Introduction

This document aims to give additional information to aid use of these data beyond the information already available on the dataset’s CEDA data catalogue page. This document gives further information including:

- How Station Data are connected in MIDAS (Station source id)
- Extracting data into columns in Excel
- State indicators
- Quality control (QC) guide
- Met element name _j

For information about accessing data, geographic and temporal coverage, please refer to the MIDAS data collection catalogue and the related individual dataset pages using the link above. The link below is an introduction demonstration to MIDAS.

http://badc.nerc.ac.uk/web_guides/midas_intro_demo_skin.swf

How Station Data are connected in MIDAS (Station source id)

The Met Office collates data from a number of observation networks both in the UK and internationally for various purposes, including storing these data in their long term database – MIDAS. As a station may issue meteorological data within a variety of networks, within which it may have a different identifier (e.g. WMO station ID), the Met Office allocate a specific station ID for all data for that station – known as the src_id. The Met Office also allocates an Id type for identifying a recording instrument and a message type giving an indication of the range observations that are recorded.

Source id (src_id) is an integer running from 1 upwards unique to all stations. All observations in the database are stored with their source identifier.

Id_type identifies a recording instrument, e.g. a raingauge (RAIN), an anemometer (WIND), a range of weather recording instruments at a UK Met Office Station (DCNN), a range of recording instruments at an airport (ICAO), a range of recording instruments at a WMO station (WMO), etc.
**Met Domain or message types** give an indication of the range of observations that can potentially be recorded at the station. For example, WADRAIN is the message type or Met Domain which will be used to send out daily rainfall data as observed at the station of interest from a station operated by a Water Authority (hence the WA prefix). The distinctive message types indicate that the stations are reporting within different networks and each may include different parameter set.

This document shows where all the elements from the SYNOP, NCM, HCM, SREW, METAR and DLY3208 (daily climate) message types are stored in MIDAS. One notable exception is the marine met domains (SHIP, RIG etc.) - all of these met domains have their data stored in the marine_ob (and possibly marine_ice_ob) table only.

**Finding relevant stations**

This link below is used to search for the relevant station source id using the name, postcode, county, interactive map or Google Earth. The table below shows the strengths and weaknesses of using each methods.

[http://badc.nerc.ac.uk/search/midas_stations/](http://badc.nerc.ac.uk/search/midas_stations/)

Please note the Met Office has used historic county borders, therefore places may appear in different counties. The link below shows the county boundaries used.

[http://badc.nerc.ac.uk/search/midas_stations/traditional_counties_map.html](http://badc.nerc.ac.uk/search/midas_stations/traditional_counties_map.html)

<table>
<thead>
<tr>
<th>Search option</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Likely to match against a name for the area</td>
<td>Some stations may be given a different name within MIDAS than would be expected.</td>
</tr>
<tr>
<td>Postcode</td>
<td>All the stations in the postcode area will appear in the options list.</td>
<td>This will miss nearby stations in neighbouring postcode areas</td>
</tr>
<tr>
<td>County</td>
<td>All the stations in the county will appear in the options list.</td>
<td>The Met Office has used historic county borders therefore places may appear under different county names than their present location: <a href="http://badc.nerc.ac.uk/search/midas_stations/traditional_counties_map.html">http://badc.nerc.ac.uk/search/midas_stations/traditional_counties_map.html</a>. May miss nearby stations in neighbouring county. Limited to UK coverage only.</td>
</tr>
<tr>
<td>Interactive Map</td>
<td>All the stations in the location can be seen on a map. Stations can be searched by both message types and geographical areas. Non-UK stations also displayed.</td>
<td>The stations will not appear in a list, they will appear on a map.</td>
</tr>
<tr>
<td>Google Earth</td>
<td>All the stations in a location can be seen on a map. The stations can be searched by message type or geographical area seen on the map in different colours. Non-UK stations also displayed.</td>
<td>Files have to be downloaded according to your search requirements and required Google Earth. The stations will not appear in a list, they will appear on a map.</td>
</tr>
</tbody>
</table>
These links below link to the web guide demonstrations:

http://badc.nerc.ac.uk/web_guides/midas_find1_demo_skin.swf
http://badc.nerc.ac.uk/web_guides/midas_find2_demo_skin.swf

The image below shows the page of the different search options.

Please note that not every station reports the same data or uses the same instruments and observations. The MIDAS system will bring together all the data from a particular station regardless of which network it is reporting on.

Once the relevant station has been found, click on it to display more information. The record below shows the information about Birmingham University.

A list of the different message types and descriptions can be found in the link below.

http://badc.nerc.ac.uk/data/ukmo-midas/met_domain.html
The message types are important as they represent each type of observation and measurements that are taken at the different stations. For example, Birmingham University (shown above) has message type AWSHRLY, which means it is an Automatic Weather Station recording hourly rainfall, temperature and weather observations. It is therefore possible to get the suitable data wanted under different message types.

**Extracting data for stations**

The MIDAS yearly data tables can be downloaded directly, but, these files are too large therefore a guide has been produced. To extract the data CEDA has developed a tool within the CEDA Web Processing Service (WPS), by using this link below, the MIDAS data can be extracted for one or more weather stations at a time. This process is strongly advised as there are data from a large number of stations within the MIDAS collection.

http://wps-web1.ceda.ac.uk/ui/home

Guidelines on how to use the CEDA WPS to extract MIDAS data is available at:

http://badc.nerc.ac.uk/data/ukmo-midas/WPS.html

A step by step guide

1. Login in

![CEDA OGC Web Services](image)

The CEDA Web Processing Service (WPS) provides a set of useful functions (proces

2. Select processes
3. Click on extract UK Station Data

4. Fill in the following fields with the appropriate parameters. It is optional to fill in bounding box, counties and station source ids, but one option must be filled in as highlighted in the diagram below.
5. Press submit and an estimated time and volume is displayed

![Information about your asynchronous job]

These links below define the column headings depending on which data has been extracted.

<table>
<thead>
<tr>
<th>Data extracted</th>
<th>Webpage to the relevant column headings</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK Daily Temperature</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/td_table.html">http://badc.nerc.ac.uk/data/ukmo-midas/td_table.html</a></td>
</tr>
<tr>
<td>UK Soil Temperature</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/st_table.html">http://badc.nerc.ac.uk/data/ukmo-midas/st_table.html</a></td>
</tr>
</tbody>
</table>

6. Then the data can be downloaded

![Output Size 0.01 MB]

**OUTPUT FILES**

The following file outputs are available from your job.

- ouput_58173f2f163883c63fc50cfda6e6eb09d_20150721_093152.zip [Download] 0.01 MB
- station_data-201401010000-201407181617.csv 0.01 MB
- disclaimer.txt 0.01 MB
- request_inputs.txt 0.01 MB
<table>
<thead>
<tr>
<th>Data Type</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK Daily Rainfall data</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/RD_Table.html">http://badc.nerc.ac.uk/data/ukmo-midas/RD_Table.html</a></td>
</tr>
<tr>
<td>UK Daily Weather Observation data</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/WD_Table.html">http://badc.nerc.ac.uk/data/ukmo-midas/WD_Table.html</a></td>
</tr>
<tr>
<td>UK Hourly Rainfall data</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/RH_Table.html">http://badc.nerc.ac.uk/data/ukmo-midas/RH_Table.html</a></td>
</tr>
<tr>
<td>UK Hourly Weather Observation data</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/WH_Table.html">http://badc.nerc.ac.uk/data/ukmo-midas/WH_Table.html</a></td>
</tr>
<tr>
<td>UK Mean Wind data</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/WM_Table.html">http://badc.nerc.ac.uk/data/ukmo-midas/WM_Table.html</a></td>
</tr>
<tr>
<td>UK Soil Minimum Temperatures (1959-1970 only)</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/TMSL_Table.html">http://badc.nerc.ac.uk/data/ukmo-midas/TMSL_Table.html</a></td>
</tr>
<tr>
<td>UK Sub-hourly Rainfall data only to April 2005</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/RS_Table.html">http://badc.nerc.ac.uk/data/ukmo-midas/RS_Table.html</a></td>
</tr>
<tr>
<td>Global Radiation Observations</td>
<td><a href="http://badc.nerc.ac.uk/data/ukmo-midas/RO_Table.html">http://badc.nerc.ac.uk/data/ukmo-midas/RO_Table.html</a></td>
</tr>
</tbody>
</table>

**Handling extracted data**

*Extracting data into columns in Excel*

The link below demonstrates how to convert a text file to Excel

http://badc.nerc.ac.uk/web_guides/csv_to_excel_demo_skin.swf

1. Highlight all the data
2. Click on the data tab and into Text to columns

3. Choose the option that best describes your data
4. Choose a delimiter (either comma or tab)

5. Select finish
Cleaning up the data

There may be more than 1 version of the line where every single digit is the same. While the Met Office MIDAS system overwrites the existing entry in their database the CEDA’s MIDAS entries do not, as we take snapshots of the MIDAS database from time to time leading to duplicate entries occurring in our archive. When CEDA carry out our monthly extractions of MIDAS data from the Met Office’s MIDAS system we typically go back over the previous 12 months to try and capture any late arriving data/late changes to the data. Then when the data arrives at CEDA they are merged into the existing files. This should NOT allow duplicate lines to be generated. However, earlier versions of the script that carried out this code may not have been so perfect.... although we think this highly unlikely.

1. To remove the duplicates select all the data and select Data tab, in the Data Tools group, click Remove Duplicates (Highlighted in yellow).

2. Do one or more of the following:
   - Under Columns, select one or more columns.
   - To quickly select all columns, click Select All.
   - To quickly clear all columns, click Unselect All.
3. Click OK. A message is displayed indicating how many duplicate values were removed and how many unique values remain, or if no duplicate values were removed. Click OK.

**Version Number and Quality control**
The data arrives at the Met Office and it undergoes quality control to check that the data are correct and consistent with the surrounding data points. Whether the process has occurred or not is indicated by the version number. 0 indicates that the data line is given BEFORE any quality control process has happened; a value of 1 indicates that the QC code has run on the data. If version 0 exists it *may* be OK to use, BUT needs to be treated with caution!

Once you have removed all duplicates, you should then filter your data using the _q and _j values to determine which data to keep, which to rule out and which you could keep if you carry out your own checks and they look reasonable.

**Quality control (QC) guide**
Quality control is when the data are checked to make sure it is correct and consistent with the surrounding data points.

These data are displayed with a QC code composed of up to 5 digits e.g. 1006, 15006 or 6

Each of the 5 digits represents one letter MESQL.

The letters stand for:

- M- marker
- E- estimate
- S- status
- Q-query
- L-level

**Example 1**

<table>
<thead>
<tr>
<th>M</th>
<th>E</th>
<th>S</th>
<th>Q</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>
The numbers represent different information for each letter given in full below. The example of 1006 is shown below.

For example

```
M E S Q L
1 0 0 6
```

E (1) - Estimate/correction derived automatically from a program with no manual intervention
S (0) - Observed and not suspect
Q (0) - Original value is/was not queried, or no information available
L (6) - Final (or only) areal or buddy job run and queries processed

Full listing is available in the link below:
http://browse.ceda.ac.uk/browse/badc/ukmo-midas/metadata/doc/QC_J_flags.html

Met element name _J_
This attribute is a single character code which either describes the method of measurement, or further qualifies the meteorological values. The meaning of any value depends on the element being qualified.

For example

This data below is taken from the UK Daily rainfall data

005349, RAIN, 1920-01-01 00:00, 1, DLY3208, 900, 1, 310, 1001, 14, , 0, , ,
D

This therefore means we need to look under the Precipitation codes. The D therefore implies the rainfall has been converted from inches.
State Indicators

A state indicator is an attribute (rec_st_ind) of each table that is used to describe the current stage in the life of a particular record, from creation to deletion.

Full listing is available in the link below:
http://browse.ceda.ac.uk/browse/badc/ukmo-midas/metadata/doc/state_indicators.html