on the ground for TOPO 10M. A more conservative estimate of 4 000 metres for the standard deviation is used in any data quality information on this product. Alternative and equal ways of expressing this error are:

- Not more than 10% of well defined points are in error by more than 6 400 metres.
- The worst case error for the data is 12 000 metres.

To test this error budget analysis, the coordinates of 59 well-defined points over the map area were compared to the coordinates for the same points derived from 1:250 000 scale maps. The standard deviation of radial error resulting from these measurements was 3 400 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	4000
Contour, Isobath	50	4000

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 allocated range.	ArcInfo Program	full population	no errors
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 plot and screen	Statistical subset	1%
The 'state' attribute of mainland and islands is correct.	ArcInfo Program	Full population	No errors
The Sea attribution is correct (ie. not named)	ArcInfo Program plot and screen	Full population	No errors
Islands have the correct island name	ArcInfo Program	Statistical subset	5%
The entire tile framework (border) is coded as tile edge.	ArcInfo Program	Full population	1%
Watercourses have the correct name	On screen	Statistical subset	5%
Roads 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 gauge is correct	On screen	Statistical subset	5%
Road classification is correct	On screen	Statistical subset	5%
Road formation is correct	On screen	Statistical subset	5%
National route numbers on roads are correct and complete	On screen	Statistical subset	5%
Localities have the correct name	On screen	Full population	2%
Names of watercourses are carried the entire length where intended	ArcInfo Program On screen	Statistical subset	2%
Names of roads are carried the entire length where intended	ArcInfo Program On screen	Statistical subset	2%
National route numbers of roads are carried the entire length where intended	ArcInfo Program On screen	Statistical subset	2%
Status of railways is carried the entire length where intended	ArcInfo Program On screen	Full population	2%
Gauge of railways is carried the entire length where intended	ArcInfo Program On screen	Full population	2%
Population ranges for Localities is correct	On screen UNIX	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 3.52 degrees long	ArcInfo and UNXI Program	Full population	5%
Line segments are not less than 0.0044 degrees long	ArcInfo and UNXI Program	Full population	0.5%
The ArcInfo coverages can be generated, with attributes attached, and can be 'built'	ArcInfo program	Full population	no errors
There are no coincident line segments within a single coverage or 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 arcs broken by related point features	ArcInfo program	Full population	2%
There are no overshoots, i.e. arc overhangs at intersections	ArcInfo and UNIX Program	Statistical subset	1%
There are no undershoots, i.e. arcs failing to meet at intersections	ArcInfo program	Statistical subset	0.5%
There are no polygons smaller than minimum specified area	ArcInfo program	Full population	5%
There are no linear features shorter than the minimum length	ArcInfo program	Full population	5%
No arcs separate polygons with identical attributes, except for the UFI, 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	ArcInfo Program and on screen	Statistical subset	5%
Junction line segment numbers correct	ArcInfo Program	Full population	5%
Separate covers have exactly coincident lines where intended	ArcInfo Program and on screen	Statistical subset	5%
There are no spot elevations in the sea	ArcInfo Program	Full population	1%
There are no spot elevations in waterbodies	ArcInfo Program	Full population	1%
Features coded as tile edge appear only on the framework of the map	ArcInfo Program	Full population	1%
Features labelled as junction features occur only between valid features	On screen	Statistical subset	1%

Logical consistency check	Test procedure	Test sample size	Allowable error
Connector features occur only within a mainland waterbody	ArcInfo Program	Full population	1%
Features labelled as islands or reef are completely surrounded by sea	ArcInfo Program	Full population	1%
Datum shifting has occurred and is correct	ArcInfo Program	Full population	No errors
Features labelled as waterbodies, i.e. lake etc., except sea, 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 waterbodies	ArcInfo Program	Full population	1%
Localities do not fall within the sea	ArcInfo Program	Full population	1%
Localities do not fall within waterbodies	ArcInfo Program	Full population	1%
In the data quality table the data quality pointer values are unique	ArcInfo Program	Full population	No errors
Data format, projection and data type are correct	ArcInfo Program	Full population	No errors
GEODATA resolution is correct	ArcInfo Program	Full population	No errors
Coastline is cloned as zero height in Waterline and Contours coverage	ArcInfo Program	Full population	1%
State Border is consistent with COAST 100K adjustment instructions (re-densification, cloning requirements)	ArcInfo Program and 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 localities that should be shown as per the selection criteria are shown	On screen	Full population	5%
All features on the source materials have been captured	On screen	Statistical subset	2%
Revision has been applied correctly	On screen	Statistical subset	2%

6.3 Data Quality Table

The Data Quality Table is a look-up table which contains data quality information about each feature instance in the dataset. There is 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	Lake
Name	LAKE ARGYLE
Perennial	1
Data Quality Pointer	BB000019
Unique Feature Identifier	FF0011233

Data quality pointer	Feature reliability	Attribute reliability	Planimetric accuracy	Elevation accuracy
BQ000018	1996/09/01	1996/09/01	4000	9999
BQ000019	1996/09/01	1996/09/01	9999	9999
BQ000020	2001/12/01	2001/01/01	4000	9999
BQ000021	1994/07/01	1994/07/01	4000	9999

Figure 16: 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*)	Only month and year information are significant. The default will be the first day of the respective month. If the month us 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 us 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: ANZCW0703005262

Title: GEODATA TOPO 10M 2002 (Double precision)

Custodian

Custodian: Geoscience Australia

Jurisdiction: Australia

Description

Abstract:

GEODATA TOPO 10M 2002 provides a fundamental base layer of geographic information on which you can build a wide range of applications. The data has been derived primarily from *GEODATA TOPO 250K Series 1* and contains the following features: State borders, roads, railways, populated places, rivers and lakes, coastline and islands, bathymetric and land contours, and spot locations.

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: 1998-12-03

Ending date: 2002-05-20

Dataset status

Progress: Complete

Maintenance and update frequency: Not Known

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:
The key source for production of the Builtup Areas, Drainage, Framework, Localities, Rail Transport, Road Transport and Waterbodies layers was the *GEODATA TOPO-250K Series 1* product. Some revision of the Builtup Areas, Road Transport, Rail Transport and Waterbodies layers was carried out using the latest available satelite imagery. The primary source for the Spot Heights, Contours and Offshore layers was the *GEODATA TOPO 10M Version 1* product. A further element to the production of *GEODATA TOPO 10M 2002* has been the datum shift from the Australian Geodetic Datum 1966 (AGD66) to the Geocentric Datum of Australia 1994 (GDA94).

Positional accuracy:
GEODATA TOPO 10M 2002 data comply with the following statement of planimetric accuracy: 'The summation of errors from all sources represents an error of 3 400 metres on the ground for 1:10 million data. This will be rounded to 4 000 metres for the *GEODATA TOPO 10M* product. Alternative and equal ways of expressing this error are: - Not more than 10% of well-defined points are in error by more than 5 600 metres. - In the worst case, a well-defined point is out of position by 12 000 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 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 is generally tested. The sample size is determined from statistical tables using the known population size of the relevant feature. A full description of the checks on the data, the test sample size and the allowable error are provided in the *GEODATA TOPO 10M 2002* Data Quality Statement file supplied with the data.

Logical Consistency:
For the *GEODATA TOPO 10M 2002* product, logical consistency is a measure of the degree to which data complies with the technical specifications. The data were tested using a mixture of UNIX scripts and ArcInfo commands which were independent of the production system. Graphical tests were used to check such things as intersections, polygon closure, minimum size of polygons and topological relationships. A full description of the checks on the data, the test procedure, test sample size and the allowable error are provided in the *GEODATA TOPO 10M 2002* Data Quality Statement file supplied with the data.

Completeness:
All instances of a feature and its attribute values that appear on the source material have been captured unless otherwise indicated in the selection criteria. The completeness was tested by overlaying symbolised plots of the data on the source material and carrying out a visual comparison.

Contact information

Contact organisation: Geoscience Australia

Contact position: Geoscience Australia Sales Centre

Mail address: GPO 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: 2003-10-09

Additional metadata

Metadata reference XHTML: http://www.ga.gov.au/meta/ANZCW0703005262.html

Metadata reference XML: http://www.ga.gov.au/meta/ANZCW0703005262.xml

Size of dataset: 4.3 – 6.7 MB depending on the format.

Scale/resolution: 1:10 million

Projection/datum: Geographical coordinates using the Geocentric Datum of Australia 1994 (GDA94). Elevation data are supplied in metres using the Australian 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 6249 9960.

Name:.....

Phone:.....

Date:.....

Organisation:.....

Product: GEODATA TOPO 10M

Data Format: ArcInfo Export
ArcView Shapefile
MapInfo mid/mif

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
A 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 1994 (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 polygons 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
A text file which carries information about the quality of the spatial data contained in each tile of a 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 geographical 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.

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.

Raster Data
Raster data is made up of picture elements, or pixels, each having a discrete value and ordered together in a regular grid. In terms of spatial data, each pixel represents an area of the earth's surface at a specific location.

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.

Universal Transverse Mercator Projection (UTM)
A worldwide systematic application of the Transverse Mercator Projection applying to the region between 80°S and 84°N latitude. The UTM is a modified TM projection whereby the natural scale of the central meridian is scaled by a factor of 0.9996 to enable a wider area to be mapped with acceptable distortion.

Each Zone is six degrees of longitude in width with a half-degree of overlap within the adjoining zone and having a true origin at the intersection of the central meridian of that zone and the Equator.

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.