

How to Georeference a Historical Map in QGIS

Objectives

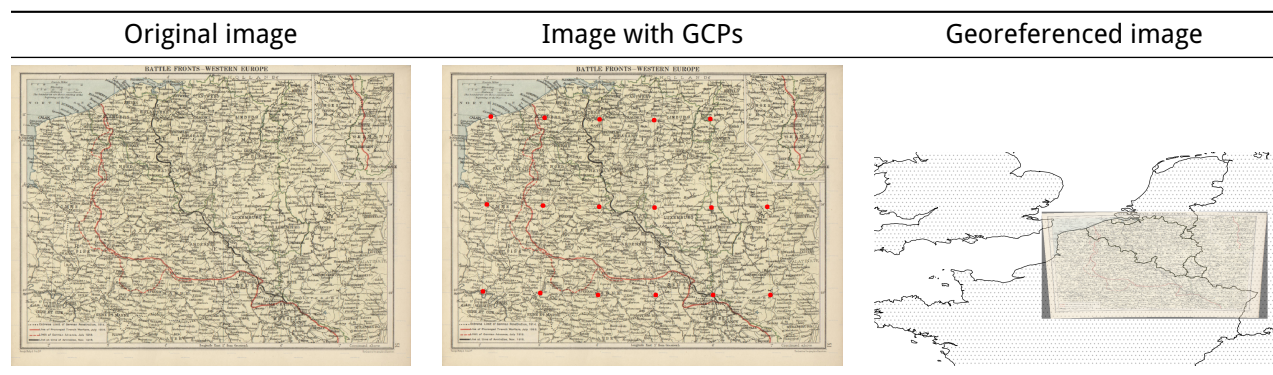
Our **immediate goal** is to assign spatial coordinates to scanned map images. The **down-stream analytical goal** is to overlay historical cartographic data onto modern base maps for comparison and statistical analysis.

Georeferencing allows digital images to be accurately positioned on Earth's surface. Think of it as teaching a computer to understand that a scanned historical map or satellite image represents a specific location on our planet, rather than just a collection of pixels. This process transforms raw imagery into actionable intelligence by identifying ground control points – recognizable features with known coordinates – that serve as anchors to mathematically align the entire image with geographic reality.

In military intelligence applications, georeferencing enables critical capabilities like pattern analysis and threat assessment, by overlaying historical reconnaissance imagery with current satellite data, to detect changes in enemy infrastructure, construction of new installations, troop movements, or weapons deployments over time. For political economy analysis, georeferencing allows researchers to study urban development patterns, analyze resource distribution changes over time, or examine how political boundaries have shifted in relation to economic infrastructure (e.g. comparing Soviet-era industrial maps with current economic activity to understand post-transition regional development). Georeferencing converts static images into dynamic analytical tools that can reveal spatial relationships, temporal changes, and other insights that would be impossible to discern from raw images alone.

The basic workflow of georeferencing involves the creation of Ground Control Points (GCPs), followed by a transformation of the map's surface to align it with real-world locations. GCPs serve as “anchor points” that allow the georeferencing algorithm to transform the unreferenced map into the correct spatial position. The accuracy and distribution of GCPs fundamentally determines the quality of the georeferenced product.

The instructions below tell you how to do this in QGIS.



Setting Up Your Project

0. Create a QGIS Project File to Back Up Your Work

- Open QGIS.
- Before starting, create a QGIS project file to save your workspace:
 - Go to Project > Save As in the main menu.

- Choose a location and name for your project file, ensuring the file is saved as a .qgz (default) or .qgs format.
- This project file will save your layer configurations, symbology, and other settings, allowing you to restore your workspace later if needed.
- Save your work periodically by clicking **Project > Save** or using the keyboard shortcut:
 - *Windows/Linux*: Ctrl+S
 - *macOS*: Cmd+S

1. Add OpenStreetMap Base Layer

- Method 1:
 - Open QGIS and navigate to **Browser Panel** (if not visible, go to **View > Panels > Browser**).
 - In the **Browser panel**, right-click on **XYZ Tiles** and select **New Connection**.
 - In the dialog box, enter "OpenStreetMap" as the name and `http://tile.openstreetmap.org/{z}/{x}/{y}.png` as the URL.
 - Click **OK**.
 - Double-click the new "OpenStreetMap" entry under **XYZ Tiles** to add it to your map canvas.
- Method 2:
 - Go to menu **Layer > Add Layer > Add XYZ Layer**.
 - In the **Data Source Manager** window, click **New** to create a new XYZ connection.
 - In the **New XYZ Connection** dialog, enter:
 - * **Name**: "OpenStreetMap"
 - * **URL**: `https://tile.openstreetmap.org/{z}/{x}/{y}.png`
 - Click **OK**.
 - Select "OpenStreetMap" from the XYZ connections list and click **Add**.
- The OpenStreetMap basemap should appear in your **Layers** panel and map canvas.

2. Set Project Coordinate Reference System

- Look for the CRS button in the bottom-right corner of the QGIS interface. It should look like a globe, with some text like "EPSG" followed by some numbers.
- If the button *does not* say "EPSG:4326", click on the **CRS** button. In the **Project Properties** dialog, search for "4326" and select "EPSG:4326 - WGS 84". Click **OK**.
- EPSG:4326, also known as WGS 84, is a geographic coordinate reference system that uses latitude and longitude to define locations on the Earth's surface. It is widely used in GPS and mapping applications for its accuracy in representing the Earth's shape as a three-dimensional ellipsoid.
- **Note**: Another popular CRS is EPSG:3857 - WGS 84 / Pseudo-Mercator, which is the de facto standard for web mapping applications. When we're creating a web map that uses tiles from services like OpenStreetMap, these tiles are delivered in EPSG:3857 by default. For georeferencing purposes — where degrees of latitude/longitude are often more convenient than meters — we need to switch to 4326 to ensure proper alignment.

3. Open Georeferencer

- Go to **Layer > Georeferencer** to open the Georeferencer window.
 - In older versions of QGIS, this is found in **Raster > Georeferencer > Georeferencer**.
 - If you cannot find the Georeferencer, you may be missing a plugin. Navigate to **Plugins > Manage and Install Plugins**. In the search box, type "georef" and select **Georeferencer GDAL**. Click **Install Plugin**. Close the plugins window and try opening the Georeferencer again.

4. Load Historical Map

- In the Georeferencer window, click the **Open Raster** button (top left) and select your historical map image file from the **Raw** directory. These will end in a file extension like .tiff, .jpg, .png or other

raster image format.

5. Configure Transformation Settings

- In the Georeferencer window, go to Settings > Transformation Settings. Set the following:
 - *Transformation type*: Polynomial 2 (for most historical maps covering large territories, second-order polynomial provides the best balance of accuracy and stability)
 - *Output file*: click on the three dots (...) next to the "File name" field to specify the exact folder/path (use the Working folder in YZRA/Data/SovietAtlas/Working) and file name for the georeferenced map (e.g., 1982_RailAtlas_Estonia_georeferenced.tif) for your working copy. Avoid saving to a temporary file location (e.g., /tmp/ or similar paths), as these files may be deleted when QGIS is closed.
 - *Resampling method*: Nearest Neighbor
 - *Target CRS*: EPSG:4326
 - Check Load in QGIS when done
 - Check Save GCP points
- Click OK.

Method A: Georeferencing with Graticule Line Intersections

Graticule lines are a network of lines on a map that represent the Earth's geographic coordinate system – specifically the *parallels of latitude* (running east-west) and *meridians of longitude* (running north-south). These lines are tied directly to the Earth's ellipsoidal shape and are always expressed in geographic coordinates (degrees of latitude and longitude).

Graticule lines should not be confused with *grid references* like A1, B2, etc., which are alphanumeric location systems that divide maps into a specified number of rows and columns, often used in atlases to help users locate features listed in an index. Grid references appear as a network of evenly spaced horizontal and vertical lines that create rectangular cells labeled with letters and numbers. In contrast, graticule lines may appear curved on projected maps as they follow the true form of meridians and parallels, and they're labeled with actual coordinate values in degrees.

The easiest way to identify graticule lines is to examine the labels along the map margins. Graticule lines will be labeled with: - *Latitude values* (e.g., 45°N, 50°N, 55°N) along the left and right edges - *Longitude values* (e.g., 10°E, 15°E, 20°E) along the top and bottom edges

These coordinate values represent actual geographic positions on Earth's surface. Grid reference systems, by contrast, will show simple alphanumeric labels (A, B, C or 1, 2, 3) that are specific to that particular map's indexing system and have no universal geographic meaning.

1: Identify Coordinate Information

- Examine the historical map for latitude and longitude markings along the edges or at intersections.
- If there are no graticule lines, skip ahead to Method B.

2: Add Control Points

- In the Georeferencer window, click the Add Point button (crosshair icon).
- Position the crosshair precisely at a graticule intersection and click to place a control point.
- Manually enter the latitude and longitude values visible on the map (e.g., x=24, y=56)
- Repeat this process for all graticule intersections you can identify.
- **Note:** The distribution of GCPs should follow several key principles:
 - GCPs should be evenly distributed across the entire map extent rather than clustered in one area.

- GCPs should be placed around the perimeter of your map area, along with multiple GCPs near the center of the domain to improve overall transformation quality.
- Avoid placing GCPs in a single localized cluster, or arranged in a straight line, as this can create instability in the transformation calculations.

3. Check GCP Errors and Residuals

- In the QGIS Georeferencer, residual values are displayed in the *GCP table* at the bottom of the window. The table shows:
 - dX (Error) and dY (Error): horizontal and vertical residuals in map units
 - Residual: total distance error calculated as the square root of the sum of squared differences in coordinate locations
 - Mean Error: overall transformation accuracy shown at the bottom
- Individual GCP residuals should typically be less than 1-2 pixels of the source image, although this can go a bit higher for large-scale maps. You should aim to keep these numbers below 1 or in the low single digits for good results.
- To remove problematic GCPs (e.g. those with high residual values), select the point on the GCP Table and right-click to choose *Delete GCP*.
- To replace the GCPs you removed with new GCPs, click *Add Point* and follow the same process as before

Method B: Using Points of Interest with Map Canvas Selection

Many (most) maps do not have graticule lines. But we can use landmarks and prominent geographic features (whose real-world locations are well-known) as a replacement or supplement to graticules. Examples include well-known buildings, crossroads, hills, cities, coastal features, curves in rivers, and borders.

Using both graticule lines and landmark features to establish Ground Control Points provides optimal georeferencing results because each method addresses different spatial and accuracy requirements. Graticule intersections offer systematic geometric distribution and precise coordinate values that can be directly entered from map margins, while landmark features fill spatial gaps where graticule coverage is inadequate and provide ground truth validation to detect potential systematic errors or projection distortions in the historical map. This complementary approach allows us to achieve a fuller perimeter and interior GCP coverage pattern, placing control points at graticule intersections around map edges where available, then using landmark features to add points in central areas and regions with sparse graticule coverage. The combination also provides the redundancy necessary for reliable transformations, typically resulting in at least 20-30 well-distributed control points that ensure both geometric precision from objective coordinate references and spatial accuracy validated through recognizable geographic features.

1: Identify Reference Feature

- Look for permanent geographic features that appear on both your historical map and the modern OSM base layer in the main QGIS window:
 - Coastal features (headlands, bays, river mouths)
 - Major road intersections
 - Mountain peaks or ridges
 - River confluences
 - Town centers or significant buildings

2: Add Control Points

- In the Georeferencer window, click the *Add Point* button and place the crosshair precisely on a recognizable feature.
- In the coordinate entry dialog, click *From Map Canvas*.

- This will switch you to the main QGIS window with your OSM base layer.
- Navigate to the same geographic feature on the modern map, zoom in for precision, and click on the exact corresponding location.
- The georeferencer will automatically populate the X and Y coordinates and return you to the Georeferencer window.
- Repeat these steps until your GCPs are about evenly distributed throughout the map surface.

3. Check GCP Errors and Residuals

- As described above for Method A, check the residual values displayed in the *GCP table*, removing and replacing GCPs with high residual values.

Running the Transformation

Once you have sufficient control points (minimum 6 for second-order polynomial, but 30+ recommended), click the **Start Georeferencing** button (green play icon) in the Georeferencer toolbar.

Quality Control and Refinement Process

1: Initial Visual Inspection

- After transformation completes, the georeferenced map will appear in your main QGIS project.
- Right-click on the georeferenced layer and select **Properties > Symbology**.
- Set the **Opacity** to about 50%, so you can see both layers simultaneously.

2: Check Alignment

- Zoom to various locations across the map and compare recognizable features between the historical map and OSM base layer.
- Look for misalignments, particularly at edges and corners.

3: Return to Georeferencer for Refinement

- If the misalignments are significant (this is a subjective assessment you will need to make as a GIS analyst), you can go back and modify the georeferencing to get a better result.
- Go to **Layer > Georeferencer** and reload your original map image.
- Your GCP points should reload automatically if you saved them.

4: Modify Control Points

- You'll want to focus on the part of the map where misalignment appeared most severe in visual inspection.
- Check that no individual GCP has a residual value more than 2-3 times the mean error, and that the overall RMSE meets your project requirements.
- Check to see if you have any gaps in GCP coverage in the problematic area.
- To remove problematic GCPs, select the point and right-click to choose **Delete GCP**.
- To move a GCP, select the point in the GCP table, right-click and choose **Move GCP**.
- To edit coordinates, double-click coordinate values in the GCP Table.
- To add new GCPs, click **Add Point** and follow the same process as before.

5: Iterative Refinement

- After making changes, click `Start Georeferencing` again and repeat the visual inspection.
- Continue this process until alignment meets minimal accuracy standards and GCP residuals are within acceptable ranges.

Save the Final Georeferenced Raster

- Once editing is complete, and you are ready to commit all changes, copy your files to the `Processed` folder:
 - The georeferenced raster file (e.g., `1982_RailAtlas_Estonia_georeferenced.tif`) can be backed up like any other file by simply copying it from the `Working` to the `Processed` folder. When you complete the georeferencing process in QGIS, the output raster (typically saved as a GeoTIFF file) contains all the spatial reference information embedded within it. When copying this file, ensure you maintain the original filename and extension.
 - The Ground Control Points are automatically saved in a separate text file with a `.points` extension that shares the same base name as your raster file. For example, if your georeferenced raster is named `"historical_map_modified.tif"`, the GCP file will be `"historical_map_modified.tif.points"`. This points file contains coordinates in the format `mapX, mapY, pixelX, pixelY`.
 - For comprehensive backup, copy both the completed georeferenced raster file and its associated `.points` file to the `Processed` directory. Keep these files together as they work as a pair – the points file references the specific pixel coordinates within the raster image.