

## Implementation Issues for OS MasterMap

### Contents

1	Introduction .....	1
2	Land-Line.Plus versus OS MasterMap: Key Issues .....	1
3	Main Options for Using OS MasterMap .....	3
4	Software .....	3
5	Infrastructure .....	4
6	Training and Documentation .....	4
7	Future Developments .....	4

## 1 Introduction

The following document discusses some of the issues associated with the introduction and use of OS MasterMap data within Digimap's Ordnance Survey Data Collection.

The implementation of OS MasterMap will have an impact in the following areas:

1. hardware - such as disk space availability, network bandwidth
2. software - such as appropriate versions of GIS and database programs
3. training - to enable users to fully exploit the data,
4. documentation - to make use of the data and to understand what it is.

This document aims to provide some reasonable estimates and guidance on the costs and issues of implementing OS MasterMap. This information is relevant both for existing subscribers, who will want to know what additional resources will be required to implement MasterMap, and also for new subscribers who will want to know what resources they will need to implement and take up the service from July 2007 onwards.

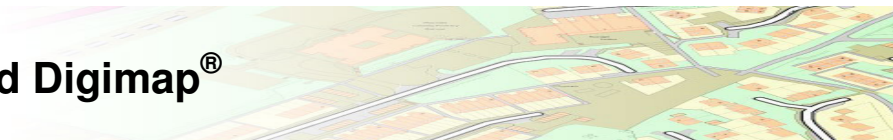
## 2 Land-Line.Plus versus OS MasterMap: Key Issues

There are a wide range of options available when dealing with OS MasterMap; from using the data like Land-Line.Plus data as flat files, to full database solutions which store and manage the data centrally. Both these options are valid (and so are numerous options in between) depending on the circumstances.

The key differences between dealing with Land-Line.Plus and dealing with OS MasterMap are:

- 1) **Volume of data** – it is estimated that **OS MasterMap data is between 3 and 5 times larger** than the equivalent Land-Line.Plus data once converted into a 'standard' data format (such as Shapefile or TAB) or imported into a Geodatabase (see Table 1).

# OS MasterMap<sup>®</sup> and Digimap<sup>®</sup>



There is also at least a 5 times compression factor between compressed and uncompressed GML.

Source Data	Compressed (.zip)	Uncomp .	MapInfo (MIF/MID)	Shape <sup>1</sup>	Shape <sup>2</sup>	ESRI Geodatabase
MasterMap GML (Tile NN34+) <sup>3</sup>	5Mb	26.3Mb	21.5Mb	6.8Mb	17.5Mb	22.1Mb
Land-Line.Plus NTF (Tile NN34) <sup>4</sup>	0.7Mb	2Mb	4.8Mb	N/A	2.8Mb	4.8Mb
MasterMap GML (Tile SU6352+) <sup>3</sup>	1.1Mb	19.1Mb	5.8Mb	5.6Mb	16.1Mb	15.3Mb
Land-Line.Plus NTF (Tile SU6352) <sup>5</sup>	0.09Mb	1.3Mb	2Mb	N/A	3.9Mb	4.5Mb

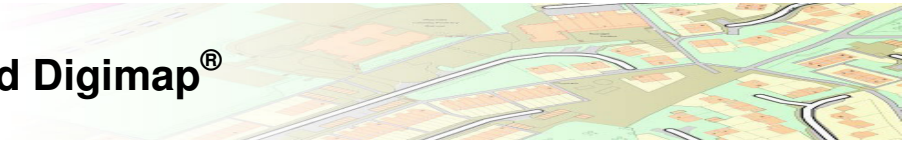
Table 1: Summary of File Sizes

Notes:

- (1) Converted to Shape using MapManager Geodatabase Filter
- (2) Converted to Shape using ArcToolbox Shapefile Export
- (3) The GML data covers a slightly larger area since it is 'hairy' i.e. objects continue outside rectangular tile boundary.
- (4) This is the equivalent of 4 x 1:10,000 tiles and 2 x 1:2500 tiles of rural area.
- (5) This is equivalent to 4 x 1:1250 urban tiles

- 2) **Data Management** – data management requirements differ considerably between Land-Line.Plus and OS MasterMap, particularly with respect to update mechanisms. OS MasterMap supports the application of Change Only Updates (COUs) to the data whereas Land-Line.Plus requires complete file replacement for updates. The concept of Change-Only Update provides a more efficient data storage solution.
- 3) **Data Content** – the content of OS MasterMap differs from Land-Line.Plus in terms of structure and classification of data. There are major differences in the way that OS MasterMap deals with polygons and every feature has a unique ID (called a TOID). OS MasterMap has its own set of feature codes that are different from Land-Line.Plus feature codes. These codes are derived from a combination of a feature's descriptive attributes. TOIDS can be used to associate other data with a particular topographic feature.
- 4) **Data Format** – OS MasterMap's native format is GML (Geography Markup Language). This is a fundamentally different format from the NTF (National Transfer Format) used by Land-Line.Plus.

## OS MasterMap® and Digimap®



- 5) **Data Structure** – OS MasterMap data is not based on a tile structure. The data is a seamless, continuous dataset of objects. This presents different challenges in identifying the data required.

### 3 Main Options for Using OS MasterMap

There are a number of ways to use and manage OS MasterMap data but there are three fundamental options available:

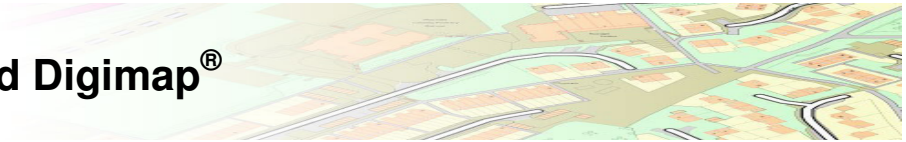
- 1) Use the raw GML data directly – some GIS applications can read GML and compressed GML (.gml and .gz) files directly without having to convert or otherwise ‘import’ the data. This is the simplest mechanism but least flexible as only some systems can read this data directly and few (if any?) can edit the data in this format.
- 2) Import the data and convert it to another ‘flat-file’ format – this can be a convenient way of dealing with relatively small amounts of OS MasterMap data initially. GML can be converted to another format (such as Shapefile or MIF/MID). However, this is not always a simple process and can take a significant amount of time depending on the volumes of data involved. It is generally not possible to update OS MasterMap data (using the COU mechanism) once it has been converted to flat files. Essentially, a full re-supply and the associated data conversion is required again.
- 3) The third mechanism for dealing with OS MasterMap is to use a database solution to manage a repository of data. This is necessary when dealing with large volumes of data or where use of COU is required. The choice of database storage depends largely on the client software to be used to access the data, but common solutions require some sort of SQL database such as Oracle, PostgreSQL, MS Access etc. The use of a relational database can bring into play a number of additional factors including installation, security, training, maintenance and support; generally speaking, the use of a relational database to support multiple users in a network would require the services of a Database Administrator.

### 4 Software

From experiments with the data, it appears that for simple cases of downloading and using small amounts of OS MasterMap, the two most common GIS software packages (ArcGIS and MapInfo) work well. For both ArcGIS and MapInfo there are tools available to convert and read the GML into each systems’ native formats (see the GIS Software Options document for more detailed information).

However, there are problems with the update mechanism in both systems unless a database solution such as the Personal Geodatabase or other relational database system is used. A Personal Geodatabase must be used with current ESRI products. It is not

## OS MasterMap<sup>®</sup> and Digimap<sup>®</sup>



possible to use OS MasterMap in ArcView 3.x products directly, without first converting to Shapefile. Older versions of MapManager, such as 6.2, do not support OS MasterMap.

The use of a database management system for storage is required if there is a need to manage Change Only Updates (COU), or to manage a large store of data centrally that is to be used as a reference dataset for multiple users.

If using ArcGIS 9.1, a 2Gb Personal Geodatabase should be sufficient for average users, who are unlikely to exceed this size (this equates to approximately 400 1:1250 Land-Line tiles) for class or project work. NOTE - future releases of ArcGIS will remove the size limitation on Personal Geodatabases.

There will be cases (e.g. for research) where downloading very large quantities of data is required. In these cases a relational database will be required and more sophisticated software is needed to process the data.

## 5 Infrastructure

Systems need to be of sufficient specification to run the chosen applications, including the client application software, server software and database (if required). There also needs to be sufficient storage and internal network capacity to handle the increased volume of data expected with MasterMap use. There will be a requirement for sufficient internet bandwidth to support the download of larger volumes of data. The examples and cases studies show that the current hardware requirements will probably impose some additional cost on institutions to ensure that more than the minimum requirements are met, however each institution will have to assess those themselves depending on their existing infrastructure.

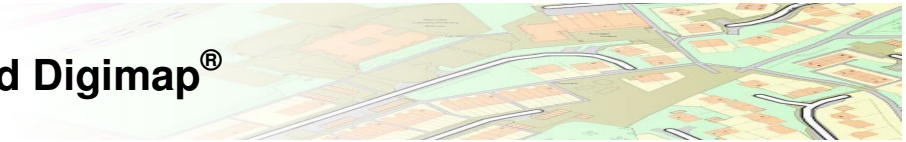
## 6 Training and Documentation

A key issue in the take up and use of OS MasterMap will be the requirement for training and familiarisation with the data. Additional training and documentation will be required for institutions taking up the use of OS MasterMap data. This additional training and documentation requirement will affect the cost to institutions of take-up. Training will be required throughout institutions from site reps to end-users in the use, management and potential applications of OS MasterMap. Additional documentation will be required for sites to enable them to make use of OS MasterMap efficiently and effectively.

## 7 Future Developments

The software available for handling OS MasterMap is in a state of ongoing development since it is still a relatively new data product. The following are some of the developments that have been announced or are likely to appear in the near future.

## OS MasterMap<sup>®</sup> and Digimap<sup>®</sup>



- ArcGIS 9.2 (due for release late 2006) will no longer need to use a Microsoft Access database to hold a Personal Geodatabase – this also removes the size limitation of 2Gb per database. This will make the use of Personal Geodatabases for OS MasterMap data more convenient. It is also claimed that this will improve performance with larger databases.
- The use of ESRI Geodatabase XML as a 'transfer format' could be an interesting development as it could enable users to manage a single Geodatabase and then transfer data to individual users. A compressed XML form of the database is about the same size as the original GML file zipped.