**Arctic-HYCOS Station Attributes**

Contents

[1 Introduction 1](#_Toc499136322)

[1.1 WIGOS Metadata Standard 2](#_Toc499136323)

[1.2 USGS Metadata Standard 2](#_Toc499136324)

[2 Metadata for the Arctic-HYCOS Database 2](#_Toc499136325)

[2.1 GRDC Metadata 2](#_Toc499136326)

[2.2 Arctic-HYCOS specific metadata 3](#_Toc499136327)

[3 Extended Arctic-HYCOS Metadata Definitions 5](#_Toc499136328)

[3.1 Drainage Area 5](#_Toc499136329)

[3.2 Regulation 5](#_Toc499136330)

[3.3 Land Use Changes 6](#_Toc499136331)

[3.4 Discharge Data Quality Flag 6](#_Toc499136332)

[3.4.1 Existing Data Quality Assessments 7](#_Toc499136333)

[3.4.2 Arctic-HYCOS Discharge Data Quality Flag 8](#_Toc499136334)

[4 References 9](#_Toc499136335)

# Introduction

Meaningful use of scientific observations (data) requires knowledge about the conditions under which the observations were made; metadata is what provides that supporting information. Metadata is supplementary information about an observation that may include its location, the conditions under which it was collected, the quality procedures applied to the observation, details of the instrument used to make the observation, and more. Including metadata with datasets makes observations discoverable and accessible, and helps enable appropriate use of data based on certain criteria and uncertainties.

The Arctic Hydrological Cycle Observing System (Arctic-HYCOS) project, a component of the [World Hydrological Cycle Observing System (WHYCOS)](http://www.whycos.org/) network, has created a database of discharge data from hydrometric stations across the expansive, transnational basin of the Arctic. This network includes two subsets: stations that account for flow to the Arctic Ocean, and stations that represent all hydrological regimes in the Arctic basin. These two subsets are intended to aid in the evaluation of freshwater flux to the Arctic Ocean and Seas and to study changes in Arctic hydrological regimes relative to climate change.

To allow the discharge data, and future other hydrological variables, to be discoverable and accessible by researchers and the public, metadata associated with each station-dataset is required. The below document discusses the metadata proposed to be implemented for the Arctic-HYCOS database.

## WIGOS Metadata Standard

“Metadata of two complementary types are required. The first of these is **discovery metadata** – information that facilitates data discovery, access and retrieval. These metadata are WIS (WMO Information System) metadata and are specified and handled as part of WIS. The second type is **interpretation/description or observational metadata** – information that enables data values to be interpreted in context. These latter metadata are the subject of the WIGOS metadata standard, which provides a WIGOS standard for the interpretation metadata required for the effective utilization of observations from all WIGOS component observing systems by all users (WMO, 2017).”

## USGS Metadata Standard

The USGS “[GAGES II: **G**eospatial **A**ttributes of **G**ages for **E**valuating **S**treamflow, version II](https://pubs.er.usgs.gov/publication/70046617)” dataset provides geospatial data and classifications for 9,322 stream gauges maintained by the USGS. This dataset has two purposes:

(1) to provide users with a comprehensive set of geospatial characteristics for a large number of gauged watersheds, particularly for gauges with long flow record, and

(2) to provide a determination of which of those watersheds represent hydrological conditions which are least disturbed by human influences ("reference gauges"), compared to other watersheds within ecoregions.

With respect to metadata, the GAGES II documents identify 27 worksheets which represent general types of variables. See page 9 of the [GAGES II report 1](ftp://ftpext.usgs.gov/pub/er/va/reston/NAWQA_ENS/GIS_DATA/gages/gagesii-updates/gagesII_sept30_2011_report.docx) (USGS, 2011).

# Metadata for the Arctic-HYCOS Database

## GRDC Metadata

The following metadata fields are included in the Global Runoff Data Centre (GRDC) database online, and provide information about all GRDC stations, including those in the Arctic-HYCOS curated subset when available. The last metadata attribute in the list, “ArcHycos”, identified stations within the GRDC that belong in the Arctic-HYCOS subset.

|  |  |  |
| --- | --- | --- |
| **Column key of "grdc\_metadata"** | | |
| 1 | grdc\_no | GRDC station number |
| 2 | wmo\_reg | WMO region |
| 3 | sub\_reg | WMO subregion |
| 4 | mix\_reg | Combination of row 2+3 |
| 5 | nat\_id | national station ID |
| 6 | river | river name |
| 7 | station | station name |
| 8 | country\_code | country code (ISO 3166) |
| 9 | lat | latitude ° |
| 10 | long | longitude ° |
| 11 | area | catchment size km2 |
| 12 | altitude | height of gauge zero above sea level m |
| 13 | ds\_stat\_no | GRDC\_No of next downstream GRDC station |
| 14 | w\_level | water level data available in addition to discharge |
| 15 | d\_start | daily data available from |
| 16 | d\_end | daily data available until |
| 17 | d\_yrs | # years of daily data |
| 18 | d\_miss | percentage of missing values (daily data) |
| 19 | m\_start | monthly data available from |
| 20 | m\_end | monthly data available until |
| 21 | m\_yrs | # years of monthly data |
| 22 | m\_miss | percentage of missing values (monthly data) |
| 23 | t\_start | totally earliest data available |
| 24 | t\_end | totally latest data available |
| 25 | t\_yrs | # years (maximum using daily and monthly data) |
| 26 | lta\_discharge | mean annual streamflow m3/s |
| 27 | r\_vol\_yr | mean annual volume km3 |
| 28 | r\_height\_yr | mean annual depth mm |
| 29 | proc\_tyrs | percentage of t\_yrs used for calculation of lta\_discharge, r\_vol\_yr and r\_height\_yr (only years with > 9 month i.e. > 75 % of monthly data are considered) |
| 30 | proc\_tmon | percentage of # of month in proc\_tyrs \* t\_yrs years used for calculation of lta\_discharge, r\_vol\_yr and r\_height\_yr (minimum 75 % are considered) |
| 31 | f\_import | date of first import |
| 32 | f\_im\_yr | year of first import |
| 33 | l\_import | date of last change |
| 34 | l\_im\_yr | year of last change |
| 35 | provider\_id | provider number (to be resolved with address database) |
| 36 | ACSYS | 1 if station belongs to the subset of the Arctic Runoff Data Base in the WCRP Arctic Climate System Study |
| 37 | statmouth | 1 if station is used for the GRDC Product "Long Term Mean Annual Freshwater Surface Water Fluxes into the World Oceans" |
| 38 | GEMS | 1 if station corresponds to a GEMS/Water water quality station |
| 39 | GCOS\_GTNH | 1 if station belongs to the subset of the Global Terrestrial Network for River Discharge within GTN-H |
| 40 | ltchars | 1 if station statistics is available in GRDC Product "Long Term Mean Monthly Discharges and Annual Characteristics of Selected GRDC Stations" |
| 41 | PristineBasins | 1 if station belongs to the subset "Pristine Basins" |
| 42 | GRDCReferenceDataset | 1 if station belongs to the subset "GRDC Reference Data set" |
| 43 | AdaptAlp | 1 if station belongs to the subset "AdaptAlp" |
| 44 | ArcHycos | 1 if station belongs to the subset "ArcticHYCOS" |

Table 1: Column key for Global Runoff Data Centre metadata

## Arctic-HYCOS specific metadata

Following a discussion at the 2016 Arctic-HYCOS project steering committee meeting, it was decided that the following additional geospatial attributes, or discovery metadata, should be included for all stations in the database, as a start (Table 2 and Table 3).

|  |  |
| --- | --- |
| **Metadata Attribute** | **Description** |
| HYCOSID | Identifier for the HYCOS project |
| StationID | Identifier for the station from the operator’s database |
| GRDC-ID | Global Runoff Data Centre (GRDC) identifier, if available |
| GRDC acsys | 1 if station belongs to the subset of the Arctic Runoff Data Base in the GRDC database |
| Country | Country from which the data originates |
| Institute | Institute responsible for data provision nationally |
| StationName | Hydrometric station name |
| Latitude | Station location |
| Longitude | Station location |
| Darea | Total catchment area in square kilometres (km2) |
| DAreaEffective | Effective catchment area in square kilometres (km2) |
| DatumAltitude | Height of gauge zero above sea level (m) |
| StartYear | First year of data available |
| EndYear | Last year of data available |
| NYears | Total number of data years available |
| Status | Active (A) or Not active (N) |
| Operational | Yes (Y) or No (N) |
| RealTime | Is continuous near real-time data available |
| FlowtoOcean | 1 if this station is considered the furthest downstream for calculation of flow to the ocean |
| FreelyAvailable | Is the data available for free to the public? |
| EasilyAvailable | Is the data easily available to the public? |
| NOTES | Text field |

Table 2: Arctic-HYCSO metadata modified from existing GRDC metadata

|  |  |  |
| --- | --- | --- |
| **\*new\* Extended metadata** | **New metadata variables** | **Definition** *(See expanded definitions below)* |
| Availability of a vector shapefile of the total station drainage area | Drainage\_vector (#45) | * Available (1) = if a vector shapefile of the station's drainage area is available; * Not Available (0) = if a shapefile is not available. |
| Regulation: “no significant regulation” vs “regulated” | Regulation (#46) | * Regulated (1) = Basins with structures providing significant flow regulation, based on the country’s standards, are considered regulated. * No significant regulation (0).   The definition of “significant regulation” should be included in the metadata field “regulated\_definition”; if no standard is available, significant regulation should be defined as basins with structures controlling more than 5% of the basin area. |
| Definition of significant regulation | Regulation\_defintion | Text field for a country’s definition of “significant regulation” |
| Regulation start date | Regulation\_startdate | What date did regulation start, if applicable |
| Regulation end date | Regulation\_enddate | What date did regulation end, if applicable |
| Land Use Change: significant land use changes that result in changes to the flow regime | Land\_use\_chng (#47) | * Significant (1) = if >10% of the surface area of the basin has been modified from natural conditions; * Non-significant (0) = if <10% of the surface area has been modified. |
| Discharge Data Quality Flag 1 (under ice conditions) | Data\_quality\_ice (#48) | General assessment of the accuracy of the derived discharge from each station during ice-cover conditions (winter), based on the stability of the rating curve.   Station data quality, or accuracy, is rated as Excellent (3), Fair (2), or Poor (1). |
| Discharge Data Quality Flag 2 (open water conditions) | Data\_quality\_open (#49) | General assessment of the accuracy of the derived discharge from each station during open water conditions (summer), based on the stability of the rating curve.   Station data quality, or accuracy, is rated as Excellent (3), Fair (2), or Poor (1). |

Table 3: Arctic-HYCOS required metadata

# Extended Arctic-HYCOS Metadata Definitions

## Drainage Area

“The term "drainage area" is defined as the land area where precipitation falls off into creeks, streams, rivers, lakes, and reservoirs. It is a land feature that can be identified by tracing a line along the highest elevation between two areas on a map, often a ridge. The drainage area for a river basin is measured in a horizontal plane enclosed by the drainage divide outlining the basin. In some cases river basins may have non-contributing sub-basins, or commonly called enclosed basins, where the runoff stays within the basin and not contributing to the larger basin surrounding the enclosed basin (USGS, 2017).”

For the Arctic-HYCOS database, please indicate in the variable “Drainage\_vector” whether a vector shapefile for the total drainage area corresponding to the hydrometric station is available (1) or not (0).

## Regulation

This attribute is intended to identify stations in basins with and without “significant” regulation or diversions in the river system. This designation reflects only the physical structures within the waterways upstream of the site; it does not reflect the land use within the basin. The “natural” designation does not infer pristine conditions, but it does infer that there are no control structures upstream (Environment Canada, 1999).

Regulated (1): Basins with structures providing significant flow regulation, based on the country’s standards. The definition for “significant regulation”, as used by each country, should be included in the metadata field “regulated\_definition”. If no standard is available, regulated should be defined as basins with structures controlling more than 5% of the basin area.

No Significant Regulation (0): There is no significant regulation or diversions within the river system, based on the country’s definition. This generally represents “natural”, if not “pristine”, conditions.

## Land Use Changes

A “significant” amount of land use change in a basin is defined as greater than 10% of the surface area of the station’s drainage area having been modified in some fashion (Environment Canada, 1999). In the case of the Canadian RHBN, stations all have a non-significant degree of basin development and are considered to represent pristine, or as a minimum, stable land-use conditions (Environment Canada, 1999).

For the Arctic-HYCOS database, please indicate in the variable “Land\_use\_chng” if there has been significant land use change in the basin (1) or if there is no significant change (0).

## Discharge Data Quality Flag

Iceland: Good, fair, estimated, suspect, unchecked, missing

Russia: good and bad – good data is measured within a 5% accuracy

The discharge data quality flag is intended to be a general and qualitative assessment of the accuracy (and uncertainty) of the derived discharge from each station. The suggested technique is to form a quality assessment for a station’s discharge timeseries based on the stability of the station’s rating curve (the model used to derive continuous discharge water level). Aspects to consider include the slope of the rating curve, how well the curve is defined in relation to the number of measurements, the spread of the measurements over the range of the curve, and the stability of the curve over time (weed growth, ice conditions, erosion, deposition, etc.). This data quality flag is meant to be a general accuracy assessment of the derived data from a station overall, not for individual data points or for the measurements themselves.

Historically around the world, data quality has been determined via a qualitative assessment by a local expert (such as the station technologist) based on knowledge of hydraulic conditions at each site (Environment Canada, 1999). More advanced techniques are currently being developed to calculate discharge uncertainty in a quantitative way, for example under the leadership of the WMO’s “Project X” (WMO, 2017). Data quality may vary throughout the year for many northern sites, such as poor quality derived discharge data during the winter (due to ice cover) but excellent data during open water conditions, or vice versa (Environment Canada, 1999). Therefore, the accuracy assessment used by the Arctic-HYCOS should be split by time of year into two metadata categories - (i.e. open-water versus ice-cover conditions).

### Existing Data Quality Assessments

Numerous stage-discharge rating uncertainty tools exist and are in use by hydrological agencies around the world.

* Qualitative methods:
  + In 2015, the Water Survey of Canada developed a tool to evaluate the complexity of station rating curves and rate each as “simple”, “medium” or “complex” (Kondakow 2015). This was then compared to expert opinion for many stations (Kondakow and Jamieson, 2016). The rating complexity was only evaluated for open water measurements and ratings produced for discharge estimated during non-ice cover conditions (by assuming a time period of open water conditions, such as October to May).
  + Also in Canada, the Reference Hydrometric Basin Network (RHBN), which is designed to provide hydrometric time series data to be used in the detection, monitoring, and assessment of climate change, uses metadata to characterize hydrometric stations based on six criteria, including “accuracy of the data” (Environment Canada, 1999). Initially, the accuracy was qualitatively assessed by local experts based on their knowledge of hydraulic conditions at each site, such as the stability of the control and the accuracy of the rating curve. They assigned a score from 1 to 5 (representing excellent to poor quality data), for both open-water and ice-cover conditions. Unfortunately, the records of the results have not been updated and are not available to the public at this time.
  + The New Zealand NEMS uses a “Data Quality Schema”, which consists of quality codes for rating curves. This is a sophisticated technique and is likely more applicable to measurements than derived discharge data (NEMS, 2016).
  + Other work in the USA (e.g. Rundell et al., 2015), and in France (e.g. Morlot et al., 2014) and by the USGS (e.g. Mason et al 2016).
  + The WMO Manual on Stream Gaging indicates that an assessment based on a range of criteria allows the expert to rate measurement accuracy on a 4-step scale as Excellent, Good, Fair or Poor, based on USGS methodology. However, this rating is in regards to the actual measurement of discharge, not the on-going continuous estimate of discharge (WMO 2010; see page I10-10).
* Quantitative methods:
  + An EGU lead group is currently working together to assess rating curve uncertainty methods (see **Figure 1**). Many of these methods are currently being used for research but not applied operationally.
  + The Water Survey of Canada is testing BaRatinAGE at a few sites/stations. This is a method and software tool from France (also listed in Figure 1), but the WSC’s application so far is very preliminary.

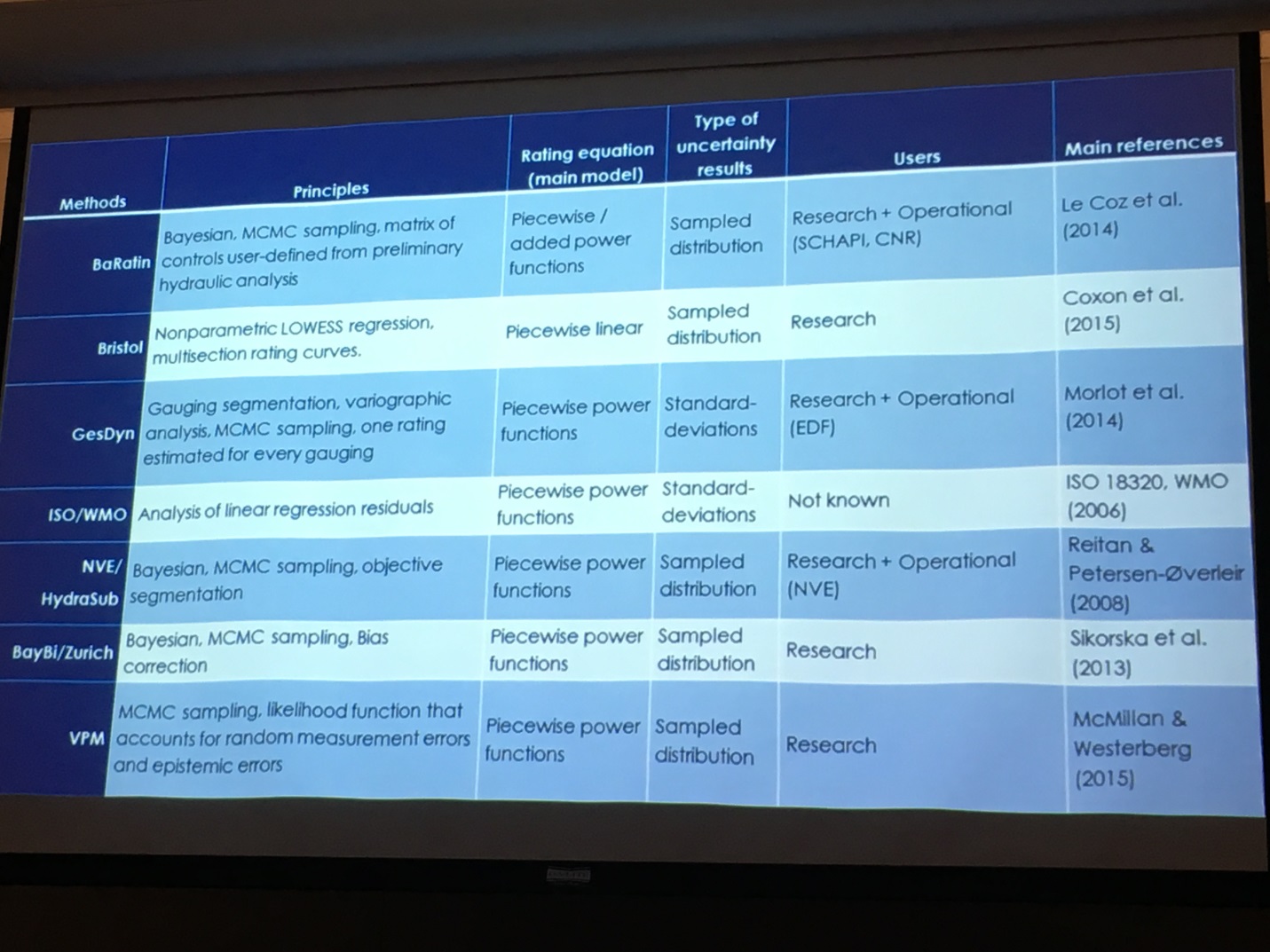


Figure 1: Slide from a conference presentation on rating curve uncertainty, 2017.

### Arctic-HYCOS Discharge Data Quality Flag

For the Arctic-HYCOS database, it is proposed that the station data quality is rated as Good (3), Fair (2), or Poor (1), for both open-water versus ice-cover conditions, based on expert opinion of the stability of the rating curve in each country’s station list.

# References

Environment Canada, 1999. Establishment of the Reference Hydrometric Basin Network (RHBN) for Canada. Ottawa, ON. *(Available on* [*Arctic-HYCOS google drive “Rating Stability at WSC” folder*](https://drive.google.com/drive/folders/0B2nFCFMIADgNc3prTk16LXZVbnc)*).*

Kondakow, 2015. Water Survey Canada Rating Complexity Study: Assessment of the complexity of Water Survey Canada discharge rating curves. Water Survey of Canada, Environment and Climate Change Canada, Ottawa, Canada. *(Available on* [*Arctic-HYCOS google drive “Rating Stability at WSC” folder*](https://drive.google.com/drive/folders/0B2nFCFMIADgNc3prTk16LXZVbnc)*).*

Kondakow and Jamieson, 2016. An automated method for assessing Canada’s rating curves. Presented to the Canadian Water Resources Association (CWRA) National Conference, Montreal, Quebec, Canada, May 26, 2016. *(Available on* [*Arctic-HYCOS google drive “Rating Stability at WSC” folder*](https://drive.google.com/drive/folders/0B2nFCFMIADgNc3prTk16LXZVbnc)*).*

NEMS, 2016. National Environmental Monitoring Standard Rating Curves. Construction of stage-discharge and velocity-index ratings, Verson 1.0. February 2016. [www.lawa.org.nz](http://www.lawa.org.nz). *(Available on* [*Arctic-HYCOS google drive “Discharge Data Quality Flag” folder*](https://drive.google.com/drive/folders/0B2nFCFMIADgNX0lxTU1qM3ZLWnc)*)*

Mason et al., 2016. Rating curve uncertainty: A comparison of estimation methods. River Flow 2016. Taylor & Francis Group, London, ISBN 978-1-138-02913-2. *(Available on* [*Arctic-HYCOS google drive “Discharge Data Quality Flag” folder*](https://drive.google.com/drive/folders/0B2nFCFMIADgNX0lxTU1qM3ZLWnc)*)*

Morlot et al., 2014. Dynamic rating curve assessment for hydrometric stations and computation of the associated uncertainties: Quality and station management indicators. Journal of Hydrology 512 (2014), 173-186. *(Available on* [*Arctic-HYCOS google drive “Discharge Data Quality Flag” folder*](https://drive.google.com/drive/folders/0B2nFCFMIADgNX0lxTU1qM3ZLWnc)*)*

Rundall et al., 2015. Towards a New Classification of Rivers based on Generic Gage Height – Discharge Rating Curves for Low-Cost Estimation of Stream Discharge. Hydrology Days 2015. *(Available on* [*Arctic-HYCOS google drive “Discharge Data Quality Flag” folder*](https://drive.google.com/drive/folders/0B2nFCFMIADgNX0lxTU1qM3ZLWnc)*)*

United States Geological Survey (USGS), 2017. USGS Current Conditions for Washington. “Drainage Area.” National Water Informationn System: Web Interface. <https://waterdata.usgs.gov/wa/nwis/current?type=basinda>, accessed August 11, 2017.

World Meteorological Organization (WMO), 2010. Manual on Stream Gaging: Volume I – Fieldwork. WMO: Geneva, Switzerland. WMO-No. 1044. <http://www.wmo.int/pages/prog/hwrp/publications/stream_gauging/1044_Vol_I_en.pdf>, accessed August 11, 2017.

World Meteorological Organization (WMO), 2017. WIGOS Metadata Standard, 2017 edition. WMO: Geneva, Switzerland. WMO-No.1192. <https://library.wmo.int/opac/doc_num.php?explnum_id=3653>, accessed August 11, 2017.

WMO, 2017. “Assessment of the Performance of Flow Measurement Instruments and Techniques”. World Meteorological Organization, Geneva, Switzerland. Accessed October 6, 2017. <http://www.wmo.int/pages/prog/hwrp/Flow/index.php>