Appendix: HydroJSON Input Type

2015-08-25

Overview

The HydroJSON time series format is designed to facilitate exchange of time series data. The file format specification and additional information are available on the following website:

https://github.com/gunnarleffler/hydroJSON

HydroJSON can be written by the TSTool `WriteTimeSeriesToHydroJSON()` command but currently cannot be browsed within TSTool. An overview of the format from the TSTool perspective is as follows:

- The file follows standard JSON conventions. Nice formatting of the file is provided as an option by TSTool; however, for web data exchange formatting is typically minimized to improve performance.
- JSON does not allow comments. Refer to the specification to understand the data format.
- Many HydroJSON data properties in the specification do not directly translate to internal time series properties. Therefore software must translate or omit some features. For example, the TSTool `WriteTimeSeriesToHydroJSON()` command provides features to map internal time series data to the HydroJSON specification.
- The data are organized by stations first, and then time series within the stations.

HydroJSON Files and Standard Time Series Properties

This section is important when reading HydroJSON files. Currently TSTool does not provide a `ReadTimeSeriesFromHydroJSON()` command. Additional detail will be added in the future.

The standard time series identifier for HydroJSON files is as follows:

```
Location.DataSource.DataType.Interval.Scenario~DateValue~PathToFile
```

Limitations

HydroJSON files have the following limitations:

- Need to understand the specification better in order to document limitations.
- Should the HydroJSON version number be included in the file?

Commented [kt1]: Yes, I've added this as an issue on the github page

Format Versions

Changes to the HydroJSON specification are provided on the GitHub repository:

https://github.com/gunnarleffler/hydroJSON/blob/master/CHANGELOG.md

Examples

The following is an example file for day interval data. Highlights indicate questions that need to be resolved in the TSTool implementation of the format. The TSTool implementation attempted to follow the standard illustrated in the following link, but some things are not fully explained:

https://github.com/gunnarleffler/hydroJSON/blob/master/hydroJSON.json

```
{
  "stationList": [
    {
      "name": "abei",
      "responsibility": "noaa",
      "coordinates": {
        "latitude": 42.953333,
        "longitude": -112.82667,
        "datum": ""
      },
      "huc": "",
      "elevation": {
        "value": 1341.0,
        "accuracy": 0.0,
        "datum": "m",
        "method": "agrimet_map"
      },
      "timezone": "US/Mountain",
      "tz_offset": "",
      "time_format": "",
      "active_flag": "",
      "location_type": "agrimet",
      "timeseries": [
        {
          "tsid": "abei.PN.sr.Day",
          "values": [
            {
              "timestamp": "2014-06-10",
              "value": 784.82,
              "quality": ""
            },
          ],
          "site_quality": [],
          "hash": "",
          "quality_type": "",
          "parameter": "mm",
          "duration": "",
          "interval": "",
          "units": "degF",
          "count": 435,
          "min_value": -6.03,
          "max_value": 80.17,
          "start_timestamp": "2014-06-10",
          "end_timestamp": "2015-08-18"
        }
      ]
    }
  ]
}
```

Commented [s2]: The spec seemed to indicate that this was the station ID. That seems odd. I added a top-level list element or it could be "station:" at the top rather than "STATION1" as in the spec. I can try to figure out how to change but it seems that you would want a list of elements of a constant name, and have the identifier as a property of the station.

Commented [s3]: I am using ${ts:siteid} here but perhaps there should be an "id" and a "name", where the latter is longer.

Commented [s4]: What should this contain? It is not documented.

Commented [s5]: This is not in Pisces. It could be added through a GIS intersect of the coordinate and HUC polygons and added after reading from Pisces.

Commented [s6]: I used the vertical datum from Pisces but this needs to be defined in HydroJSON. Is it units? Or a spatial datum?

Commented [s7]: Should this be pulled out of Pisces? Many are blank.

Commented [s8]: Time format for what? The timestamps shown in this file? I have a bit of a problem that TSTool uses % specifiers for some built in time series properties so I need to evaluate how to handle.

Commented [s9]: This seems consistent with Pisces but lots of blanks. If blank, does that equate to T?

Commented [s10]: There is no property for missing value. Does the list of values below include ALL values or only non-missing. How does Pisces handle missing values – no entry in database? null? NaN? Special value like -999?

Commented [s11]: The specification has "TSID1" as an element under "timeseries". I made it a data property under the time series in the list. I'm defaulting to the time series alias or TSID (whichever is specified) but it can be indicated as a command parameter in TSTool.

Commented [s12]: Specification could be clearer as to what the integer is and whether the string should include time zone, etc. If an integer, how are more precise time parts handled (set to 00, etc.?). TSTool generally only uses as much of a date/time as it needs and does not display extra when it is not appropriate. Showing the timestamp is useful but eats up a lot of filespace.

Commented [s13]: This needs to be documented better. I don't understand what it is. Should I remove? Also, why are these properties after the values? It makes it easier to read the file if they are at the top.

Commented [s14]: Not sure what this is or how computed. Do I need to compute. There are different ways to get a hash.

Commented [s15]: Not sure what this is – needs documented.

Commented [s16]: This needs to be defined, preferably with example – don't know how to get from Pisces.

Commented [s17]: Need to know how this is defined. Is it "Daily", "Day", etc.?

Commented [s18]: Is this the count of nonmissing (or nonmissing + missing) between start_timestamp and end_timestamp, inclusive?

Commented [s19]: Specification seems to indicate that this should also have a timestamp to go with min and max. The specification needs to be clarified.

Commented [s20]: TSTool allows start and end to be specified when reading. If not specified all available data are returned.