# Serving Collections of Forecast Model Runs with the THREDDS Data Server

John Caron (caron@ucar.edu)
University Corporation for Atmospheric Research / UNIDATA
IN13A-1164

**Collections of model data** are stored on disk, typically each hour's output in a file, or all the output from one run in a file. The output can be in any format understood by the CDM; GRIB (1 or 2) and netCDF are common.

This example uses NCEP RUC2 GRIB2 files distributed through Unidata's Internet Data.

**NetCDF-Java library / Common Data Model**

The data collection is described using NcML ( *NetCDF Markup Language*), a declarative language for modifying and aggregating netCDF/CDM files. In this example, *Forecast Model Run Collection aggregation* is used to create a virtual dataset with two time dimensions (the run time and the forecast/valid time), out of the collection of files.

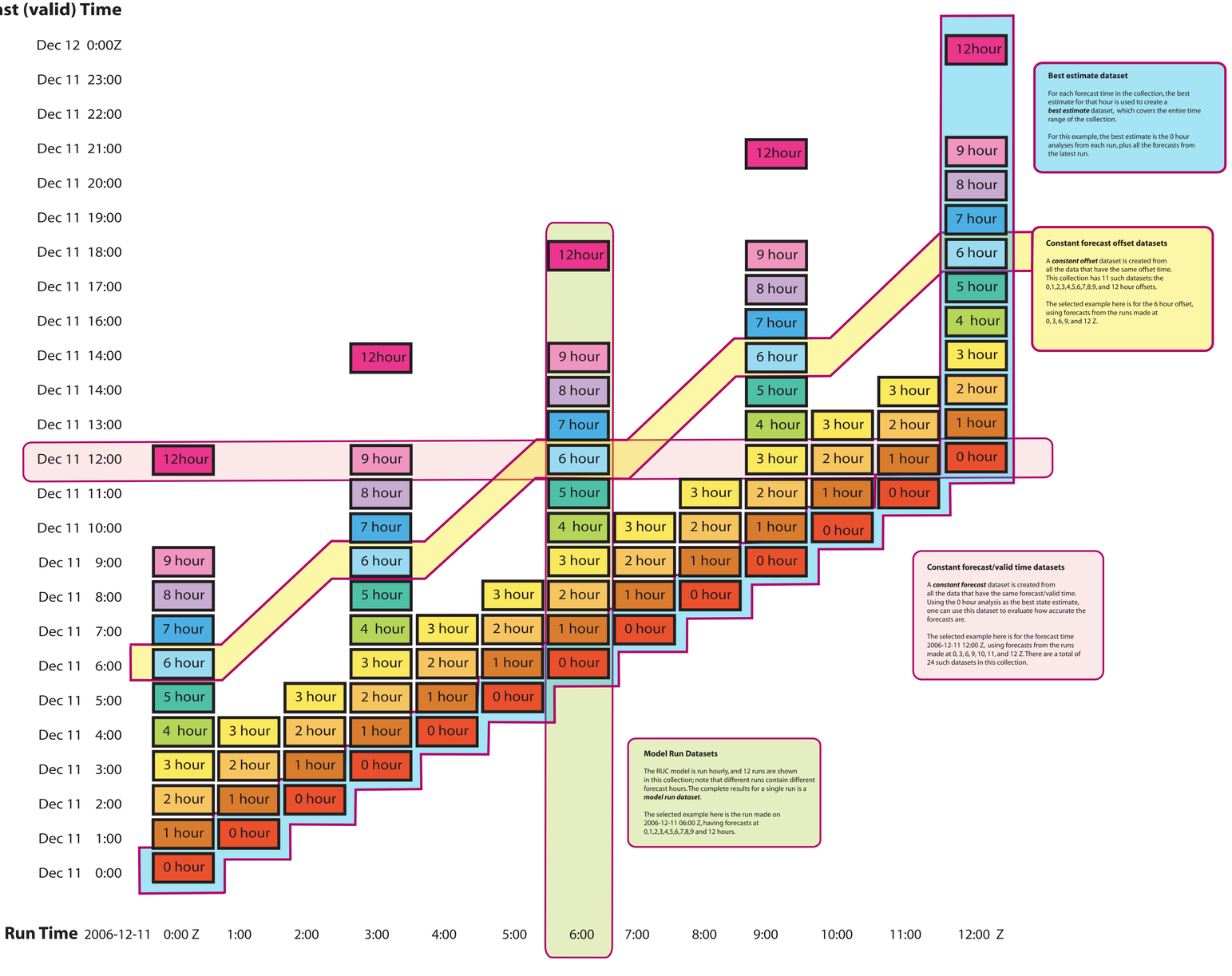
This special dataset is then sliced and diced into many virtual datasets.

**THREDDS DATA SERVER**

The TDS is a Java Servlet-based server for scientific data, using the NetCDF Java library to read data. The data can then be remotely accessed in several ways, including OPeNDAP and the OGC Web Coverage Service. Metadata can be added to the THREDDS catalogs to enable discovery services like GCMD, DLESE, and NSDL. The goal is to help data providers publish, and incrementally improve, data to other researchers and educators.

The TDS uses NcML embedded in the catalog that describes the collection of model output files, to automatically create and serve the virtual datasets described here.

OPeNDAP WCS




**Best estimate dataset**

For each forecast time in the collection, the best estimate for that hour is used to create a ***best estimate*** dataset, which covers the entire time range of the collection.

For this example, the best estimate is the 0 hour analyses from each run, plus all the forecasts from the latest run.

**Constant forecast offset datasets**

A ***constant offset*** dataset is created from all the data that have the same offset time. This collection has 11 such datasets: the 0,1,2,3,4,5,6,7,8,9, and 12 hour offsets.

The selected example here is for the 6 hour offset, using forecasts from the runs made at 0, 3, 6, 9, and 12 Z.

**Constant forecast/valid time datasets**

A ***constant forecast*** dataset is created from all the data that have the same forecast/valid time. Using the 0 hour analysis as the best state estimate, one can use this dataset to evaluate how accurate the forecasts are.

The selected example here is for the forecast time 2006-12-11 12:00 Z, using forecasts from the runs made at 0, 3, 6, 9, 10, 11, and 12 Z. There are a total of 24 such datasets in this collection.

**Model Run Datasets**

The RUC model is run hourly, and 12 runs are shown in this collection; note that different runs contain different forecast hours. The complete results for a single run is a ***model run dataset***.

The selected example here is the run made on 2006-12-11 06:00 Z, having forecasts at 0,1,2,3,4,5,6,7,8,9 and 12 hours.

## NCEP/RUC2 CONUS 40 km model runs