

Advanced Methods for the Estimation of an Unknown Projection from a Map

Tomas Bayer — bayertom@natur.cuni.cz

Department of Applied Geoinformatics and Cartography, Faculty of Sciences, Charles University in Prague

Abstract

This article presents three new methods (M5, M6, M7) for the estimation of an unknown map projection and its parameters. Such an analysis is beneficial and interesting for historic, old, or current maps without information about the map projection; it could improve their georeference. The location similarity approach takes into account the residuals on the corresponding features; the minimum is found using the non-linear least squares. For the shape similarity approach, the minimized objective function ϕ takes into account the spatial distribution of the features, together with the shapes of the meridians, parallels and other 0D-2D elements. Due to the non-convexity and discontinuity, its global minimum is determined using the global optimization, represented by the differential evolution. The constant values of projection φ_k , λ_k , φ_1 , λ_0 , and map constants R , ΔX , ΔY , α (in relation to which the methods are invariant) are estimated. All methods are compared and the results are presented for the synthetic data as well as for 8 early maps from the Map Collection of the Charles University and the David Rumsey Map Collection. The proposed algorithms have been implemented in the new version of the `detectproj` software.

Keywords: digital cartography, map projection, optimizing, differential evolution, least squares, BFGS, early maps, analysis, georeferencing, cartographic heritage, history of cartography.

1 Introduction

The map projection analysis, together with the estimation of the best constant values of the projection, belongs to the new methods of the cartographic research of early maps. It represents a process of the recognition and reconstruction of the geometric relationship, between the early map content and the present representation of the Earth. The early maps are a source of interest for many scientific fields. For serious analysis of the map content, the correct position of the map in the projected coordinate system needs to be established. However, without any information about map projections, this may be complicated. In this context, the importance of the map projection analysis is primarily referred to the refinement of spatial georeference for medium- and small-scale maps, to the analysis of the knowledge about the world, and to the appropriate cataloging of maps, which is discussed in Sec. 4. In georeferencing small-scale maps, it is not correct to transform the analyzed map into a distinct projected coordinate system and neglect the influence of the different map projections. Fig. 1, illustrates this issue, when the similarity transformation is utilized.

There are many requirements imposed on the developed solution. It should be robust to gross errors, paper aging, sufficiently accurate, provide results in real

time, and invariant to the map constants. Even their partial acceptance represents a serious challenge. To improve the results in terms of reliability, accuracy, computational time, relevancy, and robustness against gross errors, three new methods M5, M6, and M7 are suggested. Compared to the approach presented in [5], some additional parameters are determined. The more accurate and on-the-fly analysis as well as a significantly lower amount of iterations allow use for maps of all scales and types.

Two approaches measuring the map similarity are presented. The location similarity estimates the map projection parameters minimizing the squared sum of residuals using the non-linear least squares solution. The shape similarity, where the objective function ϕ takes into account the spatial distribution of the analyzed features together with the shapes of the meridians and parallels, is minimized using differential evolution (DE). Due to its non-convexity, this problem is NP-hard. It will be also investigated whether, for the location similarity, it is possible to replace global optimizing techniques with the local optimization.



Figure 1: Georeferencing of a small-scale early map in the Bonne projection to the Mercator projection using the similarity transformation; the transformation key is determined from the identical points.

2 Related Work

Interest in early maps and their analysis has grown significantly, during the last 10 years. Therefore, the International Cartographic Association (ICA) established a Commission on Digital Technologies in Cartographic Heritage (2005). There are several early papers, dealing with the assessment of early maps [33], [44], their accuracy [69], [34], [56], the correspondence between patterns in geographical maps [74], the analysis of an ancient map projections [75], [76]. The projection analysis of the general map of Britain can be found in [68], the accuracy of the first world map in [79]. However, the development of computer sciences has provided new visualization techniques [38], [64], analytical methods and procedures [3], [4], [8], [51]. The analysis of the American civil war maps can be found in [73], the Ptolemy's map of Greece in [49], portolan maps in [60], the Gough map of Britain in [47], [52], the Mercator map of Slavonia, Croatia, Bosnia, and Dalmatiae in [67], the Vogt's map of Bohemia in [6]. There are software tools focused on georeferencing and the analysis of early maps [14], [43], [65], [18], widely used in various institutions, including the British Library [41], [65], and the New York Public Library [43]. So far, they do not allow automated detection of a map projection.

The method, measuring the map projection similarity from the residuals of corresponding points, was described in [77], the bidimensional regression for the comparison of geographic phenomena in [78]. Some simple methods for identifying a projection can be found in the Esri Knowledge Base [19], [20]. The `prjfinder` software [23] represents a more sophisticated tool for Arc Map, which searches for the best matching coordinate system. A different detection method, based on 2D transformations [36], was developed in [37], the algorithms were implemented in MapAnalyst open-source software. According to author knowledge, it is the only solution for the projection analysis that works well. Another approach based on the Nelder-Mead optimization of the objective function can be found in [5]. The method proposed in this paper is more robust; it supports the analysis of elements of different spatial dimensions, enables the determination of additional parameters of the map projection in all aspects (normal, transverse, oblique).

The detection process consists of different procedures that are frequently used in computer sciences. Given the wide range of the paper, not all relevant articles could be referenced. The detailed description of local optimizing methods for problems with a known gradient is in many books: [29], [27], [28], [7], [40], [58], [59]. The

BFGS method was introduced in [24], [31], [70], [10], [11]; the efficient combination of the globally convergent Gauss-Newton and BFGS methods was described in [45], [42]. Very powerful hybrid method combining the Gauss-Newton and BFGS methods for solving non-linear least squares (NLS) was introduced in [53], [54], [81]. The BFGS formula has been found to be one of the most efficient optimizing approaches; this fact was discussed in [15], [46], [55].

Differential evolution (DE) represents a global optimizing method in the continuous search domain; it was proposed in [71] as the replacement of simulated annealing. Several modified versions with a new mutation scheme were described in [13], [21], [22], [16]. Due to the small amount of input parameters, the differential evolution is easily adjustable [9], [66]. Appropriate setting of the DE parameters is crucial; this step has been extensively tested by many researchers: [63], [62], [48], [30], [80], [61]. A detailed overview can be found in [66]. The slow convergence can be accelerated by modifying the parameters [16], [35] or changing the evolution scheme [32], [39].

3 Map projection

In terms of our problem, let us put the simplification that only projections between the sphere and the plane, will be considered. Let S^2 be a sphere in \mathbb{R}^3 (reference surface, Earth), given by $x^2 + y^2 + z^2 = R^2$, $A = (0, 0, R)$, $B = (0, 0, -R)$ North and South Poles of S^2 , and σ a plane. For a current point $Q = [\varphi, \lambda] \in S^2$, different from A, B , and its image $P' = [x, y] \in \sigma$, the projection \mathbb{P} is

$$\mathbb{P} : S^2 - \{A, B\} \rightarrow \sigma : Q \rightarrow P'.$$

The projection \mathbb{P} is defined by the coordinate functions F, G , continuous with their first and second partial derivatives and finite

$$\begin{aligned} X(R, \varphi_1, \lambda_0, \Delta X) &= F(\varphi, \lambda), \\ Y(R, \varphi_1, \lambda_0, \Delta Y) &= G(\varphi, \lambda), \end{aligned}$$

where φ_1 is the standard parallel, λ_0 is the central meridian and $\Delta X, \Delta Y$ are the shifts between P, P' . If $\lambda_0 \neq 0$, the latitude is reduced

$$\lambda = \lambda - \lambda_0. \quad (1)$$

Map projections are frequently used in the transverse or oblique aspects; so $\lambda_0 \neq 0$ can be compensated by the coordinate λ_k , see Sec. 5.4. For the image $Q' = [\varphi', \lambda'] \in S'^2$ of Q on a sphere S'^2 in \mathbb{R}^3 concentric to S^2 , the transformation \mathbb{T}

$$\mathbb{T} : S^2 \rightarrow S'^2 : Q \rightarrow Q'$$

turns the North Pole A into the new meta pole $K = [\varphi_k, \lambda_k]$. The transformation is given by the functions H, I

$$\varphi'(\varphi_k, \lambda_k) = H(\varphi, \lambda), \quad \lambda'(\varphi_k, \lambda_k) = I(\varphi, \lambda), \quad (2)$$

which express the laws of the spherical trigonometry. The coordinate functions F, G can be rewritten in the following, more complex, form of the composite functions

$$\begin{aligned} X(R, \varphi_k, \lambda_k, \varphi_1, \lambda_0, \Delta X) &= F(H(\varphi, \lambda), I(\varphi, \lambda)), \\ Y(R, \varphi_k, \lambda_k, \varphi_1, \lambda_0, \Delta Y) &= G(H(\varphi, \lambda), I(\varphi, \lambda)). \end{aligned}$$

From the laws of cosine for φ' and sine for λ' are

$$\begin{aligned} \sin \varphi' &= \sin \varphi_k \sin \varphi + \cos \varphi_k \cos \varphi \cos(\lambda - \lambda_k), \quad (3) \\ \sin \lambda' &= \frac{\sin(90^\circ - \varphi) \sin(\lambda - \lambda_k)}{\sin(90^\circ - \varphi')}. \end{aligned}$$

Due to the ambiguity in $\lambda' \in [-180^\circ, 180^\circ]$, the second computation based on the five-part rule is needed

$$\cos \lambda' = \frac{\cos \varphi \sin \varphi_k \cos(\lambda - \lambda_k) - \sin \varphi \cos \varphi_k}{\cos \varphi'}.$$

To avoid the quadrant adjustment, the both equations are divided each other, and the final formula is independent of φ'

$$\tan \lambda' = \frac{\cos \varphi \sin(\lambda - \lambda_k)}{\cos \varphi \sin \varphi_k \cos(\lambda - \lambda_k) - \sin \varphi \cos \varphi_k}. \quad (4)$$

If the prime meridian does not correspond to Greenwich but to a different meridian (Ferro, Paris, Lisbon) of the longitude λ_0 , the longitude λ needs to be reduced subtracting λ_0 . For the current point Q_i , and its image P'_i , the projection equations are written as follows

$$\begin{aligned} X_i(R, \varphi_k, \lambda_k, \varphi_1, \lambda_0, \Delta X) &= F(H(\varphi_i, \lambda_i), I(\varphi_i, \lambda_i)), \\ Y_i(R, \varphi_k, \lambda_k, \varphi_1, \lambda_0, \Delta Y) &= G(H(\varphi_i, \lambda_i), I(\varphi_i, \lambda_i)), \end{aligned}$$

or, when a rotation involved

$$\begin{aligned} X_i(R, \varphi_k, \lambda_k, \varphi_1, \lambda_0, \Delta X, \alpha) &= F(H(\varphi_i, \lambda_i), I(\varphi_i, \lambda_i)), \\ Y_i(R, \varphi_k, \lambda_k, \varphi_1, \lambda_0, \Delta Y, \alpha) &= G(H(\varphi_i, \lambda_i), I(\varphi_i, \lambda_i)). \end{aligned}$$

Determined parameters of the projection. During the analysis, the following parameters of the projection \mathbb{P} , affecting the graticule shape, are estimated:

- *Transformed pole position* φ_k, λ_k

The transformed pole position $K = [\varphi_k, \lambda_k]$ of the analyzed projection has a significant influence on the shape of the graticule.

- *Standard parallel* φ_1

The latitude of the standard parallel φ_1 , along which the nominal scale is preserved, is determined. Conic projections in a secant form specify two standard parallels φ_1, φ_2 , $\varphi_1 \neq \varphi_2$, representing intersections of the cone and the secant plane; or

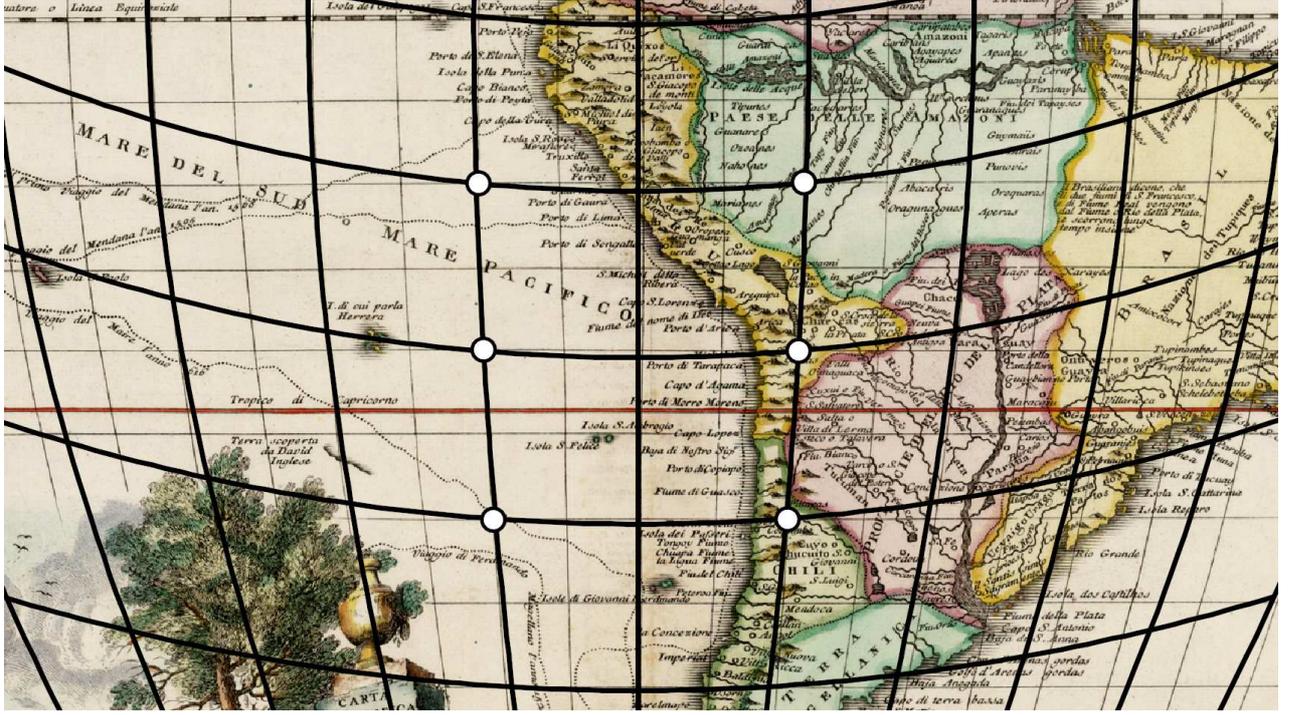


Figure 2: An insufficient amount of analyzed features on parallels leads to the detection of a different type of projection.

$\varphi_1 = \varphi_2$ for its tangent form. For cylindrical projections, the secant form uses two standard parallels $\varphi_1 = -\varphi_2$, but for the tangent form, is $\varphi_1 = \varphi_2 = 0$. Some pseudo- or polyconic projections require the standard parallel φ_1 , but for others, it is not used. During the analysis, the cartographic model is *simplified* to $\varphi_1 = \varphi_2$. However, not all projections define the standard parallel, then, $\varphi_1 = 0$.

- *Longitude λ_0 of the central meridian*

To minimize the distortion and provide a true depiction of the mapped region, the central meridian is shifted by λ_0 .

Determined constants of the map. The proposed solution is invariant to the following constants of the analyzed map M , which do not influence the graticule shape:

- *Auxiliary sphere radius R'*

The analyzed and reference maps may have different scales; the auxiliary sphere radius $R' \equiv R$ ensures that the analyzed map fits best to the reference one.

- *Shifts $\Delta X, \Delta Y$*

Both the analyzed and projected reference coordinate systems may be shifted each other. Sometimes, the shifts are small, especially for the

scanned maps, otherwise, become larger values (UTM, shifted origin, $\Delta Y = 500,000,000$ m).

- *Angle of rotation α*

This optional parameter treats the inappropriate insertion of the map into the scanner, or its additional rotation on the page.

The parameters are stored in the vector x and determined so as to minimize the objective function ϕ . While the estimation of $R', \Delta X, \Delta Y$ leads to the convex optimization, the remaining parameters result in the non-convex optimization. Instead of the Nelder-Mead (NM) approach [57], [5], the differential evolution (DE), and the non-linear least squares (NLS), are utilized.

4 The concept and utilization of the analysis

The map projection analysis represents a challenging, but conceptually difficult, problem. Compared to the solution based on the Nelder-Mead optimization [5], the methods proposed in this paper are better in several aspects (accuracy, reliability, convergence, time consumption, and robustness). Let us give a brief description of the basic concepts and the practical applications.

Test and reference features. Consider a set of the test features $P = \{P_1, \dots, P_n\} \subset \mathbb{R}^2$, ($1 \leq n \leq \infty$), where

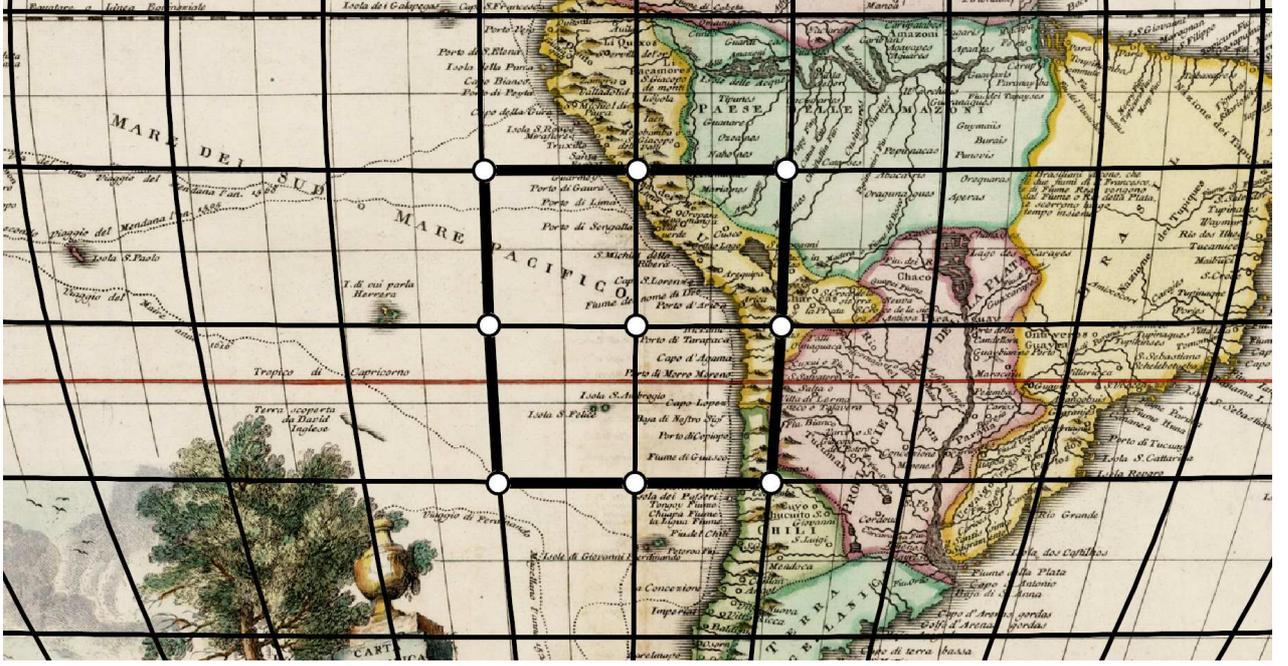


Figure 3: The analyzed features covering a small part of the analyzed territory; the estimated graticule fits the original only inside the convex hull.

P_i represents a point, a line segment, or a closed area on the analyzed map, and analogously, a set of the reference features $Q = \{Q_1, \dots, Q_n\} \subset \mathbb{R}^2$, ($1 \leq n \leq \infty$) on the sphere. Assume that the elements in P and Q do not intersect each other, and that any element P_i corresponds to Q_i .

Overview of the analysis. Our approach is based on a comparison of the subset on the analyzed map and the corresponding subset on the reference surface. The Cartesian and geographical coordinates of each analyzed feature of both subsets are known. Let \mathbb{P} represent the analyzed map projection with the determined parameters x . Suppose the dissimilarity δ

$$\delta = \phi(\mathbb{P}_x(Q), P),$$

of features on the analyzed map P and the reference surface Q projected with \mathbb{P} , measured by the objective function ϕ . A projection \mathbb{P}

$$\mathbb{P}_x : Q \rightarrow P'_x,$$

and its constant values x need to be found such that, satisfying

$$\delta = \phi(P, P'_x) = 0,$$

which is fulfilled if $P = P'_x$. In general, the elements P, Q are affected by the errors, therefore

$$\phi(P, P'_x) \neq 0, \quad P \neq P'_x.$$

The unconstrained global optimizing problem is being solved, where the vector of its best constant values \hat{x} is

given by

$$\phi(P, P'_{\hat{x}}) = \min, \quad \hat{x} = \arg \min_{\mathbb{P}} (\phi(P, P'_x)). \quad (5)$$

Under all analyzed projections, the optimal projection $\bar{\mathbb{P}}$, and its the best constant values \tilde{x} hold

$$\tilde{x} = \arg \min_{\forall \mathbb{P}} (\phi(P, P'_{\tilde{x}})). \quad (6)$$

The projection with the smallest objective function value is assigned to the analyzed map. Instead of the global optimization, which is computationally expensive, the problem may be solved using the non-linear least squares. There are many ways to propose the objective function. The most simple approach is the *location similarity*, when the objective function with the known gradient $\nabla\phi$ is represented by the squared sum of the residuals. The objective function ϕ may be more complex, or discrete, when the *shape similarity* is taken into account. For such a function its gradient $\nabla\phi$ may not be available for computation.

The most important factors affecting the results will be discussed; their impact on the projection analysis will be shown.

4.1 Spatial distribution of features

A crucial role is played by the uniform distribution of the analyzed features. The following rules should be respected:

- To capture the curvature, at least three points should be placed over each analyzed meridian or parallel (i.e., north-south and east-west directions). Fig. 2 illustrates the situation when the meridians are well reconstructed but the parallel shape was not recognized.
- The analyzed features should preferably be distributed over the entire analyzed map. The estimated projection parameters fit well inside the convex hull of the analyzed set; any extrapolation outside the analyzed set is not supported and leads to the wrong results, see Fig. 3.
- The analyzed features should be placed “symmetrically”, on both sides of the equator and the central meridian.

4.2 The graticule and the map content accuracy

In the early maps created before the 17th century, the graticule is the geometric construct. However, the map content was not measured seriously; it is not too accurate, some territories are completely missing (Australia, the polar regions of North America). The reconstructed graticule generated from the determined parameters can be shifted or rotated to the original graticule. For early maps, the graticule as well as the map frame are significantly more accurate than the map content, which was discussed in [50]; they are also not affected by the generalization. There is also far less chance of outliers, so the estimated parameters and reconstructed graticule fit better to the original.

4.3 Analysis and the territory size

The success of map projection analysis is greatly affected by the size of the analyzed territory. If there are no measurable distortions (small depicted territories, territories near the equator, central meridian, poles, or large scale maps) representing the *projection footprint*, there is a little chance that map projection could be detected; the analyzed territory has a similar shape in multiple projections.

4.4 Analysis and georeference

Georeferencing assigns a spatial information to each pixel of the map so as it aligns to a projected coordinate system and obtains a correct geographical location. The problem arises when the analyzed map and the destination coordinate system are based on different projections. A similarity relationship between two sufficiently small territories may be established; the impact of the map projection may be neglected. For large territories, the

linear transformation of the analyzed map to the destination coordinate system omitting the influence of different projections is undesirable. Increasing the order of transformation does not lead to any reasonable results; the higher order transformations cause the secondary deformation and twist of the map content.

Another approach, splitting the analyzed small-scale map into tiles, applying a transformation to each tile, with the restoration of the continuous raster image, is time-consuming and laborious.

The correct method is to determine the analyzed map projection and re-project the map from the Cartesian coordinates to geographical coordinates using inverse projection formulas. This is followed by projecting the geographical coordinates to the destination map’s coordinate system. Suppose that our analyzed small-scale early map, created in the thus far unknown projection \mathbb{P}_1 , needs to get the spatial reference in the coordinate system given by different projection \mathbb{P}_2 representing the national grid, heterogeneous to the map’s coordinate system. Therefore, the best recommended solution seems to be:

1. *The determination of the analyzed map projection \mathbb{P}_1*

Using the proposed solution, the vector \tilde{x} of the analyzed map projection \mathbb{P}_1 parameters, is determined.

2. *The inverse re-projection*

Inverse projection formulas are applied; the analyzed map is reprojected to the sphere $\mathbb{P}_1^{-1} : P' \rightarrow Q'$. This step may be solved using the Proj.4 library.

3. *Projection into destination coordinate system*

The map re-projected on the sphere is projected to the destination coordinate system given by \mathbb{P}_2 , so $\mathbb{P}_2 : Q' \rightarrow P''$. Nevertheless, possible residuals may be fixed using a transformation.

4. *Corrections of shifts and rotation*

Despite the proposed solution, the additional shifts and rotation may occur between the reconstructed and reference maps at the destination coordinate system. They may be corrected with the zero residual transformation.

The procedure is shown in Fig. 5, the georeferenced map in the Bonne projection, reprojected to the Mercator projection, is illustrated in Fig. 4. This problem was discussed in [37], where an analogous scheme was presented.

4.5 Analysis and incorrectly drawn map content

For early maps created before the 17th century, the lack of a solid geometric and geodesic basis is typical. Maps

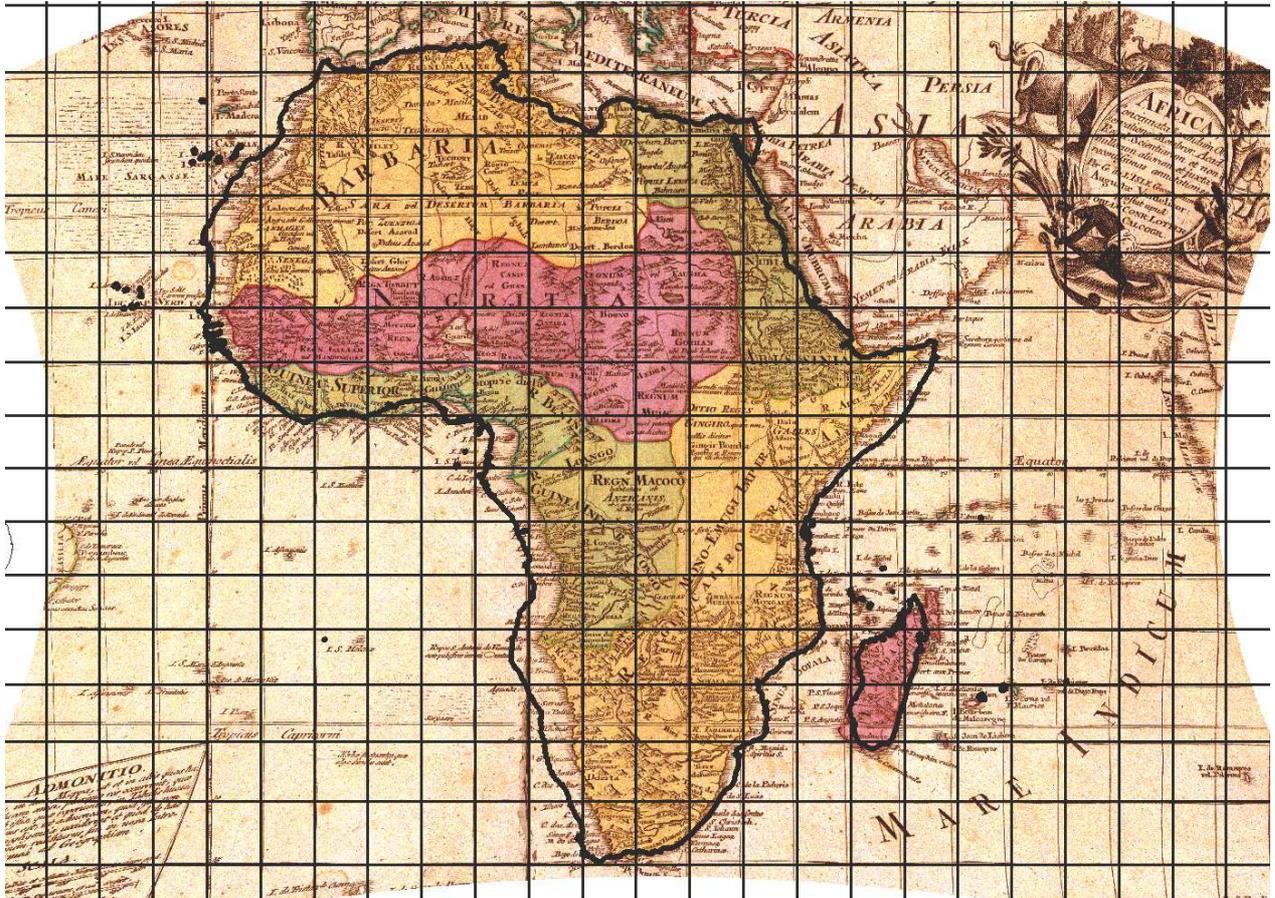


Figure 4: Georeferencing of a small-scale early map in the Bonne projection using the proposed solution: the re-projection to the Mercator projection. Only slight discrepancies between the original and reconstructed graticules are visible.

without a geometric basis, affected by many errors, were more like charts. Their inappropriate geometric properties, especially the length, aerial and shape distortions, randomly changing depending on the geographic position, makes the detection process more difficult and ambiguous. The map content was not seriously measured, and the only geometric basis can be found on the graticule, as well as on the map frame. This feature may be utilized for further analysis of the map, which gives an answer, as to which territories are drawn more or less accurately. Many territories are well-placed with the distorted shape, or well-shaped with the systematic shifts, or a combination of both factors may occur. When superimposing the early map on the current state, the spatial or location dissimilarities become clearly visible, see Fig. 4. For the analysis of the incorrectly drawn elements, M-estimates with the Huber function, solved by the iteratively reweighted least squares (IRLS), were used.

4.6 Cataloging of maps

Currently, many map collections have been digitized, or the process is still in progress. To be easily accessi-

ble, they must be sorted, organized, and stored in the database, represented and organized as a catalog. The cataloging of maps creates the need for information about the map projection, which form a part of the cartographic metadata. 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), and 342 (Geospatial Reference Data); see Fig. 6. It is evident that some parameter values may be visually estimated better (projection family), worse (projection aspect), or the visual estimation is impossible (standard parallel). 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.

5 Detection based on non-linear least squares

The first strategy is based on solving the weighted non-linear least squares problem (NLS). NLS is globally con-

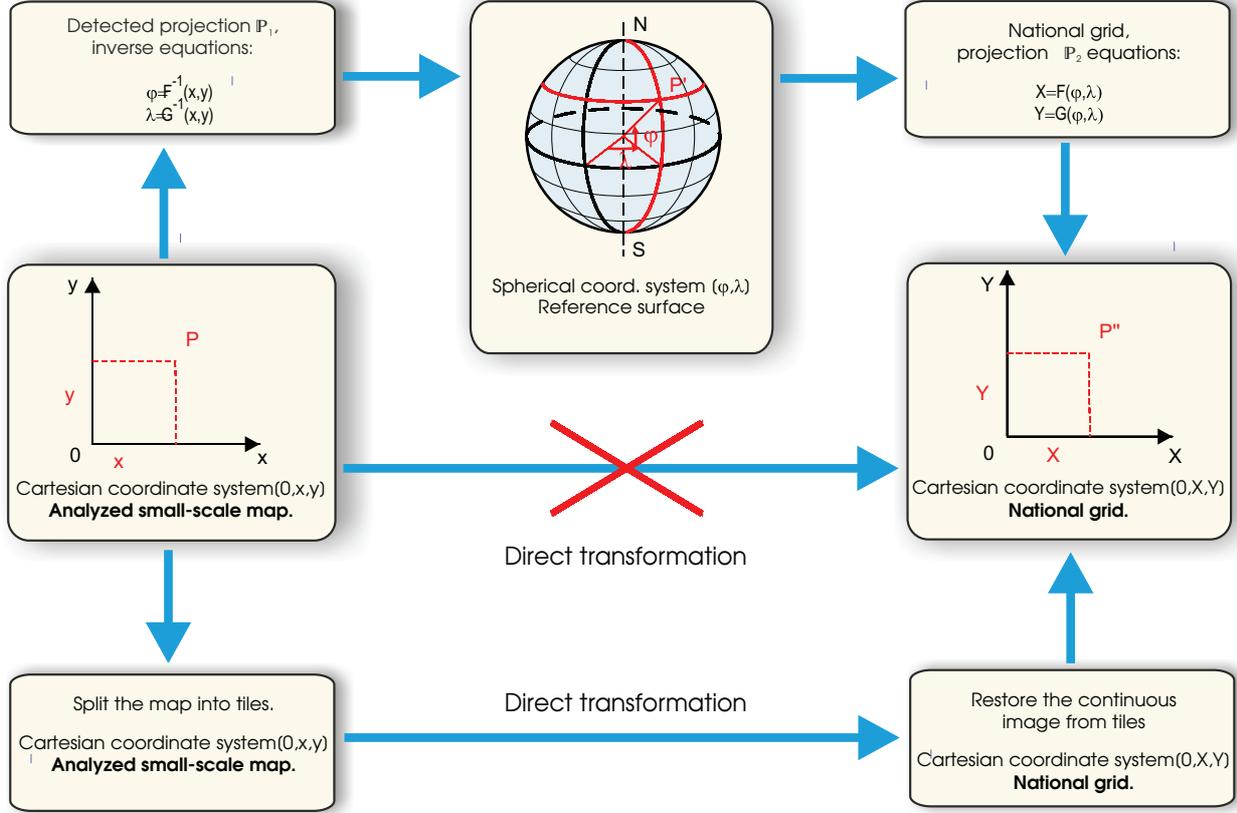


Figure 5: Georeferencing of a small-scale early map to the national grid, several approaches are illustrated.

vergent to the local minimum, but finding the global minimum is not ensured. Despite this fact, it provides fast and reliable results, which for this problem, are not significantly worse than the global minimizer. It is supposed that the objective function $\phi(x)$ is available for the gradient computations. The NLS approach finds a minimizer \hat{x}

$$\hat{x} = \arg \min(\phi(x)), \quad (7)$$

of the sum of the squares of m non-linear functions $r_i(x)$

$$\phi(x) = \frac{1}{2} \sum_{i=1}^m r_i^2(x) = \frac{1}{2} r^T(x)r(x),$$

where x and r are m -dimensional and n -dimensional vectors of the variables and residuals, and $n \geq m$. For the i -th item of the given data b_i , $i = 1, 2, \dots, m$, and the model function $g(x_i)$, the residual $r(x_i)$ is given by

$$r(x_i) = g(x_i) - b_i.$$

Three methods, denoted as M7, M6, and M5 minimizing Eq. 7, are presented. They differ in the number of determined parameters and the robustness, but their computational demands are similar.

5.1 The 7-parameter method

The first method (denoted as M7) is based on the estimation of 7 parameters

$$\hat{x}^T = [R' \quad \varphi_k \quad \lambda_k \quad \varphi_1 \quad \lambda_0 \quad \Delta X \quad \Delta Y].$$

The model function $g(x)$ is represented by the coordinate functions of the projection \mathbb{P}

$$g(x) = \begin{bmatrix} F(H(\varphi_1, \lambda_1), I(\varphi_1, \lambda_1)) \\ \dots \\ F(H(\varphi_n, \lambda_n), I(\varphi_n, \lambda_n)) \\ G(H(\varphi_1, \lambda_1), I(\varphi_1, \lambda_1)) \\ \dots \\ G(H(\varphi_n, \lambda_n), I(\varphi_n, \lambda_n)) \end{bmatrix},$$

the vector of residuals r , between sets P, P'_x , is written as follows

$$r(x) = \begin{bmatrix} X_1 - x_1, \\ \dots \\ X_n - x_n \\ Y_1 - y_1 \\ \dots \\ Y_n - y_n \end{bmatrix}. \quad (8)$$

While the set P has a fixed position, the size and position of the P'_x set changes from iteration to iteration. The

```

<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">(025°44`03&quot;;--179°34`41&quot;; /081°09`00&quot;; --077°44`44&quot;;)</subfield>

```

Figure 6: A part of an XML file containing cartographic meta data.

residuals are continuously decreasing, until a suitable solution is found. Taking into account the 7 unknown parameters, at least $m = 4$ analyzed features must be collected by the user. Due to the problematic determination of the initial values $\Delta X, \Delta Y$, this method proved to be less robust and reliable, especially for large shifts, which is discussed in Sec. 7.

Jacobian matrix. The Jacobian matrix $J(2n, 7)$ of the coordinate functions F, G is

$$J = \begin{bmatrix} \dot{j}_{1,1} & \dot{j}_{1,2} & \dot{j}_{1,3} & \dot{j}_{1,4} & \dot{j}_{1,5} & \dot{j}_{1,6} & \dot{j}_{1,7} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dot{j}_{n,1} & \dot{j}_{n,2} & \dot{j}_{n,3} & \dot{j}_{n,4} & \dot{j}_{n,5} & \dot{j}_{n,6} & \dot{j}_{n,7} \\ \dot{j}_{1+n,1} & \dot{j}_{1+n,2} & \dot{j}_{1+n,3} & \dot{j}_{1+n,4} & \dot{j}_{1+n,5} & \dot{j}_{1+n,6} & \dot{j}_{1+n,7} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dot{j}_{2n,1} & \dot{j}_{2n,2} & \dot{j}_{2n,3} & \dot{j}_{2n,4} & \dot{j}_{2n,5} & \dot{j}_{2n,6} & \dot{j}_{2n,7} \end{bmatrix}$$

where the partial derivative of the first coordinate function according to determined parameters are

$$\begin{aligned} \dot{j}_{i,1}(x, \varphi, \lambda) &= \left(\frac{\partial X}{\partial R'} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)}, \\ \dot{j}_{i,2}(x, \varphi, \lambda) &= \left(\frac{\partial X}{\partial \varphi_k} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)}, \\ \dot{j}_{i,3}(x, \varphi, \lambda) &= \left(\frac{\partial X}{\partial \lambda_k} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)}, \\ \dot{j}_{i,4}(x, \varphi, \lambda) &= \left(\frac{\partial X}{\partial \varphi_1} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)}, \\ \dot{j}_{i,5}(x, \varphi, \lambda) &= \left(\frac{\partial X}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)}, \\ \dot{j}_{i,6}(x, \varphi, \lambda) &= \left(\frac{\partial X}{\partial \Delta X} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)} = 1, \\ \dot{j}_{i,7}(x, \varphi, \lambda) &= \left(\frac{\partial X}{\partial \Delta Y} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)} = 0, \end{aligned}$$

and the partial derivative of the second coordinate function according to the determined parameters are

$$\begin{aligned} \dot{j}_{i+n,1}(x, \varphi, \lambda) &= \left(\frac{\partial Y}{\partial R'} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)}, \\ \dots &\dots \dots \\ \dot{j}_{i,n+6}(x, \varphi, \lambda) &= \left(\frac{\partial Y}{\partial \Delta X} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)} = 0, \\ \dot{j}_{i,n+7}(x, \varphi, \lambda) &= \left(\frac{\partial Y}{\partial \Delta Y} \right)_{(x=x^{(k)}, \varphi=\varphi_i, \lambda=\lambda_i)} = 1. \end{aligned}$$

In other words, each row $J(i, :)$ stores the values of the partial derivatives for the point Q_i .

5.2 The 5-parameter method

The next method proposed in this paper (denoted as M5) determines the following parameters

$$\hat{x}^T = [R' \quad \varphi_k \quad \lambda_k \quad \varphi_1 \quad \lambda_0].$$

It aligns both sets P, P'_x in their centers of mass $C = [X_C, Y_C]$, $c = [x_c, x_c]$

$$\begin{aligned} X_C &= \frac{1}{n} \sum_{j=1}^n X_j, & Y_C &= \frac{1}{n} \sum_{j=1}^n Y_j, \\ x_c &= \frac{1}{n} \sum_{j=1}^n x_j, & y_c &= \frac{1}{n} \sum_{j=1}^n y_j, \end{aligned}$$

so it eliminates the need to determine the $\Delta X, \Delta Y$ shifts. Let us denote the model function from the previous method M7 with the upper index ⁽¹⁾. The model function $g(x)$ is

$$g(x) = \begin{bmatrix} g_1^{(1)}(x) - \frac{1}{n} \sum_{l=1}^n F(H(\varphi_l, \lambda_l), I(\varphi_l, \lambda_l)) \\ \dots \\ g_n^{(1)}(x) - \frac{1}{n} \sum_{l=1}^n F(H(\varphi_l, \lambda_l), I(\varphi_l, \lambda_l)) \\ g_{n+1}^{(1)}(x) - \frac{1}{n} \sum_{l=1}^n F(H(\varphi_l, \lambda_l), I(\varphi_l, \lambda_l)) \\ \dots \\ g_{2n}^{(1)}(x) - \frac{1}{n} \sum_{l=1}^n F(H(\varphi_l, \lambda_l), I(\varphi_l, \lambda_l)) \end{bmatrix}.$$

The vector of residuals r , between aligned sets \bar{P}, \bar{P}' in coordinate systems (C, X', Y') and (c, x', y') , minimized with the least squares adjustment, is written as follows

$$r(x) = \begin{bmatrix} (X_1 - X_c) - (x_1 - x_c) \\ \dots \\ (X_n - X_c) - (x_n - x_c) \\ (Y_1 - Y_c) - (y_1 - y_c) \\ \dots \\ (Y_n - Y_c) - (y_n - y_c) \end{bmatrix} = \begin{bmatrix} X'_1 - x'_1 \\ \vdots \\ X'_n - x'_n \\ Y'_1 - y'_1 \\ \vdots \\ Y'_n - y'_n \end{bmatrix}, \quad (9)$$

where the reduced coordinates are

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

The M5 method, set as the *primary analysis tool*, brings a better convergence; $\|J\|$ is significantly smaller. At least, 3 points are necessary. During the iteration process, the aligned set \bar{P} has a fixed position and size. The properties of \bar{P}' are continuously changing, depending on x_k , the residuals between \bar{P}, \bar{P}' , decrease; see Fig. 7, illustrating the analysis of the Eckert V projection in the normal aspect. The fast convergence rate is clearly visible. M5 is not robust against a possible rotation of the analyzed map. A situation for a map rotated by $\alpha = 1.7^\circ$ due to inappropriate insertion into the scanner is shown in Fig. 8.

Jacobian matrix. Let us denote the derivatives from the previous method M7 with the upper index ⁽¹⁾. The new Jacobian matrix $J(2n, 5)$ is

$$J = \begin{bmatrix} \dot{j}_{1,1} & \dot{j}_{1,2} & \dot{j}_{1,3} & \dot{j}_{1,4} & \dot{j}_{1,5} \\ \dots & \dots & \dots & \dots & \dots \\ \dot{j}_{n,1} & \dot{j}_{n,2} & \dot{j}_{n,3} & \dot{j}_{n,4} & \dot{j}_{n,5} \\ \hline \dot{j}_{1+n,1} & \dot{j}_{1+n,2} & \dot{j}_{1+n,3} & \dot{j}_{1+n,4} & \dot{j}_{1+n,5} \\ \dots & \dots & \dots & \dots & \dots \\ \dot{j}_{2n,1} & \dot{j}_{2n,2} & \dot{j}_{2n,3} & \dot{j}_{2n,4} & \dot{j}_{2n,5} \end{bmatrix},$$

where the partial derivatives of the first coordinate function according to the determined parameters are

$$j_{i,1}(x, \varphi, \lambda) = j_{i,1}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial R'} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)},$$

$$j_{i,2}(x, \varphi, \lambda) = j_{i,2}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \varphi_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)},$$

$$j_{i,3}(x, \varphi, \lambda) = j_{i,3}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \lambda_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)},$$

$$j_{i,4}(x, \varphi, \lambda) = j_{i,4}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \varphi_1} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)},$$

$$j_{i,5}(x, \varphi, \lambda) = j_{i,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)},$$

and analogously, for the second coordinate function,

$$\begin{aligned} j_{i+n,1}(x, \varphi, \lambda) &= j_{i+n,1}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial R'} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)}, \\ \dots \dots \dots & \\ j_{i+n,5}(x, \varphi, \lambda) &= j_{i+n,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)}. \end{aligned}$$

5.3 The 6-parameter method

The last method proposed in this paper represents an extension of the M5 method, but it supports the additional rotation α of the map. In other words, it is invariant to all map constants. The vector of determined parameters is

$$\hat{x}^T = [R' \quad \varphi_k \quad \lambda_k \quad \varphi_1 \quad \lambda_0 \quad \alpha].$$

Analysis will be performed in local coordinate systems (C, χ', γ') and (c, x', y') , with origins c, C , where C is also the rotation center; this approach is analogous to the similarity transformation. Taking into account the residual of i -th element

$$\begin{bmatrix} r_{x_i} \\ r_{y_i} \end{bmatrix} = \Re \begin{bmatrix} X_i - X_c \\ Y_i - Y_c \end{bmatrix} - \begin{bmatrix} x_i - x_c \\ y_i - y_c \end{bmatrix},$$

where \Re represents the rotation matrix

$$\Re = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix},$$

the vector of the residuals is given by

$$r(x) = \begin{bmatrix} (X_1 - X_c) \cos \alpha - (Y_1 - Y_c) \sin \alpha - (x_1 - x_c), \\ \dots \\ (X_n - X_c) \cos \alpha - (Y_n - Y_c) \sin \alpha - (x_n - x_c) \\ (X_1 - X_c) \sin \alpha + (Y_1 - Y_c) \cos \alpha - (y_1 - y_c) \\ \dots \\ (X_n - X_c) \sin \alpha + (Y_n - Y_c) \cos \alpha - (y_n - y_c) \end{bmatrix}.$$

Denoting

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

the residuals may be rewritten as

$$r(x) = \begin{bmatrix} \chi'_1 - x'_1, \\ \dots \\ \chi'_n - x'_n \\ \gamma'_1 - y'_1 \\ \dots \\ \gamma'_n - y'_n \end{bmatrix}. \quad (10)$$

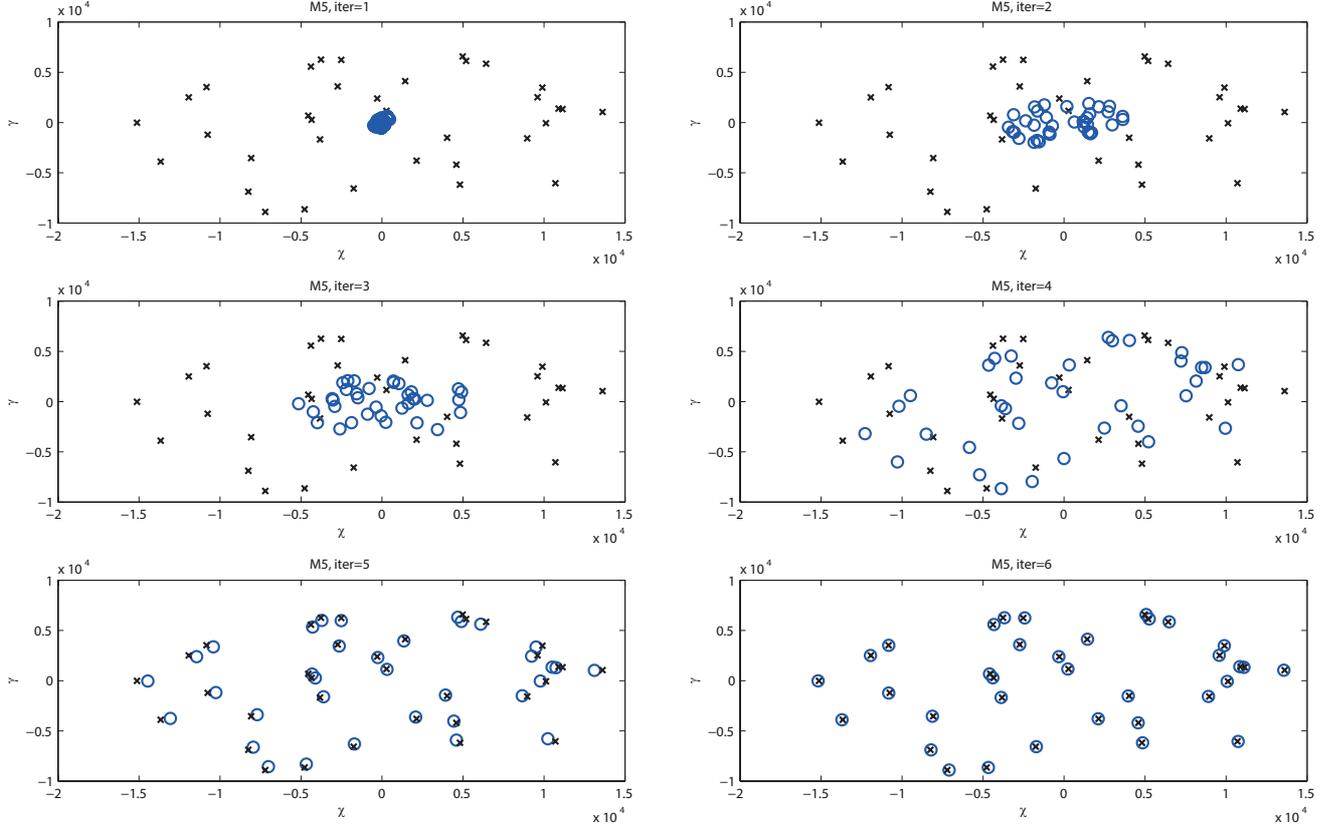


Figure 7: Detection of the Eckert V projection in the normal aspect using the M5 method, a random set of analyzed features, iterations 1-6. The residuals between \bar{P}, \bar{P}' are continuously decreasing.

Jacobian matrix. The Jacobian matrix $J(2n, 7)$ has the form of

$$J = \begin{bmatrix} \dot{j}_{1,1} & \dot{j}_{1,2} & \dot{j}_{1,3} & \dot{j}_{1,4} & \dot{j}_{1,5} & \dot{j}_{1,6} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dot{j}_{n,1} & \dot{j}_{n,2} & \dot{j}_{n,3} & \dot{j}_{n,4} & \dot{j}_{n,5} & \dot{j}_{n,6} \\ \hline \dot{j}_{1+n,1} & \dot{j}_{1+n,2} & \dot{j}_{1+n,3} & \dot{j}_{1+n,4} & \dot{j}_{1+n,5} & \dot{j}_{1+n,6} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dot{j}_{2n,1} & \dot{j}_{2n,2} & \dot{j}_{2n,3} & \dot{j}_{2n,4} & \dot{j}_{2n,5} & \dot{j}_{2n,6} \end{bmatrix},$$

$$j_{i,3}(x) = \begin{bmatrix} j_{i,3}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \lambda_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \cos \alpha - \begin{bmatrix} j_{i+n,3}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \lambda_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \sin \alpha,$$

$$j_{i,4}(x) = \begin{bmatrix} j_{i,4}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \varphi_1} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \cos \alpha - \begin{bmatrix} j_{i+n,4}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \varphi_1} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \sin \alpha,$$

where the partial derivatives of the first coordinate χ' according to the determined parameters are

$$j_{i,1}(x) = \begin{bmatrix} j_{i,1}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial R'} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \cos \alpha - \begin{bmatrix} j_{i+n,1}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial R'} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \sin \alpha,$$

$$j_{i,2}(x) = \begin{bmatrix} j_{i,2}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \varphi_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \cos \alpha - \begin{bmatrix} j_{i+n,2}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \varphi_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \sin \alpha,$$

$$j_{i,5}(x) = \begin{bmatrix} j_{i,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \cos \alpha - \begin{bmatrix} j_{i+n,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \end{bmatrix} \sin \alpha,$$

$$j_{i,6}(x) = - \left(X_i - \frac{1}{n} \sum_{l=1}^n X_l \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \sin \alpha - \left(Y_i - \frac{1}{n} \sum_{l=1}^n Y_l \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \cos \alpha,$$

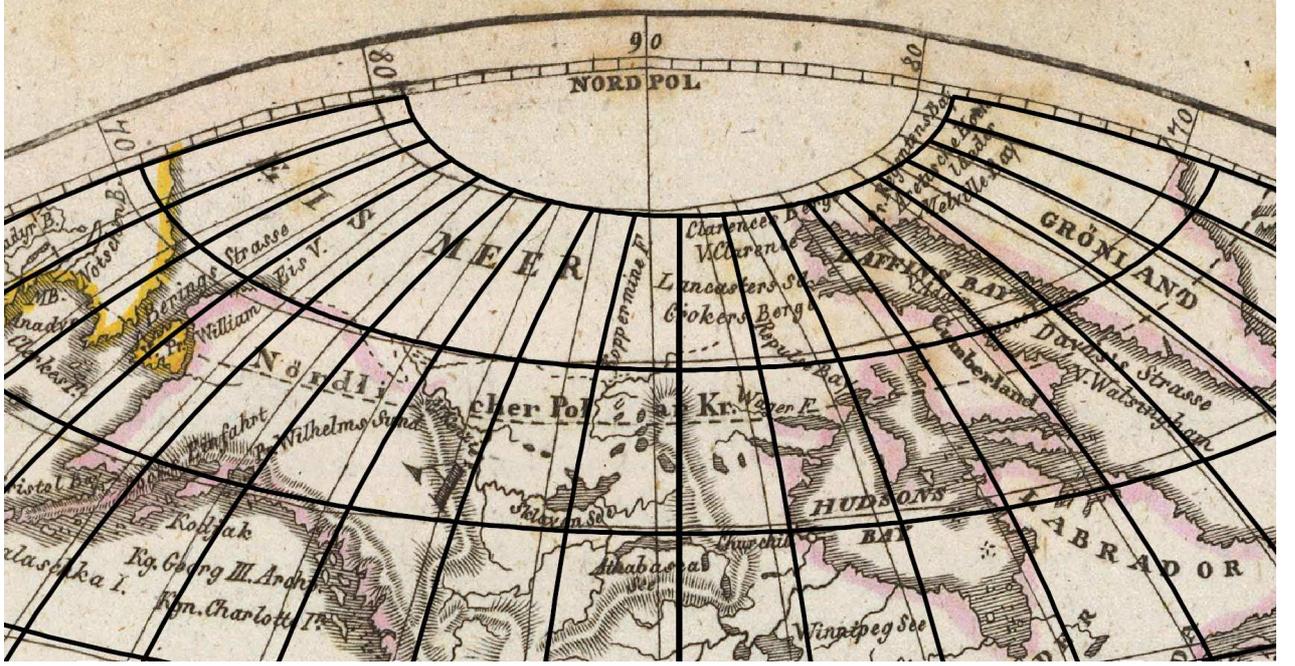


Figure 8: A superimposition of the original and reconstructed graticules; method M5, map rotation $\alpha = 1.7^\circ$ is not involved.

and the partial derivatives of the second coordinate γ' 2D transformation according to the determined parameters are

$$j_{i+n,1}(x) = \left[j_{i,1}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial R'} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] \sin \alpha + q_k = (A_k^T W_k A)^{-1} A_k^T W_k l, \quad q_k = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix},$$

$$\left[j_{i+n,1}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial R} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] \cos \alpha, \text{ where}$$

$$\dots \dots \dots$$

$$j_{i+n,5}(x) = \left[j_{i,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] \sin \alpha + \alpha_k = \arctan \frac{q_2}{q_1}, \quad R'_k = R'_0 \|q_k\|_2, \quad \|q_k\|_2 = \sqrt{q_1^2 + q_2^2},$$

$$\left[j_{i+n,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] \cos \alpha,$$

A_k is the design matrix

$$j_{i+n,6}(x) = \begin{pmatrix} X_i - \frac{1}{n} \sum_{l=1}^n X_l \\ Y_i - \frac{1}{n} \sum_{l=1}^n Y_l \end{pmatrix}_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \begin{matrix} \cos \alpha - \\ \sin \alpha. \end{matrix} \quad A_k = \begin{bmatrix} X'_1 & -Y'_1 \\ \dots & \dots \\ X'_n & -Y'_n \\ Y'_1 & X'_1 \\ \dots & \dots \\ Y'_n & X'_1 \end{bmatrix}, \quad l = \begin{bmatrix} x'_1 \\ \dots \\ x'_n \\ y'_1 \\ \dots \\ y'_n \end{bmatrix}.$$

Analogously, each row of the Jacobian matrix $J(i, :)$ stores the partial derivatives at the point Q_i .

Modified M6 method. To improve the robustness and keep the solution \hat{x}

$$\hat{x}^T = [\varphi_k \quad \lambda_k \quad \varphi_1 \quad \lambda_0 \quad | \quad q_1 \quad q_2],$$

invariant to the map constants R', α , their increments are not determined from NLS, but they are computed directly, from the coefficients q_1, q_2 of the weighted Helmert

The residuals are determined from

$$r(x_k) = A_k q_k - l.$$

The modified M6 method is more robust to the possible errors as well as to the far initial guess x_0 . The Jacobian matrix $J(2n, 4)$ is analogous to the previous case

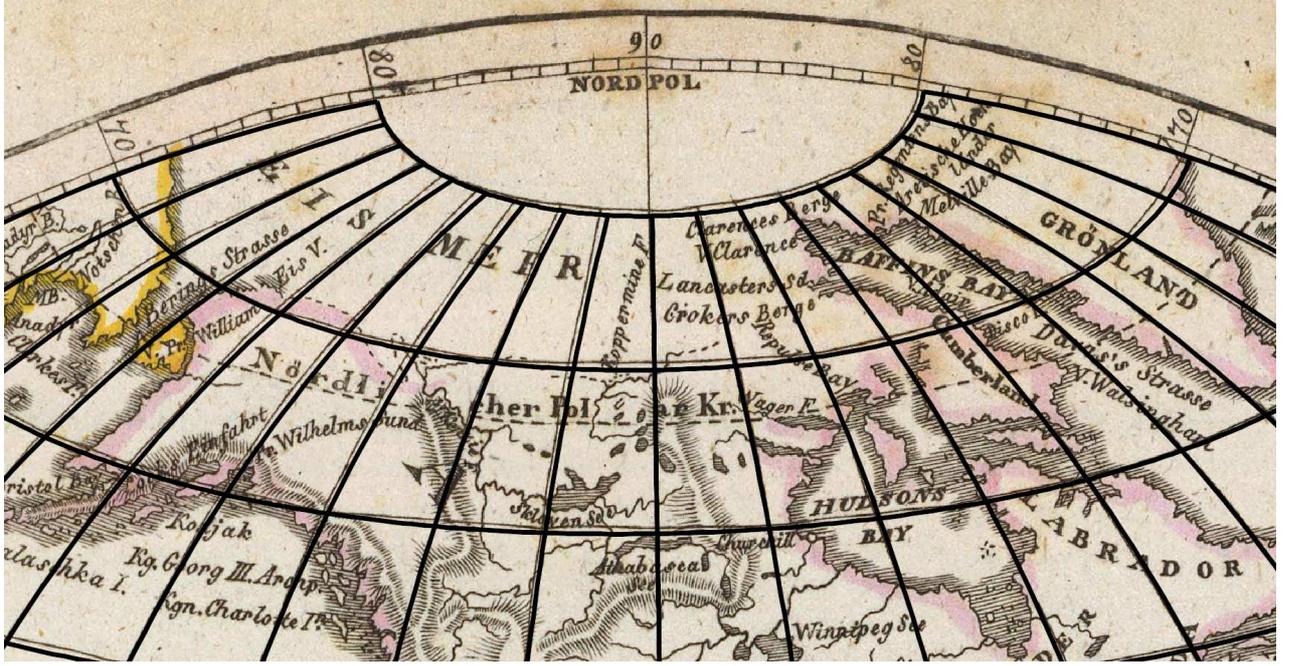


Figure 9: A superimposition of the original and reconstructed graticules; method M6, map rotation $\alpha = 1.7^\circ$ is involved.

but $\cos \alpha$ is replaced with q_1 , and $\sin \alpha$ with q_2

$$\begin{aligned}
 j_{i,1}(x) &= \left[j_{i,2}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \varphi_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] q_1 - \\
 &\quad \left[j_{i+n,2}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \varphi_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] q_2, \\
 \dots &\dots \dots \\
 j_{i,4}(x) &= \left[j_{i,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] q_1 - \\
 &\quad \left[j_{i+n,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] q_2,
 \end{aligned}$$

and

$$\begin{aligned}
 j_{i+n,1}(x) &= \left[j_{i,2}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \varphi_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] q_2 + \\
 &\quad \left[j_{i+n,2}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \varphi_k} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] q_1, \\
 \dots &\dots \dots \\
 j_{i+n,4}(x) &= \left[j_{i,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial X}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] q_2 + \\
 &\quad \left[j_{i+n,5}^{(1)} - \frac{1}{n} \sum_{l=1}^n \left(\frac{\partial Y}{\partial \lambda_0} \right)_{(x=x^{(k)}, \varphi=\varphi_l, \lambda=\lambda_l)} \right] q_1.
 \end{aligned}$$

A situation for the M6 method and a map rotated by $\alpha = 1.7^\circ$ due to inappropriate insertion into the scanner is shown in Fig. 9. Involving a rotation, a significantly better fit between the original and reconstructed graticules is obvious.

5.4 Partial derivatives

Some partial derivatives stored in the Jacobian matrix are projection-specific, others are independent of the projection, or constant.

Derivatives $\frac{\partial X}{\partial \varphi_k}, \frac{\partial X}{\partial \lambda_k}$ and $\frac{\partial Y}{\partial \varphi_k}, \frac{\partial Y}{\partial \lambda_k}$. Both F, G are the composite functions, the chain rule must be applied

$$\frac{\partial X}{\partial \varphi_k} = \frac{\partial X}{\partial \varphi'} \frac{\partial \varphi'}{\partial \varphi_k} + \frac{\partial X}{\partial \lambda'} \frac{\partial \lambda'}{\partial \varphi_k},$$

$$\frac{\partial Y}{\partial \varphi_k} = \frac{\partial Y}{\partial \varphi'} \frac{\partial \varphi'}{\partial \varphi_k} + \frac{\partial Y}{\partial \lambda'} \frac{\partial \lambda'}{\partial \varphi_k},$$

and analogously for $\frac{\partial X}{\partial \lambda_k}, \frac{\partial Y}{\partial \lambda_k}$, etc. Let us put

$$\begin{aligned}
 A &= \cos \varphi \cos(\lambda - \lambda_k), \\
 B &= \cos \varphi \sin(\lambda - \lambda_k), \\
 C &= \sin \varphi_k \sin \varphi + A \cos \varphi_k, \\
 D &= \sin \varphi \cos \varphi_k \cos(\lambda - \lambda_k) - \cos \varphi \sin \varphi_k, \\
 E &= \cos \varphi_k \sin \varphi - A \sin \varphi_k.
 \end{aligned}$$

For the oblique aspect of the projection, the derivatives are

$$\frac{\partial \varphi'}{\partial \varphi_k} = \frac{E}{\sqrt{1-C^2}}, \quad \frac{\partial \lambda'}{\partial \lambda_k} = \frac{B \cos \varphi_k}{\sqrt{1-C^2}},$$

where $1 - C^2 > 0$. Analogously,

$$\frac{\partial \lambda'}{\partial \varphi_k} = -\frac{BC}{B^2 + E^2}, \quad \frac{\partial \lambda'}{\partial \lambda_k} = \frac{D \cos \varphi_k}{B^2 + E^2},$$

where $B^2 + E^2 \neq 0$. The transformed latitude and longitude φ', λ' are frequently arguments of other functions. Let us briefly state the most common cases, when φ', λ' are functions of $\sin(\cdot)$ and $\cos(\cdot)$. Then

$$\begin{aligned}\frac{\partial \sin \varphi'}{\partial \varphi_k} &= E, & \frac{\partial \sin \varphi'}{\partial \lambda_k} &= B \cos \varphi_k, \\ \frac{\partial \cos \varphi'}{\partial \varphi_k} &= -\frac{CE}{\sqrt{1-C^2}}, & \frac{\partial \cos \varphi'}{\partial \lambda_k} &= -\frac{BC \cos \varphi_k}{\sqrt{1-C^2}},\end{aligned}$$

where $1 - C^2 > 0$ and analogously for other composite functions of φ', λ' . For the transverse aspect, the φ_k coordinate is fixed, $\varphi_k = 0^\circ$. Then, derivatives $\partial \cdot / \partial \varphi_k$ are zero

$$\frac{\partial \varphi'}{\partial \varphi_k} = \frac{\partial \lambda'}{\partial \varphi_k} = 0,$$

the both $\partial \cdot / \partial \lambda_k$ derivatives are equal to the oblique aspect. In the normal aspect, $\varphi_k = 90^\circ$, $\lambda_k = 0^\circ$, all derivatives $\partial \cdot / \partial \varphi_k$ and $\partial \cdot / \partial \lambda_k$ are zero

$$\frac{\partial \varphi'}{\partial \varphi_k} = \frac{\partial \lambda'}{\partial \varphi_k} = 0, \quad \frac{\partial \varphi'}{\partial \lambda_k} = \frac{\partial \lambda'}{\partial \lambda_k} = 0.$$

Derivatives $\frac{\partial X}{\partial \varphi_1}, \frac{\partial Y}{\partial \varphi_1}$. The derivatives are only the analyzed projection \mathbb{P} function. They appear in some projection equations (conic, cylindrical, pseudoconic), otherwise, they are zero.

Derivatives $\frac{\partial X}{\partial \Delta X}, \frac{\partial X}{\partial \Delta Y}$ and $\frac{\partial Y}{\partial \Delta X}, \frac{\partial Y}{\partial \Delta Y}$. They do not depend on the analyzed projection \mathbb{P}

$$\frac{\partial X}{\partial \Delta X} = \frac{\partial Y}{\partial \Delta Y} = 1, \quad \frac{\partial X}{\partial \Delta Y} = \frac{\partial Y}{\partial \Delta X} = 0,$$

they are constant.

Derivatives $\frac{\partial X}{\partial \lambda_0}, \frac{\partial Y}{\partial \lambda_0}$. With respect to Eq. 1, both derivatives depend on the projection equations. Moreover, λ may be a composite function. The λ_0 shift appears, when the analyzed projection uses a prime meridian different from Greenwich (e.g., Ferro), or, if the map has a shifted central meridian passing a center of the mapped territory (it minimizes distortions on the boundaries of the map). Otherwise, $\lambda_0 = 0$.

For both transverse and oblique aspects, the situation is slightly complicated. Replacing λ with $\lambda = \lambda - \lambda_0$ in Eqs. 3, 4, it is obvious that changing the position λ_k of the meta-pole may be compensated by the same shift λ_0 in the opposite direction. There are infinitely many pairs λ_0, λ_k holding the following condition

$$\lambda_k + \lambda_0 = \text{const},$$

and λ_k, λ_0 are dependent. Therefore, for the transverse and oblique aspects it is not necessary to determine λ_0 . We set $\lambda_0 = 0$, and

$$\frac{\partial X}{\partial \lambda_0} = \frac{\partial Y}{\partial \lambda_0} = 0.$$

Alternatively, some numerical methods of differentiation can be used; for example, the Stirling method, which is efficient.

5.4.1 Mercator projection, example

The above-mentioned principles will be illustrated on the Mercator projection

$$X = R\lambda \cos \varphi_1, \quad Y = R \ln \left[\tan \left(\frac{\varphi}{2} + \frac{\pi}{4} \right) \right],$$

proposed in the oblique aspect, so $\varphi \equiv \varphi'$, and $\lambda \equiv \lambda'$. Only the crucial steps of the derivation of the Jacobian matrix elements for the M7 method, are shown. The following partial derivatives are easy to compute

$$\begin{aligned}j_{i,1} &= \frac{\partial X}{\partial R'} = \lambda' \cos \varphi_1, \\ j_{i,4} &= \frac{\partial X}{\partial \varphi_1} = -R' \lambda' \sin \varphi_1, \\ j_{i,5} &= \frac{\partial X}{\partial \lambda_0} = 0, \\ j_{i+n,1} &= \frac{\partial Y}{\partial R'} = \ln \left[\tan \left(\frac{\varphi'}{2} + \frac{\pi}{4} \right) \right], \\ j_{i+n,4} &= \frac{\partial Y}{\partial \varphi_1} = 0, \\ j_{i+n,5} &= \frac{\partial Y}{\partial \lambda_0} = 0,\end{aligned}$$

and $R' \equiv R$. For the remaining derivatives, the chain rule must be applied. First, the following derivatives need to be evaluated

$$\begin{aligned}\frac{\partial X}{\partial \varphi'} &= 0, & \frac{\partial X}{\partial \lambda'} &= R' \cos \varphi_1, \\ \frac{\partial Y}{\partial \varphi'} &= \frac{R'}{\cos \varphi'}, & \frac{\partial Y}{\partial \lambda'} &= 0.\end{aligned}$$

Taking into account the previous derived formulas for $\partial \varphi' / \partial \varphi_k, \partial \varphi' / \partial \lambda_k, \partial \lambda' / \partial \varphi_k, \partial \lambda' / \partial \lambda_k$, the partial derivatives of the coordinate function X are

$$j_{i,2} = \frac{\partial X}{\partial \varphi_k} = \frac{\partial X}{\partial \varphi'} \frac{\partial \varphi'}{\partial \varphi_k} + \frac{\partial X}{\partial \lambda'} \frac{\partial \lambda'}{\partial \varphi_k} = R' \frac{BC \cos \varphi_1}{B^2 + E^2},$$

and

$$j_{i,3} = \frac{\partial X}{\partial \lambda_k} = \frac{\partial X}{\partial \varphi'} \frac{\partial \varphi'}{\partial \lambda_k} + \frac{\partial X}{\partial \lambda'} \frac{\partial \lambda'}{\partial \lambda_k} = R' \frac{D \cos \varphi' \cos \varphi_1}{B^2 + E^2}.$$

Analogously, for the second coordinate function Y ,

$$\begin{aligned}j_{i+n,2} &= \frac{\partial Y}{\partial \varphi_k} = \frac{\partial Y}{\partial \varphi'} \frac{\partial \varphi'}{\partial \varphi_k} + \frac{\partial Y}{\partial \lambda'} \frac{\partial \lambda'}{\partial \varphi_k}, \\ &= \frac{R'}{\cos \varphi'} \frac{E}{\sqrt{1-C^2}}, \\ j_{i+n,3} &= \frac{\partial Y}{\partial \lambda_k} = \frac{\partial Y}{\partial \varphi'} \frac{\partial \varphi'}{\partial \lambda_k} + \frac{\partial Y}{\partial \lambda'} \frac{\partial \lambda'}{\partial \lambda_k}, \\ &= \frac{R'}{\cos \varphi'} \frac{B \cos \varphi_k}{\sqrt{1-C^2}}.\end{aligned}$$

5.5 Initial cartographic parameters

The iteration process reliability depends on accurate initialization, relatively close to the stationary point. Let us denote x_0 as the initial guess

$$x_0^T = [R' \quad \varphi_k \quad \lambda_k \quad \varphi_1 \quad \lambda_0 \quad \Delta X \quad \Delta Y \quad \alpha]_0.$$

The correct estimation of the initial value R'_0 of the auxiliary sphere is a step of high importance. A wrong setting may cause a failure of the iterative process, which was confirmed in practice. The initial guess R'_0 is

$$R'_0 = \frac{R}{s}, \quad s = \sqrt{q_1^2 + q_2^2}$$

where s is the scale factor of the Helmert transformation of $T(P, P'_x)$, R the initial Earth radius (currently $R \doteq 6,380,000$ m), and q_1, q_2 are the transformation coefficients. The standard parallel of latitude φ_1 , preserving a nominal scale, is centered along the analyzed region

$$(\varphi_1)_0 = (\underline{\varphi} + \overline{\varphi})/2,$$

where $\underline{\varphi}$ represents the minimum and $\overline{\varphi}$ the maximum values of the geographic extent. Even the approximate determination of the initial values $\Delta X, \Delta Y$ for the oblique aspect represents a serious problem. The first approach sets $\Delta X_0 = 0, \Delta Y_0 = 0$, the second approach uses shifts between centroids c, C of both systems

$$\Delta X_0 = X_c - x_c, \quad \Delta Y_0 = Y_c - y_c.$$

If the actual shifts are too large, the first method may fail, due to the inaccurate initialization. The initial value of the rotation angle may be determined from the Helmert transformation of $T(P, P'_x)$

$$\alpha_0 = \arctan \frac{q_2}{q_1}.$$

The approximate values R'_0, α_0 may differ significantly from their true values; especially if P, P'_x have a circular shape, they are affected by gross errors, the initial guess of the pole position is too far, or a conformal projection is analyzed. Suppose the analyzed region to be located along the equator and the cylindrical projection in the transverse aspect is applied so as the standard parallel is centered above the region. Rotating the tangent cylinder in the east direction by the Earth axis so that equator of the rotated system is aligned with a point $[0, \overline{\lambda}]$, the intersection $K = [0, \lambda_k]$, of the sphere and the cylinder axis representing the pole is $\lambda_k = \overline{\lambda} \pm 90^\circ$. Analogously, rotating the cylinder in the west direction, is $\lambda_k = \underline{\lambda} \pm 90^\circ$. Then,

$$(\lambda_k)_0 = (\overline{\lambda} + \underline{\lambda})/2 + 90^\circ.$$

For conic projections $\overline{\lambda}_k < \overline{\lambda} + 90^\circ$, and $\underline{\lambda}_k > \underline{\lambda} - 90^\circ$; the middle of the λ interval is set as the initial value

$$(\lambda_k)_0 = (\overline{\lambda} + \underline{\lambda})/2.$$

The same approach is used for azimuthal projections, where the extreme values are $\overline{\lambda}_k = \overline{\lambda}$, and $\underline{\lambda}_k = \underline{\lambda}$. The the second pole coordinate is

$$(\varphi_k)_0 = (\overline{\varphi} + \underline{\varphi})/2.$$

The central meridian shift is set to $\lambda_0 = 0$.

For the transverse aspect, an analogous approach is used. The first coordinate of the meta pole is fixed, $\varphi_k = 0^\circ$, while λ_k uses the same rules as the oblique aspect. In the normal aspect of the projection, the pole coordinates are fixed, so $\varphi_k = 90^\circ, \lambda_k = 0^\circ$. The central meridian of longitude λ_0 is set so that it passes approximately through the center of an analyzed region

$$(\lambda_0)_0 = \frac{1}{n} \sum_{i=1}^n \lambda_i.$$

5.6 Hybrid BFGS method

BFGS is one of the most popular and versatile quasi-Newton methods. For small residual problems, it has a slower convergence rate. A possible way to overcome this issue is the ability to switch between first order and second order methods; this approach was studied in [1], [17], [25], [26], [53], [54], [81]. The advantage of all hybrid methods, solving the least squares problem, is the scaling unnecessary. A combination of the Gauss-Newton and BFGS methods, is efficient; for non-zero residual problems, it converges superlinearly, and quadratically for zero residual problems. The initial approximation of the Hessian matrix

$$B_1 = J^T(x_0)W_0J(x_0),$$

is based on the Gauss-Newton method. In next iterations, the following test, proposed in [26],

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

indicates which actualization step will taken. If $\tau > \tau_{min}$, the nearly zero residual problem is solved, in which the Gauss-Newton method

$$B_{k+1} = J^T(x_k)W_kJ(x_k),$$

is more efficient. If $\tau < \tau_{min}$, and $y_k^T s_k > 0$, the large residual problem is solved. Here the BFGS update

$$B_{k+1} = B_k + \frac{y_k^T y_k}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k^T}{s_k^T B_k s_k},$$

where

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

works better. For practical computations, the recommended value in [54] is $\tau_{min} = 0.0001$. The new solution is

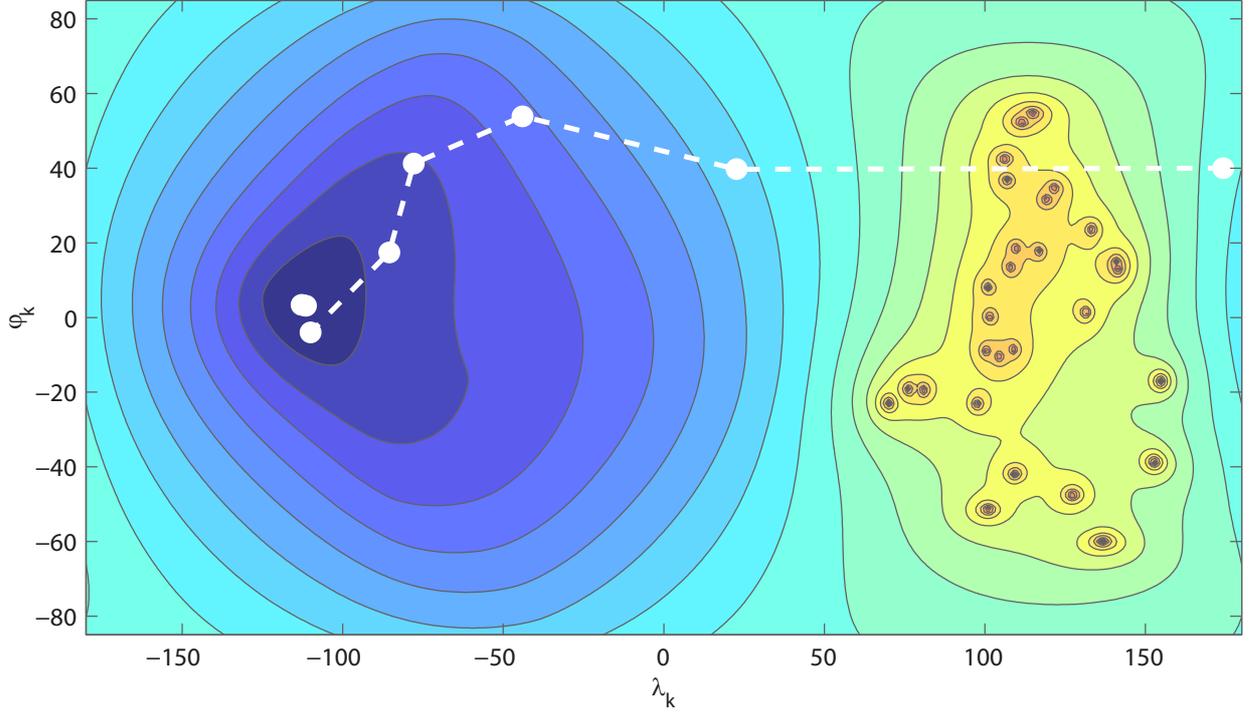


Figure 10: A convergence of the vector x to the global minimum $\varphi_k = 2.2^\circ N$, $\lambda_k = 109.1^\circ W$ (Map 4) of $\phi(x)$ using hybrid BFGS; iterations 1-8, together with contour lines, are shown.

$$x_{k+1} = x_k + \alpha_k h_k, \quad h_k = -B_{k+1}^{-1} \nabla \phi(x_k),$$

the length step α_k satisfies the sufficient decrease condition

$$\phi(x_k + \alpha h_k) \leq \phi(x_k) + \alpha c_1 \nabla f_k^T(x) h_k, \quad (11)$$

where $\alpha_k \in [0, 1]$, and $c_1 \in [0, 1]$. Because of the faster convergence (see Fig. 10), hybrid BFGS seems to be efficient for this problem.

Terminal condition. The following stopping conditions indicating a local minimum proximity, widely described in [81], have been involved. Their numerical values have been adapted to our problem:

- $\|\nabla \phi(x)\|_2 < 10^{-8}$,
- $|\phi(x_k) - \phi(x_{k+1})| < 10^{-10} \cdot \max(1, \phi(x_{k+1}))$,
- $\phi(x_k) < 10^{-8}$,
- iterations < 200 .

If there is no convergence after 200 iterations, the computation is terminated.

The detection algorithm. The algorithm for the map projection analysis based on the hybrid BFGS may be expressed as follows:

1. Initialize

Initialize x_k , $k = 0$, compute $J(x_k)$. Apply the projection formulas: $\mathbb{P}_{x_k} : Q \rightarrow P'_x$. Detect the outliers, and form the weight matrix W_k of all elements. Compute the residuals $r(x_k)$, gradient $\nabla \phi(x_k)$, the objective function $\phi(x_k)$, and the approximation of the Hessian matrix B_{k+1}

$$\begin{aligned} \nabla \phi(x_k) &= J(x_k) W_k r(x_k), \\ \phi(x_k) &= 0.5 r^T(x_k) W_k r(x_k), \\ B_{k+1} &= J^T(x_k) W_k J(x_k). \end{aligned}$$

2. Main loop

Until the terminal condition is true, do the following steps:

(a) Solve normal equations

Solve normal equations, compute the trial point x_t

$$h_k = -B_{k+1}^{-1} \nabla \phi(x_k), \quad x_t = x_k + h_k.$$

(b) Determine the length step α_k

For given $\alpha = 1$, and the trial point x_t , repeat

$$\alpha = \alpha/2, \quad x_t = x_k + \alpha h_k,$$

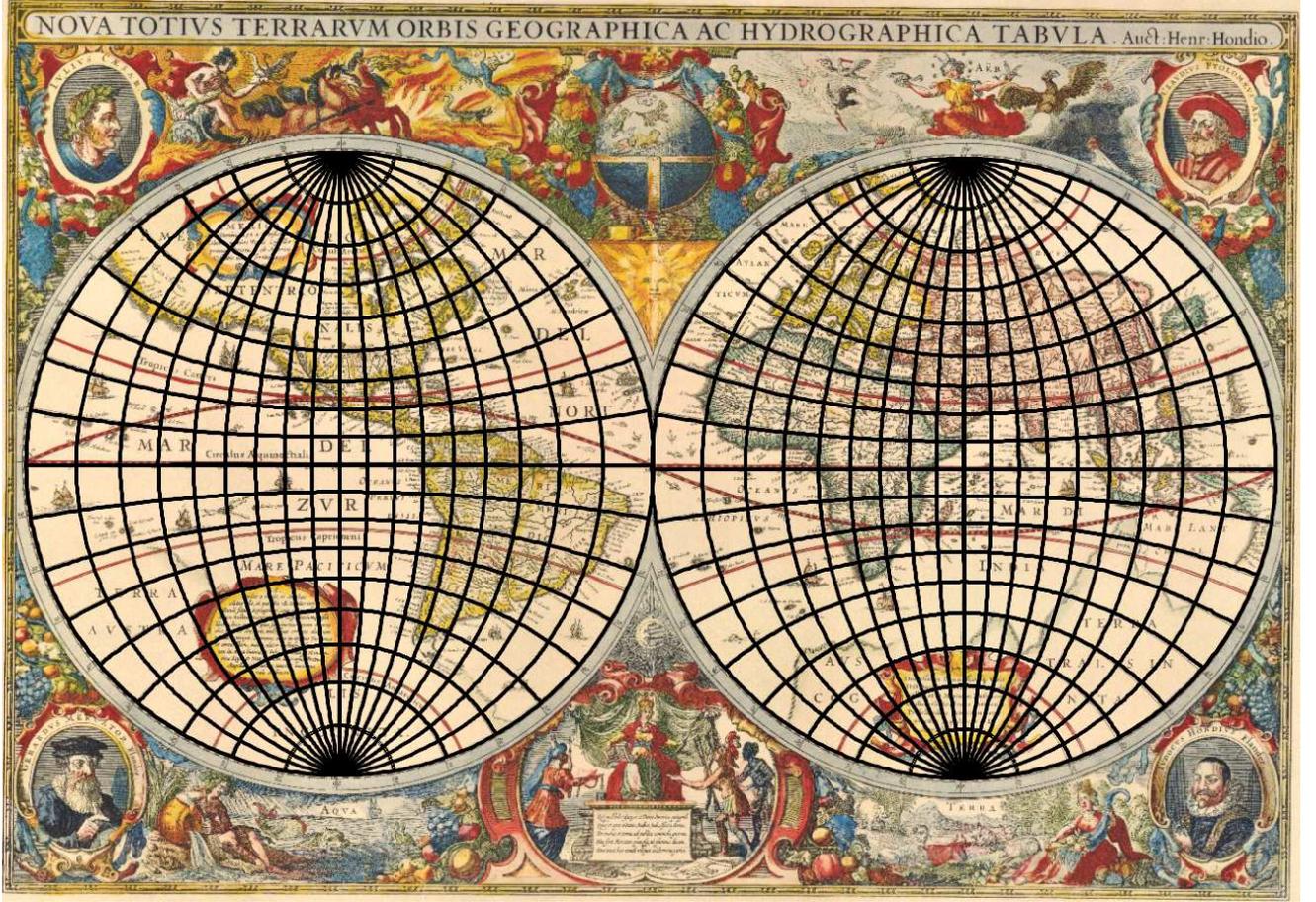


Figure 11: Maps 3,4: A superimposition of the graticule reconstructed from the estimated parameters and the early map, stereographic projection in the transverse aspect, Western and Eastern Hemispheres.

and determine α as the largest number holding the sufficient decrease condition given by Eq. 11. Compute the new solution

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

(c) *Outlier detection*

Apply the projection formulas: $\mathbb{P}_{x_{k+1}} : Q \rightarrow P'_x$. Detect and remove the outliers between P, P'_x using IRLS, form the weight matrix W_{k+1} of all elements.

(d) *Compute new residuals*

Determine new Jacobian matrix $J(x_{k+1})$, the residuals $r(x_{k+1})$, the gradient $\nabla\phi(x_{k+1})$, and the objective function $\phi(x_{k+1})$

$$\begin{aligned} \nabla\phi(x_{k+1}) &= J(x_{k+1})W_{k+1}r(x_{k+1}), \\ \phi(x_{k+1}) &= 0.5r^T(x_{k+1})W_{k+1}r(x_{k+1}). \end{aligned}$$

Check the terminal conditions for $\nabla\phi(x_{k+1})$ and $\nabla\phi(x_k)$ and stop the iteration process, if necessary.

(e) *Selection criterion*

Evaluate the selection criterion τ for $\phi(x_{k+1})$ and $\phi(x_k)$. If $\tau > \tau_{min}$, or $y_k^T s_k < 0$, actualize B_{k+1} from the Gauss-Newton method. Otherwise, use the BFGS update for B_{k+1} .

(f) *Assign values*

Assign old values

$$B_k = B_{k+1}, \nabla\phi(x_k) = \nabla\phi(x_{k+1}), \phi(x_k) = \phi(x_{k+1}).$$

The algorithm reliability can also be controlled by setting the β value. Decreasing $c_1 = 0.0001$, the efficiency improved to 98%; see Sec. 7.

6 Detection based on the differential evolution

The differential evolution, probably the most popular genetic algorithm, is successfully used for solving large or NP-hard optimizing problem, where deterministic methods mostly fail, and the gradient $\nabla\phi$ is not available for

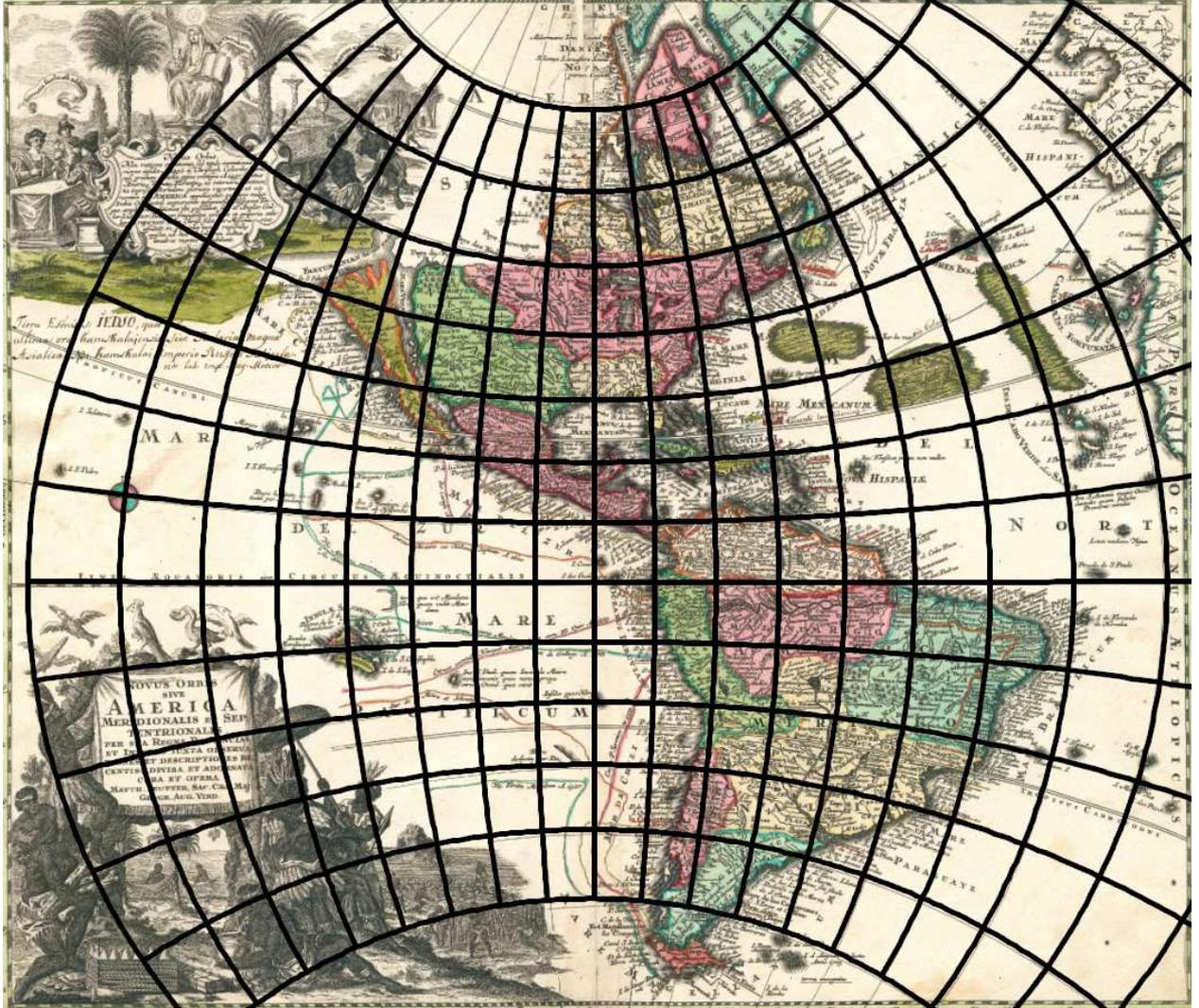


Figure 12: Map 5: A superimposition of the graticule reconstructed from the estimated parameters and the early map, stereographic projection in the transverse aspect.

computations. Compared to NLS, they overcome a dependency on the initial solution x_0 . Instead of a single starting point, an initial population of individuals covers the search space (Fig. 14).

6.1 Methods M5-M7

For the location similarity approach, the model function $g(x)$, residuals $r(x)$, and the objective function $\phi(x)$ are analogous to the NLS approach; the differential evolution may be used to verify the quality of the NLS solution. Due to the lack of partial derivatives, the Jacobian matrix $J(x)$ cannot be evaluated. Considering the shape similarity approach where the global minimizer \hat{x}

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

is required, and $\phi(x)$ is the non-continuous, non-convex, or complex function (the shape-matching invariant), DE represents the natural solution. Frequently, L2 norm of the solution is minimized, see [5]. For the M5-M7 methods, \hat{x} has the analogous form, but $g(x)$, and $r(x)$ are not used and replaced with the shape-matching criteria. Due to the translation invariance of the shape-matching functions, the M5, M6 methods are more appropriate, while the M7 method depends on the shifts $\Delta X, \Delta Y$.

6.2 Shape similarity, objective function

The shape similarity approach, taking into account the change of the shape between two corresponding features, expressed with the dissimilarity δ , leads to the global optimization. There is a wide set of available distance functions, starting from the simple geometric invariants, and

proceeding to the complicated shape matching functions. Some of them are well-known from computer graphics or machinery perception. Let us give the formal definition of the objective function $\phi(P, P'_x)$ as the mean of the objective functions under all dimensions

$$\phi(P, P'_x) = \frac{1}{3} (\phi(P, P'_x)_{0D} + \phi(P, P'_x)_{1D} + \phi(P, P'_x)_{2D}).$$

Objective function for 0D features. The shape similarity cannot be directly measured for 0D entities [5]. Suppose, an analysis of corresponding faces $F(P_i)$, $F(P'_i)$ formed by merged Voronoi cells $\mathcal{V}(P_i)$, $\mathcal{V}(P'_i)$, generated by sets P, P'_x . The objective function $\phi(P, P'_x)$ expresses the average dissimilarity measured by the distance function $d_2(P, P'_x)$ over all corresponding faces

$$\phi(P, P'_x)_{0D} = \frac{1}{n} \sum_{i=1}^n d_2(F(P_i), F(P'_i)).$$

The faces are constructed from k adjacent and subsequently merged Voronoi cells $\mathcal{V}(P_i)$, $\mathcal{V}(P'_i)$ as follows

$$F(P_i) = \bigcup_{j=1}^k [[\mathcal{V}(P_i) \vee \mathcal{V}(P_j)] \wedge [\mathcal{V}(P_i) \wedge \mathcal{V}(P_j) \neq 0]],$$

$$F(P'_i) = \bigcup_{j=1}^k [[\mathcal{V}(P'_i) \vee \mathcal{V}(P'_j)] \wedge [\mathcal{V}(P'_i) \wedge \mathcal{V}(P'_j) \neq 0]].$$

Only cells, whose generators P_i, P'_i satisfy the condition $P_i \notin \partial\mathcal{H}(P) \wedge P'_i \notin \partial\mathcal{H}(P')$, e.g., they are not on the convex hulls $\mathcal{H}(P), \mathcal{H}(P')$, can be merged. The algorithm for the $\bar{d}_2(F(P_i), F(P'_i))$ computation based on the turning function can be found in [2], [5]. All faces are to be normalized by their perimeter.

Objective function for 1D and 2D features. 1D and 2D elements, represented by the line and polygonal features, or the graticule (meridians, parallels), are important source matter for the analysis. Involving these features significantly increases the reliability; they allow the analysis of extensive parts of maps in a single step. The objective function represents the average dissimilarity measured over all n corresponding features

$$\phi(P, P'_x)_{1D/2D} = \frac{1}{n} \sum_{i=1}^n d_2(P_i, P'_i).$$

2D elements are used for larger parts of the surface (woods, lakes), their position and shape may significantly change over time, their geometric accuracy and reliability is lower. Because the acquisition of corresponding 1D and 2D features as well their analysis represent time-consuming processes, such a complex data structure is available only in rare cases; the vast majority of analyses are based on point sets. Further information can be found in [5].

6.3 Differential evolution

Differential evolution, the well-known and efficient optimizing strategy [72], was proposed in [62] as the replacement of the simulated annealing. It has small computational demands (the lower complexity and good memory utilization), a fast convergence, and may be adapted to a wider set of optimizing problems, regardless of the dimension of the problem and the type of the objective function. However, it has a significant disadvantage; its performance depends on the numerical values of three main parameters: the population size NP , the mutation factor F , and the crossover constant CR . The shape similarity approach leads to the multimodal objective function ϕ .

Initial population. The population P_0 is distributed randomly over the search space S^D ; its size $NP = 10 \cdot D$ depends on the detection method, where $D = 5, 6, 7$ is a dimension of the problem. The initial i -th random vector $x_{i,g}$, $i = 1, 2, \dots, 10 \cdot D$, with $j = 1, 2, \dots, D$ components $x_{i,g}^j$ at the generation $g = 0$ is given by

$$x_{i,g}^j = \underline{x}^j + \text{rand}(0, 1) (\bar{x}^j - \underline{x}^j), \quad (12)$$

where $\underline{x}^j, \bar{x}^j$ are their lower and upper bounds.

Mutation scheme. The differential mutation combines selection and mutation strategies. For this problem, the random strategy, represented by the DE/rand/1 scheme

$$v_{i,g} = x_{r_1,g} + F(x_{r_2,g} - x_{r_3,g}),$$

generating the new vector $v_{i,g}$ at g -th generation from three random vectors $x_{r_1,g}, x_{r_2,g}, x_{r_3,g}$, is efficient. To accelerate the convergence, the adaptive control of the differential evolution parameters was used. The MFDE technique

$$F_{i,g} = s \sqrt{d_{i,g} \cdot \text{rand}(0, 1)^2} - b. \quad (13)$$

of controlling the mutation factor F , based on the modified dithering, was introduced in [12]. For each generation g , the linear decreasing factor $d_{i,g}$ is computed

$$d_{i+1,g} = d_{i,g} - 1/n_p,$$

where $d_{1,g} = 1.2$, and $d_{n_p,g} = 0.2$. In initial steps, the algorithm adds a stronger noise, which further reduces. If the fitness between two generations does not improve, the deceleration factor $b = 1.5$ is used; otherwise, the acceleration factor $s = 0.2$, is utilized

$$F_{i,g} = \begin{cases} s \sqrt{d_{i,g} \cdot \text{rand}(0, 1)^2}, & \phi_{\text{best},g-1} - \phi_{\text{best},g} > 0, \\ \sqrt{d_{i,g} \cdot \text{rand}(0, 1)^2} - b, & \text{otherwise.} \end{cases}$$

Algorithm 1 Reflection of the individual $v_{i,g}^j$ into the search space S^D .

```

1: function Reflect( $v_{i,g}, \underline{v}_j, \bar{v}_j$ )
2:   for  $j \leftarrow 0 : m-1$  //Process all parameters of  $v_{i,g}$ 
3:     while  $(v_{i,g}^j < \underline{v}_j) \vee (v_{i,g}^j > \bar{v}_j)$  // Parameter outside interval
4:       if  $v_{i,g}^j < \underline{v}_j$  //Parameter lower than the lower bound
5:          $v_{i,g}^j \leftarrow 2\underline{v}_j - v_{i,g}^j$  //Flip
6:       else if  $v_{i,g}^j > \bar{v}_j$  //Parameter greater than the upper bound
7:          $v_{i,g}^j \leftarrow 2\bar{v}_j - v_{i,g}^j$  //Flip

```

Reflection into the search space. The mutation does not ensure that the new $v_{i,g}^j$ element will be inside the search space S^D . If $v_{i,g}^j \notin S^D$, a simple method that flips all $v_{i,g}^j$ elements to S^D is used. Let us denote $\underline{v}^j, \bar{v}^j$ the lower and upper bounds of j -th determined parameter forming the edge of the search space S^D . According to the reflection used in [71], new $v_{i,g}^j$ elements are written as follows

$$v_{i,g}^j = \begin{cases} 2\underline{v}^j - v_{i,g}^j, & v_{i,g}^j < \underline{v}^j, \\ 2\bar{v}^j - v_{i,g}^j, & v_{i,g}^j > \bar{v}^j. \end{cases} \quad (14)$$

The reflection algorithm is iterative, see Alg. 1; if $v_{i,g}^j$ is too far from a bound, one step may not be sufficient.

Cross-over scheme. Exchanging the part of the genetic information, between the trial and the current individuals keeps the population to a minimum standard. Using the binomial cross-over, a random number $p_c^j = \text{rand}(0, 1)$, expressing the probability of the mutated individual, is generated. The target individual $x_{i,g}$ is crossed with the mutated one $v_{i,g}$

$$u_{i,g}^j = \begin{cases} v_{i,g}^j, & p_c^j \leq CR \vee j = l, \\ x_{i,g}^j, & \text{otherwise,} \end{cases} \quad (15)$$

the cross-over factor is set to $CR = 0.8$.

Selection scheme. The selection process decides, which individual survives, and proceeds to the next generation $g+1$. If the trial vector $u_{i,g}$ has a lower objective value $\phi(u_{i,g})$ than the target vector $\phi(x_{i,g})$, it replaces the target vector; otherwise, the target vector is preserved

$$x_{i,g+1} = \begin{cases} u_{i,g}, & \phi(u_{i,g}) \leq \phi(x_{i,g}), \\ x_{i,g}, & \text{otherwise.} \end{cases} \quad (16)$$

Stopping criterion. The first condition is based on the sufficient decrease condition

$$|\phi_{best,g} - \phi_{best,g+1}| < \varepsilon \max(1, \phi_{best,g}), \quad (17)$$

where $\varepsilon < 1.0 \cdot 10^{-8}$. The second criterion takes into account the convergence of the population to the global

minimum, when no significant improvements in the population are found

$$|\phi_{best,g} - \phi_{worst,g}| < \varepsilon. \quad (18)$$

The third and auxiliary criterion involves an improvement during the last k generations

$$|\phi_{best,g} - \phi_{best,g+\Delta g}| < \varepsilon, \quad (19)$$

where $\Delta g = 50$. It is used if the differences between $\phi_{best,g} - \phi_{best,g+1}$ tend to be small and constant, but the objective function value continuously decreases. In accordance with the previous condition, this behavior may be falsely classified as suitable for stopping the iteration process. The last criterion takes into account the total amount of iterations

$$k < MAXITER. \quad (20)$$

If there is no convergence after 2,000 generations, the computation is terminated. Avoiding the premature convergence for Maps 1, 2 (see Sec. 7), the threshold was decreased to $\varepsilon < 1.0 \cdot 10^{-10}$.

Find the best matching projection. For each cartographic projection \mathbb{P} , the vector \hat{x} of parameters minimizing ϕ is determined. It represents the best individual of a population P_g at the generation g

$$\hat{x} \equiv x_{best,g} = \arg \min_{\mathbb{P}} \phi(x_{i,g}). \quad (21)$$

The detection algorithm. The algorithm for the map projection analysis based on the differential evolution is summarized as follows:

1. *Create the initial population*

Set the initial population size to $NP = 10D$. For $i = 1, \dots, NP$ and for $j = 1, \dots, D$ create the initial random population P_0 . Initialize $W_{i,g} = I$. For each individual $x_{i,g}$, evaluate the residuals $r(x_{i,g})$ and the objective function $\phi(x_{i,g})$.

2. *Main loop*

Until the stopping criteria ε are satisfied, do the following steps:

- (a) *Mutation step*
Initialize $d_{1,g} = 1.2$. For all $i = 1, \dots, 10D$ in the current generation: Compute the decrement factor $d_{i,g}$ and the modified mutation factor $F_{i,g}$. Pick 3 random points and compute the mutation vector $v_{i,g}$, where $F = F_{i,g}$.
- (b) *Cross-over step*
For all $i = 1, \dots, 10D$ and for all $j = 1, \dots, D$ in the current generation: perform a binomial crossover; generate a random number $r^j = \text{rand}(0, 1)$, and compute the item $u_{i,g}^j$ of the crossed vector.
- (c) *Reflection step*
For all $i = 1, \dots, 10D$ in the current generation: flip all elements $u_{i,g}^j$ into the search space S^D , if necessary.
- (d) *Outlier detection*
Apply the projection formulas: $\mathbb{P}_{u_{i,g}} : Q \rightarrow P'$. Detect and remove the outliers between P, P' using IRLS, form the weight matrix $W_{i,g}$ of all elements.
- (e) *Selection step*
To compare the trial vector $u_{i,g}$ and the target vector $x_{i,g}$, evaluate residuals $r(u_{i,g})$ and $r(x_{i,g})$ and their objective function values $\phi(u_{i,g})$, $\phi(x_{i,g})$, the better of them proceeds to the next generation. Update the population P_g .
- (f) *Stopping criteria*
Find the worst and best individuals in the current population

$$\begin{aligned} x_{best,g} &= \arg \min(\phi(x_{i,g})), \\ x_{worst,g} &= \arg \max(\phi(x_{i,g})), \end{aligned}$$

check the terminal conditions, and stop the iteration process, if necessary. Increment the generation $g = g + 1$.

The process of the map projection analysis may be visualized. The early Map 3 was analyzed, every 2nd iteration (Fig 13) and 10th generation (Fig. 14), together with the contour lines, are shown. It is apparent that the population is becoming more concentrated around the global minimum.

7 Experiments and results

All the above-mentioned methods have been extensively tested both on synthetic and real cartographic data. The computations have been performed in C++; our testing PC had the following hardware specification: CPU Intel i5-4460, 8GB RAM, Win8.1 64bit, VS 2013 compiler, single thread mode.

The following characteristics have been measured: The stationary point efficiency (EFF1), the global minimum efficiency (EFF2), the number of iterations (N_IT), the objective function value $\phi(x)$ of the correctly detected samples given by the residuals (RES_C), the objective function value $\phi(x)$ of all tests (RES_A), and the computational time (TIME). The results of DE were considered flawless after being compared to the Mathematica 8 optimizers. The test has been classified as successful if $\phi < 1.1\phi$, where ϕ was determined from a priori known parameters or from the Mathematica 8 optimizer.

7.1 Synthetic tests

The impact of the small territory formed by the spherical quadrangle of the fixed size of $\Delta\varphi = \Delta\lambda = 10^\circ$, with the center $[\varphi_c, \lambda_c]$, continuously shifting over the planisphere, was analyzed. An assumption that there are some areas on the planisphere where most projections have an analogous shape of the graticule, and the detection reliability is lower, or fail, was verified. The comparison was undertaken for Eckert V projection in the normal aspect, method M5, and four distance functions. The location similarity was represented by minimization of the squared sum of residuals (RES), and the cross nearest distance (CND); the shape similarity approach by the turning function over the graticule (GTF), and over the merged Voronoi faces (VFTF). To simulate the early map, each testing set was additionally contaminated with errors. The error ε lies inside the circles of the radius $\langle 50, 100, 150, 200, 250 \rangle$ km. Due to the small amounts of tests, the contour lines look slightly artificial, but the properties and behavior of the distance functions are clearly visible.

The Eckert V projection was reliably detected, except the areas along the equator. With the increasing level of error contamination, the reliability rapidly reduces. Up to the level of 100 km, the results are reliable. Exceeding the value, the detection becomes rather a random process, especially for the shape similarity criteria. The crucial factor is represented by the location of the analyzed territory. Minimum reliability is achieved along the equator, see Tab. 1 and Figs. 15, 16. Comparing the location and shape similarity approach, the shape similarity approach appears to be less efficient and more sensitive to additional noise.

The analogous behavior was recognized for many projections. For most distance functions, a territory located along the central meridian/equator or near the poles represents a serious problem, projections have here an analogous shape of the graticule; the projection footprint is vague. In such territories, at a higher level of noise, the detection process fails or becomes inefficient.

7.2 Early map tests

8 early maps of different scales, ages, and categories, with projections applied in all aspects, were selected. The normal aspect is easily detectable and does not bring serious information about the algorithm reliability (Maps 5, 6, and 8). The NLS, DE, and NM optimizing techniques have been compared, a robustness against the initial guess x_0 and the shifts ΔX , ΔY , was measured. For this purpose, the elements of x_0 have been randomly initialized on the whole domain, and the additional shifts, were added to the data. The test was repeated 300 times. The coordinates $[\varphi, \lambda]$, collected using Google Maps, referring to the WGS-84 ellipsoid, were not transformed to the sphere, so the influence of the ellipsoid was neglected. This issue primarily affects the large-scale maps (Maps 1 and 2), because the positional differences caused by the negligence of the different surfaces are on the level of the graphical accuracy of the map. For Maps 1 and 2, the correct map projection parameters have been a priori known; for early maps 3-8, they were unknown. The results are summarized in Tab. 2.

Map 1. “Czech topographic map”, scale 1:10,000, 16 identical points, analyzed area: $\varphi \in [47^\circ N, 51^\circ N]$, $\lambda \in [12^\circ E, 18^\circ E]$. Projection: Lambert conformal conic, $\lambda_0 = 17^\circ 40'$ (Ferro), $\varphi_k = [59^\circ 43' N, 42^\circ 32' E]$, $\varphi_1 = 78.5^\circ N$, $R = 6,380,703.611$ m, $\Delta X = \Delta Y = 0$. Two types of additional shifts were involved: $\Delta X = 1$ m, $\Delta Y = 2$ m and $\Delta X = 2,000$ m, $\Delta Y = 15,000$ m. The interesting fact that the meta pole position $[\varphi_k, \lambda_k]$ may be compensated by the height of the cone φ_1 and shifts $\Delta X, \Delta Y$ is mentioned. Many local minima and the lower reliability of the EFF2 criterion for the NLS method appeared. The overall best results were achieved by the M5 method. While the DE and NM approaches have always been successful, the least squares solution was slightly less efficient.

Map 2. “Czech topographic map”, scale 1:10,000, 16 identical points, analyzed area: $\varphi \in [47^\circ N, 51^\circ N]$, $\lambda \in [12^\circ E, 18^\circ E]$. The coordinates were measured from the digital copy of the map available on-line. Projection: transverse Mercator, $K = [0^\circ N, 105^\circ E]$, zone 33, $\Delta X = 0$, $\Delta Y = 3,500,000$ m. The determined results are: $\varphi_k = 0.4^\circ N$, $\lambda_k = 105.6^\circ E$, $\varphi_1 = 0.0^\circ N$, $\lambda_0 = 0.0^\circ E$. In several cases, the convergence to the local minimum $\varphi_k = 0.4^\circ N$, $\lambda_k = 105.6^\circ E$, $\varphi_1 = 23.2^\circ N$, $\lambda_0 = 0.0^\circ E$ was recognized. The large value of the ΔY shift seems to be quite problematic, especially for the M7 method. However, the global minimum is clearly defined and easy to recognize.

Maps 3 and 4. “Nova Totius Terrarum Orbis Geographica ac Hydrographica Tabula”, Hendrik Hondius, 1630, Atlantis Maioris Appendix, The Map Collection of the Charles University, maps of the Eastern and

Western Hemispheres. No solid geometric and geodesic bases, probably the graphical method of the construction, 25 identical points, no shifts. For both maps the stereographic projection in the oblique aspect near to the transverse aspect has been detected. The determined projection parameters for the Western Hemisphere: $\varphi_k = 2.2^\circ N$, $\lambda_k = 109.1^\circ E$, $\varphi_1 = 0.0^\circ N$, $\lambda_0 = 0.0^\circ W$. The determined parameters for the Eastern Hemisphere: $\varphi_k = 2.9^\circ S$, $\lambda_k = 55.2^\circ W$, $\varphi_1 = 0.0^\circ N$, $\lambda_0 = 0.0^\circ W$. The M5 method brought the overall best results, but the improvement is not as visible as in the previous tests. The estimated parameters were continuously visualized with the step of 2 generations; see Fig. 13. The iteration process for DE and NLS, is shown in Figs. 14, 10, the reconstructed graticule in Fig. 11.

Map 5. “Novus Orbis sive America Meridionalis et Septentrionalis.”, Matthäus Seutter, 1730, Map Collection of the Charles University, 30 identical points, map of the Western Hemisphere. The determined parameters: stereographic projection, $\varphi_k = 2.9^\circ N$, $\lambda_k = 94.4^\circ W$, $\varphi_1 = 0.0^\circ N$, $\lambda_0 = 0.0^\circ W$. The oblique aspect, close to the transverse, was detected. All methods have been successful. For the generated graticule, see Fig. 12; the prime meridian relates to Ferro.

Map 6. “Ethnografická mappa ke Slowanským starozitnostem P. J. Šafářjka”, 1839, Map Collection of the Charles University, 28 identical points, map of Europe. The map has a solid geometric basis. The determined parameters are: conical equal area projection, $\varphi_k = 90.0^\circ N$, $\lambda_k = 0.0^\circ E$, $\varphi_1 = 44.6^\circ N$, $\lambda_0 = 30.3^\circ E$. All the methods achieved almost analogous results. However, in few situations, the M7 method failed.

Map 7. “Africa Concinnata Secundum Observationes Membror. Acad. Regal. Scientiarum et nonnullorum aliorum, et juxta recentissimas annotationes”, Guillaume Delisle, 1675-1726, Map Collection of the Charles University, 28 identical points, continental map. The determined projection parameters are: Bonne projection, $\varphi_k = 90^\circ N$, $\lambda_k = 0^\circ E$, $\varphi_1 = 26.2^\circ N$, $\lambda_0 = 21.9^\circ E$. The M7 method provided worse results; its reliability is about 20% lower.

Map 8. “British Islands”, World Atlas, A. Arrowsmith, 1817, David Rumsey Map Collection, 29 identical points, country map. To illustrate the capability of the algorithm over the smaller territory, the map was involved in testing. It has a solid geometric basis; the determined parameters relate to the oblique aspect: orthographic projection, $\varphi_k = 42.3^\circ N$, $\lambda_k = -2.7^\circ W$, $\varphi_1 = 0.0^\circ N$, $\lambda_0 = 0.0^\circ E$. All the methods achieved analogous results.

There is no clear winner in the presented methods. Method M7 achieved worse results in most tests. The

discrepancies between M5, M6, and M7 are clearly noticeable for the large shifts (Map 2). In other cases, they were not apparent (Map 4). It can be confirmed that methods M5, and M6, are more robust and efficient; they require less iterations, and both residuals (RES_C, RES_A) are smaller. For the M6 method, the slightly less accurate determination of parameters R' , φ_k , λ_k , φ_1 , λ_0 is compensated by the rotation α .

Hybrid BFGS brings consistent results in terms of the reliability, residuals, robustness, convergence, and amount of iterations. The overall success rate for the random initialization is over 95%, but it can be further improved to 98%, if β is decreased (for $\beta = 0.0001$). This behavior is not surprising; BFGS provides the approximation of the second derivatives. If there is no need for on-the-fly analysis, both the DE and NM methods, having 100% and 99% success rates, may be appropriate. In all experiments, DE detects the global minimum correctly. Although DE is more robust than NM, it has a poor computational time.

Our experiments also illustrate a lower reliability of the analysis for the transverse/oblique aspects, which is due to the non-convexity of the problem.

8 Conclusion

This paper brings an overview of 3 new methods M5, M6, M7 for the estimation of an unknown map projection and its parameters. Due to the the problematic determination of the initial values of ΔX , ΔY , especially, if they are large, the methods M5, and M6 are more efficient and reliable than M7. Moreover, M6 supports the additional map rotation. An impact of the analyzed territory on the results is crucial. Small territories, territories near the equator, central meridian, or poles, where the projec-

tion footprint is not recognizable, may bring uncertainty in determining R' , φ_k , λ_k , φ_1 , λ_0 , α , the results become ambiguous.

Several optimizing techniques have been compared. DE proved to be reliable, but completely inappropriate for on-the-fly analysis; the computational time is in the order of tens of minutes. NM brings only slightly worse reliability, but it is faster; hybrid BFGS is about 5 percent less efficient, but provides the real-time analysis. Unfortunately, NLS may stuck in the local minimum, but the difference between the global optimizer is frequently bellow the graphical accuracy of the map.

The primarily importance is referred to the refinement of the spatial georeference. It may be adopted by the librarians to acquire the cartographic meta-data semi-automatically and with a higher degree of relevance which accelerates the work and saves time, as well as by cartographers, geographers, or GIS users.

All algorithms have been implemented in the new version of the `detectproj` software (the source code is available on the `github`), and in `Georeferencer` (online tool for the map analysis).

The author hopes that introduced methods extend the capabilities of the cartometric analysis, and become additional research tools for the study of the national cartographic heritage.

Acknowledgement

This article was supported by a grant from the Ministry of Culture of the Czech Republic, No. DF11P01OVV003 “TEMAP - Technology for Access to Czech Map Collections: Methodology and Software for the Protection and Re-use of the National Cartographic Heritage”.

References

- [1] Al-Baali, M., Fletcher, R.: Variational methods for non-linear least-squares. *Journal of the Operational Research Society* pp. 405–421 (1985) 5.6
- [2] Arkin, E.M., Chew, L.P., Huttenlocher, D.P., Kedem, K., Mitchell, J.S.B.: An efficiently computable metric for comparing polygonal shapes. *IEEE J PAMI* **13**(3), 209–216 (1991) 6.2
- [3] Balletti, C.: Analytical and quantitative methods for the analysis of the geometrical content of historical cartography. *International Archives of Photogrammetry and Remote Sensing* **33**(B5/1; PART 5), 30–37 (2000) 2
- [4] Balletti, C.: Georeference in the analysis of the geometric content of early maps. *e-Perimtron* **1**(1), 32–42 (2006) 2
- [5] Bayer, T.: Estimation of an unknown cartographic projection and its parameters from the map. *GeoInformatica* **18**(3), 621–669 (2014). DOI 10.1007/s10707-013-0200-4. URL <http://dx.doi.org/10.1007/s10707-013-0200-4> 1, 2, 3, 4, 6.1, 6.2, 6.2
- [6] Bayer, T., Potůčková, M., Čábelka, M.: Cartometric analysis of old maps on the example of vogt’s map. In: G. Gartner, F. Ortog (eds.) *Cartography in Central and Eastern Europe, Lecture Notes in Geoinformation and Cartography* (2010). DOI 10.1007/978-3-642-03294-3_33 2

- [7] Björck, Å.: Numerical Methods for Least Squares Problems. SIAM, Philadelphia (1996) 2
- [8] Boutoura, C., Livieratos, E.: Some fundamentals for the study of the geometry of early maps by comparative methods. *e-Perimetron* **1**(1), 60–70 (2006) 2
- [9] Brest, J., Greiner, S., Boskovic, B., Mernik, M., Zumer, V.: Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems. *Trans. Evol. Comp* **10**(6), 646–657 (2006). DOI 10.1109/TEVC.2006.872133 2
- [10] Broyden, C.G.: The convergence of a class of double rank minimization algorithms. pp. 76–231 (1970). URL <http://www.mendeley.com/research/the-convergence-of-a-class-of-doublerank-minimization-algorithms-parts-i-and-ii/> 2
- [11] BROYDEN, C.G., DENNIS, J.E., MORE, J.J.: On the local and superlinear convergence of quasi-newton methods. *IMA Journal of Applied Mathematics* **12**(3), 223–245 (1973). DOI 10.1093/imamat/12.3.223 2
- [12] Chien, C.W., Hsu, Z.R., Lee, W.P., et al.: Improving the performance of differential evolution algorithm with modified mutation factor. In: *Proceedings of International Conference on Machine Learning and Computing (ICMLC 2009)* (2009) 6.3
- [13] Corne, D., Dorigo, M., Glover, F.: *New Ideas in Optimization*. Advanced Topics in Computer Science Series. McGraw-Hill Companies (1999) 2
- [14] Crăciunescu, V., Constantinescu, S.: Eharta (2006). URL <http://earth.unibuc.ro/articole/eHarta?lang=en> 2
- [15] Dai, Y.H.: Convergence properties of the bfgs algorithm. *SIAM J. on Optimization* **13**(3), 693–701 (2002). DOI 10.1137/S1052623401383455. URL <http://dx.doi.org/10.1137/S1052623401383455> 2
- [16] Das, S., Konar, A., Chakraborty, U.K.: Two improved differential evolution schemes for faster global search. In: *Proceedings of the 2005 conference on Genetic and evolutionary computation, GECCO '05*, pp. 991–998. ACM, New York, NY, USA (2005). DOI 10.1145/1068009.1068177. URL <http://doi.acm.org/10.1145/1068009.1068177> 2
- [17] Dennis Jr, J.E., Songbai, S., Vu, P.A.: A memoryless augmented gauss-newton method for nonlinear least-squares problems. Tech. rep., DTIC Document (1985) 5.6
- [18] Erle, S., Krishnan, S., Waters, T.: World map warp (2009). URL <http://warp.worldmap.harvard.edu/> 2
- [19] Esri: Identify an unknown projected coordinate system using arcmap (2003) 2
- [20] Esri: Identify an unknown projected coordinate system using arcmap (2005) 2
- [21] Fan, H.Y., Lampinen, J.: A trigonometric mutation operation to differential evolution. *J. of Global Optimization* **27**(1), 105–129 (2003). DOI 10.1023/A:1024653025686 2
- [22] Feoktistov, V., Janaqi, S.: New strategies in differential evolution. In: I. Parmee (ed.) *Adaptive Computing in Design and Manufacture VI*, pp. 335–346. Springer London (2004) 2
- [23] Flacke, W., Kraus, B., Warcup, C.: *Working with projections and datum transformations in ArcGIS: theory and practical examples*. Points Verlag (2005). URL <http://books.google.cz/books?id=PfEsAQAAMAAJ> 2
- [24] Fletcher, R.: A new approach to variable metric algorithms. *Comput. J.* **13**(3), 317–322 (1970) 2
- [25] Fletcher, R.: *Practical methods of optimization* (2nd ed.). Wiley-Interscience, New York, NY, USA (1987) 5.6
- [26] Fletcher, R., Xu, C.: Hybrid methods for nonlinear least squares. *IMA Journal of Numerical Analysis* **7**(3), 371–389 (1987). DOI 10.1093/imanum/7.3.371. URL <http://imajna.oxfordjournals.org/content/7/3/371.abstract> 5.6
- [27] Fraley, C.: *Solution of nonlinear least-squares problems*. Ph.D. thesis, Stanford, CA, USA (1987). AAI8800933 2

- [28] Fraley, C., Laboratory, S.U.S.O.: Algorithms for nonlinear least-squares problems. Technical report (Stanford University. Systems Optimization Laboratory). Stanford University, Dept. of Operations Research, Systems Optimization Laboratory (1988). URL <http://books.google.cz/books?id=dPgEAAAAIAAJ> 2
- [29] Gill, P.E., Murray, W., Wright, M.H.: Practical optimization. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London (1981) 2
- [30] Gämperle, R., Müller, S.D., Koumoutsakos, P.: A parameter study for differential evolution. In: WSEAS Int. Conf. on Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation, pp. 293–298. Press (2002) 2
- [31] Goldfarb, D.: A Family of Variable Metric Methods Derived by Variational Means. *Maths. Comput.* **24**, 23–26 (1970) 2
- [32] Gong, W., Cai, Z., Jiang, L.: Enhancing the performance of differential evolution using orthogonal design method. *Applied Mathematics and Computation* **206**(1), 56 – 69 (2008). DOI 10.1016/j.amc.2008.08.053 2
- [33] Harley, J.B.: The evaluation of early maps: Towards a methodology. *Imago Mundi* **22**, pp. 62–74 (1968) 2
- [34] Hooke, J., Perry, R.: The planimetric accuracy of tithe maps. *The Cartographic Journal* **13**(2), 177–183 (1976) 2
- [35] Hui-rong, L., Yue-lin, G., Chao, L., Peng-jun, Z.: Improved differential evolution algorithm with adaptive mutation and control parameters. In: Proceedings of the 2011 Seventh International Conference on Computational Intelligence and Security, CIS '11, pp. 81–85. IEEE Computer Society, Washington, DC, USA (2011). DOI 10.1109/CIS.2011.26. URL <http://dx.doi.org/10.1109/CIS.2011.26> 2
- [36] Jenny, B.: Mapanalyst-a digital tool for the analysis of the planimetric accuracy of historical maps. *e-Perimetron* **1**(3), 239–245 (2006) 2
- [37] Jenny, B., Hurni, L.: Cultural heritage: Studying cartographic heritage: Analysis and visualization of geometric distortions. *Comput. Graph.* **35**(2), 402–411 (2011). DOI 10.1016/j.cag.2011.01.005. URL <http://dx.doi.org/10.1016/j.cag.2011.01.005> 2, 4.4
- [38] Jenny, B., Weber, A., Hurni, L.: Visualizing the planimetric accuracy of historical maps with mapanalyst. *Cartographica: The International Journal for Geographic Information and Geovisualization* **42**(1), 89–94 (2007) 2
- [39] Jiang, S., Cai, Z.: Faster convergence and higher hypervolume for multi-objective evolutionary algorithms by orthogonal and uniform design. In: Z. Cai, C. Hu, Z. Kang, Y. Liu (eds.) *Advances in Computation and Intelligence, Lecture Notes in Computer Science*, vol. 6382, pp. 312–328. Springer Berlin Heidelberg (2010) 2
- [40] Kelley, C.T.: Iterative Methods for Linear and Nonlinear Equations. No. 16 in *Frontiers in Applied Mathematics*. SIAM (1995) 2
- [41] Kowal, K.C., Pridal, P.: Online georeferencing for libraries: The british library implementation of georeferencer for spatial metadata enhancement and public engagement. *Journal of Map & Geography Libraries* **8**(3), 276–289 (2012). DOI 10.1080/15420353.2012.700914. URL <http://www.tandfonline.com/doi/abs/10.1080/15420353.2012.700914> 2
- [42] Krejic, N., Luzanin, Z., Stojkowska, I.: Gauss-newton-based bfgs method with filter for unconstrained minimization. *Applied Mathematics and Computation* **211**(2), 354–362 (2009) 2
- [43] Labs, M.: Map rectifier (2009). URL <http://labs.metacarta.com/rectifier/> 2
- [44] Laxton, P.: The geodetic and topographical evaluation of english county maps, 1740-1840. *The Cartographic Journal* **13**(1), 37–54 (1976) 2
- [45] Li, D., Fukushima, M.: A globally and superlinearly convergent gauss–newton-based bfgs method for symmetric nonlinear equations. *SIAM J. Numer. Anal.* **37**(1), 152–172 (1999). DOI 10.1137/S0036142998335704. URL <http://dx.doi.org/10.1137/S0036142998335704> 2

- [46] Li, D.H., Fukushima, M.: A modified BFGS method and its global convergence in nonconvex minimization. *Journal of Computational and Applied Mathematics* **129**, 15–35 (2001) 2
- [47] Lilley, K., Lloyd, C., Campbell, B.: Mapping the realm: A new look at the gough map of britain (c.1360). *Imago Mundi: The International Journal for the History of Cartography* **61**(1), 1–28 (2009). DOI 10.1080/03085690802456228 2
- [48] Liu, J., Lampinen, J.: A fuzzy adaptive differential evolution algorithm. *Soft Computing* **9**(6), 448–462 (2005). DOI 10.1007/s00500-004-0363-x. URL <http://dx.doi.org/10.1007/s00500-004-0363-x> 2
- [49] Livieratos, E.: Graticule versus point positioning in ptolemy cartographies. *e-Perimtron* **1**(1), 51–59 (2006) 2
- [50] Livieratos, E.: Graticule versus point positioning in ptolemy cartographies. *e-Perimtron* **1**(1), 51–59 (2006) 4.2
- [51] Livieratos, E.: On the study of the geometric properties of historical cartographic representations. *Cartographica: The International Journal for Geographic Information and Geovisualization* **41**(2), 165–176 (2006) 2
- [52] Lloyd, C.D., Lilley, K.D.: Cartographic veracity in medieval mapping: analyzing geographical variation in the gough map of great britain. *Annals of the Association of American Geographers* **99**(1), 27–48 (2009) 2
- [53] Lukšan, L.: Computational experience with known variable metric updates. *Journal of Optimization Theory and Applications* **83**(1), 27–47 (1994) 2, 5.6
- [54] Lukšan, L.: Hybrid methods for large sparse nonlinear least squares. *Journal of Optimization Theory and Applications* **89**(3), 575–595 (1996). DOI 10.1007/BF02275350. URL <http://dx.doi.org/10.1007/BF02275350> 2, 5.6
- [55] Mascarenhas, W.F.: The bfgs method with exact line searches fails for non-convex objective functions. *Math. Program.* **99**(1), 49–61 (2004) 2
- [56] Murphy, J.: Measures of map accuracy assessment and some early ulster maps. *Irish Geography* **11**(1), 88–101 (1978) 2
- [57] Nelder, J.A., Mead, R.: A simplex method for function minimization. *The Computer Journal* **7**(4), 308–313 (1965). DOI 10.1093/comjnl/7.4.308. URL <http://comjnl.oxfordjournals.org/content/7/4/308.abstract> 3
- [58] Nemirovski, A.: *Optimization ii: Standard numerical methods for nonlinear continuous optimization* (2013) 2
- [59] Nocedal, J., Wright, S.J.: *Numerical Optimization*, 2nd edn. Springer, New York (2006) 2
- [60] Page, Y.P.: Assigning map projections to portolan maps. *e-Perimtron* **1**(1), 40–50 (2006) 2
- [61] Pedersen, M.E.H.: Good parameters for differential evolution. Tech. rep., Hvaas Laboratories (2010) 2
- [62] Price, K., Storn, R.M., Lampinen, J.A.: *Differential Evolution: A Practical Approach to Global Optimization* (Natural Computing Series). Springer-Verlag New York, Inc., Secaucus, NJ, USA (2005) 2, 6.3
- [63] Price, K.V.: New ideas in optimization. chap. An introduction to differential evolution, pp. 79–108. McGraw-Hill Ltd., UK, Maidenhead, UK, England (1999). URL <http://dl.acm.org/citation.cfm?id=329055.329069> 2
- [64] Příklad, M.P., Žabička, P.: Tiles as an approach to on-line publishing of scanned old maps, vedute and other historical documents. *e-Perimtron* **3**(1), 10–21 (2008) 2
- [65] Pridal, P.: Georeferencer (2011). URL <http://www.georeferencer.org> 2
- [66] Qin, A.K., Huang, V.L., Suganthan, P.N.: Differential evolution algorithm with strategy adaptation for global numerical optimization. *Trans. Evol. Comp* **13**(2), 398–417 (2009). DOI 10.1109/TEVC.2008.927706. URL <http://dx.doi.org/10.1109/TEVC.2008.927706> 2

- [67] Rajaković, M., Kljajić, I., Lapaine, M.: Map projection reconstruction of a map by mercator. In: M. Buchroithner, N. Prechtel, D. Burghardt (eds.) *Cartography from Pole to Pole, Lecture Notes in Geoinformation and Cartography*, pp. 31–44. Springer Berlin Heidelberg (2014) 2
- [68] Ravenhill, W.: Projections for the large general maps of britain, 1583-1700. *Imago Mundi* **33**(1), 21–32 (1981). DOI 10.1080/03085698108592512 2
- [69] Ravenhill, W., Gilg, A.: The accuracy of early maps? towards a computer aided method. *The Cartographic Journal* **11**(1), 48–52 (1974) 2
- [70] Shanno, D.F.: Conditioning of Quasi-Newton Methods for Function Minimization. *Mathematics of Computation* **24**(111), 647–656 (1970). DOI 10.2307/2004840. URL <http://dx.doi.org/10.2307/2004840> 2
- [71] Storn, R., Price, K.: Differential Evolution- A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces. Tech. rep. (1995). URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.1.1.9696> 2, 6.3
- [72] Storn, R., Price, K.: Minimizing the real functions of the icec’96 contest by differential evolution. In: *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*, pp. 842–844. IEEE (1996) 6.3
- [73] Strang, A.: The analysis of ptolemy’s geography. *The Cartographic Journal* **35**(1), 27–47 (1998) 2
- [74] Tobler, W.R.: Computation of the correspondence of geographical patterns. In: *Papers of the Regional Science Association*, vol. 15, pp. 131–139. Springer (1965) 2
- [75] Tobler, W.R.: Medieval distortions: The projections of ancient maps. *Annals of the Association of American Geographers* **56**(2), 351–360 (1966). DOI 10.1111/j.1467-8306.1966.tb00562.x 2
- [76] Tobler, W.R.: Numerical approaches to map projections. *Beitrag zur theoretischen Kartographie, Festschrift für Erik Amberger*, hg. **1**(14), 51–64 (1977) 2
- [77] Tobler, W.R.: Measuring the similarity of map projections. *The American Cartographer* **13**(2), 135–139 (1986) 2
- [78] Tobler, W.R.: Bidimensional regression. *Geographical Analysis* **26**(3), 187–212 (1994) 2
- [79] Yerci, M.: The accuracy of the first world map drawn by piri reis. *The Cartographic Journal* **26**(2), 154–155 (1989) 2
- [80] Zaharie, D.: Control of Population Diversity and Adaptation in Differential Evolution Algorithms. In: R. Matousek, P. Osmera (eds.) *Proc. of Mendel 2003, 9th International Conference on Soft Computing*, pp. 41–46. Brno, Czech Republic (2003) 2
- [81] Zhou, W., Chen, X.: Global convergence of a new hybrid gauss-newton structured bfgs method for nonlinear least squares problems. *SIAM J. on Optimization* **20**(5), 2422–2441 (2010). DOI 10.1137/090748470. URL <http://dx.doi.org/10.1137/090748470> 2, 5.6, 5.6

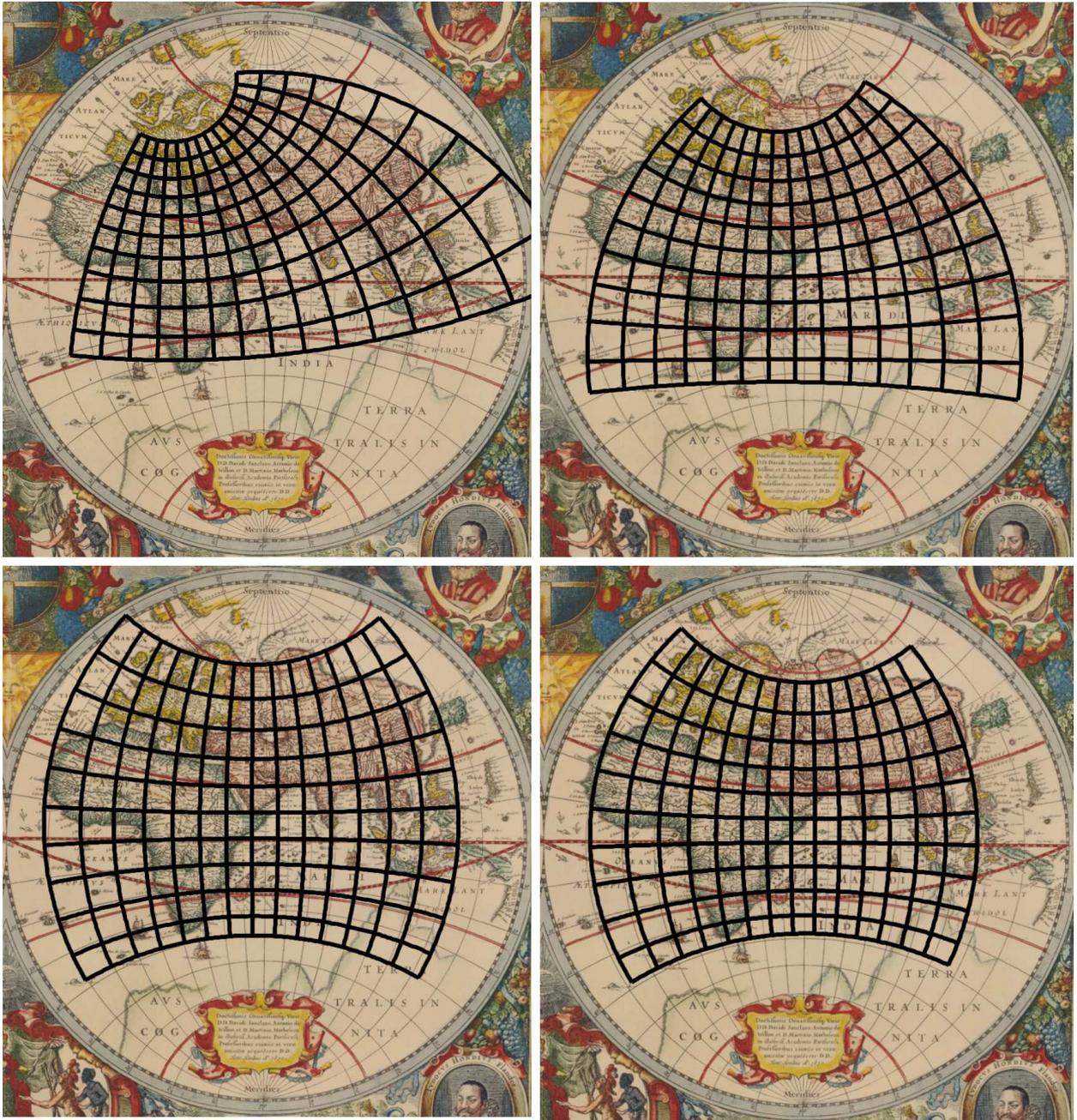


Figure 13: Stereographic projection in the transverse aspect, the estimation of parameters for the Map 3 using DE; generations 2, 4, 6, and 8 are shown.

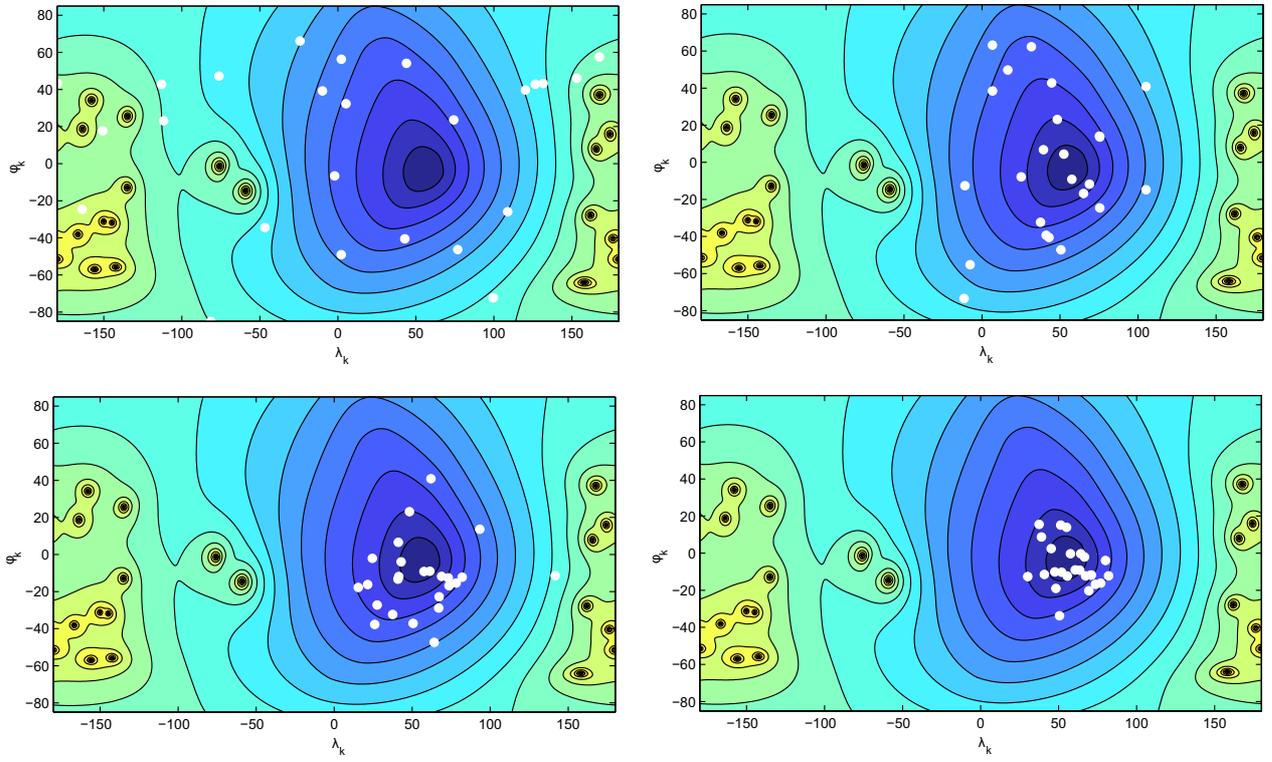


Figure 14: The convergence of the generated population x_g to the global minimum $\varphi_k = 2.9^\circ S$, $\lambda_k = 55.2^\circ E$ (Map 3) of $\phi(x)$, every 10-th generation, together with contour lines, are shown.

Table 1: Spherical territory shifting over the planisphere, Eckert V projection. The analysis reliability measured over the grid, 4 distance functions, method M5, and 5 levels of error contamination, are separated by /.

$d(P, P'_x)$	φ_c/λ_e	Eckert V projection								
		0	20	40	60	80	100	120	140	160
RES (M5)	80	90/82/78/ 73/65	91/84/76/ 75/72	90/82/77/ 73/67	90/80/76/ 68/64	90/81/79/ 73/64	93/86/74/ 72/70	90/86/82/ 74/73	87/82/80/ 77/75	97/94/86/ 75/66
	60	81/63/55/ 47/37	83/62/54/ 49/44	82/69/53/ 45/39	82/65/50/ 41/36	83/66/54/ 36/32	90/77/61/ 56/47	87/78/64/ 60/54	86/75/71/ 64/55	88/76/68/ 68/60
	40	77/50/41/ 34/23	77/62/45/ 40/29	79/53/44/ 37/27	78/60/48/3 7/31	84/65/49/ 37/29	87/71/57/ 33/26	89/79/65/ 47/27	89/81/71/ 62/49	92/86/79/ 67/53
	20	41/22/17/ 16/12	41/29/24/ 16/12	51/37/23/ 20/20	52/37/31/ 25/23	54/41/32/ 30/25	58/44/34/ 27/23	70/56/43/ 29/22	80/61/39/ 34/22	90/73/51/ 32/27
	0	15/18/8/ 12/7	23/23/8/ 15/12	21/13/24/ 15/8	24/19/18/ 17/7	33/23/15/ 8/9	30/17/16/ 19/6	32/23/14/ 11/7	38/25/18/ 12/10	41/20/19/ 14/10
CND (M5)	80	85/73/69/ 67/67	88/77/67/ 75/72	89/78/70/ 72/74	89/80/75/ 74/77	88/82/77/ 74/75	90/80/74/ 75/76	92/82/77/ 76/79	91/79/74/ 77/76	93/81/77/ 74/71
	60	73/43/42/ 47/46	86/40/35/ 52/48	83/62/42/ 49/55	83/52/32/ 54/53	82/55/49/ 48/52	83/54/37/ 54/54	88/60/48/ 50/47	86/58/55/ 58/49	89/67/55/ 60/61
	40	73/39/25/ 26/33	80/48/18/ 35/33	80/48/23/ 34/35	74/53/18/ 32/31	79/49/28/ 36/39	84/52/31/ 38/40	84/57/31/ 54/47	84/62/46/ 54/57	86/68/46/ 51/55
	20	47/24/18/ 23/19	59/34/11/ 22/30	60/43/17/ 26/26	63/43/16/ 27/28	68/44/18/ 34/29	70/51/19/ 36/30	74/54/27/ 38/37	77/64/28/ 41/35	77/65/34/ 45/47
	0	35/31/14/ 17/22	39/38/12/ 18/24	36/24/19/ 24/25	41/37/22/ 16/27	45/34/21/ 13/27	42/29/13/ 21/27	47/27/16/ 18/33	54/35/19/ 21/30	54/28/18/ 21/27
GTF (M5)	80	56/15/3/ 3/0	89/29/13/ 5/5	98/47/18/ 8/4	95/49/21/ 2/4	96/44/30/ 11/8	96/49/30/ 14/2	96/53/27/ 13/3	95/54/29/ 13/8	93/51/23/ 10/5
	60	33/18/4/ 3/2	89/45/16/ 5/1	95/58/21/ 10/6	98/67/25/ 14/5	98/70/26/ 13/5	97/63/37/ 14/6	96/69/33/ 11/9	97/67/34/ 14/8	99/67/39/ 15/14
	40	33/17/9/ 7/2	85/45/18/ 5/2	94/62/22/ 11/3	94/74/27/ 14/5	94/75/33/ 11/7	96/68/34/ 16/8	96/72/46/ 17/12	96/72/32/ 22/9	98/77/43/ 32/11
	20	29/21/14/ 5/1	73/42/17/ 6/2	91/55/28/ 10/5	95/70/25/ 8/5	96/74/31/ 13/8	99/75/40/ 18/4	95/76/39/ 19/7	100/81/31/ 17/11	98/78/49/ 23/8
	0	36/29/11/ 4/5	55/37/18/ 7/6	64/53/16/ 10/5	68/55/23/ 20/5	73/48/27/ 2/4	78/47/22/ 13/5	76/45/12/ 4/2	75/43/14/ 4/5	77/37/15/ 3/2
VFTF (M5)	80	74/40/13/ 7/5	91/40/20/ 6/8	83/39/18/ 5/4	67/24/9/ 2/3	90/40/24/ 9/7	88/40/29/ 13/2	85/39/22/ 9/2	88/48/28/ 14/8	90/43/20/ 8/5
	60	58/66/34/ 13/9	88/60/31/ 11/4	92/57/26/ 17/9	95/62/25/ 13/6	92/53/18/ 11/6	87/47/28/ 14/6	93/55/26/ 8/10	89/41/25/ 10/7	93/42/27/ 10/11
	40	48/51/28/ 16/3	79/33/22/ 7/1	95/62/30/ 16/3	95/60/27/ 15/7	96/71/32/ 13/8	97/61/28/ 13/10	96/63/41/ 14/14	93/55/20/ 15/8	96/68/37/ 25/11
	20	40/26/24/ 12/3	52/23/12/ 5/3	85/34/21/1 1/7	95/57/24/ 9/6	97/69/31/ 13/10	99/71/42/ 21/6	96/78/43/ 22/9	100/82/32/ 18/14	98/74/51/ 27/10
	0	35/16/11/ 11/7	42/9/11/ 3/2	66/27/8/ 7/3	66/41/17/ 16/6	78/48/25/ 2/4	80/48/20/ 13/7	80/50/13/ 5/2	84/48/15/ 5/5	82/41/16/ 3/2

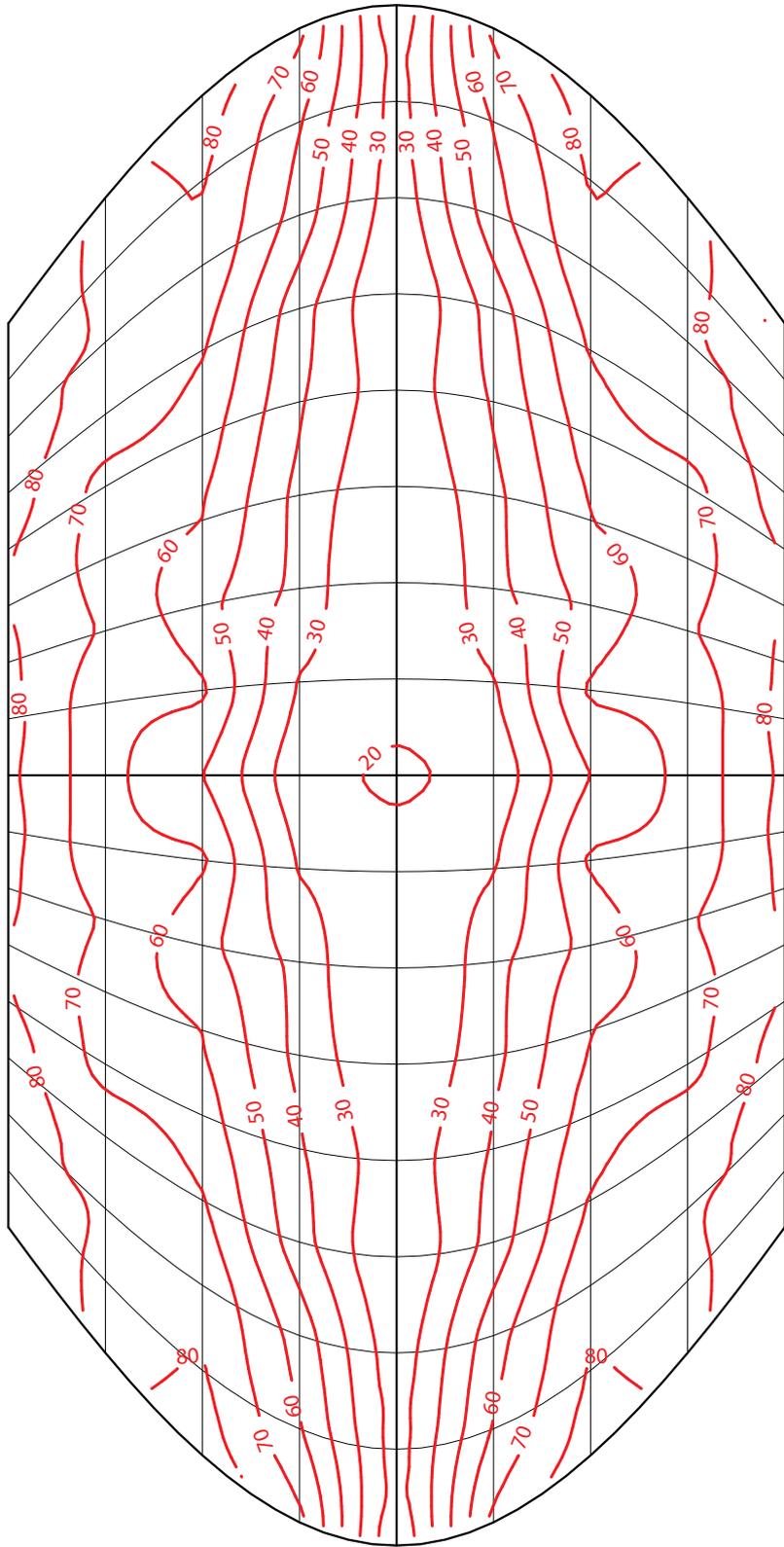


Figure 15: Eckert V projection, contour lines of the analysis reliability, RES distance function.

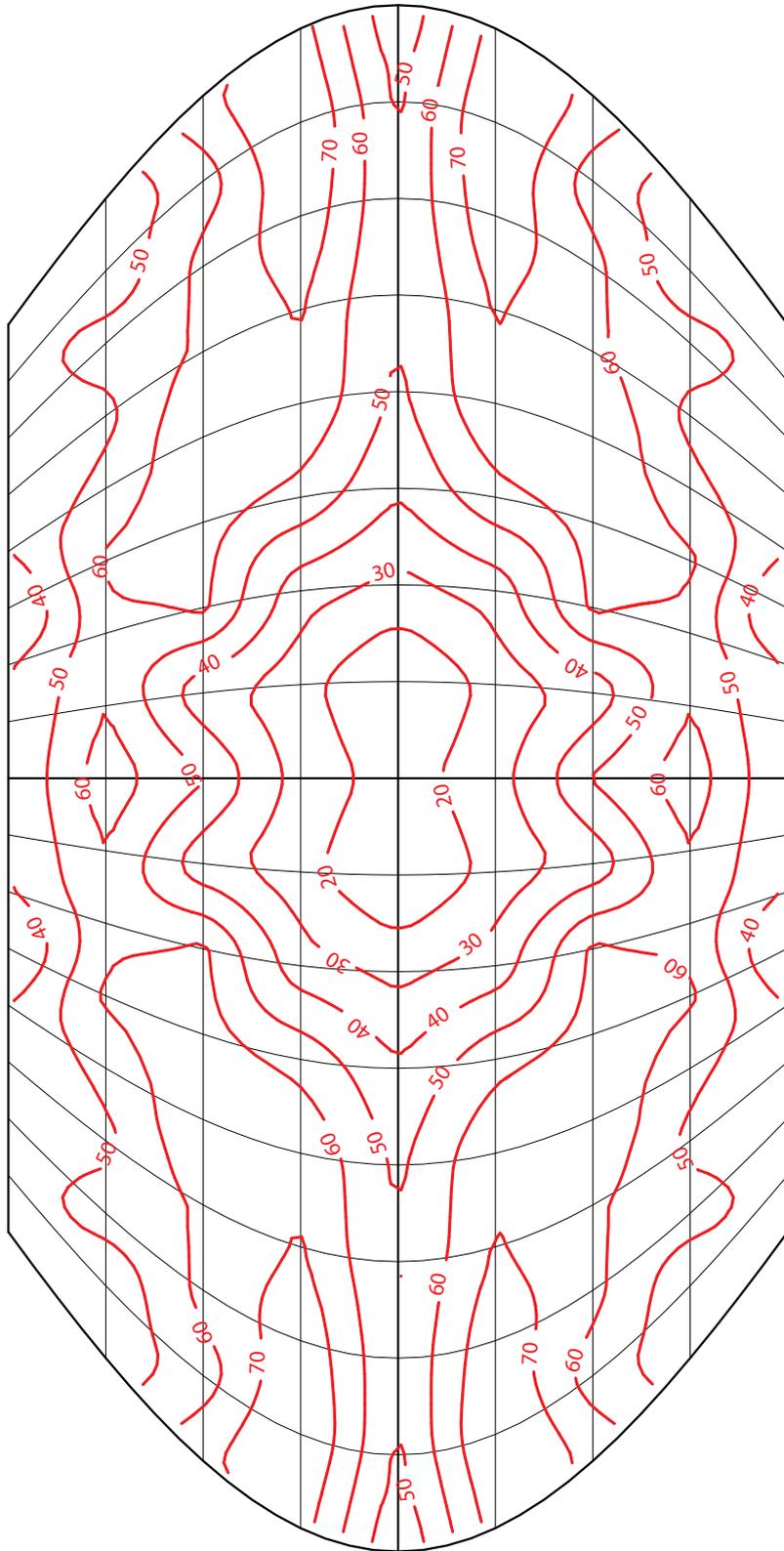


Figure 16: Eckert V projection, contour lines of the analysis reliability, VFTF distance function.

Table 2: Results of 300 tests for Maps 1-8, a comparison of methods M5-M7 and 3 optimizing techniques (NLS, DE, NM).

Map	ΔX [m]	ΔY [m]	Method	Optimiz.	EFF1	EFF2	N_IT	RES_C	RES_A	TIME
1	1	2	M7	NLS	96	12	8905	$3.23 \cdot 10^{-4}$	$3.68 \cdot 10^{-1}$	697
			M6	NLS	94	14	4890	$1.78 \cdot 10^{-4}$	$2.80 \cdot 10^{-1}$	518
			M5	NLS	99	24	4276	$8.96 \cdot 10^{-5}$	$4.91 \cdot 10^{-3}$	279
				DE	100	7	56005	$8.78 \cdot 10^{-4}$	$8.78 \cdot 10^{-4}$	23838
				NM	97	7	59587	$1.01 \cdot 10^{-4}$	$1.16 \cdot 10^{-4}$	1661
	2000	15000	M7	NLS	95	3	9387	$8.35 \cdot 10^1$	$2.42 \cdot 10^6$	1561
			M6	NLS	88	8	13499	$7.32 \cdot 10^1$	$6.32 \cdot 10^6$	1352
			M5	NLS	95	48	14422	$8.41 \cdot 10^1$	$3.25 \cdot 10^6$	1324
				DE	100	35	268852	$8.53 \cdot 10^1$	$8.53 \cdot 10^1$	114568
				NM	99	15	11845	$9.60 \cdot 10^1$	$9.86 \cdot 10^1$	3798
2	0	3500000	M7	NLS	64	64	5698	$1.76 \cdot 10^6$	$1.98 \cdot 10^{16}$	403
			M6	NLS	89	70	4470	$2.89 \cdot 10^8$	$6.01 \cdot 10^{12}$	556
			M5	NLS	95	95	2944	$2.59 \cdot 10^6$	$3.02 \cdot 10^{12}$	168
				DE	100	100	103700	$2.74 \cdot 10^6$	$2.74 \cdot 10^6$	39280
				NM	97	97	16393	$2.68 \cdot 10^6$	$1.05 \cdot 10^{12}$	3551
3E	0	0	M7	NLS	91	91	2107	$2.46 \cdot 10^{-1}$	$4.52 \cdot 10^0$	168
			M6	NLS	95	95	2261	$7.54 \cdot 10^{-1}$	$1.63 \cdot 10^0$	188
			M5	NLS	94	94	1964	$2.56 \cdot 10^{-1}$	$2.87 \cdot 10^0$	154
				DE	100	100	10864	$7.68 \cdot 10^{-1}$	$7.68 \cdot 10^{-1}$	13517
				NM	100	100	40860	$6.67 \cdot 10^{-1}$	$6.67 \cdot 10^{-1}$	1157
4W	0	0	M7	NLS	88	88	2357	$3.31 \cdot 10^{-1}$	$3.84 \cdot 10^0$	232
			M6	NLS	91	91	2700	$6.64 \cdot 10^{-1}$	$9.04 \cdot 10^0$	259
			M5	NLS	96	96	2062	$3.63 \cdot 10^{-1}$	$1.69 \cdot 10^0$	187
				DE	100	100	11474	$6.06 \cdot 10^{-1}$	$6.06 \cdot 10^{-1}$	14775
				NM	100	100	52018	$9.22 \cdot 10^{-1}$	$9.22 \cdot 10^{-1}$	1722
5	0	0	M7	NLS	100	100	3007	$8.89 \cdot 10^{-1}$	$8.89 \cdot 10^{-1}$	325
			M6	NLS	99	99	2482	$8.84 \cdot 10^{-1}$	$1.08 \cdot 10^0$	312
			M5	NLS	100	100	2400	$8.89 \cdot 10^{-1}$	$8.89 \cdot 10^{-1}$	273
				DE	100	100	11376	$8.89 \cdot 10^{-1}$	$8.89 \cdot 10^{-1}$	19868
				NM	100	100	51137	$2.18 \cdot 10^0$	$2.18 \cdot 10^0$	2255
6	0	0	M7	NLS	97	97	1932	$1.83 \cdot 10^{-1}$	$2.07 \cdot 10^2$	350
			M6	NLS	100	100	1424	$1.88 \cdot 10^{-1}$	$1.88 \cdot 10^{-1}$	210
			M5	NLS	100	100	1610	$1.88 \cdot 10^{-1}$	$1.88 \cdot 10^{-1}$	263
				DE	100	100	20992	$3.77 \cdot 10^{-1}$	$3.77 \cdot 10^{-1}$	21460
				NM	100	100	35399	$4.61 \cdot 10^{-1}$	$4.61 \cdot 10^{-1}$	2413
7	0	0	M7	NLS	82	82	2101	$6.72 \cdot 10^{-1}$	$9.78 \cdot 10^1$	355
			M6	NLS	100	100	1595	$6.08 \cdot 10^{-1}$	$6.08 \cdot 10^{-1}$	227
			M5	NLS	100	100	1664	$7.10 \cdot 10^{-1}$	$7.10 \cdot 10^{-1}$	247
				DE	100	100	20563	$1.41 \cdot 10^0$	$1.41 \cdot 10^0$	23165
				NM	100	100	33352	$7.31 \cdot 10^{-1}$	$7.31 \cdot 10^{-1}$	1160
8	0	0	M7	NLS	100	100	1814	$2.62 \cdot 10^{-1}$	$2.62 \cdot 10^{-1}$	251
			M6	NLS	100	100	1265	$4.49 \cdot 10^{-1}$	$4.49 \cdot 10^{-1}$	208
			M5	NLS	100	100	1736	$2.62 \cdot 10^{-1}$	$2.62 \cdot 10^{-1}$	234
				DE	100	100	46278	$2.61 \cdot 10^{-1}$	$2.61 \cdot 10^{-1}$	40468
				NM	99	99	49071	$2.97 \cdot 10^{-1}$	$3.88 \cdot 10^{-1}$	3990