Loading and Managing
OS MasterMap® Topography Layer

An ESRI (UK) White Paper
Confidentiality Statement

This document contains information which is confidential to ESRI (UK) Limited. No part of this document should be reproduced or revealed to third parties without the express permission of ESRI (UK) Limited.

© 2008 ESRI (UK) Ltd and its third party licensors. All rights reserved.

ESRI (UK) Ltd
Millennium House
65 Walton Street
Aylesbury
Buckinghamshire
HP21 7QG

Tel: +44 (0) 1296 745500
Fax: +44 (0) 1296 745544
Website: www.esriuk.com

Released: 10/11/2008

Version: 5.3

Document History:

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>31/08/2003</td>
<td>Document initial release.</td>
</tr>
<tr>
<td>2.0</td>
<td>16/12/2003</td>
<td>Document Updated/ Name Change.</td>
</tr>
<tr>
<td>4.0</td>
<td>05/06/05</td>
<td>OS MasterMap v6</td>
</tr>
<tr>
<td>4.1</td>
<td>11/08/05</td>
<td>Edits after Ordnance Survey review</td>
</tr>
<tr>
<td>5.0</td>
<td>19/05/2006</td>
<td>Technology updates</td>
</tr>
<tr>
<td>5.1</td>
<td>19/05/2006</td>
<td>Len Laver and Wai-Ming Lee review and edit of sections relevant to OS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MasterMap-to-Go service</td>
</tr>
<tr>
<td>5.2</td>
<td>01/04/2008</td>
<td>Updated to new ESRI (UK) branding and style.</td>
</tr>
<tr>
<td>5.3</td>
<td>10/11/2008</td>
<td>Updates for new ‘minimise load on database’ option.</td>
</tr>
</tbody>
</table>
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Introduction</td>
<td>5</td>
</tr>
<tr>
<td>2.</td>
<td>OS MasterMap</td>
<td>6</td>
</tr>
<tr>
<td>2.1.</td>
<td>What is OS MasterMap?</td>
<td>6</td>
</tr>
<tr>
<td>2.2.</td>
<td>Benefits of OS MasterMap</td>
<td>7</td>
</tr>
<tr>
<td>2.3.</td>
<td>TOIDs and TOID-Associations</td>
<td>8</td>
</tr>
<tr>
<td>2.4.</td>
<td>How OS MasterMap Topography Layer is Structured</td>
<td>9</td>
</tr>
<tr>
<td>2.5.</td>
<td>OS MasterMap Topography Layer Data Schema</td>
<td>10</td>
</tr>
<tr>
<td>3.</td>
<td>OS MasterMap Topography Layer Data Supply</td>
<td>16</td>
</tr>
<tr>
<td>3.1.</td>
<td>Initial OS MasterMap Topography Layer Data Supply</td>
<td>16</td>
</tr>
<tr>
<td>3.2.</td>
<td>Change-Only Updates for OS MasterMap Topography Layer</td>
<td>18</td>
</tr>
<tr>
<td>4.</td>
<td>Loading OS MasterMap Topography Layer Data</td>
<td>20</td>
</tr>
<tr>
<td>4.1.</td>
<td>Database Types</td>
<td>20</td>
</tr>
<tr>
<td>4.2.</td>
<td>ProductivitySuite for Initial Topography Layer Data Load</td>
<td>22</td>
</tr>
<tr>
<td>4.3.</td>
<td>Database Performance</td>
<td>22</td>
</tr>
<tr>
<td>4.4.</td>
<td>Database Sizing</td>
<td>22</td>
</tr>
<tr>
<td>4.5.</td>
<td>OS MasterMap-to-Go for Initial Load</td>
<td>25</td>
</tr>
<tr>
<td>5.</td>
<td>Enterprise Database Configuration</td>
<td>29</td>
</tr>
<tr>
<td>5.1.</td>
<td>Tablespace Sizing</td>
<td>29</td>
</tr>
<tr>
<td>5.2.</td>
<td>Load Space Requirements</td>
<td>29</td>
</tr>
<tr>
<td>5.3.</td>
<td>Archiving Space Requirements</td>
<td>30</td>
</tr>
<tr>
<td>6.</td>
<td>Using the OS MasterMap Topography Layer Repository</td>
<td>31</td>
</tr>
<tr>
<td>6.1.</td>
<td>Using a “Seamless” Database</td>
<td>31</td>
</tr>
<tr>
<td>7.</td>
<td>Updating OS MasterMap Topography Layer Holdings</td>
<td>34</td>
</tr>
<tr>
<td>7.1.</td>
<td>OS MasterMap Update Mechanism</td>
<td>34</td>
</tr>
<tr>
<td>7.2.</td>
<td>Ordering Change-only Update</td>
<td>34</td>
</tr>
<tr>
<td>7.3.</td>
<td>Data Delivery Format</td>
<td>35</td>
</tr>
<tr>
<td>7.4.</td>
<td>Processing Change-only Update with ProductivitySuite</td>
<td>35</td>
</tr>
<tr>
<td>7.5.</td>
<td>How ESRI (UK) Software Updates the Database</td>
<td>39</td>
</tr>
<tr>
<td>7.6.</td>
<td>Managing the COU Process Logistics</td>
<td>44</td>
</tr>
<tr>
<td>8.</td>
<td>Validating the OS MasterMap Holding</td>
<td>48</td>
</tr>
<tr>
<td>8.1.</td>
<td>Feature Validation Dataset</td>
<td>48</td>
</tr>
<tr>
<td>8.2.</td>
<td>Validation with ProductivitySuite</td>
<td>48</td>
</tr>
<tr>
<td>9.</td>
<td>Summary</td>
<td>50</td>
</tr>
<tr>
<td>10.</td>
<td>Appendix A – OS MasterMap Converter Log Definitions</td>
<td>51</td>
</tr>
<tr>
<td>10.1.</td>
<td>Information Logs</td>
<td>51</td>
</tr>
</tbody>
</table>
10.2. Warning Logs...............................................................................................52
10.3. Error Logs....................................................................................................53

11. Appendix B - Correcting Discrepancies in Data............................... 60
1. Introduction

This report introduces the Ordnance Survey OS MasterMap® digital data products, and describes how they can best be implemented within ESRI (UK) software and solutions. The document focuses on the Topography Layer, although many of the principles apply equally to the OS MasterMap Integrated Transport Network (ITN) and OS MasterMap Address Layers.

Topics covered in this report include:

- Introduction to OS MasterMap: its characteristics, benefits, and issues;
- Database implementation considerations: database selection and configuration;
- Loading and maintaining the OS MasterMap repository: optimising performance and keeping the database current;
- Using OS MasterMap in a live environment: using GIS applications to display and use the data;
- Troubleshooting.
2. **OS MasterMap®**

2.1. **What is OS MasterMap?**

OS MasterMap is the brand name for Ordnance Survey GB’s large-scale digital data products. OS MasterMap represents a series of seamless data layers containing features which represent ‘real-world’ objects, both man-made and natural. The concept of a ‘real-world object’ in this case is of geographic features representing discrete objects on the ground that a person could identify, such as a building, road, land parcel or an address. This type of geographic feature differs from the previous concept of features, which essentially represented only a cartographic view of the world.

OS MasterMap consists of four Layers, which may be acquired individually:

- **Topography:** Polygonised representation of more than 400 million geographic (surface) features, successor to Land-Line®
- **Integrated Transport Network (ITN):** Routable link and node transportation network; successor to OSCAR®. ITN consists of the road network and road restriction information (RRI) themes.
- **Address:** Postal addresses (including coordinates) for residential and commercial properties across Great Britain. At the time of writing, there are two address products available: Address Layer and Address Layer 2. Address Layer 2 builds on the original Address Layer product, providing additional data such as geographic alternatives, building names aliases, Welsh alternatives, address classifications and two new themes – Objects Without Postal Address (OWPA) and Multi-Occupancies Without a Postal Address (MOWPA). Non postal addresses include a range of features such as post boxes, electricity sub stations, telephone boxes, public car parking, libraries, halls etc.
- **Imagery:** National colour aerial photography at 25cm resolution.*

*At the time of writing of this document, complete national coverage of Imagery was not yet available.

The predecessor to OS MasterMap Topography Layer was and is Land-Line. Land-Line also represents a comprehensive dataset of all features on the ground, however represented only by points, lines and text (i.e., is not polygonised). One of the chief differences in OS MasterMap Topography layer, therefore, is that area features are now represented as polygons.

**Tip:** For complete information on OS MasterMap, visit the Ordnance Survey’s OS MasterMap web pages at:

[www.ordnancesurvey.co.uk/oswebsite/products/osmastermap](http://www.ordnancesurvey.co.uk/oswebsite/products/osmastermap)

Ordinance Survey publish the “OS MasterMap User Guide”, which gives detailed description of each of the products.
2.2. **Benefits of OS MasterMap**

OS MasterMap offers a number of benefits above and beyond predecessor products, including:

- Improved geographic representation of the real world and feature symbolisation, partly because of polygonisation of Topography Layer;
- Complete integration of the four Layers: Topography, Address, ITN, and Imagery;
- Improved speed and simplicity of data capture for an organisation’s own layers using polygonised Topography Layer features;
Opportunities to keep your OS MasterMap data repository current (via change-only update) using the OS Managed GB supply service;

- The ability to manage a single seamless geographic database for a specific area of interest;
- An opportunity to make better and more informed decisions about the database and the ‘real-world’ by identifying specific database changes;
- A unique identifier (TOID) for all represented features;
- The potential to access historical views by applying the change history and version information within OS MasterMap.

2.3. TOIDs and TOID-Associations

Every OS MasterMap feature has a unique identifier, known as a TOID. When the TOID is combined with the feature’s Version number, a unique reference can be made to the feature, fixed to a given point in its life-cycle.

A feature’s TOID will normally stay the same throughout the life of that feature. The feature life-cycle is documented in the OS MasterMap User Guide (www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/userguides); but in simple terms, for an area feature, the logic is that unless the land use of that area can be identified as having changed, the TOID will remain the same. Even if the area changes size, or if a piece of land is subdivided, the area where land use can be seen to be the same before and after the change will retain the original TOID.

An example could be a land parcel, part of which is redeveloped. Although the land parcel feature will have changed shape to accommodate a new building feature, its TOID will remain the same and only its Version number will increment. The building features representing the new house and gardens, etc., will get newly-assigned TOIDs, and be assigned Version ‘1’. It should be noted, however, that because of internal Ordnance Survey update procedures, a feature’s Version number may have incremented to a higher number by the time the feature becomes available for the first time.

Other relevant geographic data can be linked or associated by end-users to an OS MasterMap feature via its TOID, and this is commonly referred to as TOID-Association. In this way, organisational data which might be naturally related in some way to underlying OS MasterMap Topography Layer features, perhaps by shared geometry or data capture method, can be associated with the relevant Topography Layer features. Once associated, the life-cycle of the related feature could be managed by change events which occur to the underlying base map features. For example, a Local Authority might link a Basic Land and Property Unit (BLPU) Extent to the Topography Layer polygons to which it shares geometry; and when the geographic features change in future, be notified that this change effects the BLPU Extent. ESRI (UK)’s ProductivitySuite is required to perform these operations.

Associating data in this way can provide a modern, leveraged means of change-management for organisational data. The full application of TOID-association could ultimately even mean that it is unnecessary to store redundant geometry for features which share geometry completely with OS MasterMap Topography Layer features (some BLPU extents, for example). Instead, the related data is managed only as
attributes linked to base-map shapes. In this scenario, the geometry of a user’s features in associated data would come directly from the OS MasterMap Topography Layer features, and would update automatically if changes occurred through the change-only update process.

2.4. **How OS MasterMap Topography Layer is Structured**

OS MasterMap Topography layer features are specified in the source data file in one of several geometric feature types, representing lines, polygons, points, and annotation. Once converted or loaded into the Geodatabase, these will represent discrete Feature Classes.

The feature types within the OS MasterMap Topography layer are:

<table>
<thead>
<tr>
<th>Feature Type</th>
<th>Description</th>
<th>ESRI Feature Class</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TopographicLine</strong></td>
<td>Line features representing topographic information and inferred TopographicArea boundaries such as polygon-closing links.</td>
<td>Line</td>
</tr>
<tr>
<td><strong>TopographicArea</strong></td>
<td>Area features representing topographic information.</td>
<td>Area</td>
</tr>
<tr>
<td><strong>BoundaryLine</strong></td>
<td>Line features representing the boundaries of administrative areas.</td>
<td>Line</td>
</tr>
<tr>
<td><strong>TopographicPoint</strong></td>
<td>Point features representing topographic detail, bench marks and spot heights.</td>
<td>Point</td>
</tr>
<tr>
<td><strong>CartographicSymbol</strong></td>
<td>Point features that store information about the location and type of symbology used when rendering OS MasterMap graphically.</td>
<td>Point</td>
</tr>
<tr>
<td><strong>CartographicText</strong></td>
<td>Point features that define the content and placement of text when rendering OS MasterMap graphically.</td>
<td>Line and/or Annotation*</td>
</tr>
</tbody>
</table>

*Within the GML source file, Cartographic Text is supplied as a point feature with an angle of display stored as an attribute. By default, ProductivitySuite will convert this to a line feature type, which represents the length and rotation-angle of the text, with the text string stored as an attribute. This can then be rendered ‘on-the-fly’ within GIS. Optionally, ProductivitySuite will also convert this to Geodatabase Annotation. **Section 4.3.3** further discusses these options.*
2.5. OS MasterMap Topography Layer Data Schema

The initial conversion of the GML files for OS MasterMap Topography Layer by ProductivitySuite results in a set of tables and Feature Classes created in the Geodatabase.

2.5.1. Feature Classes:

For the basic OS MasterMap Topographic feature types, six Feature Classes are created in the Geodatabase. These represent the various feature types supplied by the OS: annotation, area, administrative boundaries, lines, points, and symbols.

During conversion, the ProductivitySuite OS MasterMap Converter requests the user to supply a text prefix to act as a prefix for the Feature Class names. This enables the end user to distinguish between data kept for two or more different geographic regions within the same Geodatabase, for example different 100km tile areas, in the case of national load. If, for example, the generic string "OS" were provided by the user as a prefix, the Geodatabase Feature Classes created would be:

- **OSAnno:** Line Layer storing Cartographic Text
- **OSArea:** Polygon Layer storing Topographic Area
- **OSBnd:** Line Layer storing Administrative Boundaries
- **OSLine:** Line Layer storing Topographic Lines
- **OSPnt:** Point Layer storing Topographic Points
- **OSSym:** Point Layer storing Cartographic Symbols

Additional Feature Classes will be created for OS MasterMap Integrated Transport Network (ITN) or Address Layers.

2.5.2. Non-Spatial Tables (Change History):

In addition to the geographic features, OS MasterMap Topography Layer also contains information about the change history of those features. For every feature in the database, a list of the dates of creation, changes that have occurred, and reasons for change is stored. Note that this is not part of the archive created by ProductivitySuite, but rather information provided by Ordnance Survey in GML. It will show all changes that have occurred to a feature from the Ordnance Survey perspective, not just the changes that the target database experiences due to its change-only update (COU) cycle. The change records can be thought of as an edit log, but they do not contain a graphical representation of the features.

These change histories are loaded into non-spatial tables in the database. Each Feature Class has an accompanying change table, so in the example of the spatial
The change history tables do not contain information about the nature of the change that took place – e.g. shape or attribute -- as this is not included in the OS MasterMap GML. The GML merely provides the date and reason for change. A reason for change might be ‘PAI’, or ‘Edit’.

Although it is impossible to know the nature of changes that took place before the initial load, the ProductivitySuite OS MasterMap Converter provides an archiving option for COU, which enables future retrieval of historical records and data views.

The change tables should not be confused with the archive layers. Change tables hold the change reason ‘edit log’ for each updated feature, where as Archive Feature Classes store departed and edited features (with shape) in a series of separate, mirror Feature Classes within the same Geodatabase. Section 7 of this document discusses managing change.

An implementation of an OS MasterMap repository containing Topography Layer, will contain up to six feature tables, and fix non-spatial (attribute) tables.

Each feature table contains the following fields:
Table Name: e.g., OSArea, OSAnno, OSLine

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Index</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>OID</td>
<td>NUMBER 38</td>
<td>38</td>
<td>Internal Geodatabase ID; maintained by ArcGIS.</td>
</tr>
<tr>
<td>TOID</td>
<td>VARCHAR 20</td>
<td></td>
<td>Name will reflect table name: e.g., OSAREATIX, OS Topographic Identifier</td>
</tr>
<tr>
<td>FEATCODE</td>
<td>NUMBER 38,0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VERSION</td>
<td>NUMBER 38,0</td>
<td></td>
<td>Name indicative of table name: e.g., OSANNOVIX, OSAREAVIX</td>
</tr>
<tr>
<td>VERDATE</td>
<td>DATE or VARCHAR 12</td>
<td></td>
<td>Data Type depends on when the GDB was created and/ or updated (*see below).</td>
</tr>
<tr>
<td>THEME</td>
<td>VARCHAR 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHANGE</td>
<td>VARCHAR 80</td>
<td></td>
<td>Most recent change date and reason</td>
</tr>
<tr>
<td>DESCGROUP</td>
<td>VARCHAR 150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAKE</td>
<td>VARCHAR 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHYSLEVEL</td>
<td>NUMBER 38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHYSPRES</td>
<td>VARCHAR 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOADDATE</td>
<td>DATE or VARCHAR 12</td>
<td></td>
<td>Date on which feature was loaded into database</td>
</tr>
<tr>
<td>SHAPE</td>
<td>NUMBER 38 or database object data type</td>
<td></td>
<td>Link to ArcSDE geometry tables (ArcSDE binary format), or binary storage of feature geometry (Spatial types, e.g. Oracle Spatial). This field represents point, line, polygon, or annotation geometry type.</td>
</tr>
</tbody>
</table>

In addition to the fields above, each of the Feature Classes contains one or more additional field unique to that class.
Table: **OSAREA** Type: **Spatial Polygon Layer**

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Index</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>...</strong></td>
</tr>
<tr>
<td>CALCAREA</td>
<td>NUMBER 38,0</td>
<td></td>
<td>OS-calculated area; distinct from Geodatabase calculated Shape_Area</td>
</tr>
<tr>
<td>BROKEN</td>
<td>NUMBER 4</td>
<td></td>
<td>OS flag for incomplete feature</td>
</tr>
</tbody>
</table>

Table: **OSLINE** Type: **Spatial Line Layer**

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Index</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCURACY</td>
<td>VARCHAR 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEIGHTDAT</td>
<td>NUMBER 38,8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DATAACCUR</td>
<td>VARCHAR 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEIGHTGRND</td>
<td>NUMBER 38,8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFTOFEAT</td>
<td>VARCHAR 20</td>
<td></td>
<td>Field indicating that feature is not also part of a polygon boundary</td>
</tr>
<tr>
<td>LININPOLY</td>
<td>NUMBER 4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Beginning at OS MasterMap v6.

Table: **OSPNT** Type: **Spatial Point Layer**

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Index</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCURACY</td>
<td>VARCHAR 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEIGHTDAT</td>
<td>NUMBER 38,8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DATAACCUR</td>
<td>VARCHAR 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEIGHTGRND</td>
<td>NUMBER 38,8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFTOFEAT</td>
<td>VARCHAR 20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: **OSANNO** Type: **Spatial Line Layer**  
(Line Feature Class with text attribute for labelling)
<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Index</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEXTANGLE</td>
<td>NUMBER 38,8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEXT</td>
<td>VARCHAR 250</td>
<td></td>
<td>Text to appear on the map</td>
</tr>
<tr>
<td>TEXT FONT</td>
<td>NUMBER 38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEXTPOS</td>
<td>NUMBER 38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEXTHEIGHT</td>
<td>NUMBER 38,8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ESRI (UK) does not recommend upgrading a DATE field to a DATE Data Type, where this field already exists as text (varchar).

New Geodatabases created with ProductivitySuite or MapManager 8.3 and above will use the DATE Data Type for Date fields.

Table: **OSBND** Type: **Spatial Line Layer**

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Index</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCURACY</td>
<td>VARCHAR 20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: **OSSYM** Type: **Spatial Point Layer**

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Index</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGLE</td>
<td>NUMBER 38,0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFTOFEAT</td>
<td>VARCHAR 20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each spatial table has a related change history table. For every unique TOID in the spatial table, there may be one or many records in the change history table describing all of the changes to that feature. The change history tables are as follows:

- OSAREACHNG
- OSLINECHNG
- OSPNTCHNG
- OSANNCHNG
- OSSYMCHNG
- OSBNDCHNG
The structure of each of these tables is the same:

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Index</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>OID</td>
<td>NUMBER 38,0</td>
<td></td>
<td>Used by ArcGIS as a Key</td>
</tr>
<tr>
<td>TOID</td>
<td>VARCHAR 16</td>
<td>Name will reflect table name: e.g., OSAREACIX</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Link from spatial table</td>
</tr>
<tr>
<td>CHANGEDATE</td>
<td>DATE or VARCHAR 10</td>
<td>Name will reflect table name: e.g., OSAREACIX</td>
<td></td>
</tr>
<tr>
<td>CHANGEDESC</td>
<td>VARCHAR 50</td>
<td>Name will reflect table name: e.g., OSAREACIX</td>
<td></td>
</tr>
</tbody>
</table>

Note the indexes listed in the table diagrams above. In the case of OS MasterMap-to-Go, these would have been created using the appropriate command as described in **Section 4.5.2**. In the case of ProductivitySuite, these will have been created automatically by the software.
3. OS MasterMap Topography Layer Data Supply

OS MasterMap layers (Topography, Address, and ITN) are supplied in OS MasterMap Geographic Markup Language (GML) format, in ziped files with a ‘.gz’ file extension, known as ‘G-Zip’ or ‘G-Zed’. GML is a variant of XML, designed for transferring geographic coordinates in addition to attribute data. An initial supply of GML will contain the full supply of data required to create an OS MasterMap data repository. A change-only update file will also be GML, but will contain only changed features.

OS MasterMap Imagery is supplied as MrSID, JPEG, or TIFF files. Imagery does not participate in change-only update; rather updates consist of complete area replacements.

3.1. Initial OS MasterMap Topography Layer Data Supply

OS MasterMap Topography Layer data is supplied by Ordnance Survey for a specified area-of-interest, according to a customer’s licence agreement. This might be a district, county, country or just a particular area digitised by the user as a polygon on the Ordnance Survey on-line service. The initial supply is a complete dataset of every feature in that area-of-interest.

Although OS MasterMap represents a seamless geographic database, it is impossible for Ordnance Survey to actually physically deliver a seamless database to a customer for all but the smallest of areas, due to the capacity limits of delivery media (e.g., DVD and FTP). Therefore, for delivery purposes the data-area is split up into separate files known as Chunks. There are two types of Chunks: geographic and non-geographic.

3.1.1. Geographic Chunks:

The more common delivery structure is the geographic Chunk. In this case, the area-of-interest of the customer’s order is extracted from the Ordnance Survey seamless database into separate files, or ‘Chunks’. The customer can choose whether Chunks are 2 km, 5 km, or 10 km in geographical size.

Geographic Chunk files contain all of the features that lie wholly or partially within that chunk. Features themselves are never split, so features which cross the Chunk boundaries will be retained in whole, resulting in Chunk files being ‘fuzzy’ around the edges. In addition, neighbouring Chunk files will always each contain all of the features which cross the shared edge, resulting in overlapping and duplicated data. ProductivitySuite software provides a mechanism for removing duplicated data at load-time.

Any two geographic Chunks may also be of considerably different file volumes, depending on the feature richness of the chunk.
3.1.2. Non-Geographic Chunks:

Non-geographic Chunks are delivery files of more or less equal volume, regardless of where within the area-of-interest the actual features are from. In other words, rather than the area-of-interest being divided into geographic chunks, the total volume of features within it is divided into equal GZ chunk volume sizes. It might be useful to imagine each chunk as a bucket of a predetermined capacity, and features are placed into the bucket until it is full, and then a new bucket is started, and so on. With non-geographic Chunks it is theoretically possible that a given delivery chunk will contain features from any part of the area-of-interest. The classic example used for illustration is that, for GB national coverage, a given chunk might contain features both from Lands End (Cornwall) and from John O’Groats (Scotland); however because the algorithm used by Ordnance Survey to fill the chunks is not actually random, such an occurrence is extremely unlikely.
3.1.3. Which to choose?

The choice of geographic or non-geographic chunks is completely up to the end-user; ProductivitySuite software supports both, and each will create an identical database. There are, however, some performance and management considerations and a balance must be struck.

Non-geographic chunks will load quicker because without the need for ‘fuzzy chunks’ there is much less redundant data to load, and no need for de-duplication. However, with non-geographic chunks, neighbouring features may not be physical neighbours in the GML supply, and therefore won’t be written to disk as neighbours. This could have the result that rendering the map display, especially when zoomed-in to large scale, will be slower because the computer must make more disk-fetches to get neighbouring features whereas with geographic the data will be already be semi-ordered.

Additionally, many users have found that geographic chunking aids the management of loading data and applying change-only update, particularly when the area-of-interest is too large to load in one session. Moreover, national customers who use a 100km tiling method (see Section 4.5) may have a requirement to choose geographic chunking.

3.2. Change-Only Updates for OS MasterMap Topography Layer

Subsequent supplies of OS MasterMap data are supplied by Ordnance Survey via change-only update (COU). In fact, the COU mechanism is itself an inseparable facet of the overall OS MasterMap philosophy. COU supply is also delivered in Chunks as GML/GZ files (geographic or non-geographic), however unlike initial supply COU files contain only those features that have changed in the area-of-interest since the last supply of the data to the customer.

Note: a COU delivery for a particular customer is unique to that customer, because it contains all the features which have changed since that customer ordered their most recent supply.

Ordnance Survey are continuously updating the OS MasterMap Topography Layer national coverage as real-world change occurs. The publication database is updated approximately every six to eight weeks.

The latest Change-only Update refresh dates can be seen at:

www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/exisingcustomer/refresdates.html
**Important:** Ordnance Survey changed in May 2005 the publication mechanism from an entirely customised one to a six-weekly scheduled data refresh. This is important particularly for customers who took an initial supply and one or more change-only updates first before this change. Please see Section 7 for more information on keeping an OS MasterMap repository current.

At the time of the customer order, Ordnance Survey will create a custom delivery for that customer current up through the most recent OS MasterMap Topography Layer data publication date. This is true of both initial supply and COU, however particular attention to dates must be applied when ordering COU, and this is discussed in detail in Section 7.
4. Loading OS MasterMap Topography Layer Data

Deployment of OS MasterMap Topography Layer is best suited to a database 'store-once–use-many' enterprise approach; dictated by its volume, complexity of attribution, change-only update business/technical opportunities. For most organisations, the best way to load and manage an OS MasterMap repository is with the ESRI (UK) ProductivitySuite product.

Establishing national geographic coverage is a more complex process and always requires a major relational database system. ProductivitySuite is suitable for managing this, however ESRI (UK) consultancy advice is recommended due to the complexity of configuring a database to hold this very large quantity of data. ESRI (UK) also offer the OS MasterMap-to-Go service. This service provides ArcSDE 9.1 export files of the national seamless MasterMap repository. Clients with ArcSDE 9.1 or later can utilise the supplied scripts and import them automatically.

4.1. Database Types

It is very important to give due consideration to the GIS database which will hold the converted OS MasterMap Topography Layer data. The first decision is whether OS MasterMap Topography Layer data will be imported into a Personal or Enterprise Geodatabase, the latter which is managed with ArcSDE.

The OS MasterMap Converter supports ArcSDE geodatabases and personal geodatabases. File geodatabases are not supported because the SQL queries used for de-duplication and change-only update are not recognised by a file geodatabase.

4.1.1. Personal Geodatabase:

A Personal Geodatabase (PGDB) uses an MS Jet/Microsoft Access file format (.mdb), and is single-editor only. The PGDB format is the native ArcGIS file-based data store, and is editable by any ArcGIS client (i.e., ArcView, ArcEditor, and ArcInfo). It is important to note that a structural 2-Gigabyte limit applies to all .mdb databases, therefore to all Personal Geodatabases.

When planning a Personal Geodatabase implementation, it is important to consider that the database will grow over time as change-only update is applied, especially if Archiving is enabled. Ordnance Survey suggest an annual rate of ‘feature churn’ (the rate at which features will change, on a national basis) is between five and fifteen percent; however this is likely to be greater in the Positional Accuracy Improvement (PAI) years.

The 2-Gb limit to Personal Geodatabase effectively eliminates this as an option for the enterprise corporate data repository for all except the smallest sites, and project groups; however, there are some ways to maximise the use of Personal Geodatabases:
- Use the Geodatabase Filter tool to strip out attribution that is not required;
- Do not enable Archiving;
- Do not choose to enable Geodatabase Annotation – choose the default Cartographic Annotation only and use the ESRI (UK) Cartographic Text Renderer;
- Divide the area of interest into sub-sections, one Geodatabase for each sub-section.

### 4.1.2. Enterprise Geodatabase:

An Enterprise Geodatabase (GDB) stores the data in a commercial relational database (e.g. Oracle or Microsoft SQL Server) and requires the use of ArcSDE for managing the data. These databases are frequently referred to as “ArcSDE Geodatabases”.

Advantages of using Enterprise Geodatabases include:

- Unlimited storage: ideal for larger geographic areas-of-interest or larger numbers of features;
- Supports multiple simultaneous editors and long transaction management;
- Allows Internet map serving (i.e., ArcIMS);
- True client-server access to geographic data;
- Joining of geographic data with other organisational data already resident in an RDBMS;
- Harnesses native security functions available in an RDBMS, including backup, rollback, and security;
- Users of UNI-form Spatial for Local Government applications.

<table>
<thead>
<tr>
<th>Data-use Profile</th>
<th>Geodatabase Type</th>
<th>Load Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Team</td>
<td>Personal</td>
<td>ProductivitySuite</td>
</tr>
<tr>
<td>Local Authority</td>
<td>Enterprise</td>
<td>ProductivitySuite</td>
</tr>
<tr>
<td>Regional Government</td>
<td>Enterprise</td>
<td>ProductivitySuite</td>
</tr>
<tr>
<td>National (e.g. Scotland, Wales)</td>
<td>Enterprise</td>
<td>ProductivitySuite or OS MasterMap-to-Go</td>
</tr>
<tr>
<td>Central Government (whole of GB)</td>
<td>Enterprise</td>
<td>OS MasterMap-to-Go</td>
</tr>
<tr>
<td>Utility, national coverage</td>
<td>Enterprise</td>
<td>OS MasterMap-to-Go</td>
</tr>
<tr>
<td>Commercial or Utility, field office or project</td>
<td>Personal or Enterprise</td>
<td>ProductivitySuite</td>
</tr>
</tbody>
</table>

**Figure 4.1-1:** Table showing examples of OS MasterMap Topography Layer data users and software and services which might apply.
4.2. **ProductivitySuite for Initial Topography Layer Data Load**

ESRI (UK)’s ProductivitySuite can be used to process the OS MasterMap Topography Layer GML data, provided by Ordnance Survey in GZ file format. It is supplied as an extension for ArcGIS Desktop, and also provides a number of tools for managing OS MasterMap Topography Layer beyond the initial data load; including data filtering, data cleaning, and creation of catalogues and metadata for the OS MasterMap Imagery layer.

ProductivitySuite operates within ArcGIS licensing constraints: ProductivitySuite within ArcView will add data to Personal Geodatabases; ProductivitySuite within ArcEditor or ArcInfo will add data to Personal and Enterprise Geodatabases.

When an initial load of OS MasterMap Topography Layer data is taken from Ordnance Survey, the ProductivitySuite OS MasterMap Converter can be used to set up the Geodatabase and load the data, all in one step. From that point, the OS MasterMap Converter will automatically manage the integration of update data into the database when change-only update is applied.

To load the data, run the OS MasterMap Converter command from the ProductivitySuite program menu or the Productivity Data Conversion toolbar in ArcCatalog. For information on how to use the OS MasterMap Converter, readers should consult the ProductivitySuite Help, which describes these processes in further detail.

4.3. **Database Performance**

Database sizes and requirements will vary tremendously according to the number and complexity of features, the computing and network environment, and client use. Implementers of ArcSDE are encouraged to consult the *ArcSDE Configuration and Tuning Guide* white paper for specifics on tuning and optimising a relational database use for geographic information, available from the ESRI, Inc. website (www.esri.com).

4.4. **Database Sizing**

It is very difficult to estimate the size of a Geodatabase. Important factors include:

- Which OS MasterMap layers and feature types are loaded
- Which annotation type is chosen – i.e. Geodatabase Annotation and/or Cartographic Text
- Whether Archiving will be exercised
- Rate of feature churn in OS MasterMap source files in customer’s given area
- Types of indexes maintained by database administrator
- Type of geometry storage, e.g. ArcSDE binary or Oracle Spatial.

As a very basic ‘rule of thumb’, for OS MasterMap Topography Layer expect the loaded GML to take at least **fifteen times (15x)** the disk space that the original GZ
files required, including the building of required indexes. For example, 100-Mb of GZ will require at least 1.5-Gigabytes of tablespace in a database system. Note that this is only to account for basic uncompressing and loading of an initial supply, and not including Geodatabase Annotation or Archiving.

Other factors which will add to the space requirements:

- Application of Geodatabase Annotation will require at least 6% more.
- Implementation of Oracle Spatial will require at least 10% more.
- Archiving, plan on 15% database growth per year.
- Additional required for OS MasterMap ITN and Address layers.

### 4.4.1. Database Sizing for Full Seamless National Coverage

Database size recommendations for storing a national seamless OS MasterMap coverage are indicated in Figure 4.3.2-1. These recommendations are based on an Oracle 9.2i database.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Recommended</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATAFILES</td>
<td>150GB</td>
<td>250GB</td>
</tr>
<tr>
<td>SPATIAL INDEX</td>
<td>120GB</td>
<td>240GB</td>
</tr>
<tr>
<td>TEMP</td>
<td>20GB</td>
<td>50GB</td>
</tr>
<tr>
<td>OTHER</td>
<td>10GB</td>
<td>15GB</td>
</tr>
</tbody>
</table>

**Figure 4.3.2-1:** Table showing the database size requirements for storing a national seamless OS MasterMap coverage for an Oracle 9.2i database.

Figure 4.3.2-1 includes provision for some temporary data files that are consumed during importing of data. However, the figures above assume the following:

1. This database is used only for viewing. No COU is to be applied on it.
2. The database does not include the GEODATABASE ANNOTATION Feature Class.
3. The space is reserved for OS MasterMap only.

### 4.4.2. Indexes

Regardless of the Geodatabase type, ProductivitySuite automatically creates important indexes on the tables, and these must be maintained. Indexes can be checked and built using ArcCatalog in the Properties of the Geodatabases. These are vital for query and especially for change-only update performance.

For each Feature Class, these include:

- Indexes on the TOID field
- Index on Version fields.
For the change tables, this includes:

- Combined index on the TOID, ChangeDate and ChangeDesc fields.

It is important for the Administrator to maintain these indexes properly and to run the ArcCatalog analyze command (ArcSDE only) on the tables after each change-only update to keep the statistics current.

### 4.4.3. Annotation

There are two methods of storing OS MasterMap Topography Layer annotation in the database, Cartographic Text and Geodatabase Annotation. This choice has an impact on database performance.

**Cartographic Text:**

The OS MasterMap Topography Layer GML includes cartographic text elements specified, each consisting of a point feature, a text string and an angle of rotation. By default, ProductivitySuite converts these into line features, rotated according to the specified angle, and with the text string attached as an attribute. This represents the ‘standard’ method of storing text information for OS MasterMap Topography Layer, and in this case the client application (ArcGIS or ArcIMS) renders the text as labels ‘on the fly’.

**Geodatabase Annotation:**

ProductivitySuite also makes available a second text option, Geodatabase Annotation, which provides the text stored as rich annotation features in an Annotation Feature Class (AFC). The AFC stores rich text information as real database elements complete with pre-specified font and formatting. The difference is that the text elements are not rendered 'on the fly', but rather are permanently stored as graphic text elements. Please note that AFC cannot by used by ArcIMS.

**Optimising for Performance:**

Choosing the text option to employ has important consequences for performance. Enabling the Geodatabase Annotation is more expensive in terms of database volume (about 6% more database space required for OS MasterMap Topography Layer) and processing performance.

To save users the labour of setting up symbolisation properties for the linear Cartographic Text feature class, ESRI (UK) have created the Cartographic Text Renderer (CTR). CTR is a free utility, downloadable from the ESRI (UK) website (www.myesriuk.com) which, when installed, provides almost identical text rendering as would be possible with the Geodatabase Annotation.

For performance reasons, ESRI (UK) recommend choosing the Cartographic Text option only for most users. The Geodatabase Annotation option would be a useful option for those interested in editing or manipulating individual text features after loading, but most users will achieve the same cartographic benefit with the CTR.
4.4.4. Archiving

ProductivitySuite provides archiving as an option when applying change-only update (COU). When COU is applied, Archiving will retire the obsolete features to separate Feature Classes, for retrieval later.

Archiving is a powerful tool for viewing and analysing historical data, however implementing it does have performance and database size costs. Particularly, the speed of applying COU with ProductivitySuite will be dramatically slowed by exercising archiving.

Before implementing archiving, it is worth spending a little time doing an analysis of its merits and drawbacks. The two key questions are:

- Will the archive actually be used? If no specific plans for use of the archive can be formulated, it may not make sense to implement it.
- How often will COU be applied? If COU is not applied regularly, there’s little use archiving.

If positive answers cannot be found for these two questions, then it can be assumed that the performance degradations to data load experienced by applying archiving are not outweighed by the benefits. Organisations that simply want an annual snapshot of the database might be better served by exporting and archiving the entire database on hard media than implementing the ‘live’ archiving of ProductivitySuite.

4.5. OS MasterMap-to-Go for Initial Load

Because the conversion of OS MasterMap Topography Layer GML files with ProductivitySuite for the whole of Great Britain would be a very involved and time-consuming task, ESRI (UK) make available the OS MasterMap-to-Go service as an alternative approach. This service provides ready-to-load ArcSDE 9.1 export files of national OS MasterMap. Sub-national or user-defined areas are planned and currently under development and testing.

The OS MasterMap-to-Go export is sourced from ESRI (UK)’s internal national OS MasterMap Topography Layer database, which is kept current according to the scheduled Ordnance Survey Managed GB supply (MGBS).

The latest schedule / calendar can be seen at:

www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/exisitingcustomer/mgsb.html

**Note:** ESRI (UK) will attempt to follow the calendar as closely as possible. It usually takes two to four weeks after the supply of the data before a suitable ArcSDE 9.1 export file is available to clients.

The data is then provided to the client as ArcSDE 9.1 export files, which can be loaded using the ArcSDE sdeimport command. A direct database import is much faster than converting the GML directly with ProductivitySuite. The ArcSDE export files can be loaded into any relational database that ArcSDE supports.
The OS *MasterMap-to-Go* service provides ArcSDE export files for the national set of OS MasterMap Topography Layer, and ArcSDE scripts that can be used to load the data. Exports are also planned for subsets of Great Britain, specified on an Ordnance Survey 100-km tile basis or user-defined polygon. This development was in progress at the time of writing this document.

### 4.5.1. The Seamless Approach

The OS MasterMap Topography Layer is designed to be seamless: the features represent ‘real world’ objects, such as buildings and land parcels, which have no respect for imposed tiles. This is a significant advantage over the Land-Line product in that features are no longer split when they happen to cross a tile boundary.

Management of a seamless national coverage continues to pose challenges beyond supply and delivery. National OS MasterMap consists of more than 450 million features. Even single tables, such as the Feature Class representing lines, contains more than 300 million records; and this presents significant challenges to hardware, software, and administration.

The approximate sizes of the ArcSDE 9.1 export files are shown in Figure 4.4.1-1.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boundary</td>
<td>240 MB</td>
</tr>
<tr>
<td>Line Annotation</td>
<td>6.0 GB</td>
</tr>
<tr>
<td>Points</td>
<td>1.2 GB</td>
</tr>
<tr>
<td>Symbology</td>
<td>840 MB</td>
</tr>
<tr>
<td>Lines</td>
<td>81 GB</td>
</tr>
<tr>
<td>Area</td>
<td>34 GB</td>
</tr>
</tbody>
</table>

**Figure 4.4.1-1:** *Table showing the approximate sizes of the ArcSDE 9.1 export files.*

The approximate sizes of the archive tables are shown in Figure 4.4.1-2. These archive tables are NOT supplied but are included for reference as they are a result of a successful application of change-only update.

<table>
<thead>
<tr>
<th>Archive Table</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS MMBNDarch</td>
<td>34 MB</td>
</tr>
<tr>
<td>OS MMANNOarch</td>
<td>490 MB</td>
</tr>
<tr>
<td>OS MMPNTarch</td>
<td>114 MB</td>
</tr>
<tr>
<td>OS MMSYMMarch</td>
<td>93 MB</td>
</tr>
</tbody>
</table>
Figure 4.4.1-2: Table showing the approximate sizes of the archive tables.

4.5.2. Loading the OS MasterMap-to-Go Exports

ArcSDE 8.1.2 or higher

The OS MasterMap-to-Go data service also provides ArcSDE scripts for loading the database. These exports are compatible with ArcSDE versions 8.1.2 or higher.

Whilst for many customers this service is carried out on site by an ESRI (UK) consultant, it is useful to know how this is achieved and the impact it may have.

Loading the data consists of a number of stages. For each layer within each 100 km tile the process is as follows:

1. Delete existing Feature Class (if it exists), using `sdetable -o delete -t`
   This removes any existing data if reloading the data holding.

2. Import the layer using `sdeimport -o create -l`
   This will import the data into a new spatial layer (Feature Class), and will build the spatial indexes for the layer.

3. Create indexes on TOID and VERSION using `sdetable -o create_index`
   This creates the same indexes that are also created by ProductivitySuite, if the data were loaded using that application rather than from the bulk load process.

4. Analyse the data tables using `sdetable -o update_dbms_stats`
   This carries out a database analysis to improve performance for using the data in a GIS and for applying change-only update (COU).
5. Delete Existing change table using `sdetable -o delete`
   
   This will remove the existing Change History table associated with the Feature Class, prior to reloading the data holding.

6. Import the change history table using `sdeimport -o create -t`
   
   This imports the non spatial change history table for the layer.

7. Create a single index on the change history table, on TOID, Changedate, Changedesc using `sdetable -o create_index`
   
   This creates a combined index on these fields, which is vital for applying COU to your data holding.

8. Analyse the change history table using `sdetable -o update_dbms_stats`
   
   To improve performance in use and for COUs of the history table.
5. Enterprise Database Configuration

The space required to store and process OS MasterMap Topography Layer data in a relational database with ArcSDE will depend on many factors, such as the database used, the extent and block sizes used, and the tablespace management strategy to name just a few. Detailed examination of database configuration and sizing is outside the scope of this document, but readers should consult ESRI’s System Design Strategies white paper (available at support.esri.com). The following are general guidelines, which an administrator should consider and research further if necessary.

5.1. Tablespace Sizing

When configuring a completely new database server, the amount of disk space required will largely depend on which disk management process is adopted, for example RAID (Redundant Array of Independent Disks). Because of the redundancy offered by some RAID configurations, double the normal disk capacity may be required.

For national GB OS MasterMap Topography Layer coverage, the space required will start at several hundred Gigabytes, and grow from there with archiving. The following guidelines can be used as general minima when sizing a database for national OS MasterMap Topography Layer coverage:

- Oracle 8/9: 500 GB of table space
- Oracle Spatial: 395 GB of table space
- SQL Server: 330 GB of table space
- DB2: 500 GB of table space

For system sizing guidelines for less than GB National coverage refer to Section 4.3: Database Performance.

5.2. Load Space Requirements

It should be borne in mind that the guidelines above refer to just the space required for storing the data. Space is also required for the data loading process. The national set of ArcSDE export files is approximately 250-GB in size, and there should be enough disk space available on the network to store these files on disk whilst loading the data.

In addition to disk space for storing the export files, the loading process will require adequate temporary table space within the database. Logging should be disabled during load for maximum performance. Typically, this may represent up to another 50-GB of storage space required for a national load. Once the data is loaded, it is good practice to take a backup of the database, which again will require sufficient space.
5.3. Archiving Space Requirements

The space requirements suggested here are just for the initial load of the data. If a customer plans to apply change-only update to the data holding and to enable archiving of the old data, then space must be allocated for this growth.

The Ordnance Survey has advised that typically there will be about 5-10% change in the database per year; however this may vary quite considerably. Currently, owing in part to the Ordnance Survey Positional Accuracy Improvement Programme, this growth has been nearer to 20%.

Even if all this available space is not required immediately, it is important to design a system that is capable of being enlarged at a later date. Ordnance Survey are adding new layers to OS MasterMap, which will increase the amount of space required when accommodating those. For example, the OS MasterMap Imagery layer will require a significant amount of disk space for national coverage. The OS MasterMap Integrated Transport Network (ITN) and Address layers are integral to OS MasterMap and there are plans for more layers to come. OS MasterMap ITN Layer, Address Layer and Address Layer 2 are fully supported by ProductivitySuite.
6. Using the OS MasterMap Topography Layer Repository

6.1. Using a “Seamless” Database

One of the main benefits of using ESRI technology for loading OS MasterMap Topography Layer data is that at the end of the process the result is a Geodatabase, whether Personal or ArcSDE, which is instantly ready to be used by GIS clients.

6.1.1. ArcGIS Desktop and ArcGIS Server

The ArcGIS Desktop applications ArcMap and ArcCatalog are suited to immediate use of the OS MasterMap Topography Layer Geodatabase using core functionality applicable to any valid data source. Once converted, this database becomes just another data resource to the clients. The OS MasterMap layers become Feature Classes in the Geodatabase, which can be previewed and managed in ArcCatalog; and analysed in ArcMap using all of the software tools available. The metadata including information on the date and order numbers of applied change-only updates can be viewed on the ArcCatalog metadata tab.

6.1.2. ArcIMS

ArcIMS is the web server technology ideally suited to serving the data across the Internet or an Intranet. For data held in an ArcSDE Geodatabase, the OS MasterMap layers can be accessed and rendered directly using the ArcIMS administration utilities for quick start-up. ESRI (UK) makes available upon request a sample ArcIMS ‘AXL’ file for map rendering in an OS MasterMap web client.

As an alternative, or for organisations holding the data in Personal Geodatabases, ArcIMS ArcMap Server can be used.

6.1.3. Using the Data in ArcMap

The converted OS MasterMap Topography Layer data can be used easily in ArcMap. Two ways of achieving this would be either having the data within an ArcIMS map service, or by utilising the group layers functionality of ArcMap.
Using a Map Service:

If ArcMap is being used in an environment in which ArcIMS is also available, then the simplest method for accessing the data is to connect to an ArcIMS image map service. Doing so will add a single map service layer into ArcMap. This service layer will have up to six sub layers, representing the six spatial layer types within OS MasterMap topography.

Figure 6.1.3.1-1: Using an ArcIMS Service in ArcMap.

Another benefit of using this method is that the rendering information and the scale thresholds can be configured by the ArcIMS administrator just once for a map service, and then these are applied automatically to all users accessing the data. Setting scale thresholds, for example, would prevent an individual user from attempting to display the entire dataset at once, which would put an unacceptable load onto the server. ArcIMS is the recommended approach for a user of National coverage.

Using Group Layers:

ArcMap supports the use of Group Layers within the Table of Contents (TOC). Group Layers are simply single TOC entries that represent more physical layers. For example, each of the OSAREA layers could be grouped into one single layer in the TOC called “Area”. The single group layer can be used to switch on and off all the layers at once.
Figure 6.1.3.2-1: Using Group Layers in ArcMap.

Group layers are best used when looking at regional quantities of data. When using a national set, these can sometimes be slow when drawing the data in a particular file.
7. Updating OS MasterMap
   Topography Layer Holdings

7.1. OS MasterMap Update Mechanism

One of the key innovations of OS MasterMap is the data update mechanism supported by the Ordnance Survey. Update and change to the dataset is provided through a programme called change-only update (COU). A delivery of COU to a customer consists of only the features that have been edited, deleted, or added by OS since the last supply.

7.2. Ordering Change-only Update

To obtain change-only update (COU) from the Ordnance Survey it is necessary to place the request online through the Ordnance Survey online service. This is the same procedure whether the initial load of OS MasterMap data was provided directly by the Ordnance Survey as GML files and loaded with ProductivitySuite, or whether it was loaded using the OS MasterMap-to-Go service.

For an explanation of the online order process see the Ordnance Survey COU web page for details:

www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/information/technical/changeonlyupdate.html

When making orders for national datasets, remember that this is a very large dataset, which takes time for the OS to process.

**Note:** Customers of the OS MasterMap-to-Go should choose 5km x 5km chunk size when defining the COU order.

**Important:** Check and understand dates thoroughly when ordering COU as if a later date than required is given, this may result in a gap of missing update data, the only fix for which is a complete reload! See the Ordnance Survey FAQ web pages for further explanation: “What date should I use when ordering COU?”

www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/faqs/service020.html

ESRI customers of the OS MasterMap-to-Go service should consult the date on their export delivery.
7.3. Data Delivery Format

Change-only update data is supplied in the same format as initial loads of OS MasterMap data: Chunks of GML data, typically 5 km square. They are usually divided into directories for each 100 km square. Each directory will contain up to 400 Compressed GML files (*.gz), which is one for each 5 km square within the tile.

Figure 7.3-1: Directory of .gz files.

Each file is named to represent the origin of the Chunk, so for example: 2936-TQ5040-5c34.gz is the Chunk with changes in the 5 km square whose lower left coordinate is 550000, 140000. This naming convention allows ease of file sorting for processing into the correct tables in the database.

7.4. Processing Change-only Update with ProductivitySuite

ProductivitySuite can be used to process the change-only update (COU) data provided by the Ordnance Survey in GML format, regardless of whether the original OS MasterMap data was loaded from GML or from OS MasterMap-to-Go. (Due to the complexity of the COU process, it is not possible for ESRI (UK) to provide an OS MasterMap-to-Go service that only supplies change.) Features from the database have to be added, modified or deleted during this process.
This section gives a general outline of how to process COU using ProductivitySuite, but for detailed instructions please consult the ProductivitySuite Help provided with the software.

### 7.4.1. Database Considerations for Change-only Update

ProductivitySuite accomplishes much of the work of change-only update (COU) by creating temporary database tables, table joins, and bulk record copies; all of which creates a sudden but temporary need for increased temporary table space. Furthermore, the greater the time-span of the COU (because a longer stretch between COUs results in more features per COU), the greater these new resource needs because more data is being processed.

Because of this, it is important that the database and system resources have enough extra capacity to handle the load created by COU. It is not possible at this time to supply generalised figures in this white paper. ESRI (UK) recommend that users contact them for consultancy advice to analyse their particular situation.

**Tip:** If the database does not have significant resources, ProductivitySuite may appear to hang when in fact the database may require the intervention of a database administrator.

### 7.4.2. Beginning Data Conversion

Begin processing the change-only update by running the OS MasterMap Converter command from the ProductivitySuite program menu or the Data Conversion Toolbar in ArcCatalog. This is the same command that is used to perform the initial load, so the key here is to select, on the Source Data step of the Wizard, that this is a change-only update.
7.4.3. Archiving

When specifying that the conversion is for processing change-only update (COU), the user also has the opportunity to enable or disable archiving. Enabling archiving means that departed and edited features are written to the archive Feature Classes in the Geodatabase (see Section 7.5), which is important for recreating historical views, or for doing historical or land-change analysis.

If archiving is specified, the user is given the option to minimise the load on the database in terms of log space and other tuning parameters. This option breaks down very large SQL queries into more manageable chunks for the database but it means that the SQL operations will take longer to execute. It is recommended when processing large areas (e.g., an area greater than a 100 km area) unless the database has been set up with enough temporary resources.

Tip: Enabling archiving will impact COU processing performance. If it is determined that access to historical data is not required, the speed of COU application can be improved significantly by disabling archiving, but doing so will mean that copies of changed features will not be retained.

7.4.4. Processing .GZ files

Because of the way ProductivitySuite processes the change-only update (COU) files, COU must be processed for the entire area-of-interest at the same time. For example, a typical Local Authority would need to apply COU to the database for the entire authority’s coverage area. Therefore, select to process all of the .gz files received from the wizard dialogue.
These batches also correspond to the two-letter ‘prefix’, e.g., OSAREA, apparent on the Feature Classes. National 100 km tiled databases are considered within this context to be a series of smaller 100 km tile-sized databases.

**Note:** ProductivitySuite is expecting specific Feature Class names to be within the Geodatabase. For example, that Topography Area features are kept in a Feature Class with “AREA” as the suffix. It is important, therefore, that these names not be changed in the Geodatabase.

**Note:** If for whatever reason the application of COU fails partway through, it will be necessary to re-apply the COU from the beginning. In the case of a 100 km –tiled coverage, it will only be necessary to restart from the beginning of the last 100 km tile area.

### 7.4.5. De-duplication

If .gz is taken in geographic chunks, there will likely be duplicates in the data between neighbouring Chunks for change-only update. There may also be duplicate records in the Change History table. ProductivitySuite handles all of this automatically by searching for and removing duplicates, as long as the ‘Check for and Remove Duplicates’ tick-box in the OS MasterMap Converter Wizard has not been disabled. Un-ticking this selection will speed up the loading process but will increase the size of the Change (CHNG) tables.

If non-geographic chunks are taken, de-duplication is not necessary, and this choice can be un-ticked.

### 7.4.6. Spatial Layers

ProductivitySuite gives the option of selecting or deselecting feature types for processing in change-only update (COU). For COU, ProductivitySuite will automatically detect which Feature Classes are already present in the Geodatabase. It is quicker to update all the feature types at the same time, but this is not a requirement. It is also possible to load or apply change to the ITN and Address Layers (if available) at the same time, or, these too can be processed separately.

There are two types of text supported by ProductivitySuite, ‘cartographic’ and ‘GDB’. Both can produce similar end results, so many may choose not to implement GDB Annotation (Cartographic Text is default and must be chosen if any text will be stored). See [Section 4.4.3](#) for more information on annotation and on this decision.

For customers of the OS MasterMap-to-Go service, the “Generate GDB Annotation” tick-box can be ignored because this spatial layer is not supported for the To Go service.
For users of ArcSDE Geodatabases, there will also be an option to enter an ArcSDE configuration keyword, which is an option for dictating database properties of the data load.

![Figure 7.4.6-1](image_url)

**Figure 7.4.6-1:** Selecting Spatial Layers in the OS MasterMap Converter wizard.

### 7.5. How ESRI (UK) Software Updates the Database

The underlying technical process that ProductivitySuite uses for updating the database with change-only update is complex, however this section attempts to describe this at a high level.

#### 7.5.1. Features in the Change-only Update Files

Change-only update data files contain the following types of feature, specified in the GML. These do not appear in the GML file in any particular order:

**New Features**: features that do not currently exist in the database and are to be added. These are supplied complete with all the attributes and geometry needed to add a new feature to the database.

**Updated Features**: like new features, the geometry and attributes are included in the GML, however they will have the same TOID as an existing feature and a higher Version number. The Version number may have incremented by more than one, because the feature may have undergone several changes since the last time the user took supply, or even within the Ordnance Survey edit cycles.

**Departed Features**: features that have 'departed’ (i.e., are to be deleted) from a particular Chunk of the database will normally be coded as 'deleted’. With geographic chunks, it is possible for a feature to leave one chunk and move to another; in this
case the GML attribute in this case specifies that the feature is ‘vacated’. Departed features of either type are provided only as a TOID; no version number, attributes, geometry or feature type is supplied.

7.5.2. The Process

If the Archiving option is selected the first time ProductivitySuite is used to apply change-only update (COU), an archive Feature Class is created for each layer that is to be updated; for example, ‘OSAreaArch’, ‘OSLineArch’, and ‘OSPntArch’.

ProductivitySuite fundamentally makes two ‘passes’ through the .gz files. The first pass reads all features which will be updated or deleted. These TOIDs are read into a large temporary table which is joined to the ‘live’ Feature Classes. On the basis of this join, features are copied to the archive Feature Classes in bulk, or deleted if no archiving is chosen.

**Important:** This temporary table can contain very many records – millions, potentially, if the area-of-interest is large enough and the time since the last COU application long enough. Join this table to the live feature class, also containing millions of records, and it is apparent that there is a tremendous load placed on the database. It is important for the system administrator to ensure that enough resource is available on the database server to accommodate this process.

In a second pass, new features and the new versions of updated features are read and placed into the live Feature Classes, and records are appended to the Change History tables. (Even new features have ‘change history’ information provided in the GML. This is because a feature may have undergone several changes at the OS before being made available publicly.)

Once all the spatial features have been processed, the final step is to remove duplicates from the change history table. A SQL query is performed on each change history table to do this. (It is important for speed of performance that the correct indexes are built on the table.)

7.5.3. The Result

Once change-only update (COU) has been run on the database for the first time, the database schema as described in Section 4, will be slightly modified. Running COU adds six Archive Feature Classes to the database.

Example set of Archive Features Classes for the OS area:

- OSANNOARCH
- OSAREAARCH
- OSBNDARCH
- OSLINEARCH
- OSPNTPARCH
These have an identical schema to the ‘main’ Feature Classes, but with the addition of three extra fields:

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Field Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>REASON</td>
<td>VARCHAR 15</td>
<td>The reason that the feature was added to the archive layer, either ‘Delete’ or ‘Edit’.</td>
</tr>
<tr>
<td>ORDERNO</td>
<td>VARCHAR 10</td>
<td>The COU order number that resulted in the feature being modified and archived.</td>
</tr>
<tr>
<td>CoUDATE</td>
<td>DATE/VARCHAR 10</td>
<td>The date that the COU was applied. This is not the same as the Version date provided in the data.</td>
</tr>
</tbody>
</table>

These tables contain all features that have been modified in any way by the COU process. This includes attribute edits, geometry edits, and deletions. Initially these tables may be quite small, but over time they will grow to contain all the features modified by the Ordnance Survey. Theoretically, they could grow to match the size of the base layers, if the Ordnance Survey eventually edited every feature.

Once COU has been applied, the base spatial layers should be identical to those produced if a complete load was applied on the same date.

7.5.4. Understanding the ProductivitySuite OS MasterMap Converter Log Files.

The ProductivitySuite OS MasterMap Converter produces a very detailed log file that describes the process of applying change-only update (COU). The log files are an important informational and diagnostic resource. The log outlines the processing steps that the ProductivitySuite OS MasterMap Converter takes, provides file size statistics, states numbers of features processed, and lists error messages.

For every COU file there will be a list of entries in the following structure*:

22/11/2004 11:38:38 Info: Converting into Personal Geodatabase: C:\TEMP\test.mdb

22/11/2004 11:38:38 Info: The following feature types will be loaded:

List of the feature classes included in the COU, such as Topological Areas etc.

22/11/2004 11:38:48 Info: Processing Change-only Update

This load is a COU.
22/11/2004 11:38:48 Info: Font Century is not present: Arial will be used instead

An alternative font will be used if the designated one is not available

22/11/2004 11:38:48 Info: no record of this order being processed previously

Checks that this COU has not previously been run

22/11/2004 11:38:51 Info: Processing Topographic Data - All Feature Types requested


22/11/2004 11:38:51 Info: Starting Change-only Update Archiving and Departed feature processing

Starting COU process

22/11/2004 11:38:52 Info: Populating Archive Tables

    Creating table of features to archive

22/11/2004 11:38:52 Info: Starting Archiving of features from OSLine

    Processing the OSLine feature class for updated followed by departed features

22/11/2004 11:38:52 Info: Copying features from OSLine to OSLineArch

    Populating the archive table


22/11/2004 11:38:55 Info: Deleting departed features from OSLine

22/11/2004 11:38:55 Info: Deleting features that will be updated from OSLine

22/11/2004 11:38:55 Info: Completed deleting features

(Log entries above repeated for each Feature Class)

22/11/2004 11:39:00 Info: Archiving and departed feature handling now complete

    (Logged when all FC’s processed)

22/11/2004 11:39:00 Info: Processing file: C:\TEMP\5192-SU6152-2l1.gz

path name of the GZ file being processed

22/11/2004 11:39:00 Info: Compressed file size is 132734 Decompressing...

Size in bytes of the GZ file before decompression
22/11/2004 11:39:00 Info: File decompressed successfully
   File must be decompressed into working folder before being processed

22/11/2004 11:39:03 Info: Flushing cursors ....

   Processing of features from GML now complete - flushing cached features to database

22/11/2004 11:39:05 Info: Completed processing current file (GML file size: 1518112)

   Processing of this GZ file now complete - size in bytes of the decompressed GZ file

22/11/2004 11:39:05 Info: Number of features processed so far:

   Number of features processed since processing all files began (cumulative)

22/11/2004 11:39:05 Info: 480 Areas, 630 Lines, 1 Boundaries, 0 Points, 37 Annotations, 0 Symbols

Note: A further selection of Info message descriptions can be found in the appendix as there are too many to include here.

This log file represents a successful conversion of the data. There may be other messages that describe any complications found in processing the data. These are not errors and do not effect the quality of the converted data. They are included for information and will be prefixed "Info:"

Examples of the information logs are included in Appendix A.

7.5.5. Troubleshooting Performance Issues

ProductivitySuite does some very complex and database-intensive work. ProductivitySuite can only complete its tasks if the running and database environments are properly tuned and configured.

If at any time the ProductivitySuite OS MasterMap Converter appears to ‘hang’, the log files should be checked to see what stage the software was at in processing. If it is in the middle of processing a file, check that the memory available on the client machine has not been exceeded, leading to excessive paging. This can happen when a very large GZ file is being processed.

**Tip:** As a rough guideline, if using 5 km geographic Chunks of supply, at least 1-GB of memory should be available on the loading client machine.

If the log file indicates that the ProductivitySuite OS MasterMap Converter appears to have hung whilst doing a de-duplication of feature classes or tables, check to confirm...
that there is activity on the server. This process can under normal conditions take
over an hour for a large feature class such OSLINE. For change-only update (COU)
performance, it is very important to ensure that correct indexes are built on the
Feature Classes. There should be a column index on each of the TOID and Version
fields. Note also that the Change History tables need to be indexed on all 3 fields,
(TOID, ChangeDate and ChangeDesc). In normal operation ProductivitySuite will
create all of the required indexes at the time that it first creates the Feature Classes.

**Important:** If any part of the system crashes while loading COU, it is
necessary to reprocess the entire COU order again from the beginning.
This precaution is necessary to make sure that departed features are
correctly handled. It will be quicker the second time, as those features
already applied will not need to be reinserted or moved to the archive.

7.6. **Managing the COU Process Logistics**

Change-only update (COU) is a technical process but also requires considerable
organisational planning, with regard to the resources and time needed to process the
data. Depending on the geographic extent of the database and the elapsed time
since the last application of COU, the update may take a considerable time. These
considerations are particularly critical for customers of OS MasterMap Topography
Layer national coverage.

7.6.1. **How Often Should the Data Be Updated?**

This depends on a number of factors, not least being the customer’s business
requirements. The more often the database is updated the quicker the update
process will be, as less change will have occurred in a smaller time interval. Although
the amount of overall change that occurs in a year is the same regardless of how
often the change is applied; it is simply a question of choosing a large number of
small applications, or just a few large ones.

A chief advantage of frequent small applications of change-only update (COU) is that
the database is never very far out of date, and if archiving, then the quality of
historical data is richer; whereas applying change less often means that potentially
many millions of real-world and Ordnance Survey changes will have taken place
before the GIS user is aware of them.

Again, this may be more critical to some users than to others. Keeping the database
very current requires constant administrative effort in the regular applications of
COU. Allowing more time between applications lessens this diligence, but means that
larger quantities of data and longer-running processes will be required to bring the
database up to date.

ESRI (UK) do not make a specific recommendation to customers on the frequency of
change application as there are many factors to consider. Moreover, Ordnance
Survey now follow a practice of scheduled and published database refreshes on an
approximately six-weekly basis. Broad experience seems to indicate that application of COU approximately every 6-12 weeks meets the needs of most users.

**Tip:** Visit [www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/existingcustomer/refreshdates.html](http://www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/existingcustomer/refreshdates.html) for the Ordnance Survey’s OS MasterMap release schedule.

Once the update frequency stretches to six months or more, the time needed to apply the updates with ProductivitySuite begins to approach that for processing a complete reload (this is because of the extra work that is done by the software for COU). Also, the longer the time between updates, the less granular the ability to view historical data through the archive.

### 7.6.2. Important COU Performance Considerations

Application of change-only update (COU) may not be appropriate for all OS MasterMap Topography Layer customers. The primary advantages of the COU system are the efficiency of data delivery (i.e., only what is new in the database) and the ability to archive old data.

A question every organisation should ask regarding archiving is, “how will we actually use the archived data?” If there is no clear answer to this, it may be worth considering not choosing the archiving option. To put it differently, archiving has COU application performance, database, and administrative ‘costs’, so a compelling argument should be made for it before that route is taken.

Another consideration is whether the Geodatabase Annotation option should be chosen. As discussed in [Section 4.4.3](#), there are two annotation options for the ProductivitySuite OS MasterMap Converter: Cartographic and Geodatabase Annotation. Choosing the latter provides rich database text objects, but increases the processing time for applying COU considerably. Moreover, the free ESRI (UK) Cartographic Text Renderer affords a graphical display of the Cartographic Text which very closely approximates that of the Geodatabase Annotation.

### 7.6.3. Updating the Database for National Coverage Customers

Applying change-only update (COU) is only one of two strategies for keeping an OS MasterMap database up to date, particularly for national coverage customers. These strategies are outlined below:

#### Apply COU using ESRI (UK) Software:

This process has been outlined earlier in this section, and will be considered by most to be the assumed method. Some of the particular benefits of this approach are:

- Existing users can carry on using the data while it’s being updated
Archive layers may be maintained to enable the viewing of historical changes in the data. There is more flexibility in the frequency with which you can update the data. The disadvantage of this approach, however, is that the process of update can consume a large amount of time and resource, particularly if the interim period between updates has been long.

Reload the data from scratch:

Reloading the database with a new set of OS MasterMap-to-Go exports can be used as another strategy for refreshing a data holding with more up-to-date data. ESRI (UK) maintains an up-to-date database with regular application of COU, and this can be used to provide a refresh of the data holding.

The benefit of this approach is that ordering and applying COU is not necessary. In addition, running OS MasterMap-to-Go exports continuously, can reload the database more quickly than COU covering a long period.

This approach does mean that the data can only be updated when a new OS MasterMap-to-Go export is released by ESRI (UK). These exports are not able to provide the Archive data layers, so no comparison between new and old datasets will be possible. Finally, reloading with exports involves deleting and recreating each layer in turn, which means that while a layer is being reloaded it is not available for use by GIS clients.

7.6.4. Updating a Personal Geodatabase Repository

OS MasterMap customers implementing the Personal Geodatabase (PGDB) have fewer logistical and management concerns when it comes to updating the database, because the data quantity is much lower.

The chief disadvantage of applying COU to a PGDB is that archiving swells the database over time, and PGDBs have a structural 2-Gigabyte size limit. Therefore, the database will run out of space eventually.

An alternative strategy is to avoid COU altogether and simply reload a fresh PGDB from scratch, at the same time that the older one is saved and ‘retired’. Although this does not use the inherent archiving functionality of ProductivitySuite OS MasterMap Converter, viewing historical data is still possible simply by viewing one of the older PGDBs that will have been saved.

This decision should be taken by considering the volume of the data holding, the anticipated rate of change from the OS, and the true archiving needs of the organisation.

7.6.5. Maintaining a Single OS MasterMap Database

For Organisations implementing a single database, the COU update process is a good option even though it can take longer than reloading, because it uses fewer resources and does not lock down the layers. It is possible to apply COU to a dataset while it is still being used for read-only access by an organisation’s GIS clients.
Running COU typically uses between 5-25% of a server’s capacity, allowing existing users to continue with a slightly reduced performance.

In order to reduce the amount of time taken to apply COU, it is possible to run several concurrent* COU processes from different client machines. This obviously puts a greater strain on the database server, reducing the performance for other users of the system, but can be a good strategy for increasing the speed of the update process.

Before beginning the COU process, it is important to make sure you have a backup of the database. This might be a complete database backup, which could be used to restore the entire database, or the creation of a set of sdeexport files for the layers to allow them to be restored relatively quickly.

7.6.6. Maintaining Dual OS MasterMap Databases

The best technical solution to the logistical challenges of change-only update (COU) is to maintain two separate OS MasterMap databases. This has obvious safety advantages for mission-critical environments in data redundancy. Maintaining two databases allows one to be updated with ProductivitySuite, while the other is in use as the ‘live’ environment. When one server has been updated, it can be switched over to serve as the production server while the other database is being updated.

Either the second database can be updated using ProductivitySuite again, or the first database can be backed up and attached to the second database. This ensures that the second database is created more quickly, and that a backup of the new database exists at all times.

It may also be beneficial to do an sdeexport of the updated layers, allowing these to be restored separately if necessary.
8. Validating the OS MasterMap Holding

8.1. Feature Validation Dataset

The Feature Validation Dataset (FVDS) enables OS MasterMap users to verify their data holding. FVDS is a set of files that can optionally be obtained from Ordnance Survey with an OS MasterMap order. It is used to ensure the data holding contains the correct set of features after loading the data with which it was supplied. FVDS cannot be ordered separately from the OS MasterMap data. It can be used after a full load, or after change-only update (COU) has been applied to the holding. FVDS can be ordered with the Topography Layer (independent polygon format only), Address or ITN Layers. FVDS itself is divided into files on a non-geographic basis but it can be used with data supplied in either geographic or non-geographic chunk file options.

Data orders with FVDS would normally take longer for Ordnance Survey to process than data orders without FVDS. It is anticipated that customer may want to periodically check their holdings are correct but may or may not want to take FVDS every time an order is placed.

Further information about FVDS can be found in the OS MasterMap user guide (www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/userguides).

8.2. Validation with ProductivitySuite

ESRI (UK)’s ProductivitySuite can be used to run checks on the OS MasterMap holding after converting the data with which the FVDS was supplied. After loading or applying change-only update with the OS MasterMap Converter, use the OS MasterMap Feature Validation tool (Figure 8-1) with the FVDS to verify that the related Feature Classes contain the correct set of features. The reports produced by the tool provide a list of TOIDs with the corresponding version number when discrepancies are found between the FVDS files and the related Feature Classes. The reports, which can be viewed on-screen, printed or saved, are split into:

- Features held in the relevant Feature Classes that are not in the FVDS files, and;
- Features in the FVDS files that are not held in the relevant Feature Classes.
The OS MasterMap Feature Validation tool can be started from the ProductivitySuite program menu or the Productivity Data Conversion Toolbar in ArcCatalog. For details on how to use the tool, readers should consult the ProductivitySuite Help, which describes the process in further detail.
9. **Summary**

OS MasterMap provides more than just a base map; it is a seamless, topological and integrated suite of geographic layers, with a delivery mechanism that ensures that customers have access to the most up to date geographic information. OS MasterMap provides polygonal topography, addresses, integrated transport networks, and imagery. Data is supplied in GML format and the updates contain only changed features since the previous update was ordered.

Among the special characteristics of OS MasterMap is the unique reference on every feature known as the TOID. The TOID can be used to link features across layers and to track feature life cycle, format and method of delivery. The TOID and change-only update (COU) mechanism can together be leveraged-upon to create an overall organisational GI change management strategy.

ESRI (UK) provide two choices for loading OS MasterMap. ProductivitySuite extension provides an intuitive OS MasterMap Converter Wizard for loading data, applying change, and archiving features. OS MasterMap-to-Go is a unique service that supplies ready-to-use ArcSDE exports for quicker implementation of national coverage databases.

Either way it is important for users to understand the challenges and opportunities that OS MasterMap offers. OS MasterMap Topography Layer will create much larger databases than Land-Line users are accustomed to, because of the feature richness and database archiving possibilities. This will require proper forethought to be given to database choice, configuration and management. ESRI (UK) provides a variety of free tools and resources to help with this.

Ordnance Survey COU process is a flexible and efficient way to keep databases current; however implementation requires proper understanding and management. Considerations such as how frequently change should be applied, and whether archiving should be implemented, are important for processing and database performance.

The goal for all users of OS MasterMap Topography Layer will be to leverage this new powerful GIS database in a production environment. ArcGIS and ArcIMS from ESRI are the ideal clients to consume and analyse the data in a local or an enterprise environment. The Geodatabase (whether Personal or ArcSDE) is an ideal intelligent storage format.

Ordnance Survey will continue to develop the OS MasterMap layers. Consultations underway between Ordnance Survey and its partners regarding potential data enhancements and improvements will continue. ESRI (UK) will ensure that its applications and services take full advantage of these changes as and when required.
10. Appendix A – OS MasterMap Converter Log Definitions

10.1. Information Logs

Info: Taking layers out of load only mode: <feature type>

This message occurs when performing an initial data load into an ArcSDE Geodatabase. The feature classes are put into a special mode that makes writing large volumes of data to them faster. The message is logged for each feature class that is written to, and indicates that the feature class is being put back into normal mode ready for use.

Info: Line with toid <toid> not found in <feature type> layer

TOID to be updated in COU was not found in original data.

Info: The Geodatabase does not support versioning

Information obtained from the Geodatabase.

Info: Removing duplicates from Change History tables

This message indicates that the Change History tables are being stripped of duplicate entries caused by the load or COU process. (If OS MasterMap Converter appears to hang at this point then check database tuning and configuration).

Info: Removed Duplicates From <feature type> Change History

This message is logged for each Change History table that is de-duplicated.

Info: Completed Removal of duplicates from all layers

This message indicates that the process of removing duplicates from all feature classes has been completed.

Info: New <feature type> found with toid <toid> and version <ver number> was deleted by previous COU

Departed features are archived and deleted in the first pass. If the second pass detects that one of those features is actually an edit rather than a true departure, then it will log this message followed by:

Info: Archived Feature has been successfully marked as edited rather than deleted
This message is logged immediately after the previous one. It indicates that the fact that the feature has been edited (and not truly departed) has been noted in the Archive Feature Class.

**Info: Area with toid <toid>: <error description>**

When some polygon geometries are built there may be issues with the direction of the rings, in terms of correctly identifying which one is the outer boundary and which are holes falling within it. If this occurs then the above message will be logged followed by:

**Info: Fixed Area with toid <toid> and successfully committed**

This message indicates that the problem described for the previous message has been rectified successfully and the corrected feature has been successfully written to the Geodatabase.

## 10.2. Warning Logs

In addition to the messages marked "Info:“, some messages are marked "Warning:“.

**Warning: Area with toid: <toid> requires line with toid: <toid> which could not be found in the data or in the existing FeatureClass**

Refer to Info message above concerning “TOID to be updated in COU not found” (Warning can occur with COU if TOID to be updated is not found).

**Warning: Timed out decompression process - file will be skipped**

This message indicates that the attempt to decompress the gzip (.gz) file has failed and processing will attempt to continue with the next file. If this problem occurs delete everything in the Working folder, check there is enough disk space to hold the particular .gz file and its decompressed representation, and then try reprocessing. In addition, it may be useful to use the Windows Task Manager to check for any gzip.exe processes that may have hung. This situation is extremely rare.

**Warning: Timed out decompression but attempting to proceed with processing file...**

If decompression has timed out, it is sometimes the case that the .gz file has still been decompressed successfully. The software will try to find the decompressed file and proceed with it.

**Warning: Unable to copy the file for decompression**

If there is insufficient disk space available in the Working folder, or if the user running ProductivitySuite OS MasterMap Converter does not have sufficient permissions to write files to the working folder then this message may be logged.

**Warning: **** Processing cancelled by user on completion of this file ******
This message indicates that the user clicked the cancel button while processing was in progress.

**Warning: *** PROCESSING ABORTED BECAUSE MORE THAN n DATABASE INSERTS FAILED IN A ROW ***

If there is a serious problem communicating with the database then this message is logged. It usually indicates that a network or database communication failure has occurred during processing.

**Warning: De-duplication was skipped**

This message is logged after the error message "ERROR: Did not process all files". It is logged only in the case of processing change-only update, and occurs when database or network communications fail and the software detects that a large number of database inserts have consequently failed. See also the description of message "ERROR: Did not process all files" in section 9.3 Error Logs.

**Warning: Failed to flush cursors: <error message>**

This message indicates a breakdown in communication with the database.

**Warning: Failed to update Archived Feature to show status as Edited rather than Deleted**

*(ref to Info message above concerning departed features being re-marked as edited)* This message is logged if an attempt to adjust the archive entry of a feature from "Delete" to "Edit" fails.

---

### 10.3. Error Logs

There are a number of messages that may be logged with "ERROR:" against them, and where possible they provide additional information of use in resolving the problem.

**ERROR: Unable to prepare decompression mechanism - no gz files (if any) will be processed**

**ERROR: Note that application install path appears correct and contains the decompression tool**

**ERROR: and the temp working folder also exists. This error could be due to write access permission**

**ERROR: issues with the temp working folder specified**

**ERROR: Application folder was "<path>**

**ERROR: Working folder was <path>**
The series of messages above indicate that the installed copy of 'gzip.exe' could not be found in the ProductivitySuite install folder, or that it could not be copied to the working folder for use during the decompression of gzip (gz) files. If the copy command failed then this may be due to a lack of disk space or user permission issues.

**ERROR: Unable to create objects from the Microsoft Scripting Component**

**ERROR: Check that SCRRUN.DLL is present and registered in the System32 folder on this machine**

**ERROR: Aborting Change Only Update**

These messages indicate that the Microsoft component used by ProductivitySuite to manipulate the file system is not present or is not registered. This could be resolved by manually locating and registering, or installing and registering the DLL indicated. Alternatively, it may be rectified by re-running the ProductivitySuite installation software.

**ERROR: Unable to open workspaces using the connection information provided**

**ERROR: Aborting processing - please check connection information is correct and try again**

These messages indicated that it has not been possible to activate the connection to an ArcSDE database. An example of where this may occur is when processing a profile saved out to a ProductivitySuite OS MasterMap Converter profile file.

**ERROR: Did not process all files**

This occurs when the software detects that database inserts have failed - which in turn indicates that the communication with the database or the database itself may have failed. In the case of a personal Geodatabase, it may mean that the database has reached its size limit of 2Gb.

**ERROR: During Processing - details follow**

This is a general error message logged in the unlikely event that an unanticipated issue arises during preparation or processing of OS MasterMap. Processing will normally attempt to continue after having logged the error, so a series of these may occur in the log file if something unusual has happened. The problem may be related to missing, unregistered or multiple registered components of the software. Depending on the messages logged it may be appropriate to uninstall and then re-install ProductivitySuite.

**ERROR: ITN Progress Monitor Initialisation Failed: <error code>**

This message indicates a problem initialising a component called mmProgressMonitor.exe. If this error is logged then the component is installed and registered correctly but the configuration to report processing progress has failed.

**ERROR: ITN Progress Monitor Attach to ITN Writer Failed: <error code>**
This message indicates a problem attaching to the process that handles the writing of ITN data (mmITNWriter.exe) to the Progress Monitor. The Progress Monitor enables the number of features processed to appear on the ProductivitySuite OS MasterMap Converter Wizard. This issue should not normally occur. If it does it may indicate the need to uninstall, clean the registry and re-install.

ERROR: ITN Initialisation Failed: <error description>

This message points out a general error when initialising the components responsible for ITN data processing. It should not normally occur and will prevent any ITN data from being processed in the input dataset.

ERROR: while shutting down ITN: <error description>

This message indicates a general error when shutting down the ITN components. It should not normally occur but may not be very serious. Unless other messages are logged indicating problems preparing to process ITN or actually processing the data, then it is likely that ITN data will have been processed successfully and there will be no serious consequences from this error having occurred.

ERROR: Failed to build the fields structure for the <FCName> feature class

This message indicates that the software has not been able to prepare the field information for use in creating the indicated Feature Class. The message is fatal and actual data processing will not be started.

ERROR: Failed to create the <FCName> feature class

The indicated Feature Class does not exist and cannot be created. It is fatal and processing preparation will be aborted at this point.

ERROR: Failed to create an insert cursor on the <FCName> feature class

This error means that the indicated feature class has been created or opened, but that it was not then possible to create a cursor on it. This could be due to database corruption, database communication failure or user permissions.

ERROR: Failed to create the <FCName> feature buffer

This error indicates that it was possible to create or open the Feature Class but a buffer for inserting into it could not be created. In the case of an ArcSDE Geodatabase, this may be a user permissions issue with the database login user used to make the ArcSDE connection to the database.

ERROR: Failed to set Load Only Mode for <FCName>. Continuing ...

It was not possible to put the indicated feature class into load-only mode. This can only occur for initial data loading and not for Change Only-Update. Processing will continue but may be slower than normal.

ERROR: SetUp<feature type> <error description> - details follow
This error indicates a general problem initialising and connecting to the Feature Class. This will be an error other than one of those described above.

**ERROR: creating the CoUDate & OrderNo fields in the ArchiveFeature classes**

The Archive Feature Classes are created by taking the field specification of the main Feature Classes and adding two fields (CouDate and OrderNo). If there is a problem adding these fields then this message is logged and the software will not proceed with the change-only update.

**ERROR: Unable to create/attach to the <FCName> dataset**

**ERROR: Aborting process (note workspace connection was successful - layer creation/attachment failed**

**ERROR: (Where appropriate check that any output disk paths have sufficient free space and user has write access**

  Failed to create or open the indicated Feature Class.

If there is a problem creating or opening a Feature Class then the software will create a separate diagnostic log file in the computer's temp folder for the current user. Under the 'temp' folder there will be a folder called "log files" containing an individual log file for each Feature Class created or opened in the last run of ProductivitySuite OS MasterMap Converter (whether or not the create/open failed). This file contains detailed information about what went wrong and may include specific database error messages that can be useful in diagnosing a problem.

**ERROR: Can't create cursor to delete feature with Toid: <toid> <error description>**

**ERROR: Can’t obtain feature from cursor in order to delete it. Toid: <toid>**

**ERROR: <error description> deleting <toid> from <FCName>**

**ERROR: (Error from delete using cursor was <error description>)**

Combinations of the above messages indicate that the software has been unable to create an update cursor on one of the Feature Classes. This may be due to database permissions or availability problems.

**ERROR: updating symbol collection: <error description>**

This indicates that the software has not been able to update the Geodatabase Annotation Features symbol collection. It could be a database permissions issue.

**ERROR: Unable to create an update feature cursor on the <feature type> feature class. Toid = <toid>**

**ERROR: Unable to retrieve features from cursor on the <FCName> feature class. Toid is <toid>**
These messages occur during change-only update when a database issue has meant that the feature to be updated could not be retrieved from the database.

**ERROR:** Failed to delete old version of `<FCName>` with toid `<toid>` Error is `<error description>`

**ERROR:** (Error from cursor delete was `<error description>`)  
This message indicates an error deleting old versions of features that need updating. It could be a database locking or permissions issue.

**ERROR:** Failed trying to create an update cursor on the archive feature class for `<Feature Type>` Features. Toid = `<toid>`

**ERROR:** `<error description>`

**ERROR:** Unable to obtain a feature from the cursor on the archive feature class for `<Feature Type>` Features. Toid = `<toid>`

These above messages point out that it has not been possible to access the indicated Archive Feature Class. It could be a database permissions or locking issue.

**ERROR:** `<Feature Type>` with toid `<toid>`

This message indicates that insertion of a feature into the indicated Feature Class has failed. The related database message will also be logged.

**ERROR:** (while trying to obtain a line feature from the layer. Toid=<`toid`>)

**ERROR:** `[<error description>]`  
There was a problem retrieving a line from the line Feature Class. This occurs where line features required to build a polygon must be retrieved from the existing data holding.

**ERROR:** Failed to create temporary working table for setting Line in Polygon Flags

**ERROR:** `<error description>`

**ERROR:** Unable to create a rowbuffer on Temp table while setting Line in Polygon Flags

**ERROR:** `<error description>`

**ERROR:** Can’t insert into Temp table while processing line in polygon flags - toid=<`toid`>

**ERROR:** `[<error description>]`  
**ERROR:** Can’t execute SQL for setting line in polygon flags on line feature class
ERROR: [<error description>]

ERROR: Can't remove the temporary table for setting line in polygon flags

These messages indicate that an attempt to create, open, access or clear a temporary table in the database has failed, possibly due to database permissions issues. The result is that processing will continue, but the line-in-polygon field will not be populated on the line Feature Class.

ERROR: Could not create annotation feature: [<error description>]

ERROR: Cartographic Text Toid is <toid>

ERROR: Couldn’t calculate size and scaling of text for annotation feature: [<error description>]

ERROR: Cartographic Text Toid is <toid>

ERROR: The text is [<text>]

These messages indicate that it was not possible to create, store or prepare a Geodatabase Annotation Feature. The reason may be database failure or permissions issues associated with the GDB Annotation Feature Class.

ERROR: newtextrecord : Annotation Class not found

ERROR: [<error description>]

The GDB Annotation feature class (**AnFC) is being reported as not supporting annotation. This could indicate an issue with the creation and configuration of the annotation Feature Class itself. It may be appropriate to check that the database connection is not using ArcSDE Direct connect and that the GDB Annotation Feature Class works and is recognised normally in ArcCatalog/ArcMap.

Failed to process GDB Annotation for text with toid <toid>

This message indicates that it was not possible to insert a prepared Geodatabase Annotation Feature into the GDB Annotation Feature Class. This could be due to problems with the creation/definition of the feature class, problems with the ArcSDE database connection (if appropriate), or it could be related to a permissions or file locking issue.

Error 6001: Cannot Create INITIAL extent

"Cannot Create INITIAL extent" is a database error indicating that the database does not have space to create the new table in the Tablespace.

The Tablespace is determined by the ArcSDE keyword specified on the ‘Features to Process’ step of the OS MasterMap Converter wizard. If none was specified then everything is going to the Default Tablespace. If a keyword was specified then it is necessary to look in the DBTune settings on ArcSDE to see which Tablespace the keyword directs ArcSDE to use.
The Tablespace needs to be increased in size or set to automatically grow (the latter may slow down data loading of large volumes of data since the database keeps having to pause to increase the size of the Tablespace). In practice, making the Tablespace bigger and then setting it to grow automatically beyond that may make sense.

If the Tablespace is large enough already then it may be on a disk that has run out of space - in which case additional space on another disk can be used, or existing data can be removed from the Tablespace to make more room on the disk (if there is anything obsolete in the TableSpace).

**ERROR: Cannot output data to an SDE Database with ArcEngine**

This error may also occur because an ArcEngine licence is accidentally obtained instead of the available ArcGIS Desktop licences from the licence server. This problem may be rectified by installing ProductivitySuite 1.3.1 or later. It can also be avoided by stopping the licence server from serving ArcEngine licences.

**ERROR: Invalid layer configuration keyword**

This indicates that an invalid ArcSDE tuning keyword was entered on the `Features to Process` step of the OS MasterMap Converter wizard.

Ensure that the ArcSDE tuning keyword exists in the DBTUNE settings.
11. Appendix B - Correcting Discrepancies in Data

If discrepancies were identified by the OS MasterMap Feature Validation Tool, please refer to the following advice on possible causes and solutions.

**FVDS and data were from different orders**

*Cause:*
- The validation was applied to data that was not from the same order in which the FVDS was supplied.
  - To check the OS MasterMap order number for the FVDS, refer to the first set of digits in the FVDS file name. To check the order numbers and load dates of the original load of OS MasterMap or subsequent COUs applied to that data, refer to the ISO metadata of the relevant Feature Class using ArcCatalog. The order information is listed under 'Resource Identification Information: Supplemental Information' section of the ISO metadata.

*Solution:*
- Rerun the validation on the correct FVDS and data holding.

**All features were deliberately not loaded**

*Cause:*
- If you have chosen not to load all the features supplied in your order, then as the FVDS lists all the TOIDs in the specified area, the TOIDs in the Feature Classes and the FVDS will not match.

*Solution:*
- If you have chosen not to load all the features supplied in your order, the feature validation cannot be used.

**Processing errors**

*Cause:*
- There were errors reported during processing.
  - To check if any errors were reported during the initial data load or subsequent COUs, review the OS MasterMap Converter log files. Unless otherwise specified in ProductivitySuite OS MasterMap Converter wizard, the log files are stored in the Log Files folder of the installation directory, which is typically C:/Program Files/ArcGIS. The OS MasterMap Converter log file is named: <date> OSMasterMapConverter Log.txt. If MapManager was used for the initial load or previous COUs, MapManager log files should also be reviewed. The default location for MapManager log files is C:\Program Files\ESRIUK\MapManager9\log files. MapManager log files are named: <date> MapManager9 Log.txt.

*Solution:*
-
Some of the common warning, info and error logs are explained in the Appendix A. Support information about errors can be found on the ESRI (UK) website - www.esriuk.com. Alternatively, contact ESRI (UK) Technical Support.

COU missing or applied wrongly

Cause:

- A change-only update (COU) ordered from Ordnance Survey has not been applied.
- A COU, which has been applied to the data, was ordered from wrong start date.
- COUs have been applied in the wrong order.

Note: It may be difficult to identify if one of the above COU issues has occurred. However, checking your own records of Ordnance Survey orders, as well as the metadata associated with the data repository, may identify some of these issues. The OS website has an FAQ section on placing COU orders, including CD pictures explaining the dates printed on COU CDs to help ensure that there will be no gap in customer data holdings when requesting updates. Follow the link below: www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/faqs/service020.html.

To check the order of the original load of OS MasterMap or subsequent COUs applied to the data repository, refer to the ISO metadata of the relevant Feature Class using ArcCatalog. The order information is listed under 'Resource Identification Information: Supplemental Information' section of the ISO metadata.

Solution:

- Order a new data supply from Ordnance Survey and reload the data; or
- Order a bulk load from ESRI (UK) OS MasterMap-to-Go service.

Duplicate Features in the data

Cause:

- ‘Duplicates’ and other issues may occur if the COU is applied with few files at a time. The COU should be applied in a single execution.

Note: The Duplicate Feature Remover tool may not remove all the features that appear to be duplicates because it only removes true duplicates and not features that have the same TOID but different version numbers.

Solution:

- If the issue has occurred with the last COU applied:
  - Data load the smallest COU file as if it was a data-supply file (at the end of this a full remove-duplicates process, including old versions, will be executed); or
  - Rerun the COU (all files at once) and, on the Profile Options step, select 'Process All Files Again as Part of Database Restoration'; or
– Restore the database and re-run the COU (all files together).

• If the issue has occurred not with the last COU but with a preceding COU:
  – Order a new data supply from Ordnance Survey and reload the data; or
  – Order a bulk load from ESRI (UK) OS MasterMap-to-Go service.

**Broken area features not loaded**

**Cause:**

• In ProductivitySuite V1.0, OS MasterMap Converter does not load broken area features. This requirement was addressed with ProductivitySuite V1.0 Hotfix 3. If the data was loaded using ProductivitySuite V1.0 without Hotfix 3 applied, it is possible that broken area features may be missing from the data load.

**Solution:**

• Reload the data with ProductivitySuite V1.1 or above (or ProductivitySuite V1.0 with Hotfix 3 applied).

Information about the latest hotfixes and releases can be found on the ESRI (UK) website - [www.esriuk.com](http://www.esriuk.com).

Contact ESRI (UK) Technical Support if further information is required about installing hotfixes and releases.

End of document