

Estimation of the Cartographic Projection and its Application in Geoinformatics-habilitation thesis presentation

Tomáš Bayer

*Department of Geomatics, Faculty of Civil Engineering, Czech Technical University
in Prague, Czech Republic
bayertom@natur.cuni.cz*

Abstract. Modern techniques for the map analysis allow for the creation of full or partial geometric reconstruction of its content. The projection $\mathbb{P}(\varphi_k, \lambda_k, \varphi_1, \lambda_0, \kappa)$ is described by the set of estimated constant values: transformed pole position $[\varphi_k, \lambda_k]$, standard parallel φ_1 , longitude λ_0 of the central meridian, and constant parameter κ . Analogously the analyzed map $M(R', \Delta X, \Delta Y, \alpha)$ is represented by its constant values: auxiliary sphere radius R' , origin shifts $\Delta X, \Delta Y$, and angle of rotation α . Several new methods denoted as M6-M9 for the estimation of an unknown map projection and its parameters differing in the number of determined parameters, reliability, robustness, and convergence have been developed. However, their computational demands are similar. Instead of directly measuring the dissimilarity δ of two projections, the analyzed map M in an unknown projection and the image M' of the sphere S^2 in the well-known (i.e., analyzed) projection \mathbb{P}_x are compared. Several distance functions for the similarity measurements based on the location as well as shape similarity approaches are proposed. An unconstrained global optimization problem poorly scaled, with large residuals, for the vector of unknown parameters \hat{x} is solved by the hybrid BFGS method. To avoid a slower convergence rate for small residual problems, it has the ability to switch between first- and second-order methods. Such an analysis is beneficial and interesting for historic, old, or current maps without information about the projection. Its importance is primarily referred to refinement of spatial georeference for the medium- and small-scale maps, analysis of the knowledge about the former world, analysis of the incorrectly/inaccurately drawn regions, and appropriate cataloging of maps. The proposed algorithms have been implemented in the new version of the detectproj software.

Keywords: Map projection; analysis; detection; history of cartography; early maps; location similarity; optimization; non-linear least squares; BFGS; georeference; M-estimators; Huber function.

1. Introduction

Maps are an important part of our history and cultural heritage; close attention is paid to their study and research. Currently, large collections of digitized maps from libraries around the world are accessible. These libraries offer a huge number of maps, atlases, or globes available for viewing online. Due to their easy accessibility, they become a subject of interest of many researchers as well as the general public.

New methods and techniques for map analysis allow for the creation of full or partial geometric reconstruction of its content. This approach belongs to the category of cartometric analysis, the capabilities of which have been significantly improved with the rapid development of the computer technology. Any serious map has the content and the geometry describing the spatial relationship between objects. Working with the map content, its geometric and spatial

characteristics cannot be ignored. Underestimating or neglecting these characteristics cause the acquired information to be flawed.

For early maps (created before the 17th century), the lack of solid geometric and geodesic bases is typical. The map content was not seriously measured; therefore, it is drawn significantly less accurately. The only geometric basis can be found on the graticule as well as on the map frame. Unfortunately, outliers have a strong influence on the results, which may be skewed.

There are many different factors affecting the results. For a successful analysis, the projection impact must be stronger than the graphical accuracy of the map. An important role will be played by the properties of the analyzed territory, especially its size and geographical position. For small territories, the impact of the projection will not significantly overcome the graphical accuracy of the map, or territories located around the equator, central meridian or near to the poles, more possible candidates may appear. An important role will also be played by the amount and spatial distribution of control points, as well as cartographic techniques like generalization.

From a mathematical point of view, the analysis will lead to unconstrained optimization. Basically, two families of methods, rotation-dependent (M6, M8), and rotation-invariant (M7, M7S, M9), will be proposed. Involving a rotation prevents the problems of an additionally rotated analyzed map (a switched orientation on the page, or inappropriate insertion into the scanner). For the computation, the Cartesian coordinates of control points on the analyzed map and the spherical coordinates of the corresponding points on the sphere are required. The process of finding the best fit projection is iterative; from iteration to iteration, the determined solution is refined.

The map coordinate system is heterogeneous to the global coordinate system, unless they are based on the same projection. For serious analysis, there is a need to perform a geometrical reconstruction of the early map involving the establishment of the correct geometric position of the map content in the projected spatial coordinate system. However, without any information about the source projections, this is almost impossible. In this context, the importance of the map projection analysis is primarily referred to the refinement of spatial georeference for the medium- and small-scale maps.

For the georeferencing of maps covering a small territory (large-scale maps), the 1st order transformation is sufficient (projection impact can be neglected). However, this approach cannot be applied to small-scale maps where the map projection influence should not be ignored. Increasing the order of transformation does not lead to any reasonable result. The current and widely applied method based on splitting the map into tiles, applying a transformation to each tile, and restoring the continuous raster image from tiles, is time consuming, tedious, and less accurate.

This approach could be improved using the proposed method, when a projection of the analyzed map is detected. Using the inverse projection equations, the map is reprojected on the sphere. Subsequently, a projection to the destination coordinate system is carried out.

Based on the results of this research, for successful map projection analysis 5-10 points are sufficient for world maps, 10 points for the medium-scale maps, and 15 for the large-scale maps.

Cataloguing early maps creates the need for additional cartographic information which is

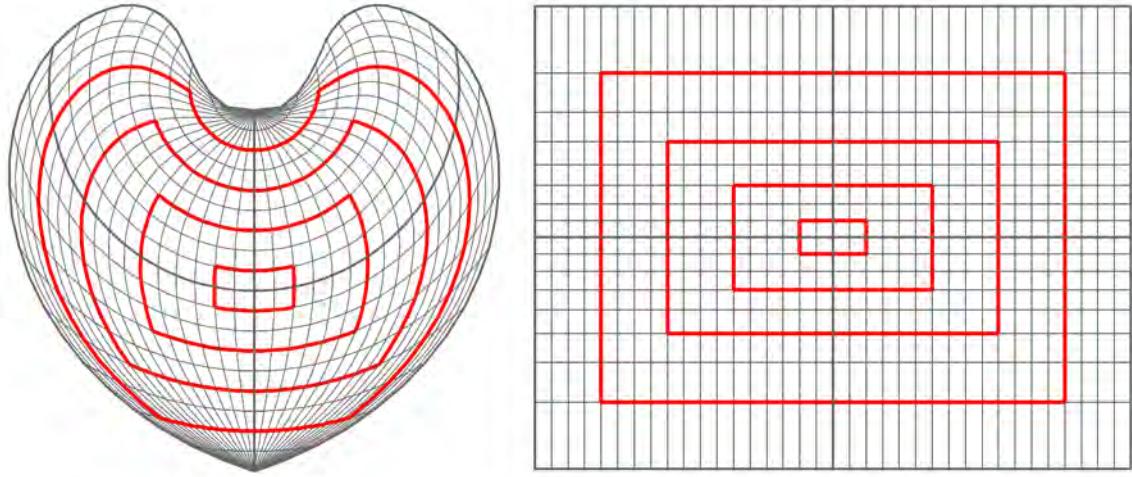


Figure 1: The impact of the territory size on its shape in two different projections: Bonne projection (left), Mercator projection (right); each projected territory has the unique shape forming the projection footprint.

part of the metadata. In particular, they include data about the geographical extent, the map projection or the map scale. The bibliographic format Marc 21 contains a detailed description of a map projection in fields 034 and 255B of the bibliographic record. Unfortunately, there was no method to determine these parameters accurately, quickly, correctly or for a large amount of maps. The real-time solution based on the non-linear least squares (NLS) method provides a tool for analyzing the map relatively quickly and with sufficient accuracy.

Using the NLS approach, only the local optimizer is guaranteed. From a wider aspect, it provides a solution of acceptable quality, but not the best. However, in most cases, the differences are below the graphical accuracy of the map.

The developed software `detectproj` supports all proposed methods and optimization techniques.

1.1. Related work

Due to the difficulty of estimating map projection parameters, especially without deep numeric analysis, this problem has not been studied in detail.

There are several early papers focused on the ancient map projections [29], the projection of a general map of Britain [25] as well as modern papers; let us mention analysis of the American Civil War maps [26], Ptolemy's map of Greece [20], the portolan maps [8], and the Gough map of Britain [19], [21]. Tobler's paper [30] emphasizes importance of numerical procedures (bivariate interpolation) for cartography, and brings the mathematical fundamentals of the projection detection, the location similarity approach is mentioned here for the first time. The new method, measuring the map projection similarity from the residuals of the corresponding points, was described in [28], the bidimensional regression for comparison of geographic phenomena in [27].

Several software tools for the projection analysis exist. The `prjfinder` software [10] searches for the best matching coordinate system. Another solution based on 2D transformations developed in [17] was implemented in the `MapAnalyst` open-source software [16] and refined in [18]. Different approach based on the Nelder-Mead optimization of the objective function describing the location/shape dissimilarities of the corresponding 0D-2D features can be found in [6], the non-linear least squares and differential evolution solutions in [3]. The proposed methods supporting the determination of additional parameters of the map projection as well as the analyzed map were implemented in open-source software `detectproj` [4]. Another detection method based on the shape of meridians/parallels assessment was described in [2]. For the projection analysis the decision trees are utilized.

2. Importance of the map projection analysis

The projection analysis, including the estimation of the best constant values of the projection \mathbb{P} and the map M constants, belongs to the new methods of the cartographic research of early maps. It represents a process of the identification, recognition, and reconstruction of the geometric relationship between the early map content and the present representation on the Earth. There is a long history of using various types of map projections, progressing from simple geometric constructions to the rigid mathematical theory, established by H. Lambert, and C. F. Gauss.

2.1. Analysis and georeference

With the increase in the amount of digitized early maps, there is a need to determine the correct geometric position, size, and dimensions in the projected coordinate system. Georeferencing assigns spatial information to each pixel of the map so as it aligns to a known projected coordinate system. The aim is to minimize the distortions and deformations that such a procedure will create.

For small territories the impact of the map projection may be neglected; see Fig. 1. The similarity relationship, between two sufficiently small territories, projected in two different projections, may be established. This approach leads to the use of linear transformations for the spatial georeferencing of large-scale maps. The current strategy of georeferencing, based on the application of different types of transformations, is not applicable to large territories (hemispheres, planispheres, continents) directly; it has only limited application for maps of small territories.

In georeferencing small-scale-maps, it is impossible to transform the analyzed map to the reference map coordinate system directly and neglect the influence of the different map projections, unless both projections are the same or very similar; see Fig. 2. While the linear transformations cannot correct the impact of the distortions, see Fig. 3, the higher-order transformation causing the secondary deformation and twist of the map content (see Fig. 4) and cannot be used; these facts are discussed in [6], [3].

So far, partitioning the analyzed map into tiles and applying a transformation to each tile with the restoration of the continuous raster image is widely applied, but it is a time-consuming and laborious approach. Avoiding the projection influence, each tile should be small enough. This may create a huge amount of tiles, each of which must be processed separately. Subsequently,

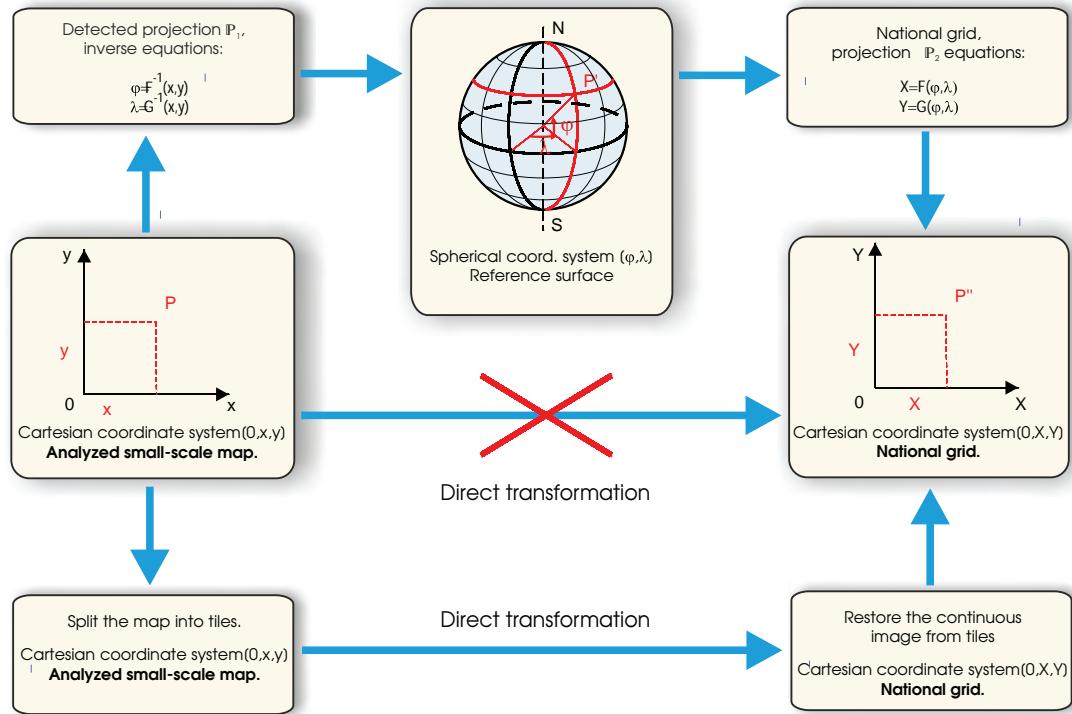


Figure 2: Georeferencing of small-scale early map to the national grid (source: [6]).

a continuous raster image needs to be restored.

The fastest, most accurate and well-defined approach is to determine the analyzed map projection and *reproject* the map from the Cartesian coordinates to spherical coordinates using inverse projection formulas. This is followed by projecting the spherical coordinates of the reprojected map to the destination coordinate system. The above-mentioned procedure is shown in Fig. 5.

2.2. Analysis and incorrectly drawn map content

The results of the map projection analysis may be utilized for the assessment of early maps. For early maps created before the 17th century, the lack of solid geometric and geodesic bases is typical. The map content was not seriously measured, and the only geometric bases can be found on the graticule, as well as on the map frame. Hence, it is drawn significantly less accurately than these map construction elements. This feature giving an answer, which territories are drawn more or less accurately may be utilized for further analysis of the map.

It is a well-known fact that frequently used shipping routes and their close territories, overseas colonies, easily accessible locations, important cities, castles, churches, and mines were drawn more accurately than places that were unattractive, trade-free, without resources, and difficult to access. Moreover, the dangerous and unexplored places on early maps stayed uncharted, with images of dragons, lines, labeled as "HC SVNT DRACONES" or "HIC SVNT LEONES" or "TERRA INCognITA". Such information is useful; it helps to increase the knowledge



Figure 3: Georeferencing of a small-scale map in Bonne projection to the destination coordinate system in the Mercator projection using the similarity transformation.

about the former understanding of the world.

Due to the lack of a geometric basis, many territories are well-placed with the distorted shape, or well-shaped with the systematic shifts, or a combination of both may occur. To reduce the impact of incorrectly drawn elements on the results, the outlier detection algorithms based on M-estimates (the Huber function) are built-in. Rejecting outliers leads mostly to the refinement of the determined parameters.

Using the results of the proposed analysis, the distorted areas are easily detectable, if the early map is reprojected to the national coordinate systems. At the superimposition of the early map and the current state, the spatial or location dissimilarities become clearly visible. Fig. 5 illustrates the situation when the east coast of Africa and Madagascar have a systematic shift, but the south coast of the Arabia and Black Sea are distorted.

2.3. Cataloging of maps

Currently, many map collections have been transformed into the digital form or the process is still in progress. To be easily accessible, they must be sorted, organized, and stored in the database, represented and organized as a catalog. Their real-time availability, together with the search options according to the various criteria accessible through cartographic metadata are beneficial. The cataloging of maps creates the need for information about the map

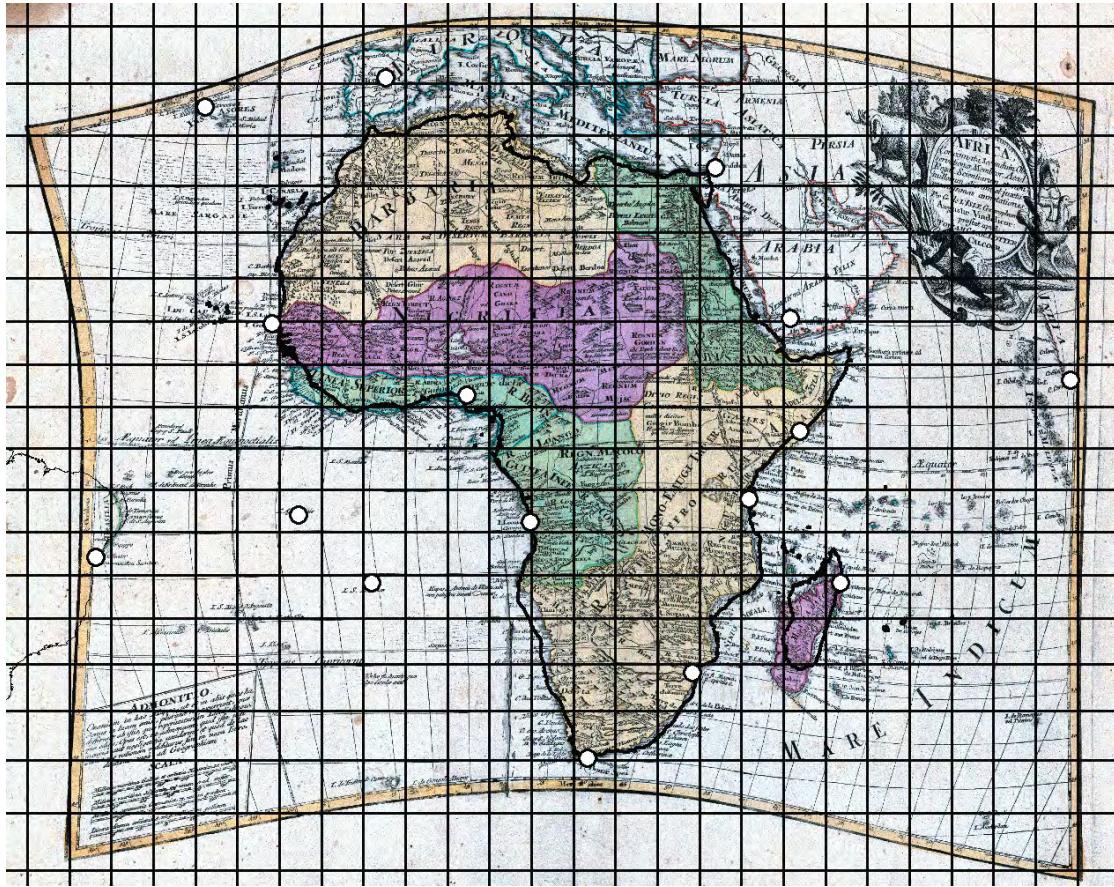


Figure 4: Georeferencing of a small-scale map in Bonne projection to the destination coordinate system in the Mercator projection using the spline transformation.

projection, which form a part of the cartographic metadata. Cartographic metadata contains a description of the mathematical, cartographic, and spatial properties of the map.

The widely used bibliographic format Marc 21 involves a detailed description of a map projection and its properties in fields 034 (Coded Cartographic Mathematical Data); 255 (Cartographic Mathematical Data); 342 (Geospatial Reference Data), see Fig. 6. However, the analogous records are also included in the INSPIRE standard.

It is clear that some parameters values may be visually estimated better (projection family) or worse (projection aspect); without additional information it is impossible (standard parallel latitude).

So far, the parameters have been estimated only visually, or the record has been left blank, if the projection description was missing.

Using the proposed solution, this step can be performed semi-automatically and with a higher degree of relevance. The tool may be useful for librarians as well as for cataloguers; it will facilitate and accelerate their work and save time.

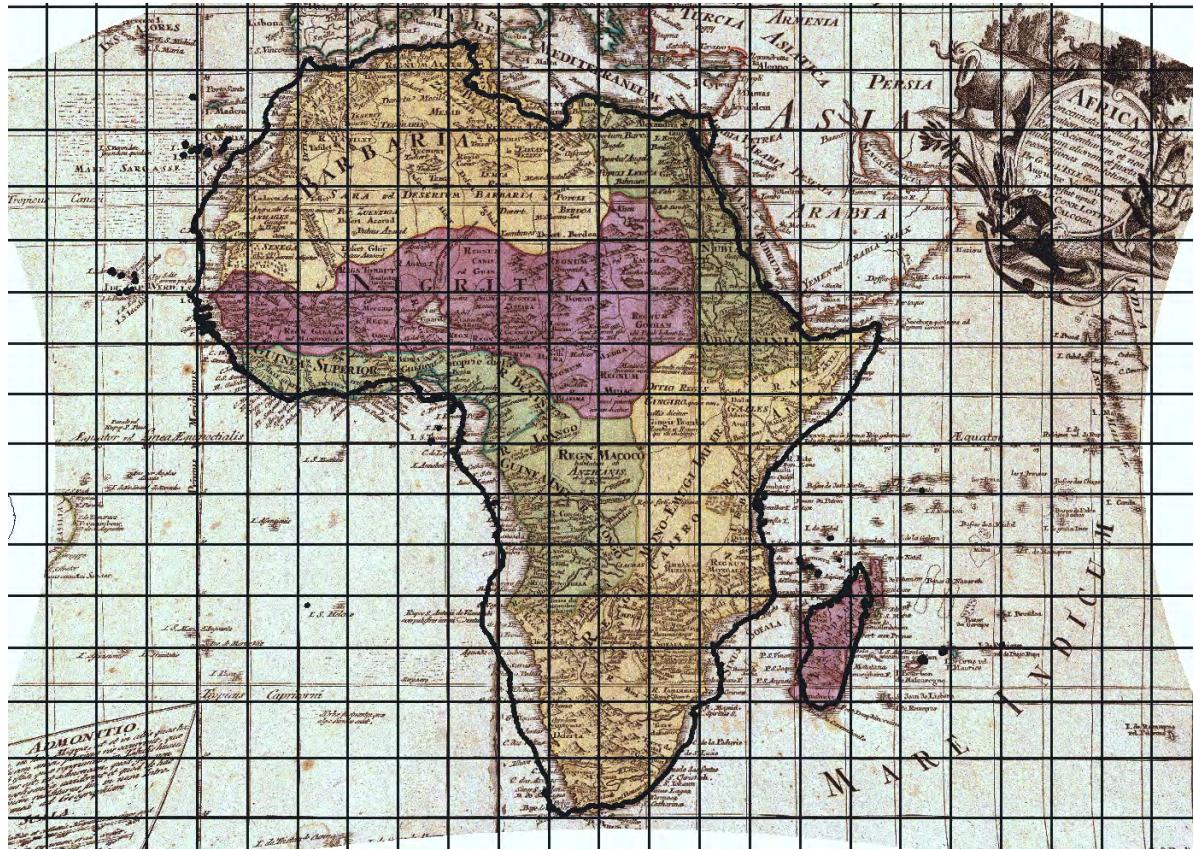


Figure 5: Georeferencing of small-scale map in Bonne projection using the proposed method to the destination coordinate system in the Mercator projection.

2.4. Factors affecting the detection

The input features of the detection process are represented by the map content or the auxiliary construction elements of the map, primarily the graticule and map frame manually collected by the user. An inappropriate choice of features negatively affects the estimated parameters and may lead to the failure of the detection process or to the assignment of a different projection. However, in many situations, the conditions cannot be met entirely.

The results of the analyses are strongly influenced by the several factors. Their neglecting may lead to the significant decrease of the analysis efficiency. Hence, it is important to be aware of their influence and try to reduce their impact on data.

Several instructions, requirements, and recommendations for handling them widely discussed in [5] are presented.

Spatial dimension of features

The proposed solution takes into account elements of different spatial dimensions. The input parameters are represented by 0D-2D features (e.g., points, lines, curves, polylines, polygons)

```
<datafield tag="245" ind1="1" ind2="1">
    <subfield code="a">[Hondius, Henricus (1597 – 1651)]</subfield>
    <subfield code="h">[Cartographic document]</subfield>
    <subfield code="b">[Nova totitus terrarum orbis geographica
        ac hydrographica tabula]</subfield>
</datafield>

<datafield tag="246" ind1="3" ind2="3">
    <subfield code="a">Henricus Hondius :</subfield>
    <subfield code="b">Stereographic projection: eastern hemisphere</subfield>
</datafield>

<datafield tag="255" ind1=" " ind2=" ">
    <subfield code="b">Stereographic projection</subfield>
    <subfield code="c">(25deg 44min 03sec; --179deg 34min 41sec;
        /81deg 09min 00sec; --77deg 44min 44sec;)</subfield>
</datafield>
```

Figure 6: A part of the XML file containing the cartographic meta-data (source: [3]).

on the map, or the auxiliary construct elements (graticule, map frame), if they are available; see Fig. 7.

The implementation of line features (rivers, roads) into the assessment process reduces the discretization and enables additional analysis, which further improves the results. Polygonal features allow the analysis of continuous and extensive parts of maps in a single step; they represent one of the best materials for the assessment process.

Errors and their distribution

Due to the proposed methods, this is one of the most important factors affecting the results. The drawn elements on the maps are contaminated by errors; there are several factors influencing both the type and the distribution of errors.

The crucial moment is that early maps, created before the 17th century, were not constructed on solid geometric and geodetic basis. In such cases, many drawn elements are contaminated by gross errors, which may not satisfy the Gaussian distribution. Hence, the automatic detection of incorrectly drawn elements, based on the M-estimators, is implemented. The suspect elements are subsequently rejected from the analysis.

In situations when maps were more like charts and no projection has been used (Ebsdorf map of the world, Gough map of Great Britain), the determined parameters represent only a geometric construct.

However, for early maps after the middle of the 17th century, these methods may be used. Based on the analysis of about one hundred early maps, an empirical limit of 20% of incorrectly drawn elements was determined. Therefore, the M-estimators with the median absolute deviation may be used; their breakdown point ε^* is higher. Considering different weights of the analyzed features, from a statistical point of view, no requirements on the input data are

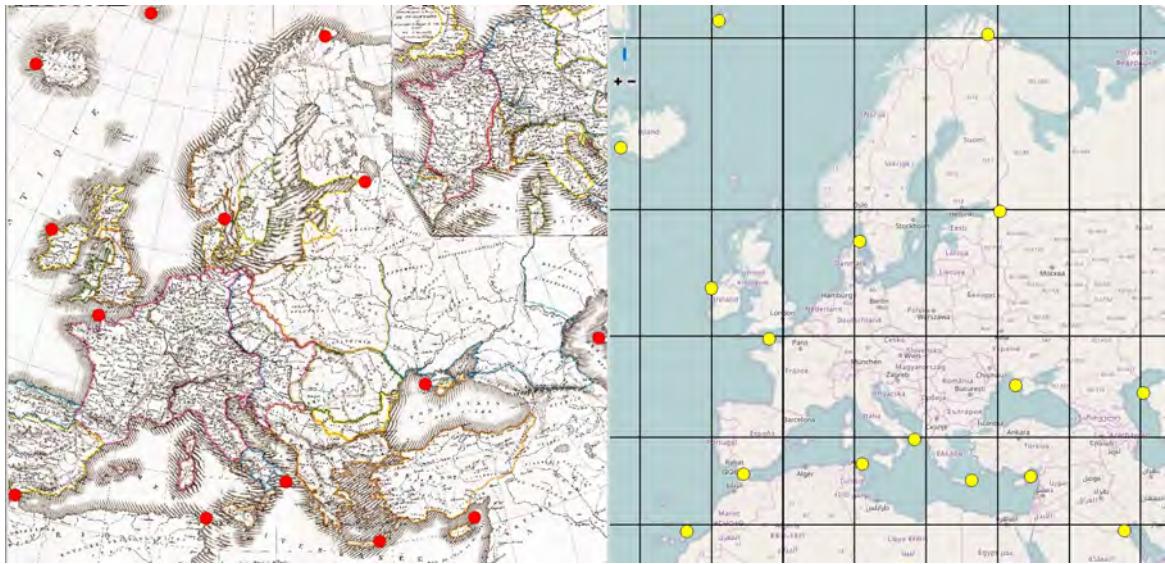


Figure 7: The set of 0D features (control points) on the analyzed and reference maps acquired in detectproj software.

imposed.

Distribution of features

The uniform distribution of analyzed features on a map also plays an important role. The proposed techniques are suitable for sets with approximately the same spatial density of features.

On the boundaries of the analyzed region, in particular, it is necessary to place enough points. Otherwise, the refined intervals of the determined parameters, computed from the extent of the analyzed features, may be set incorrectly. Hence, a proper solution may be thrown out.

The following rules should be respected:

- At least three points should be placed over each analyzed meridian or parallel. Such a locus of points allows expression of their curvature, which is necessary for the shape description. It is obvious that two points are insufficient, regardless, whether they are curved or not; see Fig. 8.
- The analyzed features should be preferably distributed over the entire analyzed map. The estimated projection parameters fit well inside the *convex hull* of the analyzed set, and thus the reconstructed graticule. No extrapolation of results outside the analyzed set is recommended; see Fig. 9.
- If the previous condition cannot be entirely fulfilled, the analyzed features should be placed “symmetrically”, on both sides of the equator and the central meridian. Omitting some quadrants or using only a single quadrant may lead to a detection of the wrong projection, see Fig. 10.



Figure 8: Insufficient amount of analyzed features on parallels (inappropriate grasp of their shape) lead to the detection of the different projection (source: [3]).

However, in many cases it is impossible to ensure the requirements, so the density of the control points may be variable. On portolan charts, the coastline and ports are drawn, but the interior of the continents is missing. Within the seas, there is also a lack of suitable points. Obviously, the irregularly spaced clusters of points, or wide territories without analyzed points, affect the results negatively.

Analyzed territory

The size and position of the analyzed territory strongly affect the reliability of the detection algorithm. The analyzed territory should have similar dimensions in the latitudinal and longitudinal directions, and should be large enough (at least $\Delta\varphi = \Delta\lambda = 3^\circ$) to ensure that the positional differences of both sets of features in the distinct map projections are not less than the graphical accuracy of the map. Then, a footprint of the projection may be recognized.

The locus of the analyzed region near the equator (a similar shape in many projections), near the pole (singular points), or near the central meridian (a similar shape in many projections), are also not recommended.

Map scale

Small-scale maps depict a large territory in detail and vice versa; the influence of the map projection on small-scale maps is stronger. Nevertheless, medium-scale and small-scale map parameters are easier to determine. Over medium-scale and large-scale maps the projection

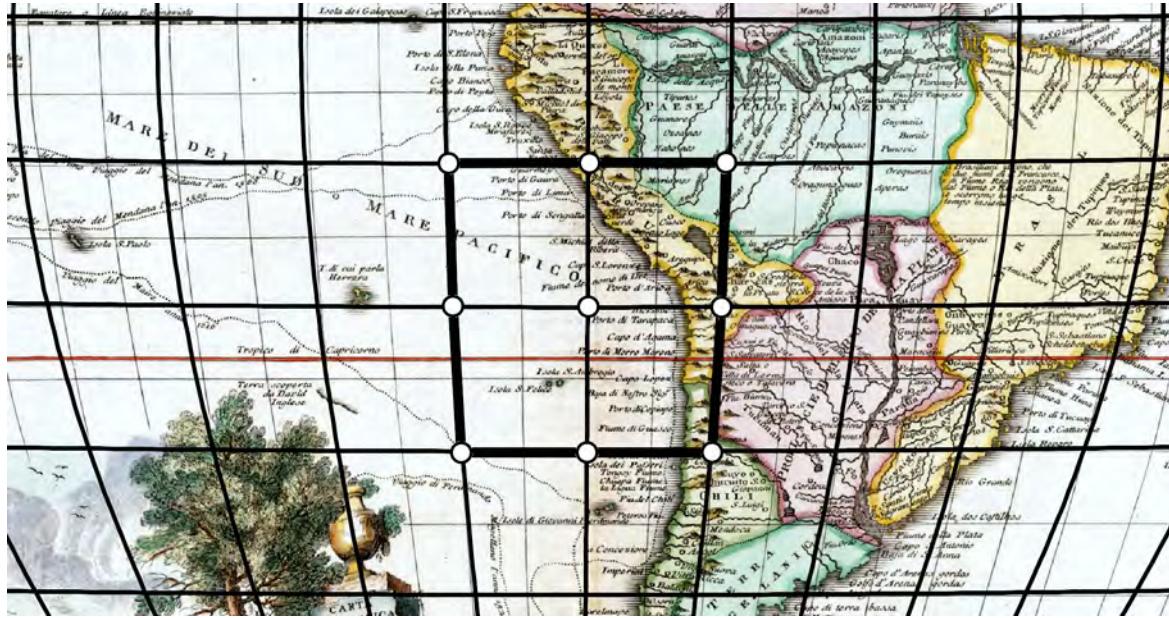


Figure 9: The analyzed features covering a small part of the analyzed territory; the estimated graticule fits into the original only inside the convex hull. Bonne projection, $\varphi_1 \neq 0^\circ$, instead of sinusoidal is detected (source: [3]).

influence becomes weaker until it disappears.

Sheets of large-scale maps up to the scale 1:10,000 are difficult to analyze. They cover a small territory, where the impact of the projection does not occur. Therefore, no reliable projection footprint exists and almost any projection may be assigned to the map.

Decreasing the map scale, the corresponding territories between two projections become less similar. Recall Fig. 1 illustrating the impact of the map scale on the territory shape in two different projections.

Map projection

Unlike the normal aspect both the oblique and transverse aspects are more difficult to detect. In general, the minimized objective function ϕ measuring the map similarity has a complicated course, with many local minima, and leads to non-convex optimization. The iteration process can become stuck in a local minimum of ϕ ; see Fig. 11.

The problem is convex for some types of projections and determined parameters, and for other parameters, it is non-convex. Some shapes of the graticule are easily recognizable (cylindrical, azimuthal projections in the normal aspect); others have a similar shape (pseudo-azimuthal).

For some projections, equations in the closed form are not available. Hence, the iterative solution, based on the Newton-Raphson method, should be involved. Including the numeric differentiation is computationally expensive, especially, if the real-time solution is required.

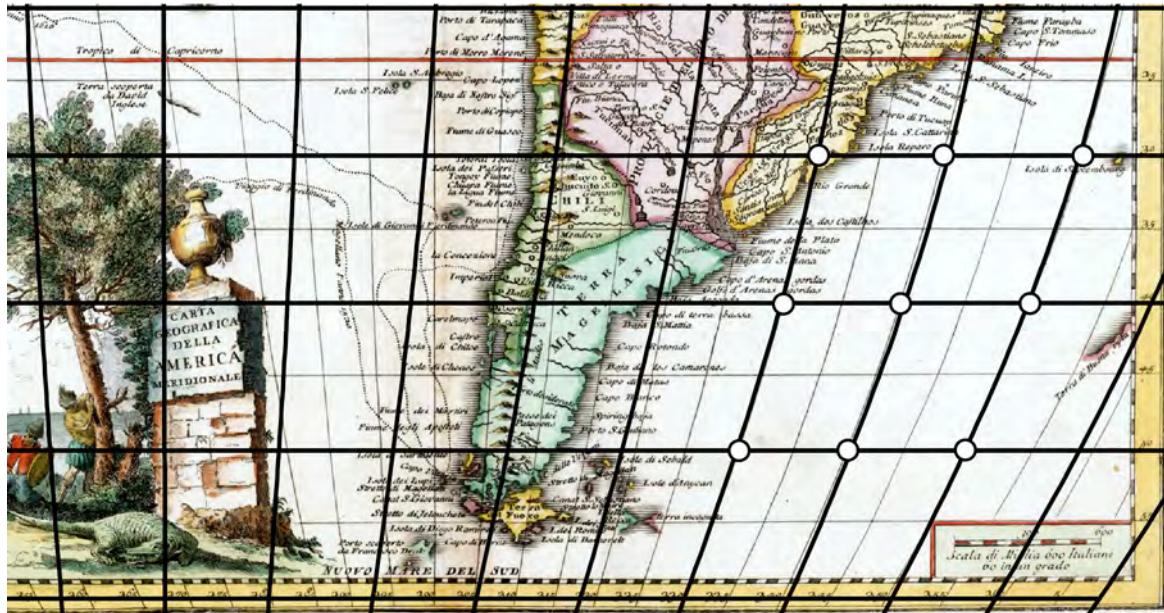


Figure 10: The analyzed features, covering only a small part of the analyzed territory, are placed in the single quadrant; the estimated graticule fits the original only inside the convex hull. Apian elliptic, instead of sinusoidal projection is detected.

Amount of the analyzed features

The total amount of the analyzed features significantly affects the results. The clearer the projection footprint is, the less the analyzed features are required. It is obvious that the amount of features depends on the map scale. Based on the results, for the small-scale maps 5 points are sufficient, for the medium-scale maps 10 points are recommended, but for the large-scale maps 15-20 points may be required.

Map sheet

The shape of the material on which the map is drawn changes over time. Map sheet distortions caused by paper aging, or organization by map fields, must be taken into account. However, using the affine transformation, the affect of the paper aging may be almost completely removed. This applies particularly to maps, where the real dimensions of the map sheet are *a priori* known (topographic maps).

Cartographic techniques

Some cartographic techniques have a strong influence on the geometric accuracy of a map. At first, a cartographic generalization must be mentioned. The map content is adjusted and simplified to maintain important geographical details in a recognizable way, but the relative position of the objects is not preserved. In other words, the relationship of the features is preferred to their geometric accuracy. Unfortunately, its effect cannot be corrected before analysis.

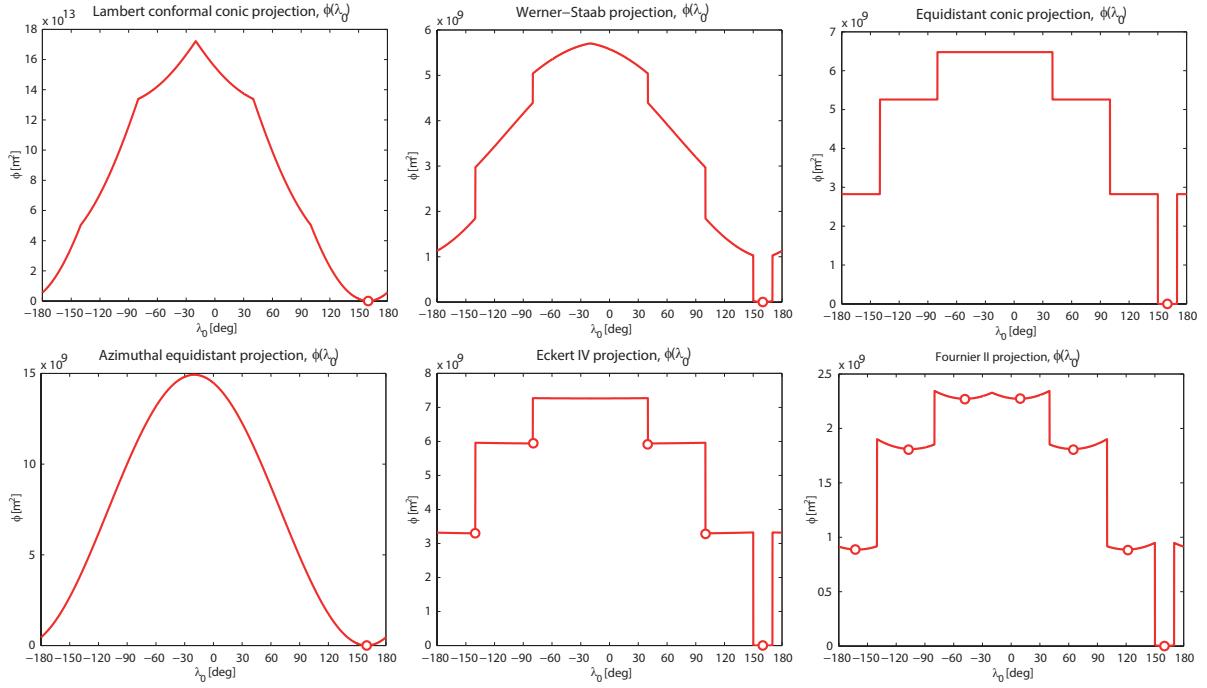


Figure 11: The objective function $\phi(\lambda'_0)$ course for 6 different projections with marked local/-global minima; the global minimum is $\hat{\lambda}'_0 = 160^\circ$.

Types of features

The selection of appropriate point features for analysis is relatively complicated. It is assumed that their position does not change significantly over time (they are stable in their geographical locations), and that they are easy to identify on a map (cities, castles, river confluences, churches). In most situations, this set of elements is sufficient for further analysis. If both the analyzed and the reference maps contain a graticule, an additional analysis of sampled meridian/parallel points is a promising method, improving the results, see Fig. 12.

Timing of maps

As mentioned above, there is no challenge in analyzing early maps created prior to the 17th century. At that time, the Earth's circumference had not been measured with sufficient accuracy, the biggest obstacle in the development of the geodetic basis necessary for accurate mapping. Some regions were measured less accurate, mapped from the horseback (a la vue method), drawn incomplete, or missing. Depending on the timing of maps, the map content as well as the geometric properties were improving.

Maps without a geometric basis affected by many errors were more like charts. Moreover, the error distribution does not satisfy the Gaussian normal laws; the resulting determined parameters must be treated with a little skepticism. The map content is heavily distorted, only graticule elements may be used as control points.

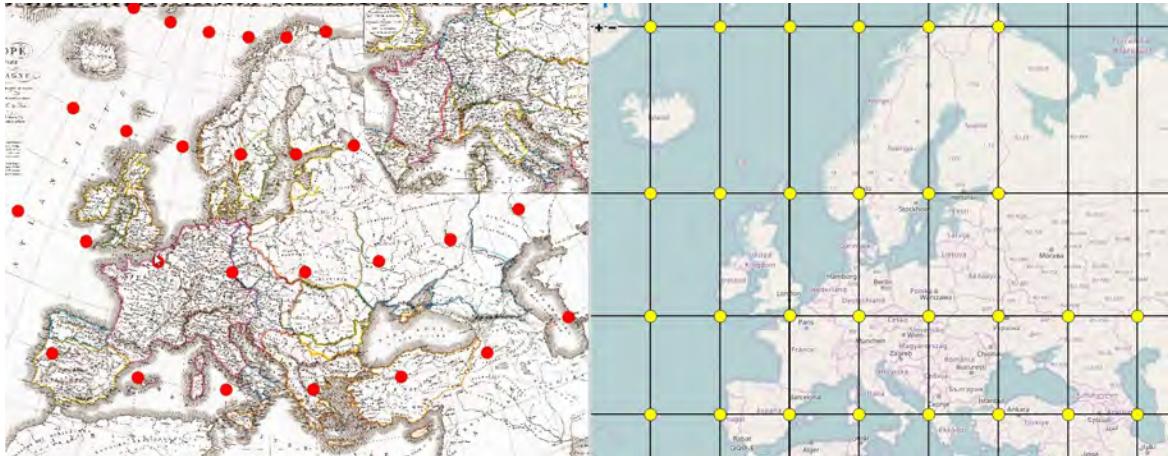


Figure 12: The set of control points on the analyzed and reference maps represented by the intersections meridians and parallels acquired in detectproj software.

Different accuracy of elements

In the early maps created prior to the 17th century, the graticule is typically the geometric construct. The map content was not measured seriously; it is not very accurate, and some territories are completely missing (Australia, the polar regions of North America).

For early maps, the graticule is significantly more accurate than the map content; it is not affected by the generalization. Hence, there is also far less chance of outliers, and the reconstructed graticule fits better; see Figs. 13, 14. Using the map content with the lack of solid geometric basis leads mostly to the wrong results, when neither the projection, nor its aspect, is recognized.

However, it should be noted that due to blunders, the meridians or parallels may be drawn incorrectly too; the map graticule may not be symmetric around the equator or the prime meridian. In most cases, these errors are not visually identifiable without magnification. They become visible when the estimated graticule is generated over the analyzed early maps.

3. Concept of the detection

The map projection analysis represents a challenging, but conceptually difficult, problem. There are many requirements imposed on the developed solution. Involving maps of various scales, sizes, and types as well as the support of different types of projections are the preferred features of any proposed detection method. Obviously, the method should be robust to gross errors and paper aging, sufficiently accurate, and providing results in real time.

Four new detection methods, denoted as M6-M9, have been proposed. They differ in the number of determined parameters, reliability, robustness, and convergence, but their computational demands are similar. They are fully or partially invariant to the map constants (scale, shifts, rotation), applicable to both approaches measuring the map similarity (the location and shape similarities).



Figure 13: The reconstructed graticule of the sinusoidal projection, analyzed features represented by the intersections of meridians and parallels.

3.1. Description of the problem

Each map projection transforms a position of the element on the curved surface into a flat surface, represented by the plane. A curved surface approximating the Earth is considered to be the sphere or ellipsoid.

From the mathematical aspect, a map projection will be seen as the function of two variables, latitude and longitude, supplied with constant values, which are subjects of analysis. A map projection \mathbb{P} is defined with the set coordinate functions F, G .

Suppose the projection $\mathbb{P}(\varphi_k, \lambda_k, \varphi'_1, \lambda'_0, \kappa)$ described by the set of constant values: transformed pole position $[\varphi_k, \lambda_k]$, standard parallel φ'_1 , longitude λ'_0 of the central meridian, constant parameter κ (may be assigned to any other value). Furthermore, suppose the analyzed map $M(R', \Delta X, \Delta Y, \alpha)$ described by its constant values: auxiliary sphere radius R' (illustrating the scale ratio), origin shifts $\Delta X, \Delta Y$, and angle of rotation α .

For the oblique aspect, the projection equations in closed form may be written as functions of the determined parameters

$$\begin{aligned} X(R', \varphi_k, \lambda_k, \varphi'_1, \lambda'_0, \Delta X, \kappa, \alpha) &= F(\varphi', \lambda'), \\ Y(R', \varphi_k, \lambda_k, \varphi'_1, \lambda'_0, \Delta Y, \kappa, \alpha) &= G(\varphi', \lambda'). \end{aligned}$$

Such equations represent most currently used projections.

Parameters of which the determination represents a convex problem are determined reliably in all situations; the local minimum also represents the global minimum. The non-convex



Figure 14: The reconstructed graticule of the sinusoidal projection, analyzed features represented by the map content.

analysis brings some uncertainty to the determined parameters; the global minimum, which is one of many local minima, may not be sufficiently detected. However, in most situations, a local minimum brings the acceptable solution.

Depending on the projection equations, determining φ'_1, λ'_0 leads to a convex/non-convex problem, but for $\varphi_k, \lambda_k, \alpha, \kappa$ it represents the non-convex problem. Determining map constants $R', \Delta X, \Delta Y$ always represents a convex problem.

3.2. Determined parameters of the projection

During the analysis, the below-mentioned constant parameters of the projection \mathbb{P} are determined. They have a strong influence on the shape of the graticule and affect the projection footprint; see [3].

Transformed pole position $[\varphi_k, \lambda_k]$

This parameter has a crucial influence on the shape of the graticule and the projection footprint. In the oblique aspect, shape of the meridians/parallels as well as angles of intersections may not be preserved. While conic, cylindrical, or azimuthal projections are currently used in the transverse/oblique aspects, for other categories of projections they are used rarely, and the normal aspect is preferred.

During the analysis, the position of the arbitrary pole $K = [\varphi_k, \lambda_k]$ is determined. Due to the non-convexity of the problem, the iteration process may get stuck in a local minimum of ϕ , which makes the analysis less efficient.

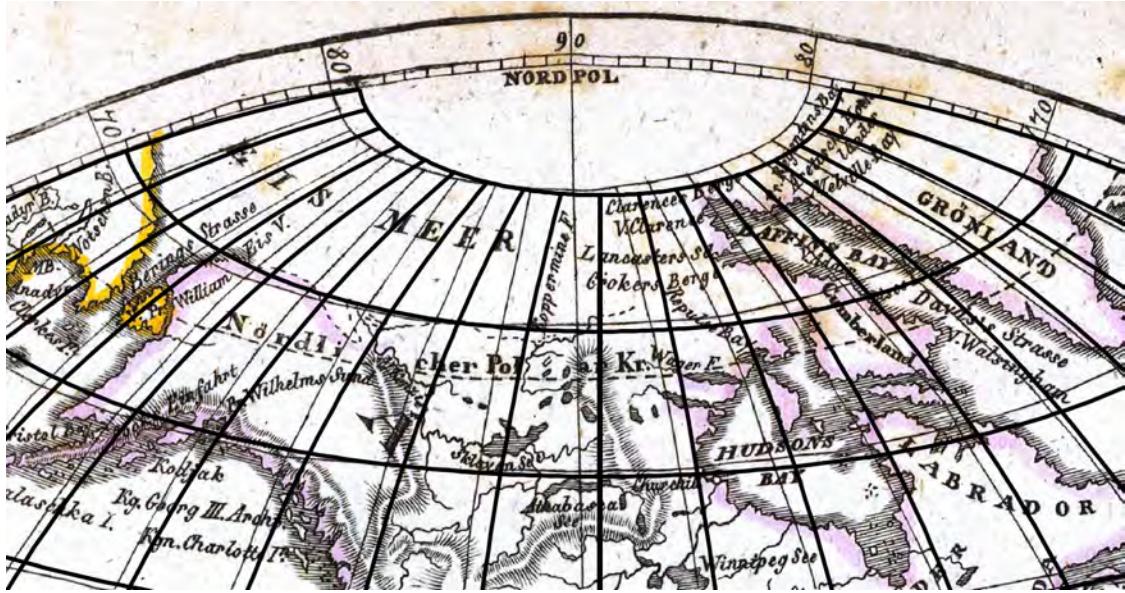


Figure 15: Map graticule reconstructed from the determined parameters. Large residuals are obvious if the analyzed map rotation $\alpha = 1.7^\circ$ was not involved (source: [3]).

Standard parallel φ'_1

Currently, the projection in a secant form specifies the latitudes of two standard parallels φ'_1, φ'_2 , $\varphi'_1 \neq \varphi'_2$, representing intersections of the sphere and the secant plane, or $\varphi'_1 = \varphi'_2$ for its tangent form. The standard parallel may also refer to the normal aspect of the projection, then $\varphi'_1 \equiv \varphi_1$, $\varphi'_2 \equiv \varphi_2$. During the analysis, the cartographic model is simplified to

$$\varphi'_1 = \varphi'_2,$$

and the latitude of the standard parallel φ'_1 , along which the nominal scale is preserved, is determined.

Longitude λ'_0 of the central meridian

To minimize the distortion and provide a true projection of the mapped region, setting the prime meridian of longitude λ'_0 as the central meridian is inappropriate. This applies in particular for the regions far east or west of the prime meridian. The central meridian is frequently chosen in the axis of the symmetry of the mapped region; it passes through the center of the region. Its longitude λ_0 represents, the next determined parameter. In general, this approach is applied to the normal aspect of the projection. Occasionally, it may be used for the transverse as well as oblique aspects and denoted as λ'_0 .

This parameter affects the shape of meridians and parallels locally. It shifts the central meridian over the analyzed territory, changes its curvature to a straight line, while the graticule shape as a whole remains unchanged.

Arbitrary parameter of the map projection κ

This may be any constant value of the map projection. It is widely used in connection with perspective projections, where the distance of the center of projection from the sphere S^2 center represents the estimated parameter. It is obvious that κ may reach a wide range of values; its determination brings a problem, especially for the simplex method.

3.3. Determined map constants

These constant values do not have any influence on the shape of the graticule, or the projection footprint; see [3]. The optimization process should be partially or fully invariant to values of map constants.

Auxiliary sphere radius R'

Both the analyzed and reference maps may have different scales. The radius R' of the auxiliary sphere S^2 (i.e., the scale factor) is determined so that the map in the estimated projection fits best with the analyzed one. Because the analyzed map dimensions are currently in centimeters (map in paper form), the determined radius R' will be small, usually in meters. The auxiliary sphere radius may be used to determine the approximate scale S of the analyzed map $S = R/R'$.

Shifts $\Delta X, \Delta Y$

Both the analyzed and reference coordinate systems may be shifted each other. This applies particularly to a scanned map in paper form, where the origin is subsequently set to the left upper pixel of the raster.

To avoid sign inconsistency in the X, Y coordinates for the mapped region and make their values more convenient, the false northing ΔY or false easting ΔX are widely used. Sometimes, the shifts are small, but they may be too high and cannot be neglected during the analysis.

Angle of rotation α

This optional parameter represents an additional rotation of the analyzed map. In most cases, it is caused by the inappropriate insertion of the paper form of the map into the scanner, or its additional rotation on the page (portrait vs. landscape), when a document is typographically processed. Involving a rotation reduces the residuals by one order of magnitude and provides a better fit; see Figs. 15, 16. However, the problem becomes non-convex, so the detection reliability may decrease.

4. Detection fundamentals

Determining the best fit projection parameters represents a complex problem leading to the unconstrained optimization. The detection methods are based on the convex/global unconstrained optimization of the objective function ϕ , describing the similarity of the analyzed and reference maps.

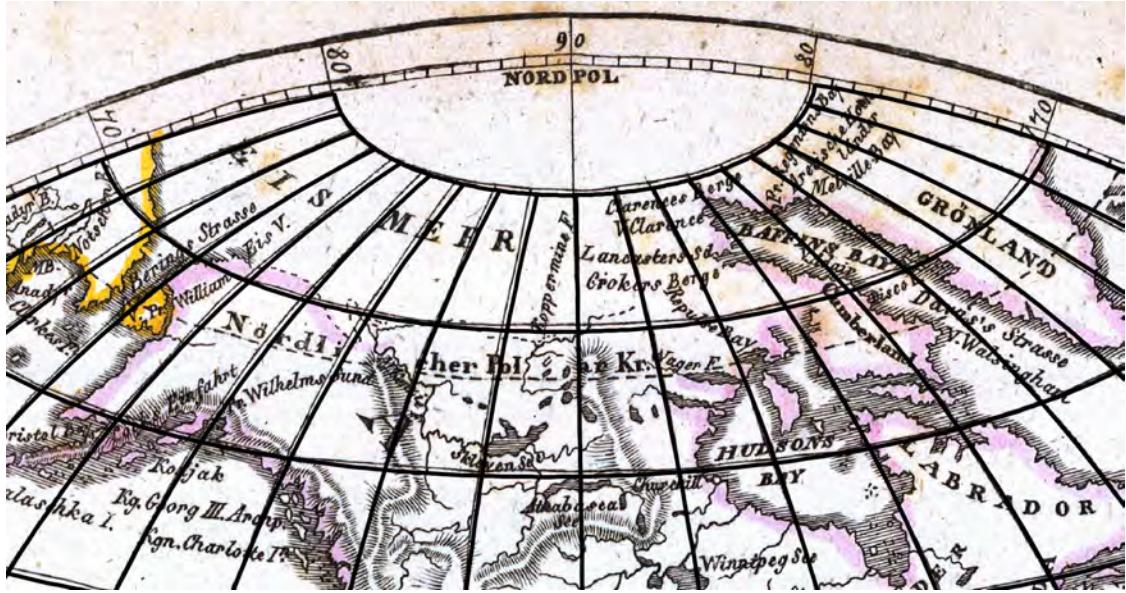


Figure 16: The map graticule reconstructed from the determined parameters. Small residuals are obvious if the analyzed map rotation $\alpha = 1.7^\circ$ is involved (source: [3]).

Let $P \in M$ and $Q \in S^2$ be the sets of features on the analyzed map M and on the sphere S^2 , $\mathbb{P}_x : S^2 \rightarrow M'$ be the analyzed projection, and $P' \in M'$ be the image of Q in \mathbb{P}_x . Then, the dissimilarity δ_x , $\delta_x \geq 0$, measured by the objective function ϕ at a point x exists, such that

$$\delta_x = \phi(\mathbb{P}_x(Q), P) = \phi(P'_x, P).$$

For each analyzed map projection \mathbb{P} , the vector of its best constant values \hat{x}

$$\hat{x} = \arg \min_{\mathbb{P}} (\phi(\mathbb{P}_x(Q), P)) = \arg \min_{\mathbb{P}} (\phi(P'_x, P)), \quad (1)$$

minimizing ϕ , may be determined. The problem may be solved using the convex optimization based on the non-linear least squares

$$\hat{x} = \arg \min_{\mathbb{P}} \left(\frac{1}{2} f^T(P'_x, P) f(P'_x, P) \right).$$

Subsequently, the optimal projection, and its best constant values \hat{x} are found and assigned to the analyzed map M .

Projection footprint. The shapes of the projected poles, meridians, parallels and their angles are unique indicators representing the projection footprint. The objective function ϕ measures the projection dissimilarity δ_x (i.e., footprint difference) indirectly, from the different parts of the maps: fragments of the graticule, regularly or irregularly distributed points, 1D features (line, polylines, curves) or 2D features (polygons, bounded areas).

Location vs. shape similarity. There are many ways to propose the objective function. The objective function ϕ may be less or more complex, discrete, its gradient $\nabla \phi$ may not

be available for computations. A simple approach based on the residuals of corresponding features (location similarity), or a more complex approach (shape similarity) utilizing the shape differences, will be used.

The location similarity takes into account the positional differences of corresponding elements, which can be easily expressed by their residuals. The most obvious objective function is represented by the sum of the squares of residuals on the corresponding elements

$$\phi(P, P'_x) = \sum_{i=1}^n (X_i - x_i)^2 + (Y_i - y_i)^2 = \sum_{i=1}^n \|P_i - P'_i\|_2^2.$$

4.1. The 7-parameter method

While the M7 method determines the map constants $[R', \alpha]$ directly using the scale coefficients q_1, q_2 of the Helmert 2D similarity transformation, the constant values of the projection $[\varphi_k, \lambda_k, \varphi'_1, \lambda'_0, \kappa]$ are estimated iteratively using the non-linear least squares procedure (hybrid BFGS). The vector \hat{x} of 7 unknown parameters

$$\hat{x} = \begin{bmatrix} \varphi_k \\ \lambda_k \\ \varphi'_1 \\ \lambda'_0 \\ \kappa \\ \hline q_1 \\ q_2 \end{bmatrix},$$

is used to determine the projected coordinates $[X_i, Y_i]$ of a point $[\varphi'_i, \lambda'_i]$

$$\begin{aligned} X_i(\varphi_k, \lambda_k, \varphi'_1, \lambda'_0, \kappa, q_1, q_2) &= F(\varphi'_i, \lambda'_i), \\ Y_i(\varphi_k, \lambda_k, \varphi'_1, \lambda'_0, \kappa, q_1, q_2) &= G(\varphi'_i, \lambda'_i). \end{aligned}$$

At least, $m = 4$ analyzed features are required. The oblique aspect transformation is performed using the laws of spherical trigonometry

$$\begin{aligned} \sin \varphi'_i &= \sin \varphi_k \sin \varphi_i + \cos \varphi_k \cos \varphi_i \cos(\lambda_i - \lambda_k), \\ \tan \lambda'_i &= \frac{\cos \varphi \sin(\lambda_i - \lambda_k)}{\sin \varphi_i \cos \varphi_k - \cos \varphi_i \sin \varphi_k \cos(\lambda_i - \lambda_k)}, \end{aligned}$$

where $K = [\varphi_k, \lambda_k]$ is the pole position, φ_i, λ_i are the geographic coordinates related to the North Pole and the selected prime meridian, and φ'_i, λ'_i are the geographic coordinates related to K ; further details can be found in [6], [3]. The coordinates of P_i, P'_i are reduced to the centers of mass $c = [x_c, y_c]$, $C = [X_c, Y_c]$,

$$\begin{aligned} X'_i &= X_i - X_c, & Y'_i &= Y_i - Y_c, \\ x'_i &= x_i - x_c & y'_i &= y_i - y_c, \end{aligned}$$

rescaled and rotated by q_1, q_2 so that

$$\begin{aligned} \chi'_i &= (X_i - X_c) q_1 - (Y_i - Y_c) q_2 = X'_i q_1 - Y'_i q_2, \\ \gamma'_i &= (X_i - X_c) q_2 + (Y_i - Y_c) q_1 = X'_i q_2 + Y'_i q_1. \end{aligned}$$

The residuals are written as follows

$$r(x) = \begin{bmatrix} \chi'_1 - x'_1 \\ \vdots \\ \frac{\chi'_n - x'_n}{\gamma'_1 - y'_1} \\ \vdots \\ \gamma'_n - y'_n \end{bmatrix} = \begin{bmatrix} X'_1 q_1 - Y'_1 q_2 - x'_1 \\ \vdots \\ \frac{X'_n q_1 - Y'_n q_2 - x'_n}{X'_1 q_2 + Y'_1 q_1 - y'_1} \\ \vdots \\ X'_n q_2 + Y'_n q_1 - y'_n \end{bmatrix},$$

or, in a more compact notation,

$$r(x_k) = A_k Q_k - l, \quad (2)$$

where A_k is the design matrix

$$A_k = \begin{bmatrix} X'_1 & -Y'_1 \\ \vdots & \vdots \\ \frac{X'_n}{Y'_1} & -Y'_n \\ \vdots & \vdots \\ Y'_n & X'_1 \end{bmatrix}, \quad l = \begin{bmatrix} x'_1 \\ \vdots \\ \frac{x'_n}{y'_1} \\ \vdots \\ y'_n \end{bmatrix}, \quad Q_k = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix},$$

and k represents the iteration. The scale coefficients are determined from the linear least squares solution

$$q_k = (A_k^T W_k A_k)^{-1} A^T W_k l,$$

where W_k is the weight matrix. To improve the convergence rate, the radius of the auxiliary sphere R'_k at iteration k is determined using the permanent scaling

$$R'_k = R'_{k-1} \|Q_k\|_2.$$

Finally, the Jacobian matrix $J(2n, 5)$ has the following form of

$$J = \begin{bmatrix} j_{1,1} & j_{1,2} & j_{1,3} & j_{1,4} & j_{1,5} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ j_{n,1} & j_{n,2} & j_{n,3} & j_{n,4} & j_{n,5} \\ j_{1+n,1} & j_{1+n,2} & j_{1+n,3} & j_{1+n,4} & j_{1+n,5} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ j_{2n,1} & j_{2n,2} & j_{2n,3} & j_{2n,4} & j_{2n,5} \end{bmatrix},$$

its elements $j_{i,k}$ may be written as

$$\begin{aligned} j_{i,k}(x) &= \left[\left(\frac{\partial X}{\partial \cdot} \right)_{(\varphi=\varphi_i, \lambda=\lambda_i)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \cdot} \right)_{(\varphi=\varphi_l, \lambda=\lambda_l)} \right] q_1 \\ &\quad - \left[\left(\frac{\partial Y}{\partial \cdot} \right)_{(\varphi=\varphi_i, \lambda=\lambda_i)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \cdot} \right)_{(\varphi=\varphi_l, \lambda=\lambda_l)} \right] q_2, \end{aligned}$$

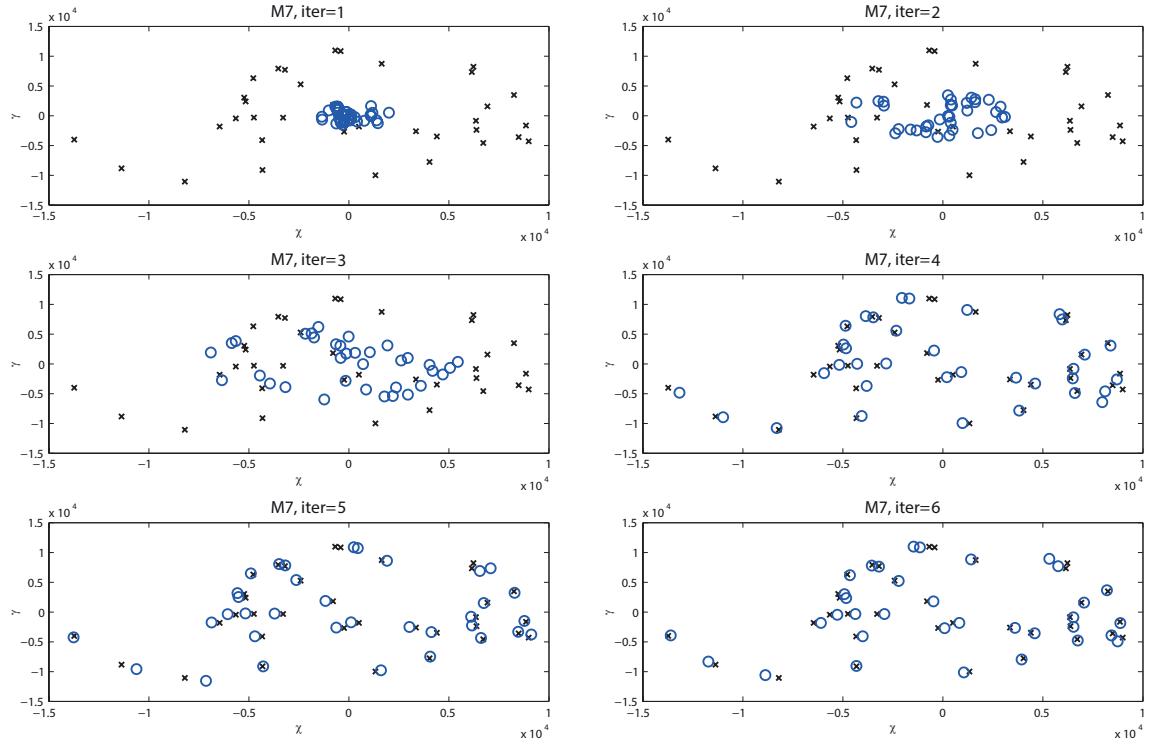


Figure 17: Detection of the Lambert azimuthal projection in the oblique aspect, rotated by $\alpha = 90^\circ$ using the M7 method, a random set of analyzed features. The residuals between test points (crosses) and reference points (circles) are continuously decreasing.

$$\begin{aligned} j_{i,k+n}(x) &= \left[\left(\frac{\partial X}{\partial \cdot} \right)_{(\varphi=\varphi_i, \lambda=\lambda_i)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \cdot} \right)_{(\varphi=\varphi_l, \lambda=\lambda_l)} \right] q_2 \\ &\quad + \left[\left(\frac{\partial Y}{\partial \cdot} \right)_{(\varphi=\varphi_i, \lambda=\lambda_i)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \cdot} \right)_{(\varphi=\varphi_l, \lambda=\lambda_l)} \right] q_1. \end{aligned}$$

It is obvious that the derivatives $\frac{\partial X}{\partial \cdot}, \frac{\partial Y}{\partial \cdot}$ are composite functions of the determined parameters $\varphi_k, \lambda_k, \varphi'_1, \lambda'_1, \kappa$ and have a complex form; for further details, see [3].

During the iteration process, the residuals between P and P'_x , decrease. The proposed method has a fast convergence, see Fig. 17, which was confirmed in practice.

For practical computations, the numerical differentiation seems to be the preferable; the Stirling formula with the step $h = 0.001$ was used.

4.2. Hybrid BFGS

The combination of the Gauss-Newton and BFGS methods is efficient for solving the non-linear least squares. It will be proposed as the line search method using a positive definite update of the Hessian matrix. For non-zero residual problems, it converges superlinearly, and quadratically for zero residual problems. This approach was studied in several early papers [1], [9], [11], [12], and later in [22], [23], [24], [31].

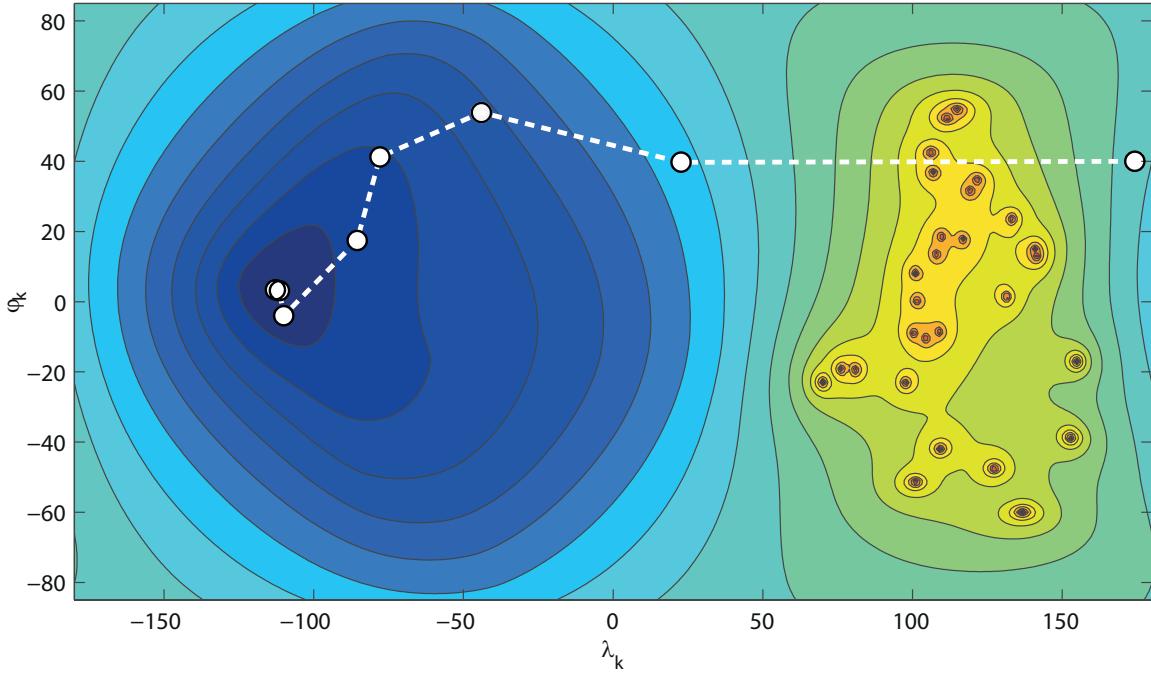


Figure 18: The fast convergence of the vector x_k to the global minimum $\varphi_k = 2.2^\circ N$, $\lambda_k = 109.1^\circ W$, of $\phi(x_k)$ using hybrid BFGS; iterations 1-8, together with contour lines, are shown (source: [3]).

The selection criterion may be rewritten to the following form of

$$\tau = \frac{\phi(x_k) - \phi(x_{k+1})}{\phi(x_k)}.$$

If $\tau > \tau_{min}$, the simple update computes B_{k+1} from the Gauss-Newton method. If $\tau < \tau_{min}$, and $y_k^T s_k > 0$, the BFGS update is used

$$B_{k+1} = \begin{cases} J^T(x_k)J(x_k), & \tau > \tau_{min}, \\ B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k^T}{s_k^T B_k s_k}, & \tau < \tau_{min} \wedge y_k^T s_k > 0, \end{cases} \quad (3)$$

where

$$s_k = x_{k+1} - x_k, \quad y_k = \nabla \phi(x_{k+1}) - \nabla \phi(x_k).$$

The solution can be found from

$$h_k = B_k^{-1} \nabla \phi(x_k),$$

with the line search

$$x_{k+1} = x_k - \alpha h_k,$$

where $\alpha \in [0, 1]$. For the BFGS method, B_k^{-1} may be found in analytic form using the Sherman-Morrison-Woodbury formula

$$B_{k+1}^{-1} = \left(I - \frac{s_k y_k^T}{s_k^T y_k} \right) B_k^{-1} \left(I - \frac{y_k s_k^T}{s_k^T y_k} \right) + \frac{s_k s_k^T}{s_k^T y_k},$$

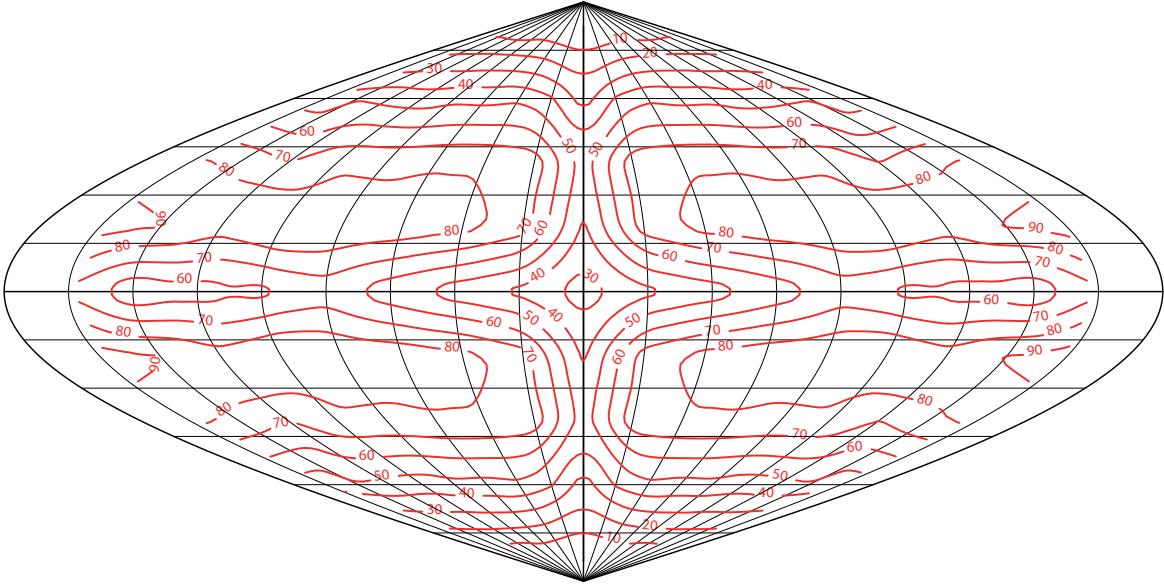


Figure 19: Sinusoidal projection, contour lines of the analysis efficiency (successful match) depending on the geographic position (source: [3]).

where $B_1^{-1} = I$ represents the identity matrix; its derivation is can be found in [5]. Otherwise, the solution based on the QR decomposition is used; see Sec. 4.3. The hybrid method proved to be very efficient in all the measured aspects, see Fig. 18. Improving properties of the Gauss-Newton method for large residual problems as well as BFGS for the zero residual problems it contains the best features of both methods. For the projection analysis, its performance is adequate and provides real-time analysis of the projection parameters.

4.3. Solving non-linear least squares using QR decomposition

The modified method for solving the non-linear least squares problem, where QR factorization $J\Pi = QR$ with pivoting determines the rank r of J from linearly independent columns and the length of the step $\|h\|_2^2$ is minimized, will be presented. This approach may be used if $\tau > \tau_{min}$; the Gauss-Newton method is solved. A similar technique can be found in [15], [14], [13], or [7].

The NLS problem finds the vector $h \in \mathbb{R}^n$ minimizing the norm

$$\phi = \|Jh + f\|_2^2 = \|Q^T(Jh + f)\|_2^2 = \|Q^T J \Pi \Pi^T h + Q^T f\|_2^2 = \min.$$

Suppose that

$$Q^T = \left[\begin{array}{c} Q_1 \\ Q_2 \end{array} \right]^T, \quad \Pi^T = \left[\begin{array}{c} \Pi_1 \\ \Pi_2 \end{array} \right]^T,$$

and let us put

$$Q^T f = \left[\begin{array}{c} a \\ b \end{array} \right], \quad \Pi^T h = \left[\begin{array}{c} c \\ d \end{array} \right].$$

Then, the determined step is given by

$$h = \Pi \left[\frac{c}{d} \right]. \quad (4)$$

The NLS problem is transformed to the form of

$$\begin{aligned} \phi &= \left\| \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} \right\|_2^2, \\ &= c^T R_{11}^T R_{11} c + d^T R_{12}^T R_{12} d + a^T a + 2c^T R_{11}^T R_{12} d + 2c^T R_{11}^T a + 2d^T R_{12}^T a + b^T b, \end{aligned}$$

for the unknown parameters c, d . The following condition $\phi(c) = \min$ leads to

$$\frac{\partial \phi(c)}{\partial c} = R_{11}^T R_{11} c + R_{11}^T R_{12} d + R_{11}^T a = 0,$$

with the c minimizer

$$c = -R_{11}^{-1}(R_{12}d + a).$$

However, the determined step

$$h = \Pi \left[\frac{-R_{11}^{-1}(R_{12}d + a)}{d} \right], \quad (5)$$

depends on the still unknown parameter d . The condition $\|h(d)\|_2^2 = \min$, leads to

$$\frac{\partial \|h(d)\|_2^2}{\partial d} = 2R_{12}^T (R_{11}^{-1})^T R_{11}^{-1} R_{12} d + 2R_{12}^T (R_{11}^{-1})^T R_{11}^{-1} a + 2d = 0,$$

and

$$d = - \left(R_{12}^T (R_{11}^{-1})^T R_{11}^{-1} R_{12} + I_2 \right)^{-1} R_{12}^T (R_{11}^{-1})^T R_{11}^{-1} a. \quad (6)$$

Instead of solving the normal equations directly, the second QR decomposition may be used. The normal equations are

$$\begin{aligned} S^T S d &= S^T g, \\ d &= (S^T S)^{-1} S^T g, \end{aligned}$$

where

$$S = \begin{bmatrix} R_{11}^{-1} R_{12} \\ I \end{bmatrix}, \quad g = \begin{bmatrix} R_{11}^{-1} a \\ 0 \end{bmatrix}. \quad (7)$$

Alternately, the solution may be found from the QR factorization of S

$$\hat{Q} \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix} = S. \quad (8)$$

in the following form of

$$d = [\hat{R}^{-1} \mid 0] \hat{Q}^T g. \quad (9)$$

Only one inverse, several multiplications, and one transposition are necessary. It is noticeable that this method is computationally cheaper than SVD; the complete derivation can be found in [5].

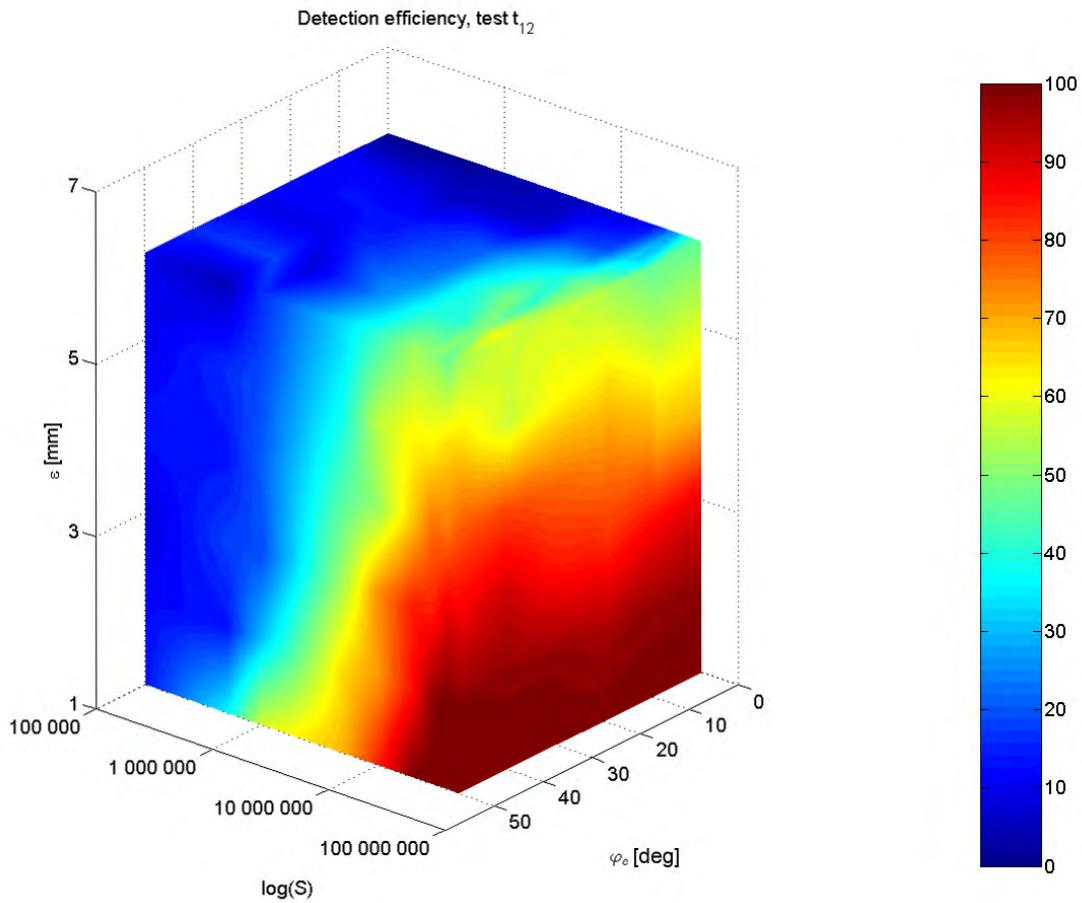


Figure 20: Efficiency of the analysis depending on the accuracy of input features, map scale, and geographical position.

5. Experiments and results

The behavior, advantages, drawbacks of the proposed methods, distance functions, and optimizing techniques, compared both on synthetic and real data, represent the important indicator, the quantitative and qualitative parameters of which will be measured.

The synthetic tests will be undertaken with several different sets of points of various spatial distributions (grid, random distribution, cluster, random meridian, random parallel, circle) in 6 a priori well-known map projections proposed in the normal, transverse, and oblique aspects.

The real data will be represented by early maps of the different scales, sizes, projections and projection aspects, created since the 18th century, and published in the David Rumsay Map Collection, Map Collection of the Charles University, or on medium- and large-scale topographic maps. For most maps, the projections, their aspects, and parameters are a priori unknown.

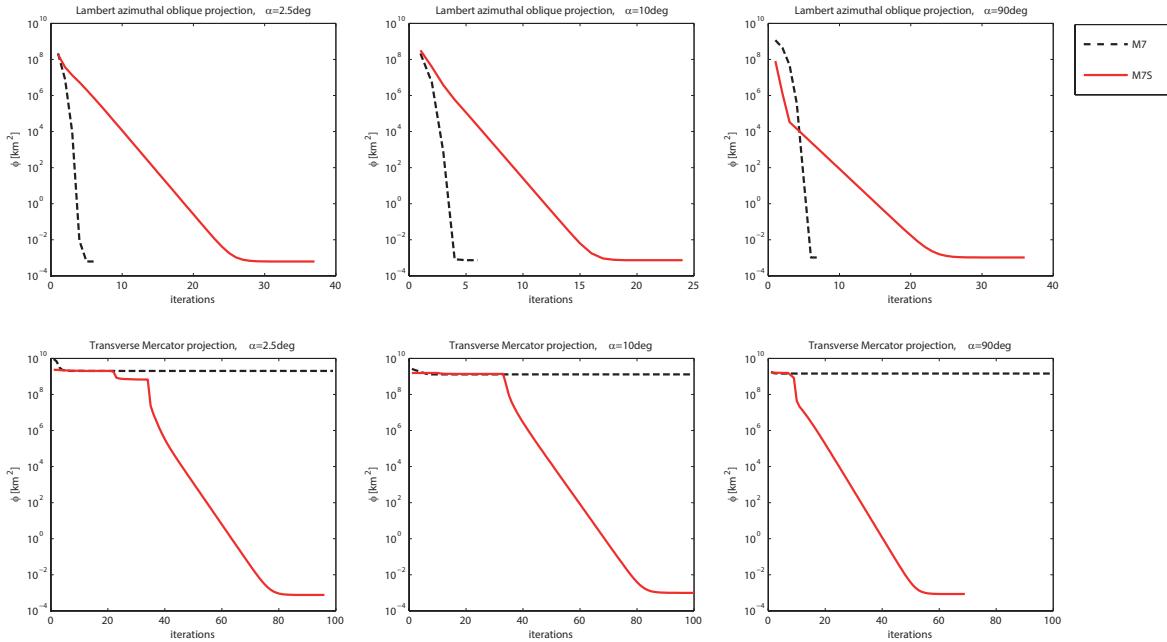


Figure 21: Comparison of M7 and M7S (scaled) methods, values of the objective function ϕ , depending on the amount of iterations: Lambert azimuthal equal-area and transverse Mercator projections.

Only the most important results will be presented. The complete list of tests can be found in [5].

5.1. Efficiency of the analysis depending on the position

This test analyzes an impact of a small territory formed by a spherical quadrangle, continuously shifting over the planisphere on the detection efficiency. It tries to verify the assumption that there are some territories on the planisphere which are more difficult to analyze. In these areas, where most projections have a similar shape of the graticule, the projection footprint is not clear, so the detection efficiency will be significantly lower.

This is typical for territories along the equator, prime meridian, or the poles; see Fig. 19 presenting the contour lines of the analysis efficiency. In most projections in the normal aspect, the central meridian formed by a straight line represents the most difficult part for the analysis. The equator may have a different form; territories along the equator are somewhat easier to analyze. The locus of the analyzed territory in these regions is not recommended, the results may be ambiguous.

5.2. Accuracy of input features

The tests illustrates the properties, behavior and dependency of the detection process on the following parameters: map scale, error contamination of input sets, spatial distribution of features, and geographic position; see Fig. 20.

From a wider aspect, the test reflects effort of collecting the control points. The user may not correctly identify the feature, or, as a result of the generalization, it may be additionally shifted. Even the points on the reference map are not entirely accurate, they are also affected by several errors.

Moving the analyzed territory in the north-south direction, the efficiency decreases slightly. However, the map scale S has a strong influence on results. It is obvious that for the scale of 1:1,000,000, the efficiency is on the verge of acceptance, close to 50%. Hence, the scale of 1:1,000,000 represents a threshold. For maps of the larger scales the results are vague. However, the mid-scale and small-scale map projections are reliably detectable.

The geometrical accuracy of the analyzed features represents the most important factor affecting the efficiency. The acceptable inaccuracy in a position of points is about 3 mm on the map. For world maps, the criterion is less strict; it is around 4 mm. The efficiency over 50% covers mid-scale and large-scale maps, used in mid-latitudes, up to the error of 3 mm in the map scale. For large-scale and partially mid-scale maps up to the scale of 1:500,000, the requirements of the precision are below 1 mm. Obviously, such a strict condition cannot be satisfied.

5.3. Impact of the scaling on the convergence

During the test both the M7 and the M7S (scaling involved) methods will be extensively tested on the rotated sets. Their properties and behavior depending on the rotation α , especially the convergence speed, and the relative error of the solution \hat{x} are the important qualitative and quantitative parameters. All elements will be additionally rotated by the angle α , $\alpha = \langle 2.5^\circ, 5.0^\circ, 7.5^\circ, 10^\circ, 90^\circ \rangle$. Methods involving the rotation are sensitive to the initial guess of α , especially, if a conformal projection is used.

The M7 method has a significantly faster convergence (an initial higher relative error decreases quickly), but it may fail (see Fig. 21, Mercator projection). It broke down for all values of α ; the stuck in a local minimum $\phi(\alpha)$ leading only to a minor improvement was recognized. The objective function $\phi(\alpha)$ value decreased only by a one-half of magnitude. Even for the smallest value of $\alpha = 2.5^\circ$, no convergence was found. Conversely, the M7S method has a slower convergence, but it is more reliable and less sensitive to the initial guess x_0 (no failure has been detected).

5.4. Early map: Seutter's map of America

This early map of North and South America (1744), assigned to Matthäus Seutter, is based on the map of the famous British cartographer John Speed. Besides, it contains some geographical attractions (California is depicted as an island; Greenland is connected to North America; Iceland is missing). For the analysis, 30 identical points, represented by intersections of meridians and parallels, where $\Delta\varphi = \Delta\lambda = 20^\circ$, were collected.

The result is presented in Fig. 22. It is obvious that the analyzed map is only slightly rotated, which was reflected by the M7S method ($\alpha < 0.4^\circ$). The reconstructed graticule in the stereographic projection, where $R' = 0.2$ m, $\varphi_k = 0.3^\circ$, $\lambda_k = -80.4^\circ$, $\varphi_1 = 0.0^\circ$, $\lambda_0 = 0.3^\circ$, fits well to the original, but some minor shifts and rotations are visible. The estimated map scale is $S = 39,381,206$.

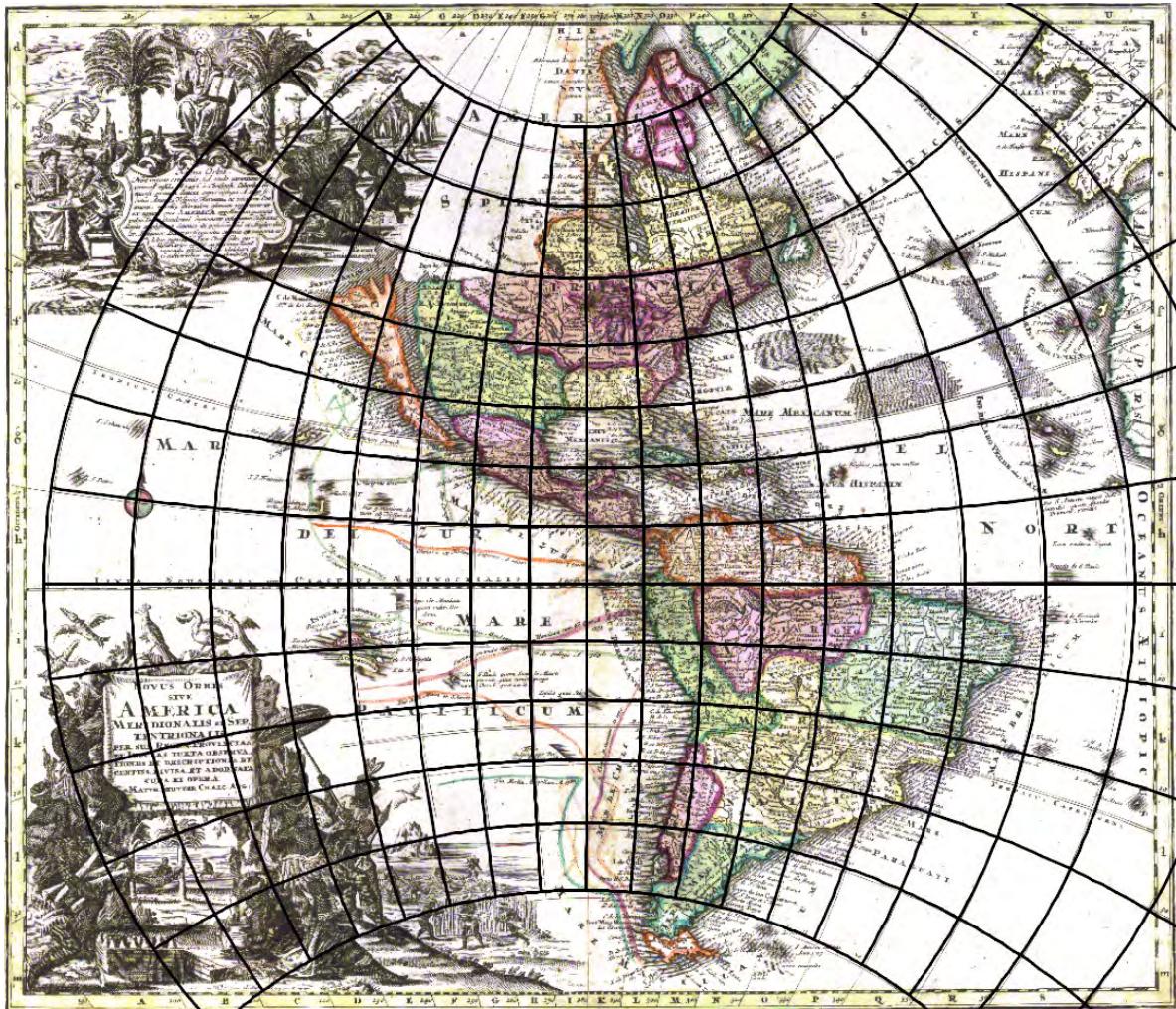


Figure 22: A superimposition of the reconstructed graticule in the stereographic projection and the original graticule on the analyzed Seutter's map of America.

5.5. Modern map: Map of air routes

From the cartographic point of view, this map represents an interesting piece of work. Meridians, as well as parallels, are complex curves; the South Pole is formed by two points, symmetrical to the central meridian. Considering the elliptical outline of the Earth, it is evident that the map was constructed in some pseudoazimuthal or pseudocylindrical projection, proposed in the oblique aspect. For the analysis, 37 identical points, intersections of meridians and parallels, where $\Delta\varphi = 30^\circ$, $\Delta\lambda = 60^\circ$, were collected.

The result is presented in Fig. 23, the Hammer projection in the oblique aspect, where $R' = 0.1$ m, $\varphi_k = 45.2^\circ$, $\lambda_k = 179.7^\circ$, $\varphi_1 = 0.0^\circ$, $\lambda_0 = 0.3^\circ$, well-known as the Nordic projection, was recognized.

The Nordic projection, proposed by the Scottish cartographer and geographer John George Bartholomew (1805-1861), is the oblique version of the Hammer projection, centered at

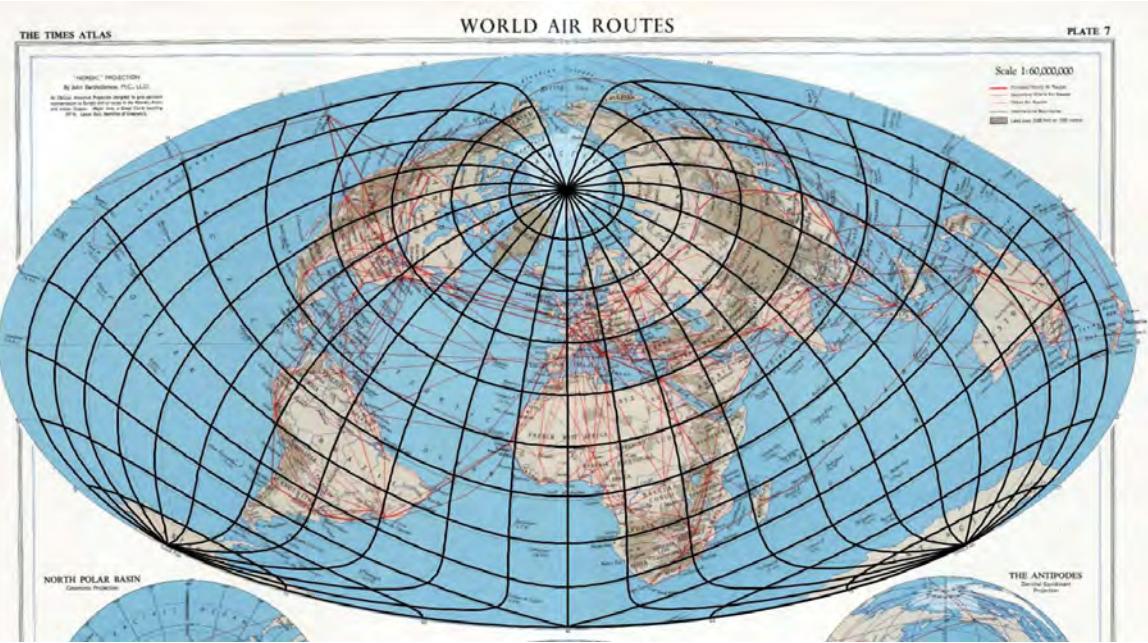


Figure 23: A superimposition of the reconstructed graticule in the Hammer projection and the original graticule on the analyzed map of air routes.

$C = [45^\circ, 0^\circ]$, which leads to the cartographic pole $K = [45^\circ, 180^\circ]$. It is obvious that the reconstructed graticule brings minor shape differences of the meridians close to the South Pole. The estimated map scale is $S = 93,046,047$.

5.6. Software *detectproj*

All proposed methods and optimizing techniques have been implemented in Java/C++ languages in the new software, **detectproj** [4] that is focused on map projection analysis. The source code, which contains approximately 25,000 lines and supports more than 100 map projections, is available in the source forge repository:

<https://sourceforge.net/projects/detectproj/>.

The software distributed under the GNU/GPL 2 license does not depend on any library; it supports more operating systems (Windows, GNU/Linux, Mac OS). It is useful for studying and analysis of different kinds of maps with the lack of information about the map projection.

Its graphical user interface is formed by two map windows side by side; see Fig. 24. As the reference map, the Open Street map client is used. Running in a separate thread the fast detection process makes the analysis comfortable. The candidate projections together with the visualization of the detected parameters (meridians and parallels) sorted by the objective function ϕ values are tabulated. The reconstructed graticule, test, and projected reference points can be extracted to the DXF file and processed by CAD or GIS software.

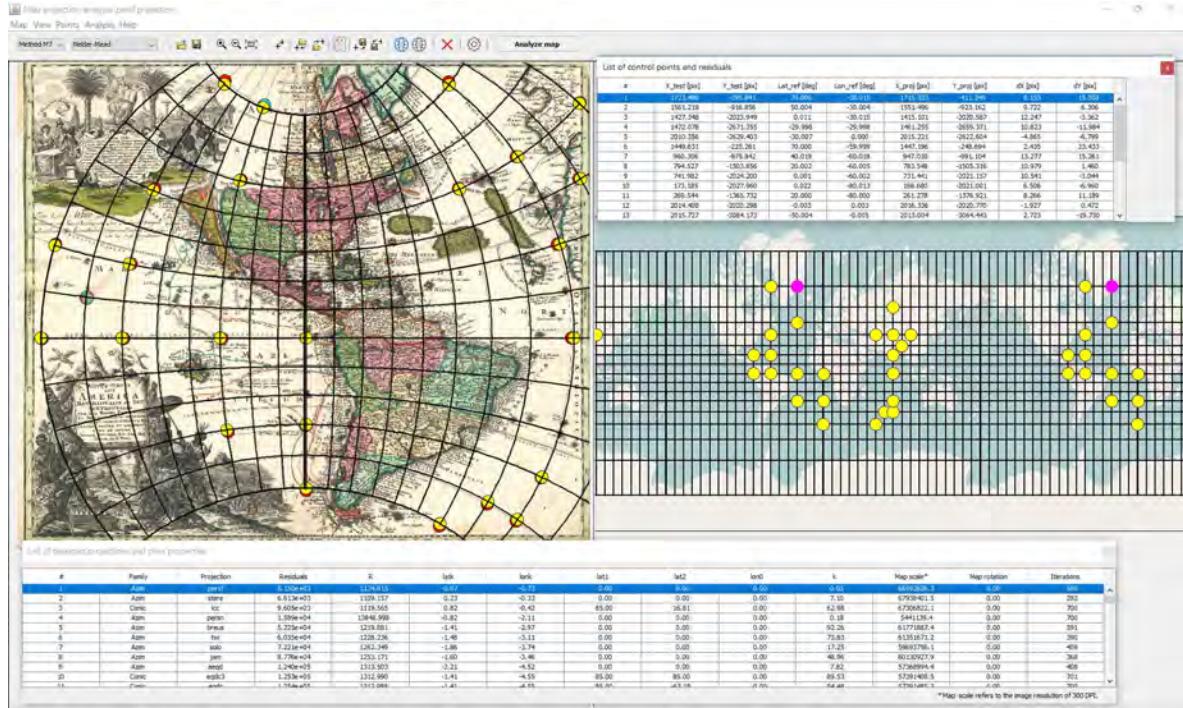


Figure 24: The detectproj software: the analyzed early map with the list of control points, the reconstructed graticule in stereographic projection, and the list of residuals on control points.

6. Conclusion

The proposed methods are focused on the detection and analysis of projections and their parameters. They are applicable to early or current maps if the information about the projection used is missing. Such an analysis is beneficial for various types of maps of different content and scale. Easy accessibility to large collections of digitized cartographic documents, where a huge amount of maps, atlases, or globes is available for viewing online, completely altered the way of working with them.

The proposed approach based on the analyzed map reprojection improves its georeference in a national grid. A threshold whether a projection may be recognized, is the geographic extent $\Delta\varphi = \Delta\lambda = 3^\circ$ of the analyzed territory. Within the size, there is no unambiguous solution; almost any projection may be set to fit well to the analyzed map.

Another important factor influencing the results is represented by the position of the analyzed territory. Territories around the equator, central parallel or near the poles suffer from the same problem; most projections have here an analogous shape of the graticule.

The uniform distribution of the analyzed features on the map also plays an important role, but cannot always be fulfilled (typically the portolan charts, where only the coastline is fully drawn). The user should collect the control points over the entire map so that they are placed symmetrically, on both sides of the equator and the central parallel.

Surprisingly, the amount of analyzed features does not play as important a role as expected.

In most cases 10-15 points are sufficient. Further increasing the amount of analyzed features does not improve the solution quality. For the world map and mid-scale maps, the accuracy of the collected points should be better than 3-4 mm on the map, which may be easily held.

Despite the efforts, several problems remain unsolved and others may be further improved. Some of them have only a minor effect on the proposed solution; others are more significant. The most important factors refer to the refined detection and rejection of the incorrectly drawn map elements, reflection of the solution to the search space, L1 norm minimization (more robust to outliers) problem, or more efficient scaling technique.

References

- [1] M Al-Baali and R Fletcher. "Variational methods for non-linear least-squares". In: *Journal of the Operational Research Society* (1985), pp. 405–421.
- [2] Ádám Barancsuk. "A Semi-automatic Approach for Determining the Projection of Small Scale Maps Based on the Shape of Graticule Lines". In: *Progress in Cartography: EuroCarto 2015*. Ed. by Georg Gartner, Markus Jobst, and Haosheng Huang. Cham: Springer International Publishing, 2016, pp. 267–288. ISBN: 978-3-319-19602-2. DOI: [10.1007/978-3-319-19602-2_17](https://doi.org/10.1007/978-3-319-19602-2_17).
- [3] Tomas Bayer. "Advanced methods for the estimation of an unknown projection from a map". In: *GeoInformatica* 20.2 (2016), pp. 241–284. DOI: [10.1007/s10707-015-0234-x](https://doi.org/10.1007/s10707-015-0234-x).
- [4] Tomas Bayer. *detectproj - software for the projection analysis*. 2017. URL: <https://sourceforge.net/projects/detectproj/>.
- [5] Tomas Bayer. *Estimation of an Unknown Cartographic Projection and its Parameters from a Map*. Habilitation thesis. Czech Technical University in Prague, 2017.
- [6] Tomas Bayer. "Estimation of an unknown cartographic projection and its parameters from the map". English. In: *GeoInformatica* 18.3 (2014), pp. 621–669. ISSN: 1384-6175. DOI: [10.1007/s10707-013-0200-4](https://doi.org/10.1007/s10707-013-0200-4).
- [7] Å. Björck. *Numerical Methods for Least Squares Problems*. Philadelphia: SIAM, 1996.
- [8] Chryssoula Boutoura. "Assigning map projections to portolan maps". In: *e-Perimetron* 1.1 (2006), pp. 40–50.
- [9] John E Dennis Jr, Sheng Songbai, and Phuong A Vu. *A memoryless augmented Gauss-Newton method for nonlinear least-squares problems*. Tech. rep. DTIC Document, 1985.
- [10] W. Flacke, B. Kraus, and C. Warcup. *Working with projections and datum transformations in ArcGIS: theory and practical examples*. Points Verlag, 2005. ISBN: 9783980846356. URL: <http://books.google.cz/books?id=PfEsAQAAQAAJ>.
- [11] R. Fletcher. *Practical methods of optimization* (2nd ed.) New York, NY, USA: Wiley-Interscience, 1987. ISBN: 0-471-91547-5.
- [12] R. Fletcher and C. Xu. "Hybrid Methods for Nonlinear Least Squares". In: *IMA Journal of Numerical Analysis* 7.3 (1987), pp. 371–389. DOI: [10.1093/imanum/7.3.371](https://doi.org/10.1093/imanum/7.3.371). eprint: <http://imajna.oxfordjournals.org/content/7/3/371.full.pdf+html>.
- [13] Gene H. Golub and Charles F. Van Loan. *Matrix computations* (3rd ed.) Baltimore, MD, USA: Johns Hopkins University Press, 1996. ISBN: 0-8018-5414-8.

- [14] Colin R. Goodall. “Computation using the QR decomposition”. In: *Computational Statistics*. Vol. 9. Handbook of Statistics. Elsevier, 1993, pp. 467–508. DOI: [10.1016/S0169-7161\(05\)80137-3](https://doi.org/10.1016/S0169-7161(05)80137-3).
- [15] P. Horata, S. Chiewchanwattana, and K. Sunat. “A comparative study of pseudo-inverse computing for the extreme learning machine classifier”. In: *Data Mining and Intelligent Information Technology Applications (ICMiA), 2011 3rd International Conference on*. Oct. Pp. 40–45.
- [16] Bernhard Jenny. *Map Analyst*. 2011. URL: <http://mapanalyst.org>.
- [17] Bernhard Jenny. “MapAnalyst-A digital tool for the analysis of the planimetric accuracy of historical maps”. In: *e-Perimetron* 1.3 (2006), pp. 239–245.
- [18] Bernhard Jenny and Lorenz Hurni. “Studying cartographic heritage: Analysis and visualization of geometric distortions”. In: *Computers & Graphics* 35.2 (2011), pp. 402–411.
- [19] Keith D. Lilley, Christopher D Lloyd, and Bruce M. S. Campbell. “Mapping the realm: A New Look at the Gough Map of Britain (c.1360)”. In: *Imago Mundi: The International Journal for the History of Cartography* 61.1 (2009), pp. 1–28. ISSN: 0308-5694. DOI: [10.1080/03085690802456228](https://doi.org/10.1080/03085690802456228).
- [20] Evangelos Livieratos. “Graticule versus point positioning in Ptolemy cartographies”. In: *e-Perimetron* 1.1 (2006), pp. 51–59.
- [21] Christopher D Lloyd and Keith D Lilley. “Cartographic veracity in medieval mapping: analyzing geographical variation in the Gough Map of Great Britain”. In: *Annals of the Association of American Geographers* 99.1 (2009), pp. 27–48.
- [22] L Lukšan. “Computational experience with known variable metric updates”. In: *Journal of Optimization Theory and Applications* 83.1 (1994), pp. 27–47.
- [23] L Lukšan. “Hybrid methods for large sparse nonlinear least squares”. English. In: *Journal of Optimization Theory and Applications* 89.3 (1996), pp. 575–595. ISSN: 0022-3239. DOI: [10.1007/BF02275350](https://doi.org/10.1007/BF02275350).
- [24] Ladislav Lukšan and Emilio Spedicato. “Variable metric methods for unconstrained optimization and nonlinear least squares”. In: *Journal of Computational and Applied Mathematics* 124.1 (2000), pp. 61–95.
- [25] William Ravenhill. “Projections for the large general maps of Britain, 1583–1700”. In: *Imago Mundi* 33.1 (1981), pp. 21–32. DOI: [10.1080/03085698108592512](https://doi.org/10.1080/03085698108592512).
- [26] Alastair Strang. “The analysis of Ptolemy’s geography”. In: *The Cartographic Journal* 35.1 (1998), pp. 27–47.
- [27] Waldo R Tobler. “Bidimensional regression”. In: *Geographical Analysis* 26.3 (1994), pp. 187–212.
- [28] Waldo R Tobler. “Measuring the similarity of map projections”. In: *The American Cartographer* 13.2 (1986), pp. 135–139.
- [29] Waldo R. Tobler. “Medieval distortions: The projections of ancient maps”. In: *Annals of the Association of American Geographers* 56.2 (1966), pp. 351–360. DOI: [10.1111/j.1467-8306.1966.tb00562.x](https://doi.org/10.1111/j.1467-8306.1966.tb00562.x).

- [30] Waldo R. Tobler. “Numerical Approaches to Map Projections”. In: *eitrage zur theoretischen Kartographie, Festschrift für Erik Amberger*, hg. 14 (1977), pp. 51–64.
- [31] Weijun Zhou and Xiaojun Chen. “Global Convergence of a New Hybrid Gauss-Newton Structured BFGS Method for Nonlinear Least Squares Problems”. In: *SIAM J. on Optimization* 20.5 (June 2010), pp. 2422–2441. ISSN: 1052-6234. DOI: [10.1137/090748470](https://doi.org/10.1137/090748470).

