British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

# An Investigation into File Formats for the use and Delivery of Large Format Images

GeoHazard Programme

Internal Report IR/04/057

BRITISH GEOLOGICAL SURVEY

INTERNAL REPORT IR/04/057

# An Investigation into File Formats for the use and Delivery of Large Format Images

L Bateson, R McIntosh.



*Key words*

Compression, Images, ECW, MrSID, GEOJP2.

*Front cover*

GeoJP2 compressed Landsat ETM image of Loch Assynt, Scotland.

*Bibliographical reference*

BATESON, L, MCINTOSH, R. 2004. An Investigation into File Formats for the use and Delivery of Large Format Images. *British Geological Survey Internal Report*, IR/04/057. 22pp.



Keyworth, Nottingham British Geological Survey 2004

## BRITISH GEOLOGICAL SURVEY

The full range of Survey publications is available from the BGS Sales Desks at Nottingham and Edinburgh; see contact details below or shop online at www.thebgs.co.uk

The London Information Office maintains a reference collection of BGS publications including maps for consultation.

The Survey publishes an annual catalogue of its maps and other publications; this catalogue is available from any of the BGS Sales Desks.

*The British Geological Survey carries out the geological survey of Great Britain and Northern Ireland (the latter as an agency service for the government of Northern Ireland), and of the surrounding continental shelf, as well as its basic research projects. It also undertakes programmes of British technical aid in geology in developing countries as arranged by the Department for International Development and other agencies.*

*The British Geological Survey is a component body of the Natural Environment Research Council.*


**Keyworth, Nottingham NG12 5GG**

☎ 0115-936 3241 Fax 0115-936 3488
e-mail: sales@bgs.ac.uk
www.bgs.ac.uk
Shop online at: www.thebgs.co.uk

**Murchison House, West Mains Road, Edinburgh EH9 3LA**

☎ 0131-667 1000 Fax 0131-668 2683
e-mail: scotsales@bgs.ac.uk

**London Information Office at the Natural History Museum (Earth Galleries), Exhibition Road, South Kensington, London SW7 2DE**

☎ 020-7589 4090 Fax 020-7584 8270
☎ 020-7942 5344/45 email: bgslondon@bgs.ac.uk

**Forde House, Park Five Business Centre, Harrier Way, Sowton, Exeter, Devon EX2 7HU**

☎ 01392-445271 Fax 01392-445371

**Geological Survey of Northern Ireland, 20 College Gardens, Belfast BT9 6BS**

☎ 028-9066 6595 Fax 028-9066 2835

**Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB**

☎ 01491-838800 Fax 01491-692345

*Parent Body*

**Natural Environment Research Council, Polaris House, North Star Avenue, Swindon, Wiltshire SN2 1EU**

☎ 01793-411500 Fax 01793-411501
www.nerc.ac.uk

# Foreword

This report is the published product of a study by the British Geological Survey (BGS) looking into suitable storage formats for large images. The scan and shaft project was established in 2003 to scan all BGS maps into digital form. The number of scans necessitates some form of compression to attain a sensible file size and amount of disc space required. Image compression also has important ramifications for Remotely Sensed images and other GIS data, therefore the format chosen must be compatible, both in terms of image viewing and retention of georeferencing information, with the GIS or image analysis systems currently in use in BGS.

# Acknowledgements

The Authors would like to thank Walt Wiley, the former president of Mapping Science Inc, for a free trial of the GeoJP2 software and his help in using it.

# Contents

**FIGURES**



**TABLES**



# Summary

The Scan and Shaft Project working group on file formats was tasked to evaluate the new breed of wavelet compression-based file formats for use with the scanned maps and field slips scanned by the project, but also for the wide range of other large format image files held in BGS e.g. remote sensing imagery, scanned aerial photography and geological photographs.

Four main requirements were specified. The format must be compatible with GIS environments in use in BGS and needs to support georeferencing; the format must be compatible with existing plotting and printing processes and software in BGS or alternatively, provide an efficient mechanism for potting with the range of BGS plotters and printers; the format must allow the efficient delivery of the images from hierarchical file servers and using web based interface and finally the software to create the files must be capable of batch processing for automated conversion of large numbers of files.

A number of formats were considered and three chosen for detailed evaluation. They were Mr. SID, a well known and established solution from LizardTech. ECW, Enhanced Compression Wavelet, an established solution from ER Mapper. JPEG2000, a fairly new format that is open source and based on the ISO/IEC 15444-1 core coding system published as a full technical standard in late 2000. GeoJP2, a fully conforming version created by Mapping Science specifically for the mapping and GIS community.

Of the three formats, Mr. SID consistently produced the poorest images. It is incapable of compressing imagery of more than three bands – a major issue with remote sensing and airborne imagery, compression is much slower than the other two formats and pricing is by cartridge, £800 per 10Gb cartridge or £16,000 per Terrabyte cartridge. This effectively takes Mr. SID out of the evaluation.

Of the two remaining file formats, ECW and JPEG2000/GeoJP2, the tests indicate that they are very close. Both produce excellent results though for absolute quality of image ECW is very slightly ahead. In practice, a slightly lower compression and slight increase in resulting file size would provide excellent quality, making JPEG2000 equal to the ECW at higher compression. Both are capable of use in GIS packages within BGS, and both can be delivered using web technology. (See the Image Web servers at http://www.earthetc.com/ (ECW) and http://www.gcs-research.net/jp2imageviewer/ (JPEG2000/GeoJP2). Unfortunately a conflict on the PC used for testing meant that the ECW and JPEG2000 had to be tested for speed of compression on significantly different PCs; a dual processor 1.2 Ghz 512 Mb RAM versus a single processor 1 Ghz. 128 Mb RAM. JPEG2000 stood up very well despite this massive handicap especially as images were up to 300 Mb each. Other anecdotal evidence suggests they are on a par or that JPEG2000/GeoJP2 encoding is slightly ahead on conversion.

On cost, it is assumed that a commercial use is the norm, as most remote sensing imagery will be used for commercial projects, and the map images will be used for sales, enquiries and other commercial projects as well as science-funded projects.

To provide ECW images, ER Mapper needs to be purchased at £3,500. This will include the tools to develop web delivery. For the image encoder JPEG2000/GeoJP2 costs $1995 (single user, single CPU) and $475 for annual support including free updates. The JPEG2000/GeoJP2 Image server for web delivery comes out in September 2003 and costs $2300. On cost JPEG2000/GeoJP2 comes out slightly ahead.

On balance, on quality, speed of compression, delivery systems, and price, ECW and JPEG 2000/GeoJP2 are very close and the working group would have no qualms about recommending either. However, looking into the future, JPEG 2000/GeoJP2 does have a number of very supportive features. Firstly JPEG2000 is a true open source product, the format is an ISO standard (ISO/IEC 15444-1 core coding system published as a full technical standard in late 2000). As such it is increasingly being adopted as a native format in a wide range of products. This trend is set to continue. Significantly, encoded in the JPEG2000 format is the ability to define and produce true lossless compression, albeit at a lower compression ratio (usually between 2:1 to 5:1 depending on image type). Because of this, many major image archives are now looking at this format as a possible replacement to Tiff uncompressed or Tiff with lossless LZW compression as an archival file format for digital preservation of ‘master images’.

## Recommendations

Because of these very important features the working group has decided to recommend JPEG2000/GeoJP2 as the file format/compression to use for images within BGS.

For the Scan and Shaft project it recommended that a detailed study of optimum compression ratios for the different types of image original is undertaken, and either a single optimum ratio is chosen for the project, or alternatively different ratios are chosen depending on the original. If the latter, a strategy for the efficient batch conversion of the images must be determined.

# 1 Introduction

BGS holds a large number of digital image objects, some have a relatively small file sizes and pose few, if any problems in use and delivery. However, significant collections of digital images exist or will soon be created that are characterized by very large file sizes. They are on the whole, too large for use in many applications and certainly too large to deliver using web technology. A number of new file formats are now available that use Discrete Wavelet Transforms. They provide a very high degree of compression without loss of quality in the image some can hold georeferencing data in the format and used efficiently in GIS software and are also capable of web delivery, usually with the aid of a plug-in.

The Scan and Shaft project created a working group to investigate these file formats and to provide a recommendation for format that should be adopted for the project and for use in a BGS as a whole. This report provides the results of this study.

The main collections of large format images in BGS include:

- Remote sensing imagery
- Scanned aerial photography
- BGS 1:10,000 geological mapping (40,000 images planned to be captured 2003-2004)
- BGS Field slips 140,000 images planned to be captured 2003-2004)
- Geological photographs – currently 20,000 images and increasing rapidly.

The four main requirements have been specified:

1. The format must be compatible with GIS environments in use in BGS.
    - They need to support georeferencing
    - They should be easy to open (i.e. supported by the GIS either natively or via a plug-in)
    - The format must allow the use of GIS processes.
2. The format must be compatible with existing plotting and printing processes and software in BGS or alternatively, provide an efficient mechanism for potting with the range of BGS plotters and printers.
3. The format must allow the efficient delivery of the images.
    - Using web technology i.e. using a web based interface, allowing the user to search for and interactively display the imagery in a web browser.
    - User access to files from hierarchical file server.

4. The file format creation software must be capable of batch processing to enable the efficient automated conversion of large numbers of files.

# 2 Formats Investigated

A number of formats have been identified:

- MrSID (stand-alone image compression software).
- Jpeg2000 (either an off the shelf solution such as GeoJP2 or a system built in house using one of the development kits, based on the Jpeg2000 standard, such as Kakadu).
- ECW (bundled with ER Mapper, but can be used as a stand-alone format).
- ZEH.
- SpectrumITech.
- Filesafe

An initial sift reduced the selection to three main formats which were taken forward for detailed study and comparison. They are:

- **Mr.SID** – a very well known and established solution from LizardTech.
- **ECW** – Enhanced Compression Wavelet, an established solution from ER Mapper.
- **JPEG 2000** – A fairly new format that is open source based on the ISO/IEC 15444-1 core coding system published as a full technical standard in late 2000. We are testing a version of this standard, **GeoJP2**, created by Mapping Science specifically for the mapping and GIS community.

# 3 Compression Methodology

For the purposes of this trial four different types of imagery were tested:

- RGB satellite imagery (8 bit)
- Greyscale satellite imagery (8 bit)
- Satellite imagery with 6 or 14 layers
- Scanned maps

.

Sample images (table 3.1) were taken into each software package in turn and compressed by various different ratios. Several criteria were noted during the compression, these were:

- The chosen compression ratio
- Actual compression ratio achieved
- Time taken to compress, size of compressed image
- The quality of the compressed image
- The speed at which it displayed.

Compression was carried out on a PC platform running windows 2000 with 512 MB of RAM and duel AMD Athlon 1.2GHz processors.

| File name | File Description | File Size | File type and |
|---|---|---|---|
| Mull_pan | Landsat 15m resolution panchromatic image of Mull, Scotland | 517 855 Kb | Greyscale (8 bit) |
| H+I_mosaic | Landsat 6 band 30m resolution image of the highlands and Islands of Scotland | 570 889 Kb | Satellite image with 6 bands (8 bit). |
| Fmny02nec128687_001.tif | Scanned 1:10 000 geology field slip of the NY 02 NE sheet. | 228 792Kb | RGB (24 bit) |
| Fmny02nwc122267_001.tif | Scanned 1:10 000 geology map of the NY 02 NW sheet. | 79 375Kb | Greyscale (24 bit) |
| H+I_mosaic_3band | Landsat 3 band 30m resolution image of the highlands and Islands of Scotland | 312 509Kb | RGB (24 bit) |

Table 3.1: Sample Imagery (before compression).

## 3.1 ECW METHODOLOGY – DETAIL

Compression is accessed and controlled through a compression wizard accessed through the ER Mapper front end. The compression allows for 3 input types: from an active window, from a file or the batch compression of multiple images. For this test images were selected from a file. For each image 3 different compression ratios were tried, with the ratio depending on the image type; greyscale images cannot be compressed by too much therefore ratios of 15, 12 and 10 to 1 were tried. For the multiband and RGB images, which can be compressed more, ratios of 40, 25 and 15 to 1 were tried.

Once the compression had finished a pop-up dialogue appeared giving information, such as time taken and actual ratio used, about the process. This feature was found to be very useful for the purpose of the trial.

Some of the imagery was previously processed in the ERDAS IMAGINE Image analysis system. Since the ECW compression software cannot directly read this format it was necessary to import the imagery in to ER Mapper. ER Mapper is another image processing suite, which comes bundled with the ECW compression routines. ECW compression can directly read Tiff files; there was no need to import files of this type.

### 3.2 MRSID METHODOLOGY – DETAIL

MrSID is capable of directly reading both ERDAS .img and the tiff formats; therefore processed imagery was taken directly into MrSID, the desired ratio chosen and the output image defined. Unlike ECW compression MrSID does not provide a summary of the process therefore it had to be timed using a watch. The same ratios, as used for ECW were tested for MrSID.

### 3.3 JPEG 2000 METHODOLOGY – DETAIL

The Mapping Science GeoJP2 encoder is, at present, a command line based process. The command used looks something like:

*msiencoder -i G:\compression\images\small_scan.tif –o*
*G:\compression\images\small_scan_25.jp2 –cr 25 –tiff*

This command takes the tiff image called ‘small_scan’ and compresses it by a 25:1 ratio and then writes out a new image called ‘small_scan_25.jp2’.

File types the are not directly supported by the software (see later section) are imported into the programme using the *–import* command, for example to compress a image by a ratio of 25:1, which is in a Imagine .img format the following is used:

*Msiencoder –I G:\compression\images\h+i_mosaic.img –o*
*G:\compression\images\h+I_mosaic_25.img –cr 25 –import*

Mapping Science includes a free GeoJP2 viewer with their compression software, however this is basic. To make the visual comparison fairer images compressed to the .jp2 format were opened in ERDAS Imagine with use of a free plug-in from Mapping Science.

It should be noted that the GeoJP2 software could not be trailed on the same machine as the other two formats. This was due to a confliction with the license server for

another piece of software. As a result GeoJP2 was trailed on a slower machine: Intel Pentium 3 1GHZ with 128Mb of RAM. Time to compress in table 4.1 may therefore not be directly comparable to times for MrSID and ECW.

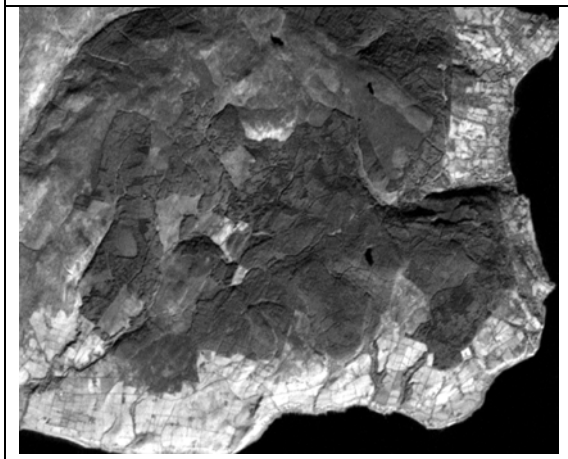
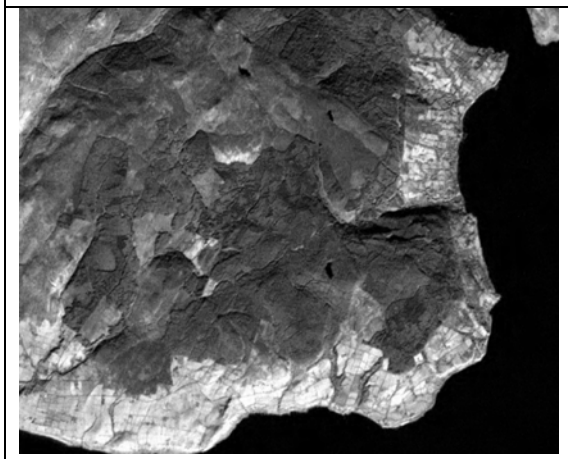
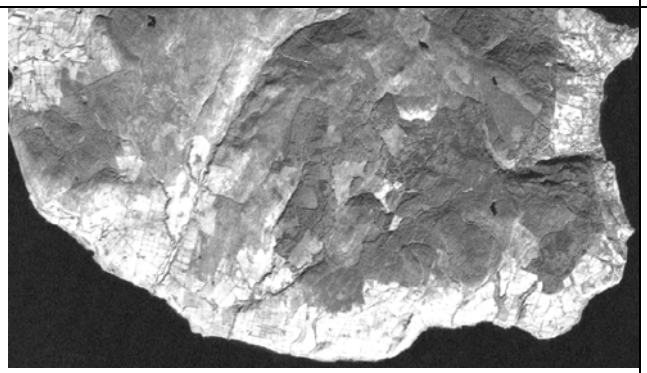
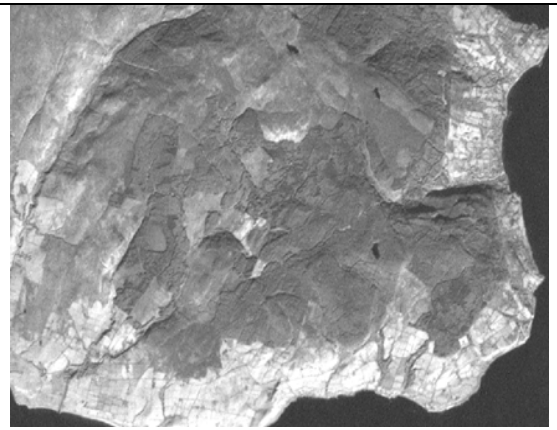
# 4 Results

| File and size | Chosen ratio | Actual Ratio | Time | Compressed size | Quality (good = v close to normal Acceptable = useable Poor = unusable) | Display speed |
|---|---|---|---|---|---|---|
| Mull panchromatic satellite image (264 143Kb) (Figure 1) | 10:1 | 12.4:1 | 46sec | 20.8Mb | Good | Fast |
| | 12:1 | 14.3:1 | 42sec | 18Mb | Good | Fast |
| | 15:1 | 17.6:1 | 41sec | 14.6Mb | Accept 50k, poor 25k | Fast |
| | 10:1 | 13.6:1 | 2m 18s | 19 435Kb | Good 50k, accept 25k | Slow |
| | 12:1 | 18:1 | 3m 10s | 14 595Kb | Accept 50k, poor 25k | Fast |
| | 15:1 | 23.5:1 | 3m 5s | 11 213Kb | Accept 50k, poor 25k | Fast |
| | 10:1 | 9.95:1 | 2min | 26 548Kb | Good | OK |
| | 12:1 | 11.93:1 | 1m 56s | 22 147Kb | Good | OK |
| | 15:1 | 14.89:1 | 1m 50s | 17 745Kb | Acceptable | OK |
| H+I mosaic (3 band) (285 455Kb) (312 509Kb) (312 509Kb) Figure 2 | 15:1 | 16.1:1 | 43sec | 17.3Mb | Good | Fast |
| | 25:1 | 24.7:1 | 44sec | 11.3Mb | good 100k, poor 50K | Fast |
| | 40:1 | 38:1 | 38sec | 7.3Mb | Accept 100k, poor 50k | Fast |
| | 15:1 | 24.8:1 | 3 min | 12 587Kb | Good 100k, accept 50k | Slow |
| | 25:1 | 41.7:1 | 2 min | 7493Kb | Good 100k, poor 50k | Slow |
| | 40:1 | 64.5:1 | 2.5min | 4 843Kb | Ok 100k, poor 50k | Slow |
| | 15:1 | 16.28:1 | 3m 20s | 19 193Kb | Good | Slow |
| | 25:1 | 27:1 | 3min | 11 577Kb | Good | OK |
| | 40:1 | 42.85:1 | 2m 55s | 7 293Kb | Poor | OK |
| H+I mosaic (6 band) (570 889) (MrSID version can't | 15:1 | 12:1 | 55sec | 23.3Mb | Good | Fast |
| | 25:1 | 18.8:1 | 51sec | 14.8Mb | Good | Fast |
| | 40:1 | 30.1:1 | 52sec | 9.2Mb | Good 150k, poor 100k | Fast |
| | 15:1 | 14.87:1 | 15m 19s | 38 383Kb | Good | OK |
| | 25:1 | 24.66:1 | 16m 55s | 23 153Kb | Acceptable | OK |
| | 40:1 | 39.14:1 | 7m 25s | 14 585Kb | Good 100k, poor 50k | OK |

| compress > 3 bands) | | | | | | |
|---|---|---|---|---|---|---|
| Fmny02nwc122267_001.tif (small_scan) (79 375Kb) Figure 3 | 25:1 | 9.7:1 | 37sec | 8Mb | Good | Fast |
| | 40:1 | 13.3:1 | 34sec | 5.8Mb | Good | Fast |
| | 60:1 | 19.2:1 | 34sec | 4Mb | Acceptable | Fast |
| | 25:1 | 39.2:1 | 1min | 2 026Kb | Poor | Fast |
| | 40:1 | 52.5:1 | 50sec | 1 513Kb | Poor | Fast |
| | 60:1 | 90.9:1 | 48sec | 873Kb | Poor | Fast |
| | 25:1 | 24.65:1 | 40sec | 3 220Kb | Good | OK |
| | 40:1 | 39.12:1 | 33sec | 2 029Kb | Good | Fast |
| | 60:1 | 58.07:1 | 30sec | 1 367Kb | Poor | Fast |
| Fmny02nec128687_001.tif (big_scan) (228 792Kb) Figure 4 | 25:1 | 27.3:1 | 36sec | 8.2Mb | Good | Fast |
| | 40:1 | 37.0:1 | 32sec | 6Mb | Good | Fast |
| | 60:1 | 49.6:1 | 31sec | 4.5Mb | Acceptable | Fast |
| | 25:1 | 35.5:1 | 2m 25s | 6 441Kb | Good | Slow |
| | 40:1 | 66.9:1 | 2m 20s | 3 418Kb | Acceptable | Slow |
| | 60:1 | 91.2:1 | 2m 18s | 2 508Kb | Just acceptable | Slow |
| | 25:1 | 24.64:1 | 1m 42s | 9 285Kb | Good | Slow |
| | 40:1 | 39.1:1 | 1m 36s | 5 851Kb | Acceptable | OK |
| | 60:1 | 58.07:1 | 1m 35s | 3 940Kb | Just acceptable | OK |

Table 4.1: Comparison of the results of the different compression formats. *Note:* in the table Times New Roman font refers to ECW compression, Comic Sans MS to MrSID compression and Arial to Jpeg2000 compression

Figure 4.1: Visual comparison of original Mull panchromatic Landsat image and its compressed versions.

| Original | | |
|---|---|---|
|  | | |
| ECW 10:1 | GeoJP2 10:1 | MrSID 10:1 |
|  |  |  |

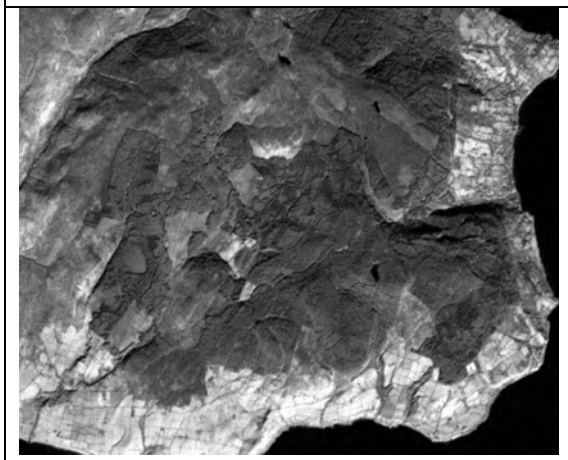
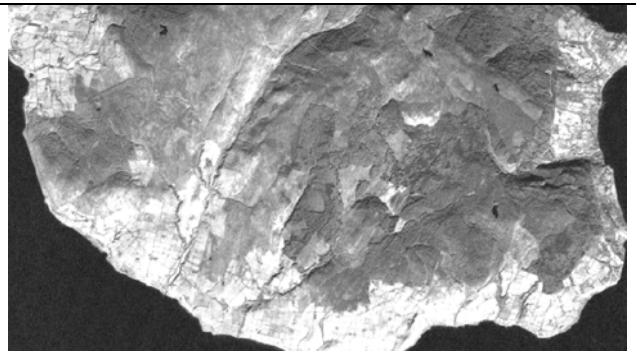
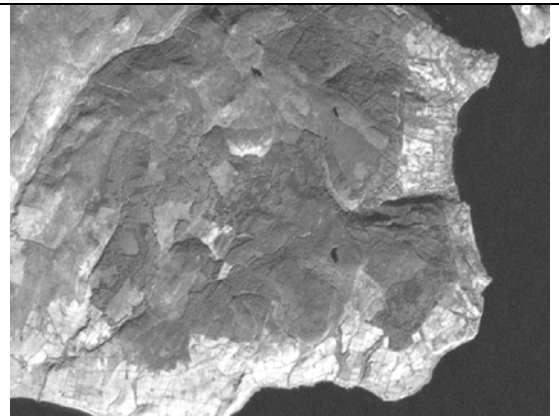
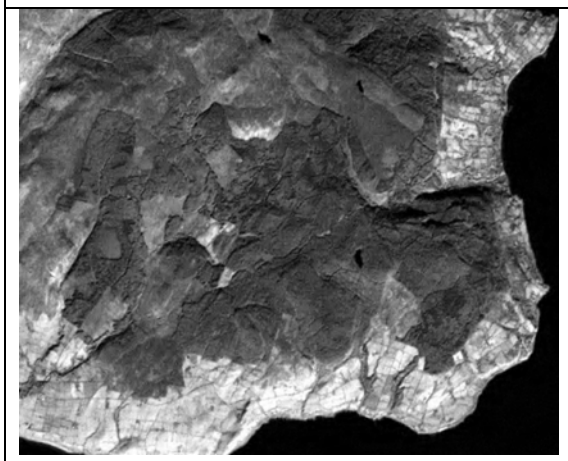
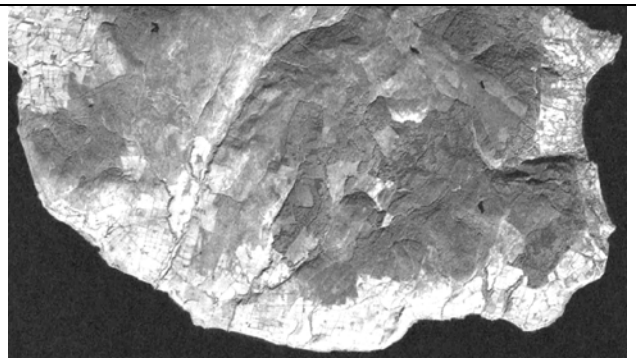
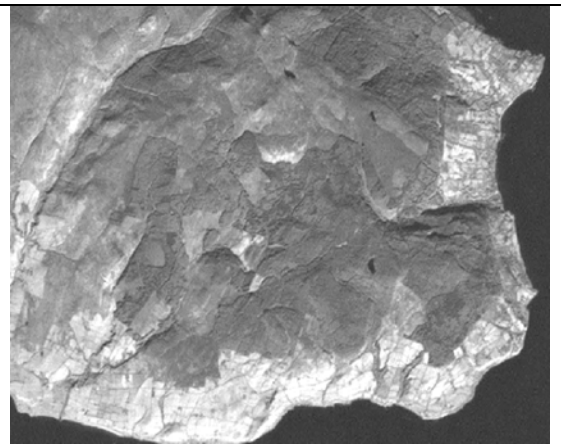
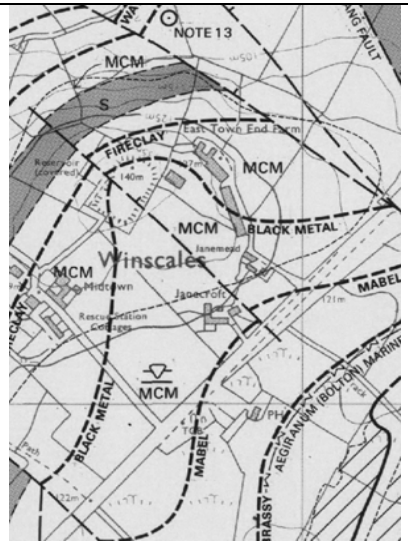
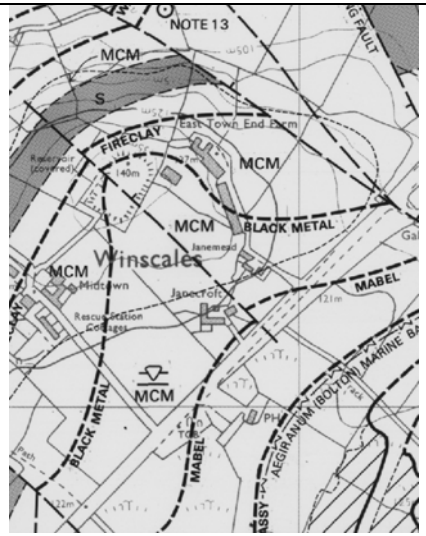
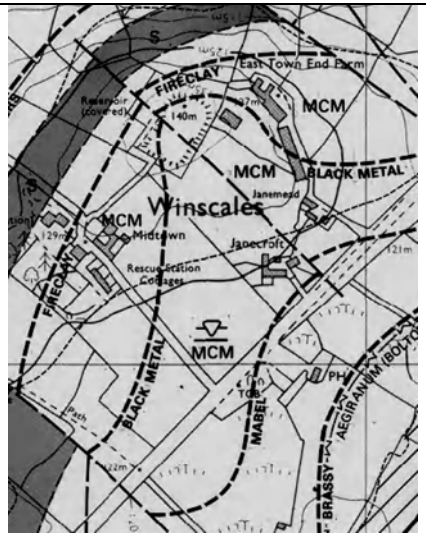
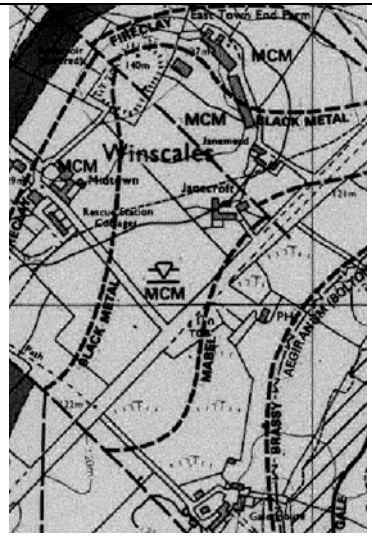
| ECW 12:1 | GeoJP2 12:1 | MrSID12:1 |
|---|---|---|
|  |  |  |
| ECW 15:1 | GeoJP2 15:1 | MrSID 15:1 |
|  |  |  |

Figure 4.2, Highlands and Islands 3 band Landsat image and compressed versions.

| Original | | |
|---|---|---|
|  | | |
| ECW 15:1 | GeoJP2 15:1 | MrSID 15:1 |
|  |  |  |

| ECW 25:1 | GeoJP2 25:1 | MrSID 25:1 |
|---|---|---|
|  |  |  |
| **ECW 40:1** | **GeoJP2 40:1** | **MrSID 40:1** |
|  |  |  |

Figure 4.3 Comparison of compressed scanned map (Fmny02nwc122267_001.tif)

| Original | | |
|---|---|---|
|  | | |
| ECW 25:1 | GeoJP2 25:1 | MrSID 25:1 |
|  |  |  |

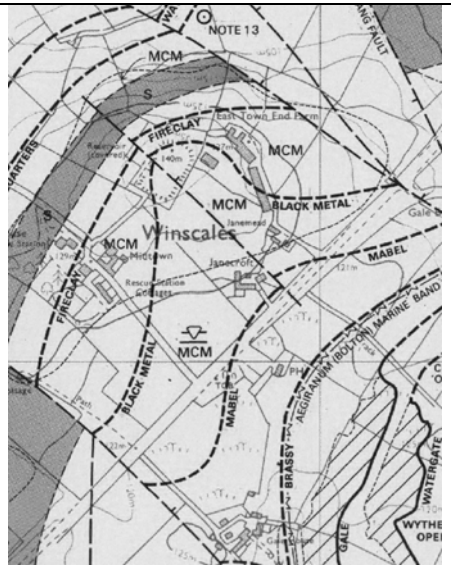
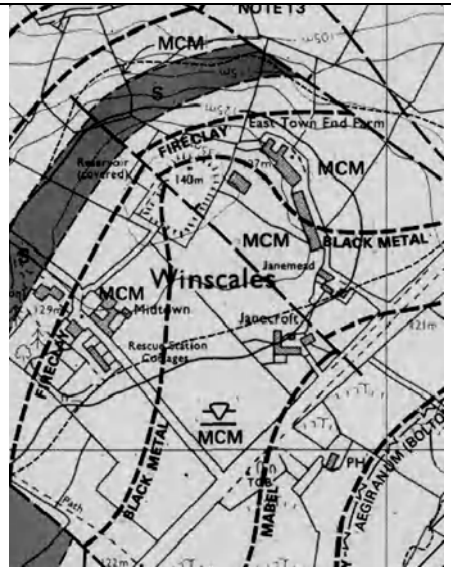
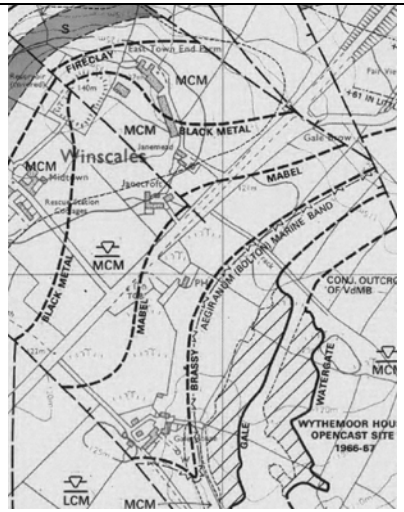
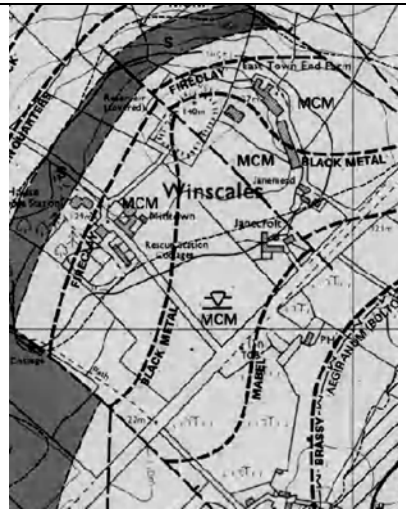
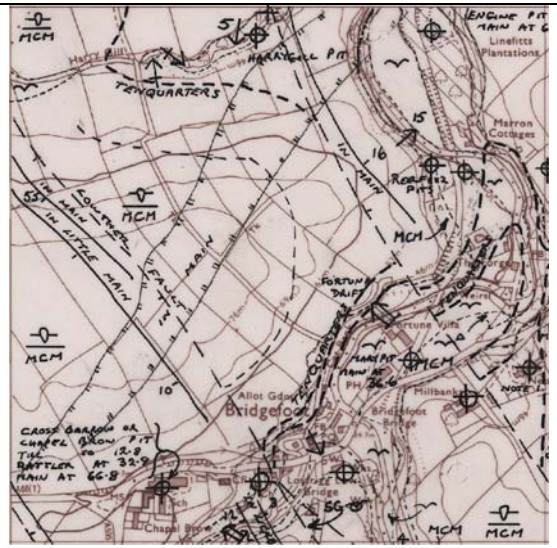
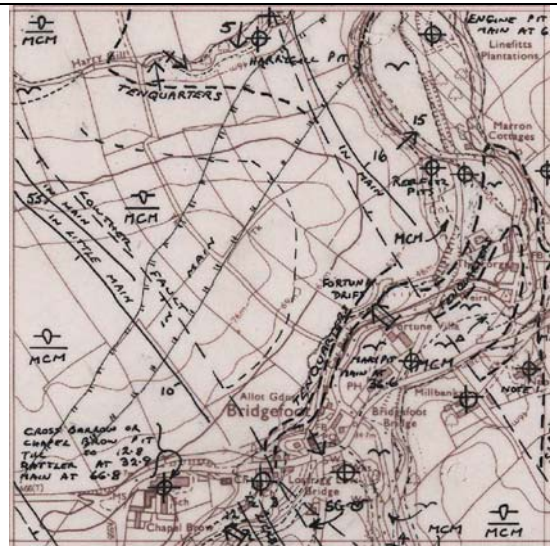
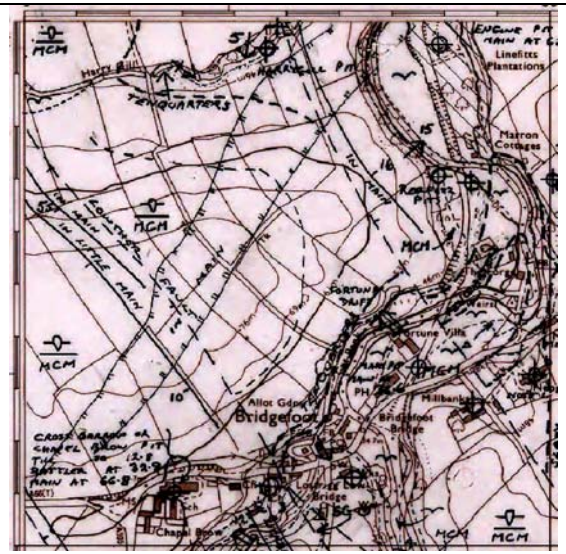
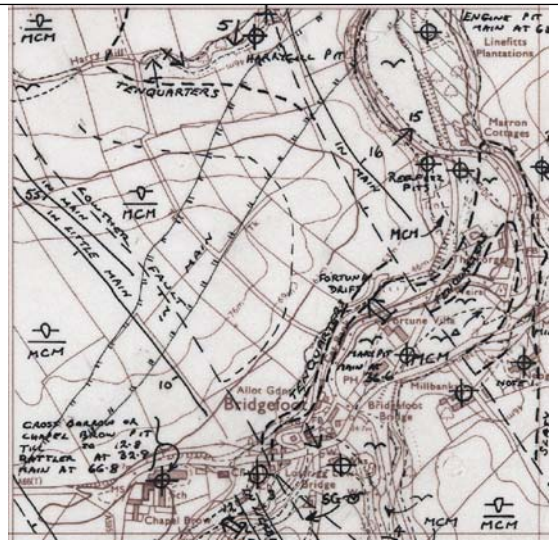
| ECW 40:1 | GeoJP2 40:1 | MrSID 40:1 |
|---|---|---|
|  |  |  |
| **ECW 60:1** | **GeoJP2 60:1** | **MrSID 60:1** |
|  |  |  |

Figure 4.4. Comparison of compressed scanned field slip (Fmny02nec128687_001.tif)

| Original | | |
|---|---|---|
|  | | |
| ECW 25:1 | GeoJP2 25:1 | MrSID 25:1 |
|  |  |  |

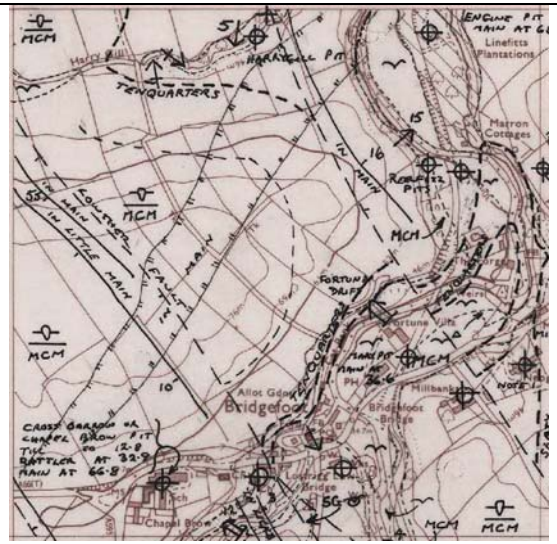
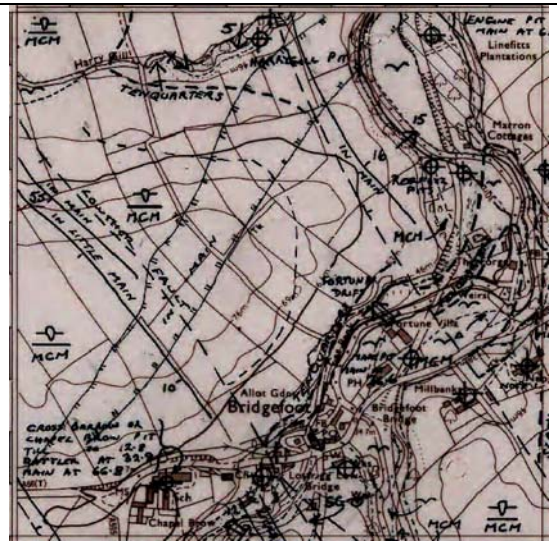
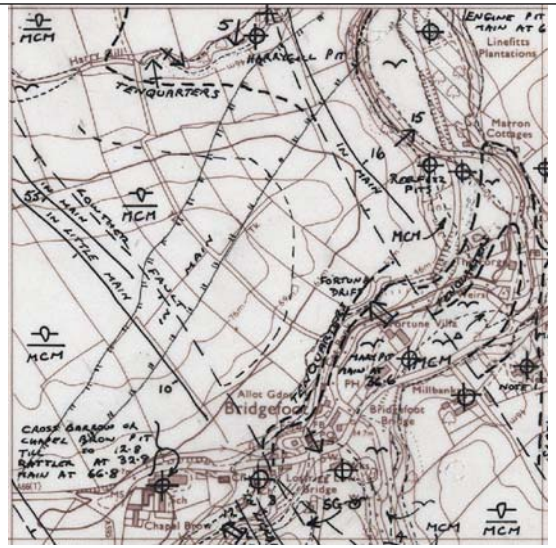
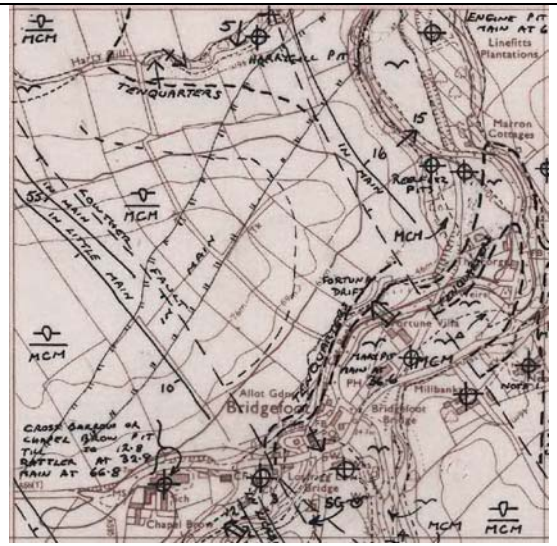
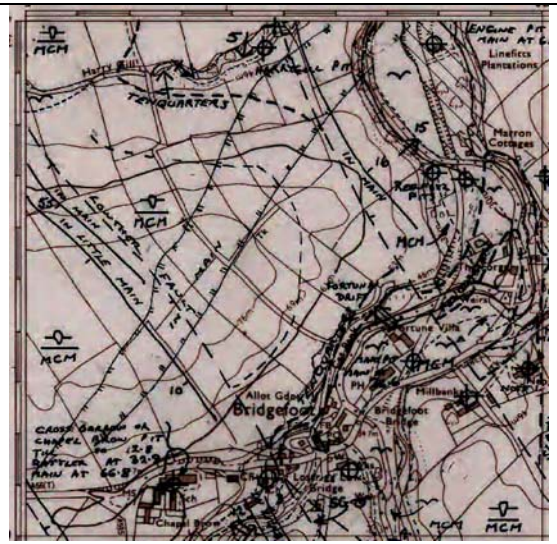
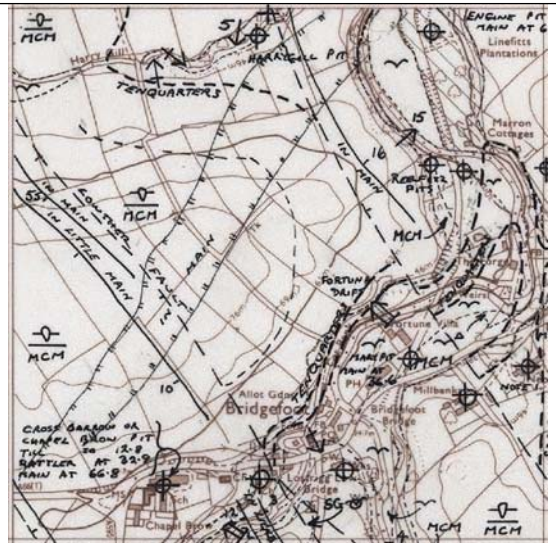
| ECW 40:1 | GeoJP2 40:1 | MrSID 40:1 |
|---|---|---|
|  |  |  |
| **ECW 60:1** | **GeoJP2 60:1** | **MrSID 60:1** |
|  |  |  |

Table 4.1 shows that in all cases ECW is at least twice as fast as MrSID and GeoJP2 for the compression of all types of imagery, in some cases it can be as much as 4 times faster. This is an important consideration when there is a large amount of imagery to compress, in our case the use of ECW would at least half the time taken to compress all the files (remember that GeoJP2 images were encoded on a slower machine).

ECW seems to be better at selecting a suitable compression ratio form the user chosen ratio, this was particularly highlighted when compressing the scanned maps. ECW would pick a ratio close to that input by the user, while MrSID would pick a ratio much higher than that stated. This lead to MrSID images of the maps that were of lower quality than the equivalent ECW images, even though MrSID took 2 to 4 times longer to compress the files. GeoJP2, on the other hand, uses a ratio, which is very close to the chosen ratio, an approach that has its advantages; users know what to expect from batch jobs and can easily predict the resultant files size.

In general the ECW files display faster (both initially and when zooming and roaming) and are of better quality than the MrSID files, a fact that is very evident with the scanned maps. The GeoJP2 files are of comparable quality to the ECW files but do take longer to display. The longer display time is only evident when opened by third party software with the use of a plug-in, files display much faster in the GeoJP2 viewer then in ERDAS Imagine.

Initial results indicate the following ratios are achievable for different types of imagery:

- 10:1 for greyscale satellite images
- 20:1 for RGB and multiband satellite imagery
- 25 to 40:1 for scanned maps (some maps could probably be compressed by about 60:1 without visually loosing visual quality – but this depends on the amount of text)

Once a format has been selected it is recommended that more detailed study is undertaken to determine the ideal ratios for each type of original. This is going to vary depending on colour space, bit depth and original file format.

# 5 Ease of use

Both ECW and MrSID are easy to use, for the ECW format the compression wizard is opened from the main menu in ER Mapper. The wizard then guides you through 3 steps. Firstly 3 choices are given; compress the file displayed in the viewing window, select a file from a drive or batch compress several files from a drive. The input and output files are defined along with the image type (greyscale. RGB, multiband), then the compression ratio is defined.

MrSID is similar except the user controls the all functions from the same window, input and output files are defined, the software automatically finds the data type and the user enters the desired ratio.

GeoJP2 is a command line based system, which to some users will detract from the software capabilities. Scripts are required to compress the imagery, these are easy to write and batch scripts can be written and saved to a text file, which is then entered at the command prompt for execution.

A GUI for the software is planned for September 2003 and will be given to current users of the software free of charge.

# 6 Cost

| | ECW | MrSID (GeoExpress) | GeoJP2 |
|---|---|---|---|
| To compress images under 50Mb | Free (ECW free compressor) | Can be encoded in ArcGIS | |
| To compress images under 500Mb | Free (ECW free compressor) | $1 500 (USD) to increase Arc capability to 500Mb | |
| To compress images of any size | £3500* | £2300 including 1 10GB data cartridge. Then £800 per 10Gb cartridge or £16 000 per Terrabyte cartridge. | $1995 (single user, single CPU)<br>$475 annual support including free updates. |
| Additional functionality | Full image processing system, image web server | Viewing software. | Viewing software. |

Table 6.1: Comparison of product pricing.

*Price for the commercial version of the software, BGS currently have several CHEST licenses, however if we wish to use the compressed imagery on a commercial project then a commercial ER Mapper license is needed.

# 7 Batch Compression

Each system is capable of batch compression; the way in which they accomplish this is a little different.

## 7.1 MRSID

With MrSID the user is required to open, in the MrSID interface, all the images they wish to compress. The images are then highlighted and the output directory specified.

## 7.2 ECW

With ECW the user is only prompted for the first file they wish to compress, the software then goes away and looks through the folder for similar images. All similar images are compressed to a user specified output directory. Batch compression of 7 tiff files, 3 of which are about 0.5Gb and the other 4 are about 1Gb in size, took approximately 10 minutes. Then the four 1Gb images were mosaiced to produce a 4Gb RGB mosaic, this was then compressed, with a 20:1 ratio in 11 minutes 57 seconds. The process of mosaicing imagery and then compressing is very easy, user friendly and fast; it only took about 13 minutes to produce a 4Gb mosaic and then compress it to a 250Mb file. Figure 7.1. compares the ECW image created in this manner with an equivalent MrSID image.


Figure 7.1. On the left is a MrSID image showing part of a Landsat mosaic of Madagascar this is image is 450Mb. On the right is a ECW image of the same mosaic this image is 250Mb. Both images are displayed at a scale of 1:100 000, even though the ECW image is 200Mb smaller than the MrSID image there is not a large difference in image quality.

## 7.3 GEOJP2

Since the version is command line driven it is necessary to write a text file containing all commands necessary to compress the files. If the same parameters are to be applied to all images then it is just a case of copying and pasting the commands down the file, replacing file names as necessary.

However this feature does allow the user to express different parameters for different images – this is a feature not possible with the other two formats.

# 8 Problems Encountered

The ECW compression takes the desired compression ratio, analyses the image and defines a different ratio that will produce the desired quality based on the user input ratio. This approach works fine for satellite imagery that has been processed and then compressed, however if an unprocessed satellite image is sent for compression the actual compression ratio tends to be much higher than the specified ratio. The reason for this lies in the contrast properties of the image; any pixel in an 8-bit satellite image can have a value between 0 and 255, this full range is very infrequently used with most of the pixel values grouped between 20 and 80. When the image is processed the values between 20 and 80 would be 'stretched' to fill the entire 0 to 255 range. If this image is then compressed it appears to the compression engine that the entire 0-255 range is used and there is little redundancy, so the compression ratio will be fairly low. If the unprocessed image is compressed it appears to the compression software that there is a lot of redundancy since the values of 0 to 20 and 80 to 255 are not used. Since large areas of the image appear to be zero it applies a much higher compression ratio, which degrades the image. This problem only seems to be apparent with satellite imagery and can easily be avoided if the user compresses the processed version of the image.

A similar problem was observed when using MrSID to compress the scanned maps. These are greyscale images but some have been scanned as RGB and the others as greyscale. Table 4.1 shows that while a compression ratio of 25:1 was entered for Fmny02nwc122267_001.tif; the actual ratio was 39.2:1. This is a more serious problem than that with ECW since we are planning to batch compress the scanned maps and therefore would want as little user interaction as possible. Satellite imagery on the other hand needs user interaction to process and any compression would take place after processing.

The above problems are not apparent with the GeoJP2 software; because the chosen ratio is applied there are no problems with over/under compression.

# 9 Product Support

It is important that the product chosen is supported by BGS IA, GIS, graphics and intra/internet applications.

## 9.1 ECW

The ECW compression engine is bundled with the ER Mapper software and as such the format is supported by this product meaning that an ECW compressed image can be opened and processed by ER Mapper. The following GIS, Image Analysis and other packages also support the ECW format, either natively or via a plug-in:

- ERDAS Imagine (plug-in)
- ArcGIS (Plug-in)
- Arc View 3.x (plug-in
- MapInfo (Plug-in)
- Paintshop Pro (plug-in)
- Photoshop (plug-in)
- MS Office (plug-in)
- PCI (built-in)
- TNTmips (built-in)
- Virtuoso (built-in)
- AutoCAD (plug-in)
- ENVI (built-in)
- PCI Geomatic's Software Suite (built in)

The plug-ins can be downloaded for free from www.ermapper.com. People distributing ECW compressed imagery are free to distribute the required plug-in with the imagery. If a user does not have or want to use any of the above then the freely available stand-alone viewer (ER viewer) can be used to read ECW image files. Once the viewer is installed OLE (Object Linking and Embedding) software such as MS Office automatically supports ECW images. It is therefore possible to embed them into Word documents etc.

A further advantage of the ECW standard is the fact that the software development SDK library is free to use for both compression and reading. This means that software providers are free to write the ECW compression routine into their own software; PCI Geomatics and ENVI have both taken advantage of this.

### 9.2 MRSID

MrSID compression is a standalone compression engine and as such does not have any other software with it. However the MrSID routines have been taken up by some of the large companies, for example ESRI's ArcGIS and ERDAS Imagine have MrSID compression routines built into the software. This enables Arc/Imagine users to compress imagery to MrSID format; unfortunately the compression will only work on images of 50Mb or less. An enhancement can be purchased for ArcGIS (for $1500 USD), which enables compression of 500Mb images.

The following support MrSID:

- ERDAS Imagine (native)
- ArcGIS (native)
- Arc View 3.x (plug-in)
- Web browsers (plug-in)

- MapInfo (plug-in)
- Microstation95 (plug-in).
- Autodesk.
- Intergraph.

(Note: plug-ins are available for free download)

A stand-alone viewer is also available for free download at http://www.lizardtech.com/download/?f=0&d=1.

LizardTech do supply decompression and compression SDK's that allow software manufactures to build in support for the product. However, unlike ECW these SDK's are charged for.

### 9.3 GEOJP2

Another stand-alone compression engine, software providers, especially those specialising in products for the mapping community, are quickly taking up this format. Plug-ins for the following are currently available:

- ERDAS Imagine
- ArcGIS
- ArcView 3.x
- Web Browsers.
- Any software offering support for Jpeg2000; PhotoShop

Since GeoJP2 and Jpeg2000 are open standards, software providers are able to integrate the format into their products. The following have done this:

- PCI Geomatics (IA)
- ENVI (IA)
- MapInfo Professional (GIS)
- MicroImages TNT (GIS)

Due to the establishment of Jpeg2000 as a new image standard, much like Jpeg, it is expected that the format will take off and hence be supported by many applications (such as the current Jpeg format).

# 10 Conclusions

In the majority of cases ECW compression is faster to compress an image than MrSID, ECW images display faster than MrSID images. The ECW compressed images tend to be better quality. This is especially true of the compression of the scanned maps (figures 3 and 4, table 2). In most cases GeoJP2 encoding speeds are slower than ECW but faster than MrSID. However encoding of a 6-band Landsat image was surprisingly slow (table 4.1), this may be due to the machine used or just the fact that 6-bands of data needed to be encoded.

A wide range of applications support the three formats, however ECW and GeoJP2 probably have the edge on MrSID since their compression and decompression SDK library's represent a free open standard, whereas MrSID charge for theirs. Having said this the corporate standards for Image analysis (ERDAS Imagine) and GIS (Arc) support MrSID as standard, and even have limited MrSID encoders included in the software. The ECW format is supported by both these products but only by the use of a free plug-in. The same is true of the GeoJP2/Jpeg2000 format, however since this is a stand-alone product with an open source it is expected that future versions of these products will natively support the format.

All formats have associated web server applications; Image web server comes bundled with ER Mapper. The server for MrSID is known as Document Express and is available at extra cost. Mapping Science are currently working on a server solution, it is unclear at present if this will come bundled with the compression software or be charged for as an additional product.

Although ECW compression initially costs more than MrSID, ECW has the added benefit that it comes with a high end Image Analysis system capable of processing aerial photography, satellite imagery and some geophysical applications, also included is the image web server. It seems as though ER Mapper represents a full package for image processing, compression and then the serving of these images over the intra/internet. MrSID on the other hand is initially cheaper, however the cost of the data cartridges, for the amount of data we need to compress, is prohibitive. GeoJP2 is the cheapest option and compares favourably with ECW.

# Glossary

*ECW* Enhanced Compressed Wavelet Images is ERMappers compression format based on wavelet technology.

*GeoJP2* Is Mapping Sciences compression format using wavelet compression.

*IA (Image Analysis)* Image Analysis system, software used to manipulate remotely sensed imagery.

*MrSID* Is Lizard Tech's wavelet based compression format.