Making Quantarctica-Friendly Datasets

Norwegian Polar Institute
Tromsø, Norway
February 2018

Part of Quantarctica's utility comes from its consistency and stability; layers are stored in common GIS formats, render quickly, are provided in the same coordinate system, and their metadata can easily be found in several consistent locations.

The Quantarctica Project Team is publishing the following guidelines for Antarctic data authors who wish to make their data available in a form that is easy for Quantarctica users to access, open, view, and analyze in their own Quantarctica/QGIS project environment.

Summary

- Formats: ESRI Shapefile (vector) or GeoTIFF (raster)
- Projection/CRS: Antarctic Polar Stereographic (EPSG:3031)
- Export .QLR (layer definition) file to retain layer styles, labels, and metadata
- Include detailed metadata in the Layer Metadata box and in a plain text file

Vector (points, lines, polygons)

Format: ESRI Shapefile, commonly displayed as ***.shp**. Note that a shapefile is actually multiple files that share the same name:

- Mandatory associated files: .shp, .shx, .dbf
- Other files that may be packaged: .aih, .ain, .atx, .cpg, .fbn, .fbx, .ixs, .mxs, prj, .qix, .qlr, .qml, .sbn, .sbx, .shp.xml

Projection: Antarctic Polar Stereographic (EPSG:3031)

- Other Antarctic projections, such as "USGS Transantarctic Mountains," "South Pole Stereographic," or "South Pole Lambert Azimuthal Equal Area" are not accepted.

Attribute Table: Individual features are encouraged to have at least one unique identifier in their attribute table, such as a Name, ID, or Station #, when applicable.

Styling and Labeling: If the data provider wants to display the dataset pre-styled or pre-labeled (i.e. not just accepting the QGIS defaults) in Quantarctica, they will need to style and label the layer themselves. Users will then see these style and layer settings as long as .QLR/.QML auxiliary files (see below) are provided.

Metadata: See the **Layer Metadata** section for information on including metadata in the QGIS layer properties and in the dataset folder.

Raster (grid)

Format: GeoTIFF, file extension ***.tif**. GeoTIFFs are TIF images which include georeferencing information in their file headers; thus, not all TIFs are GeoTIFFs. GeoTIFFs must be produced using GIS software.

Projection: Antarctic Polar Stereographic (EPSG:3031)

- NOTE: If you have re-projected your raster from a different map projection/coordinate system, then it is very likely the pixel resolution has changed. The pixel resolution is left to the judgment of the author, and there is no perfect value to choose. We recommend a value that is high-resolution enough to show important details, but not so high-resolution that it misrepresents the resolution of the original dataset. **Make a note of both the original and re-projected pixel resolution of your raster.**
- Other Antarctic projections, such as "USGS Transantarctic Mountains," "South Pole Stereographic," or "South Pole Lambert Azimuthal Equal Area" are not accepted.

***(optional) Compression**:* Applying compression to raster files can significantly decrease filesize. The data provider should judge the pros and cons of using lossless or lossy compression.

To compress GeoTIFFs, we recommend using the **gdal_translate** tool, from the command line, the QGIS Processing Toolbox, or from QGIS in *Raster > Conversion > Translate (Convert Format)…*

Documentation for gdal_translate can be found here: http://www.gdal.org/gdal_translate.html

- Lossless Compression (LZW; Deflate, PackBits, LZ77, RLE)
  - Retains original pixel values and image quality
  - LZW is the most common recommended method. The Quantarctica data package most often uses Deflate with ZLEVEL=9 for slightly smaller filesize at the cost of more processing power required to display.
- Lossy Compression (JPEG, JPEG2000)
  - Use when the original pixel values are not important (i.e. visible light imagery, scanned paper maps)
  - Smaller filesize than lossless compression
  - Often has a "Quality" setting (0-100 or 0-10) when exporting. The data provider should judge how to set this.
  - We recommend JPEG2000 compression over JPEG, as JPEG will only work for images under a certain number of pixels.

(optional) Pyramids: If your raster file takes a long time to render in QGIS, especially when panning and zooming around the map, it may make sense to create external Pyramids for the image. Pyramids are reduced-resolution images representing the raster that display when zoomed out, so that QGIS doesn't need to constantly re-render the image at full resolution. Pyramids do not alter the original raster pixel values.

Build Pyramids by Right-clicking the layer in the *Layers Panel > Properties > Pyramids (tab)*. We strongly recommend setting Overview Format to **External** so that the original GeoTIFF file is not modified. We leave all other settings up to the data provider's judgment. See more information on building pyramids in the QGIS Manual here: https://docs.qgis.org/2.18/en/docs/user_manual/working_with_raster/raster_properties.html?highlight=pyramids#pyramids-properties

You can also create pyramids in the command line using the **gdaladdo** command: http://www.gdal.org/gdaladdo.html

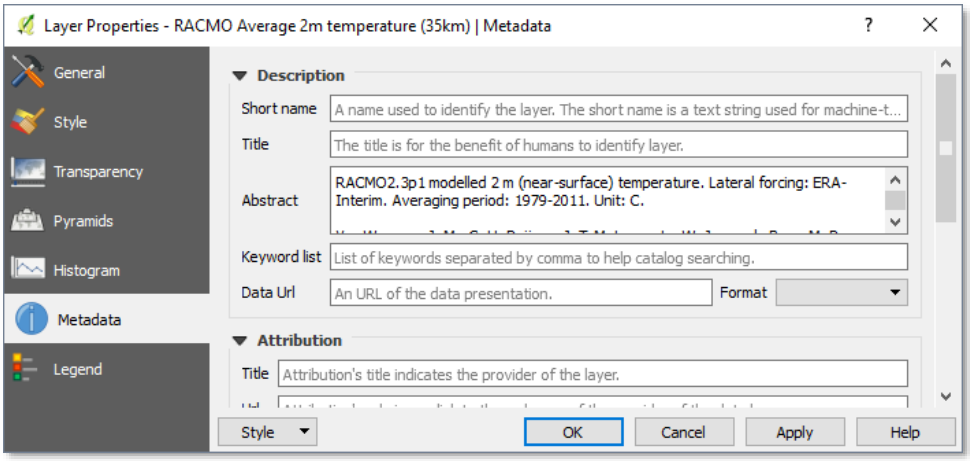
Filesize and Compression Examples: Selected raster layers included in Quantarctica 3

Layer	Resolution	Dimensions	Filesize	Compression	Bit Depth
LIMA Landsat high-resolution virtual mosaic	15 m	370000 x 310000	975 MB (1.35 GB with pyramids)	JPEG	Byte
RAMP2 DEM	200 m	28680 x 24580	196 MB (257 MB with pyramids)	LZW	Int16 signed
MEaSUREs Ice flow speed	450 m	12018 x 10693	223 MB (296 MB with pyramids)	Deflate	Float32
IBCSO Elevation model, surface	500 m	13335 x 13341	172 MB (222 MB with pyramids)	Deflate	Int16 signed
RACMO Average 2m temperature	35 km	225 x 236	209 KB (no pyramids)	None	Float32

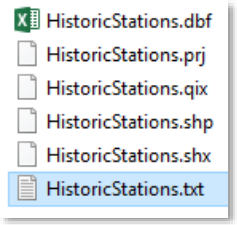
Metadata

Metadata for each layer should be written in two locations: in the QGIS layer metadata, and in the dataset folder.

QGIS Layer Metadata: *Right-click the layer in the Layers Panel > Metadata (tab) > Abstract (box)*. Remember that this metadata will only be saved if you also save a .QLR file for your layer (see section **Auxiliary Files**).



Text File in the Dataset Folder: Create a plain-text, .txt file inside the dataset folder with the same name as your layer filename, e.g. *my_layer.txt* corresponding to *my_layer.shp.*



Metadata Specifications

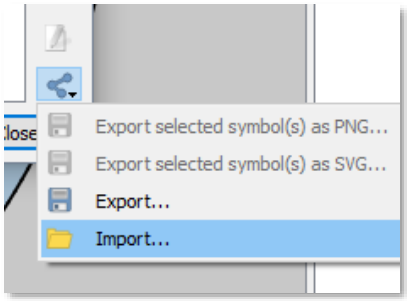
We recommend including the following pieces of information in your metadata:

1. **Layer Title:** In most cases, this will simply be the same as the name of the layer in QGIS, and is the choice of the author.
2. **Abstract:** A brief description of the dataset. The writing level should be appropriate for scientific users who may be outside your particular discipline. Include items such as the time and date of data collection, whether the dataset was observed or modelled, the instrument or platform that acquired the data, any important parameters or tuning, processing methods, averaging or binning and the **units** of the dataset.
   For raster data, list the **pixel resolution** (both original and re-projected, if applicable).
   For vector data, give a note about the **accuracy** and **precision** of the data. For example, point coordinates acquired with **minute** precision may differ from their actual locations by up to 1.8km.
3. **Citation(s):** Citation (in APA format or similar) for the dataset. Whether this citation corresponds to the entry for this dataset in a data portal, or the scientific publication/journal article containing this dataset is up to the author.
4. **URL:** A link to the data portal/download location for the original dataset, and for the original publication, if not already included in the citation.
5. **Published Date**
6. **Download/Processing Date**
7. **Author Contact Information:** Name and email address

For examples, see the Quantarctica Data Catalog: http://quantarctica.npolar.no/data-catalog/

(optional) Layer Styles and Labeling

The following section describes Quantarctica does not dictate strict style and label requirements for Quantarctica-friendly datasets, so choices for how to best display a dataset are left to the data provider. Below, we describe some general guidelines for how layers in the default Quantarctica data package are consistently styled and labeled, that we hope will serve as useful examples.



Color Ramp: In addition to the color ramps provided by QGIS, we provide our customized, colorblind-friendly Quantarctica "rainbow" color ramp in the Quantarctica folder. To import this color ramp into QGIS for convenient use, navigate to *Settings > Style Manager >* *[bottom-right of window] > Import…* and select *Quantarctica_ColourRamp.xml* from the Quantarctica3 folder.

We also use the following color ramps for other types of data:

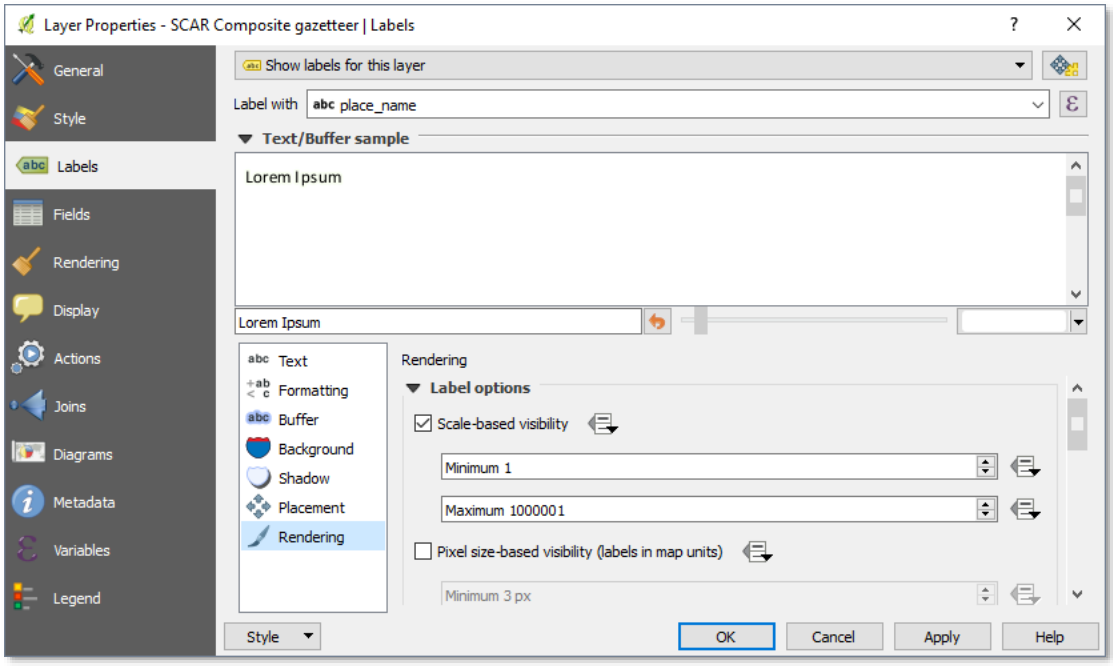
- Increasing/"low to high":
  - Custom white-to-purple (rgb 255-0-255) , white-to-orange (rgb 255-127-0), , or similar.

- Diverging/"negative and positive values where 0 is important":
  - Custom Pink-White-Orange with White at the 0 value:
  - Elevation/bathymetry: Black-blue-white (0)-green-yellow-red, includes very light blue value at -1 and very light green at +1 to create a clear break at 0.

Creating and modifying color ramps works similarly for raster and vector data.
Read more about creating and using color ramps here:
https://docs.qgis.org/2.18/en/docs/user_manual/working_with_vector/style_library.html#color-ramp

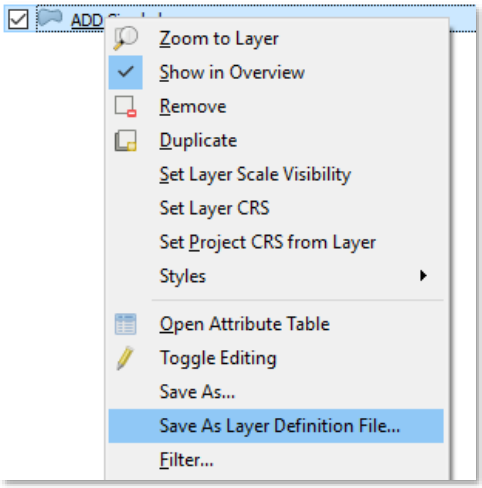
Labeling: Quantarctica uses the following labeling settings by default for vector layers.
Read more about labeling settings here:
https://docs.qgis.org/2.18/en/docs/user_manual/working_with_vector/vector_properties.html#labels-properties

- ✓ Font: Calibri
- ✓ Size: 10pt
- ✓ Text Buffer: Enabled, Size 1mm
- ✓ Drop shadow: Enabled, Offset: 135°, 1mm, 1.5mm blur radius, 30% Transparency, Blend Mode: Multiply
- ✓ Placement: Cartographic, 0.5mm distance offset From symbol bounds
- ✓ Scale Dependence: If desired, you can specify a minimum and maximum scale for labels to appear at, in *Layer Properties > Labels [tab] > Rendering [tab] > Scale-based visibility:*



Layer Definition File

Include a .QLR "Layer Definition File" with your dataset to retain layer name, style, metadata, labeling, and query filter information for raster layers, or style, metadata, and label information for vector layers when a user opens the layer in QGIS. Users can then open your layer exactly the way you intend for it to be seen.



- Save a QLR by *right-clicking the layer in the Layers Panel and selecting* ***Save As Layer Definition File…***
- Give the layer the same name and location as the shapefile, e.g. *my_layer.qlr* corresponding to *my_layer.shp*
- Users will need to open your shapefile by either dragging the .QLR file into their QGIS window, or by navigating to Layer > Add from Layer Definition File…

Checklist

- ✓ Vector format: ESRI Shapefile (.SHP)
- ✓ Raster format: Geotiff (.TIF)
- ✓ Projection/Coordinate System/CRS: EPSG:3031 Antarctic Polar Stereographic
- ✓ Include metadata (Layer Title, Abstract, Citation(s), URL, Published/Download Date, Author Contact Info) in the QGIS layer and in a .txt file in the dataset folder
- ✓ (optional) Layer Styling and Labeling
- ✓ (optional, raster): LZW/Deflate lossless compression, generate external Pyramids
- ✓ Export .QLR Layer Definition File for each layer to retain metadata and styling

You can examine datasets already included in the Quantarctica 3 data package to see ideas and standards for how to display and package your own Quantarctica-friendly data.

If you have any questions about Quantarctica or providing Quantarctica-friendly datasets, send us an email at quantarctica@npolar.no .

Thanks for your contribution!
-The Quantarctica Project Team