Lineage:

- Initial version – Radim and Rashid, Aug 2016
- Substantial Edits – Bruce,Denis, Jan 2017

User Story

6.1.2 Generate Derived Observations Statistical Data

Version: 0.2

==== User Story ====

Related components: 4.3.3, 4.5.1, 5.1, 5.3, 5.4

Related User Stories:

- 4.1 Observations data
- XXX spatial, discrete + derived grids

A Climate Data Manager requires functionality to automatically generate **derived observation statistical data**:

- For a single sensor and station:
  - at the completion of a data ingest, or Quality Control process
- on a regularly scheduled basis for all sensors within the NMHS' Observations network to ensure that all relevant statistics are kept correct and current for the observations network.

The Climate Data Manager also requires the ability for authorized staff to **manually** generate these statistics on an ad hoc basis via the Climate Data Management User Interface.

The statistics required for each sensors and for each variable (observed property) are:

- Daily, Monthly, Annual and long-term averages (mean, minimum, maximum);
- Extremes;
- Climatological Standard Normals;
- Ten-year and other Standard Normals;
- Anomalies compared with the Climatological Standard Normal (Daily, Monthly, Annual, Decadal); and
- Various Climatological Indices.

Architectural Comments:

- It is expected that the '6.1.2 Generate Derived Observations Statistical Data' analytical component will be typically called by:
  1. '5.1.1 Data Ingest' component to automatically update statistical values as appropriate.
  2. '5.3 Observations Quality Control' component to automatically update statistical values as appropriate at the completion of each successful update.

Bruce Bannerman 20/1/2017 15:19

Comment [1]:

Add when available.

Bruce Bannerman 20/1/2017 15:17

Comment [2]:

What are the rules to be used for each of these ?

We need formal and consistent rules for each of these (plus probably more) !

e.g. :

- What is a climatological day ?
- How do we handle multiple sensors per phenomena per station ?
  - Record stats for each sensor.
  - How do we determine what is the official record for a particular phenomena for a specific station.
- How are missing observations handled ?
  - hours in a day ?
  - days in a month ?
  - months in a year ?
  - years in a decade ?
- How do we calculate accumulated values when we :
  - have missing observations ?
  - one observation covers a number of days.
- What algorithms are to be used ?
- what else ?
- **Who can provide an authoritative answer for these types of issues ?**

3. An automated scheduled process to update periodic values (daily, weekly, monthly, annual etc.) at the completion of each day;

- There is also a requirement to run this process manually by suitably authorized staff via a Climate Data Management view of '7.1.4 Integrated search of climate data'.

===== end ====

Reference: WMO 1131