GCW Data Management

Øystein Godøy
Purpose of GCW Data Management

- To provide an overview of the datasets that are relevant for GCW
- To provide access to datasets
  - Real time data streams
  - Archive access
- Distributed Data Management
  - Metadata driven
  - Not hosting data
- To connect GCW with
  - WMO Information System
  - WIGOS
Heterogeneous community

- National Meteorological and Hydrological Services
- Universities
- Independent research institutions
- Varying degree of structured data management
- Varying degree of interoperability for
  - metadata
  - data
- Mutual benefit of standardisation
- Not necessarily the same objective
Challenges

- Interoperability
  - Metadata
    - Protocols (√)
    - Structures (√)
    - Semantics/terminology (-)
  - Data
    - Protocols (√)
    - Formats (-)
    - Semantics/terminology (-)
    - Common data model (-)
DIKW chain

• How to transition from data to knowledge and understanding...
  - The illustration is a common redrawing of Russ Ackoff “From Data to Wisdom”
• Take home...:
  - Take care of data for the future
  - Data is the basis for knowledge
    • Now and in the future
  - Knowledge based management depends on national, regional and global interaction

http://www.easterbrook.ca/steve/2012/09/what-is-climate-informatics/
Metadata and data standards

• Discovery
  – Information encoding in
    • GCMD DIF
      – Also keywords
    • ISO19115
  – Exchange through
    • OAI-PMH
    • (OGC CSW)
    • (OpenSearch)

• Use
  – Formatting
    • NetCDF4 classic
  – Documentation standard
    • Climate and Forecast
  – Exchange through
    • OPeNDAP
      – Allows transformation into BUFR

• Discovery metadata is required for every physical file unless aggregation is used
• Updating interoperability guidelines when collaborating with contributing data centres
• Must also align WIGOS and CF efforts
Interoperability with CryoNet stations

- Dedicated effort with Davos (WSL/SLF)
  - Software for discovery and data interoperability developed in Davos
  - Solution capable of integration a wide range of input streams, including RDBMS
- Takes care of data from measurement to published data where it can be picked up by services
Why OPeNDAP and NetCDF/CF?

- Do not start from scratch
- WMO formats not used nor accepted within the science community
- NetCDF/CF well developed concerning annotation and self-explaining data encoding
- OPeNDAP provides software that allows you to access data over the internet,
  - from programs that weren't originally designed for that purpose,
  - as well as some that were.
- Allows connecting to data streams
  - Delivers data, not files
  - Segmenting data in time and space(s)
  - Bridges the message approach of operational data with the long time series approach of climate analysis
Data management services

Relies quite heavily on:
- OAI-PMH for federated search
- OPeNDAP for integration of data
- OGC WMS for visualisation of gridded data

The GCW Data Portal is being restructured
- Modularised
- Service Oriented Architecture
- Enabling distributed services on distributed data
Challenges and next steps

- Harvesting information from legacy systems
  - No standardisation at the data level
  - But can show what is available when datasets relevant to cryosphere are identified
    - Gaps in discovery metadata

- Need to get data from CryoNet stations
  - Must be harmonised to service downstream purposes
    - Common annotation
  - Varying degree of existing data management capabilities and understanding
  - Building software and services to help stations
  - Selected datasets to be integrated with WMO GTS

- Need to get descriptions of CryoNet stations
  - WMDS required
    - Transparency in e.g. glossary development
  - Facilitating lightweight implementation
  - Discussing governance aspects
  - Joining forces with GAW

- Semantic annotation
  - Need to align WIGOS and CF keywords
System update
LAST UPDATED: NOVEMBER 29, 2018

The software running the catalogue integration in the test portal has been upgraded. This upgrade brings relationships for datasets as well as to define collections within the system. The content in the portal is still relevant. This will be changed once this portal moves from test to operations.
Tags:
System update

New functionality
LAST UPDATED: MAY 23, 2018

A validation service for NetCDF/CF files have been added. You find this interactive service under the menu "Support". It will also check ACDD elements if requested. The purpose of this service is to help data providers to check their products before they are made available. The service is available at https://gcw-test.met.no/dataset_validation/form.
Tags:
Services

https://gcw-test.met.no/node/2/
Dataset Validation

Webform for validation of netCDF files based on the IOOS compliance checker.

Select the test you want to run:
- CF-1.6
- ACDL

Upload Your File

You can only upload a single netCDF file with "nc" extension with a maximum size of 1500M. You need to upload a bigger file, take contact with the website support directly.

Submit
Questions?
What can we do in the front end?
Data Access

Full text search
Search words

Data collection period
Start date
1966-01-01
End date
yyyy-mm-dd

Bounding box
Institutions
Investigator
Topics and variables
sea

Geographical search
Arctic Sea Ice Concentration charts based on a manual interpretation of different satellite data. The main satellite sensor used are the SAPI sensor (Synthetic Aperture Radar) supplemented by visual and infrared sensors and data from passive microwave sensors. As part of the Copernicus project the sea ice concentration product is provided to a 1km spatial resolution and conveyed in a NetCDF format. The concentration interval follows the World Meteorological Organization (WMO) total concentration standard. A new product is delivered every weekday around 13:00 UTC.

Sea Ice Extent estimates based upon EUMETSAT Ocean and Sea Ice SAF products. Both reprocessed and operational datasets have been used. See http://encyclopedia.wmo.int/satellite-observations for details.

The product is based on a manual interpolation of available satellite data and in situ observations and provides a gridded map. It is a continuation of the previous sea ice chart which basically identified the ice edge.
## Available Metadata

<table>
<thead>
<tr>
<th>Metadata key</th>
<th>Metadata value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE</td>
<td>Arctic Sea Ice Scavenging</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>Sea ice concentration charts based on a manual interpretation of different satellite data. The main satellite sensor used are the SAR sensor (Synthetic Aperture Radar) supplemented by visual and infrared sensors and data from passive microwave sensors. As part of the Copernicus project the sea ice concentration product is grouped to a 1 km spatial resolution and converted to a NetCDF format. The concentration intervals follow the World Meteorological Organization (WMO) total concentration standard. A new product is delivered every weekday around 1500 UTC.</td>
</tr>
<tr>
<td>PERSONNEL ROLE</td>
<td>Investigator</td>
</tr>
<tr>
<td>PERSONNEL ORGANISATION</td>
<td>Norwegian Meteorological Institute</td>
</tr>
<tr>
<td>TEMPORAL EXTENT START DATE</td>
<td>2015-01-04T12:00:00Z</td>
</tr>
<tr>
<td>DATA ACCESS RESOURCE</td>
<td>HTTP: &quot;<a href="http://freda.met.no/freda/my/scene/SW-TAG/ice-metro-svalbard.html">http://freda.met.no/freda/my/scene/SW-TAG/ice-metro-svalbard.html</a>&quot;</td>
</tr>
<tr>
<td></td>
<td>GPICAP: &quot;<a href="http://freda.met.no/freda/stock/Cryos_ice-SW-METNO-ARC-SEASICE_HR_OBS/ice_conc_svalbard_aggregated">http://freda.met.no/freda/stock/Cryos_ice-SW-METNO-ARC-SEASICE_HR_OBS/ice_conc_svalbard_aggregated</a>&quot;</td>
</tr>
<tr>
<td></td>
<td>ISO WMS: &quot;<a href="http://freda.met.no/freda/arc/Cryos_ice-SW-METNO-ARC-SEASICE_HR_OBS/ice_conc_svalbard_aggregated">http://freda.met.no/freda/arc/Cryos_ice-SW-METNO-ARC-SEASICE_HR_OBS/ice_conc_svalbard_aggregated</a>&quot;</td>
</tr>
<tr>
<td>BBOX</td>
<td>ENVELOPE: (-0.34627896464055,-85.300246600775,61.1025712155825,54.30654936923047)</td>
</tr>
<tr>
<td>ACTIVITY TYPE</td>
<td>Space borne instrument</td>
</tr>
<tr>
<td>DATA ACCESS WMS LAYERS WMS LAYER</td>
<td>Ice_concentration</td>
</tr>
<tr>
<td>DATASET PRODUCTION STATUS</td>
<td>In Work</td>
</tr>
<tr>
<td>ISO TOXD CATEGORY</td>
<td>icecon</td>
</tr>
</tbody>
</table>
Transform dataset

Title (discovery metadata): Arctic Sea Ice Svalbard

Abstract (discovery metadata): Sea ice concentration charts based on a manual interpretation of different satellite data. The main satellite sensor used are the SAR sensor (Synthetic Aperture Radar) supplemented by visual and infrared sensors and data from passive microwave sensors. As part of the Copernicus project the sea ice concentration product is gridded to a 1km spatial resolution and converted to a NetCDF format. The concentration intervals follow the World Meteorological Organization (WMO) total concentration standard. A new product is delivered every weekday around 1500 UTC.
Transform dataset

Title (discovery metadata): Arctic Sea Ice Svalbard

Abstract (discovery metadata): Sea ice concentration charts based on a manual interpretation of different satellite data. The main satellite sensor used are the SAR sensor (Synthetic Aperture Radar) supplemented by visual and infrared sensors and data from passive microwave sensors. As part of the Copernicus project the sea ice concentration product is gridded to a 1km spatial resolution and converted to a NetCDF format. The concentration intervals follow the World Meteorological Organization (WMO) total concentration standard. A new product is delivered every weekday around 1500 UTC.

- The e-mail address to send the results to
- Select spatial extent
  - Degrees north
  - Degrees south
  - Degrees east
  - Degrees west
- Select temporal extent
- Select variables
Transform dataset

Title (discovery metadata): Arctic Sea Ice Svalbard

Abstract (discovery metadata): Sea ice concentration charts based on a manual interpretation of different satellite data. The main satellite sensor used are the SAR sensor (Synthetic Aperture Radar) supplemented by visual and infrared sensors and data from passive microwave sensors. As part of the Copernicus project the sea ice concentration product is gridded to a 1Km spatial resolution and converted to a NetCDF format. The concentration intervals follow the World Meteorological Organization (WMO) total concentration standard. A new product is delivered every weekday around 1500 UTC.

- The e-mail address to send the results to
- Select spatial extent
- Select temporal extent
- Select variables

<table>
<thead>
<tr>
<th>Name</th>
<th>Standard name</th>
<th>Long name</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>lat</td>
<td>latitude</td>
<td></td>
<td>degrees_north</td>
</tr>
<tr>
<td>lon</td>
<td>longitude</td>
<td></td>
<td>degrees_east</td>
</tr>
<tr>
<td>ice_concentration</td>
<td>sea_ice_area_fraction</td>
<td>sea ice concentration</td>
<td>%</td>
</tr>
<tr>
<td>concentration_range</td>
<td></td>
<td>concentration range</td>
<td>%</td>
</tr>
</tbody>
</table>

- Select map projection

Submit  Back to results
Basket

Operations

- Choose an operation -

- Choose an operation -

Download

<table>
<thead>
<tr>
<th>Item added</th>
<th>IID</th>
</tr>
</thead>
<tbody>
<tr>
<td>35 sec</td>
<td>4</td>
</tr>
<tr>
<td>30 min 7 sec</td>
<td>2</td>
</tr>
</tbody>
</table>

Transform

<table>
<thead>
<tr>
<th>Item added</th>
<th>IID</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 min 7 sec</td>
<td>3</td>
</tr>
</tbody>
</table>

Visualize

<table>
<thead>
<tr>
<th>Item added</th>
<th>IID</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 min 7 sec</td>
<td>3</td>
</tr>
</tbody>
</table>
Search results

- **Dataset name**: Snow height layer at the WestSvalbard research site, Davos, Switzerland
  - **Institutions**: Metzner, Visalea
  - **Project**: ASCI
  - **Keywords**: EARTH SCIENCE > CRYOSPHERE > SNOW ICE
  - **Abstract**: Snow pillow instrument measuring snow water equivalent at the Weissefjellsbreen experimental site, Davos, CH. The experimental site is at the Weissefjellsbreen (WfJ), 46.83 N, 9.81 E and is located at an altitude of 2545 m in the Swiss Alps near Davos. During the winter months, most precipitation falls as snow at this altitude. As a consequence, a continuous seasonal snow layer builds up every winter, with a maximum snow height ranging from 150-360 cm over the period 1984-2012. The measurement site is located in an almost flat part of a southeast-oriented slope. LICENSE: This dataset is made available under the Open Database License: http://opendatacommons.org/licenses/odbl/1.0. Any rights in individual contents of the database are licensed under the Database Contents License: http://opendatacommons.org/licenses/dbcl/1.0. (C) Copyright WSL Institute for Snow and Avalanche Research (SLF).
  - **Collection period**: 2017-02-04T12:00:00Z to 2017-04-11T12:00:00Z

- **Dataset name**: Lysimeter at the WestSvalbard research site, Davos, Switzerland
  - **Institutions**: Metzner, Visalea
  - **Project**: ASCI
  - **Keywords**: EARTH SCIENCE > CRYOSPHERE > SNOW ICE > TERRESTRIAL HYDROSPHERE > SURFACE WATER
  - **Abstract**: Lysimeter measuring snow water runoff at the Weissefjellsbreen experimental site, Davos, CH. "Lysimeter" values are the number of bucket fills (0.8 l bucket,Sfcm). Lysimeter current is a 10min-value, Snow Water Equivalent (SWE) values are calculated as [Lysimeter/0.8]. SWE-values are in mm. **The experimental site at the Weissefjellsbreen (WfJ), 46.83 N, 9.81 E and is located at an altitude of 2545 m in the Swiss Alps near Davos. During the winter months, almost all precipitation falls as snow at this altitude. As a consequence, a continuous seasonal snow cover builds up every winter, with a maximum snow height ranging from 150-360 cm over the period 1984-2012. The measurement site is located in an almost flat part of a southeast-oriented slope. LICENSE: This dataset is made available under the Open Database License: http://opendatacommons.org/licenses/odbl/1.0. Any rights in individual contents of the database are licensed under the Database Contents License: http://opendatacommons.org/licenses/dbcl/1.0. (C) Copyright WSL Institute for Snow and Avalanche Research (SLF).
  - **Collection period**: 2017-02-04T12:00:00Z to 2017-04-11T12:00:00Z

- **Dataset name**: Meteorological station at the WestSvalbard research site, Davos, Switzerland
  - **Institutions**: Metzner, Visalea
  - **Project**: ASCI
  - **Keywords**: EARTH SCIENCE > ATMOSPHERE > PHYSICS > ATMOSPHERE > ATMOSPHERIC TEMPERATURE
  - **Abstract**: Meteorological stations at the Weissefjellsbreen experimental site, Davos, CH. The experimental site at the Weissefjellsbreen (WfJ), 46.83 N, 9.81 E and is located at an altitude of 2545 m in the Swiss Alps near Davos. During the winter months, almost all precipitation falls as snow at this altitude. As a consequence, a continuous seasonal snow cover builds up every winter, with a maximum snow height ranging from 150-360 cm over the period 1984-2012. The measurement site is located in an almost flat part of a southeast-oriented slope. LICENSE: This dataset is made available under the Open Database License: http://opendatacommons.org/licenses/odbl/1.0. Any rights in individual contents of the database are licensed under the Database Contents License: http://opendatacommons.org/licenses/dbcl/1.0. (C) Copyright WSL Institute for Snow and Avalanche Research (SLF).
  - **Collection period**: 2016-10-17T12:00:00Z to 2017-04-11T12:00:00Z
Time series plot

Meteorological data timeseries for the Weissfluhjoch versuchsfeld meteo mast station

X-axis variable: time

Y-axis variable: surface temperature (K)

Data points every Nth point:

Output file format:
ASCII data download

CSV

<table>
<thead>
<tr>
<th>Standard name</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>air_temperature</td>
<td>K</td>
</tr>
<tr>
<td>relative_humidity</td>
<td>t</td>
</tr>
<tr>
<td>surface_temperature</td>
<td>K</td>
</tr>
<tr>
<td>time</td>
<td>minutes since 2016-01-01 00:00:00</td>
</tr>
<tr>
<td>wind_direction</td>
<td>degree</td>
</tr>
<tr>
<td>wind_speed</td>
<td>m/s</td>
</tr>
<tr>
<td>wind_speed_of_gust</td>
<td>m/s</td>
</tr>
</tbody>
</table>

Output format:
CSV
# Search results

Number of datasets found: 196

<table>
<thead>
<tr>
<th>Dataset name</th>
<th>Institutions</th>
<th>Project</th>
<th>Keywords</th>
<th>Abstract</th>
<th>Collection period</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNOP data from station EVENSTAD DH</td>
<td>Norwegian Meteorological Institute</td>
<td></td>
<td>Atmospheric, Clouds, Precipitation, Temperature</td>
<td>Synoptic meteorological measurements from EVENSTAD DH extracted from the WMO Global Telecommunication System (GTS). Data are not quality controlled after extraction from GTS.</td>
<td>2018-01-01T12:00:00Z to</td>
</tr>
</tbody>
</table>
Time series plot

SYNOP data from station VEGA-VALLSJO

X-axis variable
- time

Y-axis variable
- air_temperature

Plot every nth data point

Output file format
- CSV

Submit
<table>
<thead>
<tr>
<th>Standard name</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>air_pressure</td>
<td>hPa</td>
</tr>
<tr>
<td>air_pressure_at_sea_level</td>
<td>hPa</td>
</tr>
<tr>
<td>air_temperature</td>
<td>Celsius</td>
</tr>
<tr>
<td>dew_point_temperature</td>
<td>Celsius</td>
</tr>
<tr>
<td>precipitation_amount</td>
<td>kg m^-2</td>
</tr>
<tr>
<td>relative_humidity</td>
<td>percent</td>
</tr>
<tr>
<td>sea_surface_temperature</td>
<td>Celsius</td>
</tr>
<tr>
<td>thickness_of_snowfall_amount</td>
<td>cm</td>
</tr>
<tr>
<td>time</td>
<td>days since 1970-01-01 00:00:00</td>
</tr>
<tr>
<td>wind_from_direction</td>
<td>degree</td>
</tr>
<tr>
<td>wind_speed</td>
<td>m/s-1</td>
</tr>
<tr>
<td>wind_speed_of_gust</td>
<td>m/s-1</td>
</tr>
</tbody>
</table>