ECMWF releases software for observational data

PIOTR KUCHTA

After many years of development and the successful use in operations at ECMWF and other meteorological centres, ODB API, ECMWF's software for the efficient processing of observational data, is being released to the general public under the open source licence Apache 2.0.

ODB (Observational Database) API, which has been developed with support from the UK Met Office, includes an SQL (Structured Query Language) filtering and statistics engine, command line tools and an API (application programming interface) for C/C++, Fortran and Python. It has been designed for the efficient processing of observational data archived in MARS, ECMWF's meteorological data archive system.

The ODB software developed between 1998 and 2004 proved to be highly efficient as an observational data store for the Integrated Forecasting System (IFS). An important novelty was the introduction of ODB/SQL, a subset of SQL, for observational data processing. However, the early ODB file format was not well suited for archiving.

In 2008, an Observation Handling and Monitoring Taskforce was set up in order to improve the archiving, monitoring, visualisation and diagnostics of observations. Its outcome was the design of a MARS extension for archiving observational feedback, including data layout and indexing. In addition, a data format was designed for the observational feedback archive, with unique features that make it especially useful for archiving but also for stream processing. A new implementation of the SQL processor was integrated with the library, and thus ODB API was born. Since November 2011, ODB feedback data has been archived operationally in MARS as OFB (ODB feedback) and MFB (Monitoring feedback).

Main features of ODB API

The software has been written in portable C++. It provides public APIs for C/C++, Fortran and Python. The package also contains various command line tools for working with ODB API files, including executing SQL queries to filter data or calculate statistics, examining ODB API files as well as importing and converting to other formats. The file format ODB API works with fulfils the following requirements:

- simple, well defined, portable and machine independent
- self-described: metadata stored in files along with the data itself
- extendable: new compression algorithms and data types can be added in form of codecs
- new data can be appended to existing files
- data can be decoded in a streaming fashion to facilitate efficient processing of long times series.

Documentation and examples of using Python, Fortran and C APIs are available on the ODB API website at https://software.ecmwf.int/wiki/display/ODBAPI/ODB+API+Home.

Main uses of ODB API

ODB API is an essential tool for many scientists working at ECMWF and beyond:

- ODB feedback produced by experiments at ECMWF is archived in MARS in the ODB API format, so researchers working on scientific studies based on observations need to work with this data.
- Scientists working on reanalysis have been early adopters of ODB API, and their requirements were one of the main drivers of the ODB API design.
- ECMWF's Monitoring (OBSTAT) and Alarm system uses ODB API to calculate statistical information on the quality of observations used or monitored by ECMWF, primarily in order to improve the usage of observations within ECMWF's data assimilation system.
- The Metview desktop and the Magics data visualisation package developed at ECMWF can both read and plot observational data encoded in the ODB API format. Additionally, Metview has extensive support for examining metadata of ODB API datasets, can retrieve data directly from MARS or the ODB Server, and can execute SQL to filter datasets before plotting them.
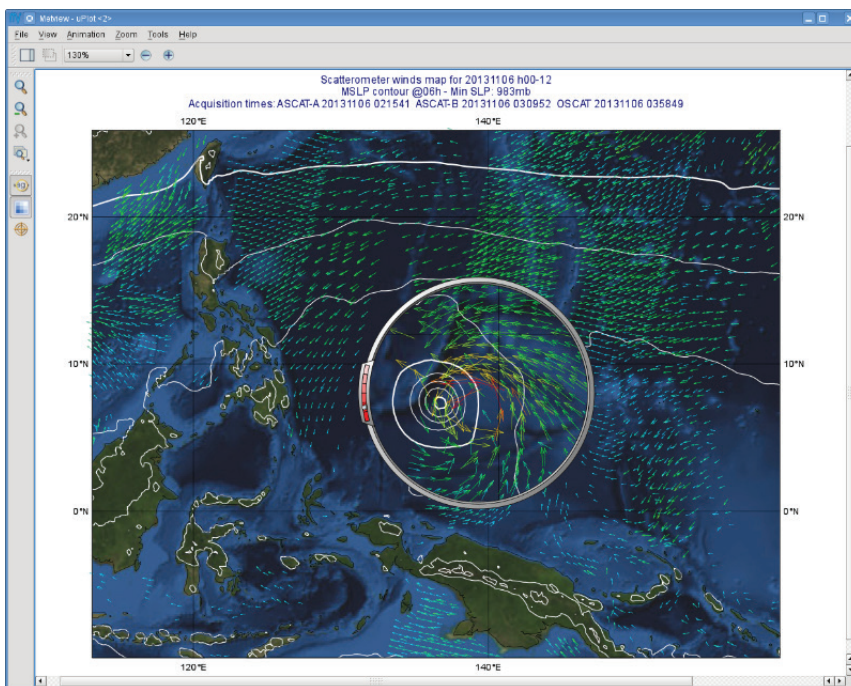
Metview screenshot. This screenshot shows how Metview can plot observation feedback data stored in MARS, thanks to ODB API.