British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

Session: 5.1
Date: Monday 26th August

YUCATÁN MEXICO 2013 IAGA meeting

Geomagnetism Data Portal - a new service from the World Data Centre for Geomagnetism, Edinburgh

Ewan Dawson[1] (ewan@bgs.ac.uk), Ikechukwu Nkisi-Orji[2], Sarah Reay[1] and Susan Macmillan[1]
[1]British Geological Survey, West Mains Road, Edinburgh EH9 3LA, United Kingdom
[2]Robert Gordon University, Schoolhill, Aberdeen, AB10 1FR, United Kingdom

http://wdc.bgs.ac.uk/dataportal

The Geomagnetism Data Portal

Since 2007, the World Data Centre (WDC) for Geomagnetism in Edinburgh has made geomagnetic observatory datasets available via its website and FTP server. This website is widely used by scientists looking for definitive geomagnetic observatory data, and has serviced more than 1 million data requests since first opening. With current data holdings spanning more than 150 years and 350 observatories, this data centre is the most comprehensive resource for observatory-standard global geomagnetic observations available.

Since mid-2013, BGS has been building a new website to provide users with access to the WDC data catalogue. This new Geomagnetism Data Portal (http://wdc.bgs.ac.uk/dataportal) is an interactive web application, which offers a number of new features such as search, visualisation, and data format conversion, to enable users to find and download data from the WDC data catalogue more easily than before.

Technologies Used

The client-side component of the Data Portal web application runs in the user's browser, and was built using HTML5 and the JavaScript technologies AngularJS and jQuery. These technologies allow us to offer the user a rich, interactive interface, compatible with all modern web browsers.

The server-side component is a RESTful web service running in BGS' Edinburgh data centre. The web service was created using the Java framework Restlet, which is a software implementation of the REST client-server architectural style. The server software runs within Apache Tomcat containers running on Linux web server nodes.

Using the Data Portal

Users can discover data that are of interest to them using a combination of search filters.

The data catalogue may be filtered using the following criteria:

- Station name or IAGA code
- Year (range from earliest to latest)
- Dataset type (hourly or 1-minute time series)
- Location of observing station

The map view shows the location of all the stations with datasets that match the search criteria. In addition, the user can draw a bounding box on the map to filter the results to include only datasets from stations within a certain region.

The search criteria are applied to the data catalogue, and the matching datasets are listed in the right-hand pane. Results may be grouped by station code or by year.

Using the checkboxes, the user can select which datasets to download. The choice of desired data format is made using the drop-down menu below the results list. Available formats are IAGA2002, WDC, CSV, XML and JSON.

http://wdc.bgs.ac.uk/dataportal

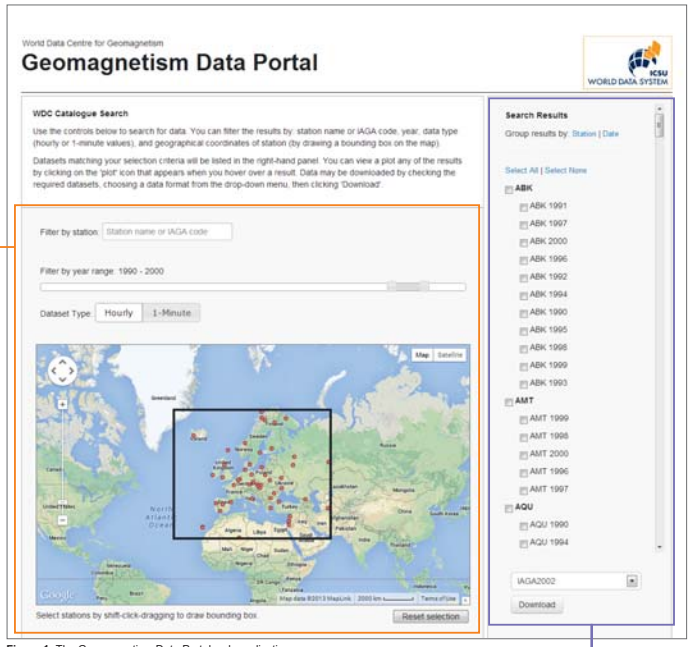

Figure 1: The Geomagnetism Data Portal web application.

The user can plot any dataset that appears in the results list by hovering the pointer over the list entry, and clicking on the 'plot' icon that appears. This causes a popup window to open, containing an interactive plot of the dataset.

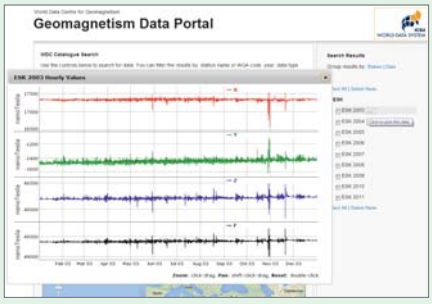

Figure 2: The interactive plotting functionality of the Data Portal.

The user may investigate the dataset further by zooming in (by click-dragging a region of the plot) and panning the plot (by shift-click-dragging on the plot). To reset the plot, the user double-clicks anywhere on the plot.

As the user moves the pointer over the plot, the value of the data point nearest the pointer is displayed in the legend for that component series.

Web Service Integration

The Geomagnetism Data Portal web application is powered by a web service, which provides catalogue search and data retrieval functionality via a simple HTTP-based interface. This web service is publicly available, meaning it is possible to access WDC resources directly, without having to go though the Data Portal.

This allows for direct integration of WDC data with any software capable of making HTTP requests, e.g. applications software such as *Excel*, *R* and *Matlab*, command-line utilities found in UNIX environments such as *wget* and *cURL*, and of course any programming language that has an HTTP client library.

Each resource may be retrieved in a number of formats, such as CSV, XML and JSON. The geomagnetism time series dataset resources may also be downloaded in the IAGA2002 and WDC formats. The desired data format is selected using the *media* query parameter. For example, to download data in IAGA2002 format, one would append media=iaga2002 to the dataset URL.

In Figure 3, we see how the web service can be used directly from within the 'R' statistical computing software. In this case we pass the URL of the Eskdalemuir 2011 hourly mean values dataset into the read.csv() function, which then directly retrieves the data from the WDC web service in CSV format, ready for further manipulation within R.

Resources available from the WDC web service at **http://app.geomag.bgs.ac.uk/wdc**:

- **/stations** - Retrieve a list of all stations known to the WDC
- **/stations/{code}** - Retrieve basic metadata for the station with the given IAGA code
- **/catalogue/search** - Query the WDC data catalogue. Results are filtered using the query parameters *minLatitude*, *maxLatitude*, *minLongitude*, *maxLongitude*, *minYear*, *maxYear*, *stations* and *frequencyTypes*. So for example, the request **http://app.geomag.bgs.ac.uk/wdc/catalogue/search?minYear=2000&maxYear=2013&stations=LER,ESK&frequencyTypes=MINUTE** would retrieve a list of links to all 1-minute datasets from Lerwick and Eskdalemuir observatories between the years 2000 and 2013.
- **/datasets/{frequency}/{id}** - Retrieve a particular dataset, for example /datasets/hour/esk1985 is the path for the dataset of hourly mean values recorded at Eskdalemuir in 1985.
- **/datasets/download** - Making a GET request to this resource retrieves an HTML form that then allows the client to make a POST request to download a large number of datasets in bulk. The server responds with a ZIP file containing the requested datasets in the requested format.

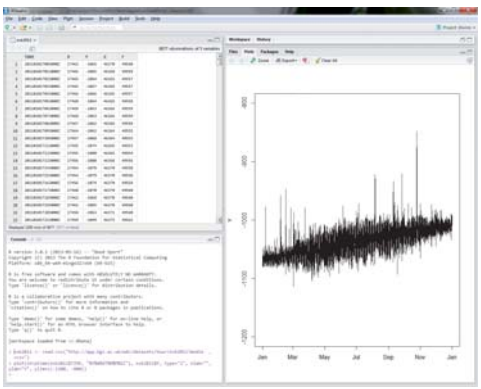

Figure 3: A screenshot of RStudio, showing direct access to WDC data from within R.