

RESEARCH ARTICLE

Computation and visualisation of the accuracy of old maps using differential distortion analysis

Manuel Claeys Boùuaert^{a*}, Bernard De Baets^b, Soetkin Vervust^a, Tijs Neutens^a,
Philippe De Maeyer^a, Nico Van de Weghe^a

^a *CartoGIS, Dept. of Geography, Ghent University,
Krijgslaan 281-S8, 9000 Ghent, Belgium*

^b *KERMIT, Dept. of Mathematical Modelling, Statistics and Bioinformatics,
Ghent University,
Coupure links 653, 9000 Ghent, Belgium*

(received July 2015)

* Corresponding author. Email: manuel.claeysbouuaert@ugent.be

The accuracy of old maps can hold interesting historical information, and is therefore studied using distortion analysis methods. These methods start from a set of ground control points that are identified both on the old map and on a modern reference map or globe, and conclude with techniques that compute and visualise distortion. Such techniques have advanced over the years, but leave room for improvement, as the current ones result in approximate values and a coarse spatial resolution. We propose a more elegant and more accurate way to compute distortion of old maps by translating the technique of differential distortion analysis, used in map projection theory, to the setting where an old map and a reference map are directly compared. This enables the application of various useful distortion metrics to the study of old maps, such as the area scale factor, the maximum angular distortion and the Tissot indicatrices. As such a technique is always embedded in a full distortion analysis method, we start by putting forward an optimal analysis method for a general-purpose study, which then serves as the foundation for the development of our technique. Thereto, we discuss the structure of distortion analysis methods and the various options available for every step of the process, including the different settings in which the old map can be compared to its modern counterpart, the techniques that can be used to interpolate between both, and the techniques available to compute and visualise the distortion. We conclude by applying our general-purpose method, including the differential distortion analysis technique, to an example map also used in other literature.

Keywords: Historical cartography, Thin plate spline, Differential distortion analysis, Area scale factor, Maximum angular distortion, Tissot indicatrices.

1. Introduction

Old maps are precious historical documents and often beautiful works of art. They carry historical information, both through their geographical content and through the techniques used to acquire and display that content. In recent years, libraries, archives and geographical agencies have increasingly made their rich cartographic collections digitally available to the world (e.g. David Rumsey Map Collection, Mapire and Cartesius). Through the work of researchers and crowdsourcing initiatives, many historical maps are currently being annotated and georeferenced, greatly facilitating their ability to be queried and compared. Thanks to this evolution, it is becoming customary to study the *geometric accuracy* of old maps using dedicated software applications, and use this information as a source for historical studies. Geometric accuracy is one of three types of map accuracies (as defined by Blakemore and Harley (1980)) and encompasses the question whether distances, angles and areas are reproduced correctly on the map. Such software applications have improved systematically over the years by adding more visualisation methods, supporting more map projections and focusing on local distortion. However, the techniques they employ to compute distortion leave room for improvement. They often result in approximate values and a coarse spatial resolution, and feature undesirable extra parameters and requirements. Hence, they do not manage to unveil all of the historical information encoded in a map's distortion characteristics.

At first sight, it may seem like the presence of geometric distortions reduces the amount of information in old maps, but in reality, their location, spatial variation, magnitude and type can help uncover the story of the map's production process. Every change

in accuracy must exist for a reason: different surveying teams may have been assigned to different areas of the map, different surveying methods may have been used, the amount of care applied may have varied by area, the cartographer may have used different existing maps or triangulation data as a starting point, or may have applied small tweaks to enforce accordance with known maps of neighbouring regions. It becomes clear that studying these distortions taps into an extremely rich source of historical information, and that we should therefore strive towards developing analysis methods that can describe distortion patterns down to the smallest detail.

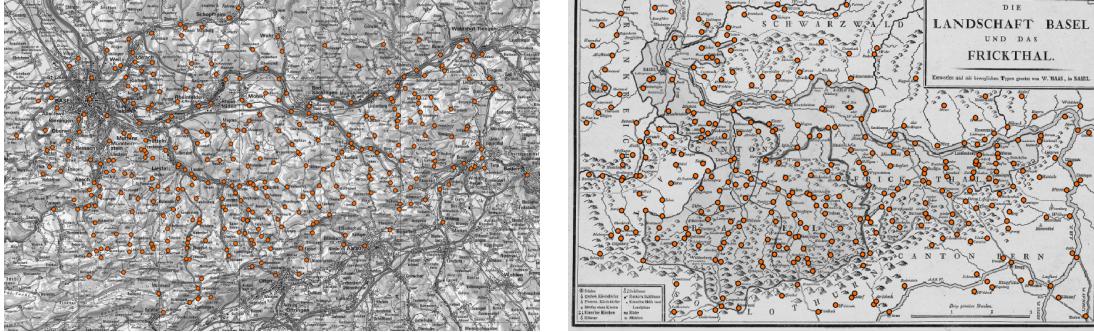
For centuries, distortion has been treated in the most precise way possible in the context of map projection theory, through an analytic study of the map projection function and its partial derivatives. As high-detail georeferencing of old maps is becoming more frequent, there is much potential for applying a similar approach to the interpolation function generated during the georeferencing procedure. This paper proposes and discusses such an improved technique.

State of the art

Distortion analysis methods, used to study the accuracy of old maps, can be seen as an extension of *georeferencing*: a popular three-step method used to optimally align raster data (images) with a modern reference map (Hill 2009). Georeferencing is often used for aerial and satellite imagery, but can be applied to images of old maps as well, as shown in Figure 1. The georeferencing method starts with the introduction of a well-chosen modern reference map. Then, using features that can be identified on both the reference map and the image, a set of *ground control points* is constructed: a list of coordinates on both the reference map and the image. Finally, employing an interpolation technique, these ground control points are used to set up a function that maps a point on the reference map to its corresponding point on the image. Once this method is applied, one can leverage the information in the image by vectorising it and porting it to the reference map using the interpolation function. Most often, however, the image is first warped to align it with the reference map (by mapping and resampling its pixels) and then vectorised directly in its new setting. To study the distortion of old maps, the procedure of georeferencing is typically extended with an extra step where the accuracy of the old map is examined, yielding a complete distortion analysis method.

Researchers have used different varieties of distortion analysis methods. A comprehensive overview of commonly used methods as well as a list of papers that employ them has been published by Forstner and Oehrli (1998). The techniques range from very basic manual approaches that simply compare a limited set of distances on the old and modern map to more advanced computational methods that are superior in accuracy and scalability. Most of the more recent studies use the software ‘MapAnalyst’ by Jenny, available since 2005 (see Jenny and Hurni 2011). A non-exhaustive list of such studies can be found on the website of MapAnalyst (Jenny 2015b), along with a list of publications on the techniques for computation and visualisation of the accuracy of old maps and computation of statistical measures (Jenny 2015a).

The introduction of the free and open source ‘MapAnalyst’ software around 2005 significantly catalysed the research in this domain due to its general applicability and ease of use. The technicalities of this software are explained in a general paper (Jenny and Hurni 2011), and details on the computation of scale and rotation isolines can be found in a separate paper (Jenny *et al.* 2007). The software loads the old map and a modern reference map whose projection can be chosen by the user. Both maps are shown side by side, allowing the user to build a set of ground control points. Then, an interpolation is



(a) Ground control points on a modern reference map: the Swiss National Map 1:200 000 (swisstopo 2015) in the Swiss Oblique Mercator projection.

(b) Ground control points on the old map under investigation: a 1798 map of the Basel and Frickthal region (Haas 1798).

Figure 1.: An old map can be georeferenced using a set of *ground control points* (displayed in orange) that indicate corresponding features on a well-chosen modern reference map and on the old map. The georeferencing procedure can be extended to yield a distortion analysis method. The distortion of this map will be analysed in Section 4.

constructed from the old to the modern map using multiquadratic interpolation (without smoothing). The user can choose to visualise the distortion using a distortion grid, displacement vectors or isolines of scale and rotation. The computation of local scale factors and rotation angles at a number of grid cells enables the isoline visualisation. This focus on the study of ‘local’ distortion is an important part of the progress contained in Jenny’s work. Still, we show that his technique leaves room for improvement.

Two authors have, to some degree, mentioned or used variants of the approach we propose in this paper. In his paper on bidimensional regression, Tobler discusses interpolation and regression between two sets of point coordinates with the prime example of ground control points on an old and a modern map (Tobler 1994). He describes linear bidimensional regression techniques such as the affine and Helmert transformation as well as different types of curvilinear transformations. In the search for a good general interpolation, he opts for an approach with optimal smoothness, which is equivalent to the concept of a thin plate spline. In his examples, he uses various distortion visualisation techniques, including Tissot indicatrices, one of the techniques from map projection theory. However, he does not elaborate on the steps needed to compute them, even though sufficiently detailed instructions are desirable for a wider adoption of this visualisation. Furthermore, Tobler discusses two important aspects of a distortion analysis method (interpolation techniques and visualisation techniques), but leaves other aspects unmentioned – largely because the focus of the paper is on bidimensional regression in general, and old maps are only treated as an example. Additionally, Tobler hints at using techniques from map projection theory to analyse the accuracy of mental maps (Tobler 1976).

Beineke also proposed to use techniques from the field of map projection theory to analyse distortion of historical maps (Beineke 2007). He describes a distorted map as resulting from a perfect projection from the globe to a map, followed by a ‘distortion mapping’ accounting for the inaccuracies. He then estimates this distortion mapping from a set of ground control points using multiquadratic interpolation (with non-zero smoothing). He analyses this total mapping from the globe to the distorted old map using classical error metrics from map projection theory, such as the area scale factor,

the maximum angular distortion and Tissot indicatrices. Beineke's work has a larger focus on the use of differential distortion analysis techniques. However, a disadvantage is that some important choices are made a priori with regard to the distortion analysis method, without presenting the reader with all available options: he makes the case for multiquadratic interpolation, but does not mention alternative interpolation techniques, and, more importantly, he also implicitly assumes the researcher wants to investigate how well the distances and angles on the map compare to the corresponding values on a globe, and does not mention comparison to values on a reference map in a certain projection.

Paper outline

In this paper, we propose an improved technique to compute and visualise distortion of old maps. However, such a technique is always applied as part of a distortion analysis method, and hence can only be discussed within this context. Since analysis methods can be set up in a variety of ways, we start by presenting a discussion of the main options the researcher has regarding the composition of a distortion analysis method in Section 2. In order to properly embed our proposed technique, we set out to determine the optimal general-purpose distortion analysis method, which will then serve as the foundation for the development of the technique we propose. Thereto, we conclude the discussion of each choice by selecting the optimal option for our general-purpose method, and review our method at the end of this Section. In Section 3, we introduce the foundations of map projection theory and its analytical technique of 'differential distortion analysis'. Then, in order to apply this technique within our general-purpose method, we translate it to the setting where an old and reference map can be directly compared. This results in expressions for distortion measures that are readily applicable. Finally, in Section 4, we demonstrate our general-purpose method, including our improved technique for distortion computation and visualisation, by performing an accuracy study of a 1798 map of the Basel and Frickthal region.

2. Design of the distortion analysis method

We have already explained how a distortion analysis method is composed of several steps and can be viewed as an extension to georeferencing. Here we provide an overview of the options available during these steps, and point out their advantages and disadvantages. Sections 2.1 and 2.2 discuss two important settings in which the inaccuracies can be viewed. Section 2.3 reviews the most relevant interpolation techniques. Section 2.4 gives an overview of the techniques to compute and visualise distortion. We conclude each section by choosing the setting or technique most suited for a general-purpose distortion analysis method, and review this method in its entirety in Section 2.5.

2.1. *Map-to-map vs. globe-to-map setting*

Most authors, such as Tobler (1994) and Jenny and Hurni (2011), examine the deformation of an old map by determining where distances, angles and areas on the old map differ from their corresponding values on a modern map ('map-to-map'). Some authors, such as Beineke (2007), investigate how these quantities on the old map differ from their values on the surface of the earth ('globe-to-map'). These settings are different because

the curvature of the earth's surface causes at least some of these quantities to change when mapped from a globe to a flat plane.

Generally, the question is how well the mapper succeeded in his task of mapping the area under consideration on a flat surface (using a specific projection and datum), and so a map-to-map analysis is the logical choice. The distortions caused by the projection are not considered, and the focus is on the distortions of the old map compared to the modern reference map. Of course, to allow for this comparison, it is important to use a modern reference map of the same projection and datum as the old map. For many maps, the projection and datum are stated by the author or can be accurately estimated. If the map concerns only a small area and the projection is unknown, a conformal projection (i.e. a projection which preserves angles locally), for example the one used by the local official mapping agency, will suffice.

Problems arise, however, when confronted with a map of a large area if it is impossible to determine which projection was used by the cartographer (if any). We cannot simply propose a projection, since the size of its distortions could be comparable to that of the map's imperfections. In this specific case, the globe-to-map setting, as used by Beineke (2007), is the more interesting option for distortion analysis. The distorted map projection $(\lambda, \phi) \rightarrow (x, y)$ is decomposed into a perfect map projection $(\lambda, \phi) \rightarrow (x', y')$ and a distortion mapping $(x', y') \rightarrow (x, y)$. For the first step, a projection must be proposed (e.g. the equirectangular projection), even though this choice will not affect the final result. The distortion mapping of the second step, from the reference map in the chosen projection to the old map, is computed from the set of ground control points using an interpolation technique (see further). The distortions resulting from a globe-to-map analysis are caused by both the curvature of the earth's surface and the errors in the old map, whereas analysis in the map-to-map setting visualises only distortion specific to the old map.

We end this discussion with an important note regarding a recent trend in digital cartography. Nowadays, crowdsourcing initiatives for georeferencing and distortion analysis are becoming increasingly popular. In these initiatives, contributors are asked to identify ground control points on an old map and on a modern reference map, both serviced through the web. The reference map is most often in the WGS84/Pseudo-Mercator projection EPSG:3857, the de facto standard for web mapping applications. This 'default', 'conformal' projection has several undesirable properties if the goal is to perform distortion analysis in the map-to-map setting. Firstly, it is slightly non-conformal (Battersby *et al.* 2014), and, secondly, like any Mercator projection, its area deformation grows at a high rate when moving away from the equator. For maps of small areas not too close to the poles the impact of these properties will be limited. In case of distortion studies where a high accuracy is required, and especially if the area under consideration is large and close to the poles, the coordinates of the ground control points on the reference map should be transformed to a more suited projection before computing distortions.

For our general-purpose method, it is most likely that we want to know how well the mapper succeeded in representing distances and angles on a flat surface using a specific projection, and that either the employed projection is known or the area is small enough to use a local conformal projection as reference. Hence, we choose the **map-to-map setting** for our method.

2.2. Old-to-modern vs. modern-to-old setting

An analysis method can either compute and visualise how the modern map or globe must be warped in order to fit on the old map, or it can perform the inverse process. The inversion is straightforward in most cases. The modern-to-old setting seems to be more popular in literature (Tobler 1994, Beineke 2007, Jenny and Hurni 2011), but when comparing multiple old maps, their distortions are often visualised on the same modern reference map.

We chose the **modern-to-old setting** for our general-purpose method.

2.3. Interpolation techniques

The problem of predicting a value at new locations, in between points with known values, is one that appears in a variety of research fields and settings. Hence, many different interpolation techniques have been developed. Typically, a distinction is made between *regression* or *curve fitting* techniques, on the one hand, where a function is sought that passes near the known data points while being optimal in some sense, and *true interpolation* techniques on the other hand, where an optimal function must pass exactly through all known data points. We use the term ‘interpolation’ to cover both types. Regression functions embody the general trend in the data, and examples include affine transformations and polynomial regression. Some examples of true interpolation techniques are multiquadratic interpolation and thin plate spline interpolation.

In this section, we briefly present some of the most common interpolation techniques applicable when both the codomain and domain (i.e. the dependent and independent space) are two-dimensional. We assume a modern-to-old setting (the inversion is straightforward) and either a map-to-map setting or the second step of a globe-to-map setting. Given a set of n ground control points, coordinates in the domain and codomain are denoted as (u_i, v_i) and (x_i, y_i) , respectively. Some techniques discussed in this section were originally constructed for the case of a one-dimensional codomain and a two-dimensional domain. They are used for a two-dimensional codomain in what follows by applying them to both dependent variables.

This section does not intend to give a complete overview of these techniques. We refer the reader to the review articles of Franke (1979) and Hardy (1990) for a discussion on the mathematical foundation of most of these techniques, experimental data on their computational costs and a comparison of their capabilities to replicate mathematical surfaces from sparse, scattered data. Here, we present those interpolation techniques which these reviews claim yield the best results under general circumstances, together with those that are frequently used in literature (Forstner and Oehrli 1998).

Affine transformations

An affine transformation is a regression technique, that is, the two-dimensional equivalent of fitting a regression line. Ground control points with coordinates (u_i, v_i) in the domain and (\hat{x}_i, \hat{y}_i) in the codomain are assigned new coordinates (x_i, y_i) in the codomain using a six-parameter transformation obtained by minimising the sum of squared residuals. The same transformation can then be applied to arbitrary points (u, v) . The expression for both dependent variables contains a linear term for both independent variables plus

a constant term, yielding two regression planes:

$$x(u, v) = \alpha_1 + \beta_{11}u + \beta_{12}v \quad (1a)$$

$$y(u, v) = \alpha_2 + \beta_{21}u + \beta_{22}v. \quad (1b)$$

The parameters can be computed from the system of linear equations that expresses Eq. (1) for all ground control points. There are (generally) many more equations than unknowns, and the solution is expressed using a Moore-Penrose inverse (Jenny and Hurni 2011). A special case is the four-parameter Euclidean transformation – also called *Helmert transformation* in geodetic literature – where $\beta_{11} = \beta_{22}$ and $\beta_{12} = -\beta_{21}$, such that both dimensions scale uniformly and the shape of objects is preserved. From its four parameters, one can obtain the total global translation (α_1, α_2) , rotation $\alpha = \arctan(\beta_{21}/\beta_{11})$ and scaling $m = \sqrt{\beta_{11}^2 + \beta_{21}^2}$ needed to transform the set of points in the domain to optimally overlay them on the points in the codomain, while preserving the shape of the point set. Helmert transformations can be used to make the coordinates of the ground control points in both spaces similar (sometimes as a pre-processing step for another interpolation technique) without introducing extra distortions, and are also used to estimate the global scaling and orientation angle between two maps.

Polynomial regression and local regression

The regression surfaces of affine transformations are limited to the shape of a plane. With polynomial regression techniques, the regression surfaces have more degrees of freedom and are more flexible in adapting to the trend in the data. Polynomials of second and third order are commonly used, but very high orders should be avoided due to a risk of overfitting. If non-parametric techniques are preferred, one could use two-dimensional versions of local regression techniques such as LOWESS or LOESS (Cleveland 1979), which respectively fit a weighted linear or quadratic least squares regression to a localised subset of data.

Multiquadratic interpolation

Hardy introduced a new interpolation technique (Hardy 1971, 1990), which is now often used in distortion analysis of old maps. First, a matrix $\tilde{\mathbf{D}}$ of smoothed distances between the n ground control points is constructed, with elements $\tilde{d}_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2 + \Delta^2}$, where Δ is a smoothing factor. In matrix notation, the coordinates in the codomain are expressed as $\mathbf{x} = \tilde{\mathbf{D}}\mathbf{a}$ and $\mathbf{y} = \tilde{\mathbf{D}}\mathbf{b}$. The vectors \mathbf{a} and \mathbf{b} contain the interpolation coefficients and can be found by solving these two systems of equations. For a new point (u, v) in the domain, the coordinates (x, y) in the codomain are computed as

$$x(u, v) = \sum_{i=1}^n a_i \tilde{d}_i(u, v) \quad (2a)$$

$$y(u, v) = \sum_{i=1}^n b_i \tilde{d}_i(u, v) \quad (2b)$$

with $\tilde{d}_i(u, v) = \sqrt{(u_i - u)^2 + (v_i - v)^2 + \Delta^2}$. Multiquadratic interpolation is clearly a true interpolation, independent of the value of Δ . Note, however, that for $\Delta = 0$ the partial derivatives of the functions in Eq. (2) do not exist at the ground control points. Indeed, when plotting the interpolated surface of one of the dependent variables, conic spikes are observed at the control points. In many cases this is undesirable, which is one of the reasons some authors apply smoothing, i.e. $\Delta \neq 0$. The partial derivatives then exist everywhere, but the function's behaviour near the control points is highly sensitive to the chosen value of the smoothing parameter. Many authors have sought to determine an optimal value for Δ , and some heuristic expressions have been proposed, but no definitive statements seem to exist (Hardy 1990). One can thus choose to follow a heuristic or leave it up to the user to use the value he or she deems best. From Eq. (2), one can see that multiquadratic interpolation is a 'radial basis function'-based technique (Buhmann 2003).

Thin plate spline interpolation

Thin plate splines, introduced by Duchon (1977), are the principal techniques from the class of spline interpolation functions applicable to data with a one-dimensional codomain and a two-dimensional domain. Out of all functions that go through the known data points, they yield the function with minimal 'bending energy', in the sense that it minimises the integral of the squares of the second derivatives:

$$E_{\text{tps}}[f] = \iint_{\mathbb{R}^2} f_{uu}^2(u, v) + 2f_{uv}^2(u, v) + f_{vv}^2(u, v) \, du \, dv. \quad (3)$$

The solution to this minimisation problem uses the solution to the biharmonic equation $\Delta^2 f = 0$, which gives the (radially symmetric) Green's function $G(r) = \frac{1}{8\pi} r^2 \ln(r)$ (Wahba 1990). The minimiser of the bending energy is a linear combination of Green's functions centred at the datapoints, plus a linear polynomial term:

$$x(u, v) = \sum_{i=1}^n a_i G(d_i(u, v)) + (\alpha_1 + \beta_{11}u + \beta_{12}v) \quad (4a)$$

$$y(u, v) = \sum_{i=1}^n b_i G(d_i(u, v)) + (\alpha_2 + \beta_{21}u + \beta_{22}v), \quad (4b)$$

where $d_i(u, v) = \sqrt{(u_i - u)^2 + (v_i - v)^2}$. The parameters are obtained by solving a linear system. The linear term corresponds to an affine transformation and represents the kernel of the bending energy operator (3), in the sense that an affine transformation does not alter the bending energy. Thin plate splines can thus be viewed as being composed of a global affine transformation that optimally relocates the data without bending it, and a local non-affine component that accounts for the additional local displacements. Note that this affine transformation is not necessarily a shape-preserving Helmert transformation. Thin plate splines result in true interpolation, are optimal with respect to smoothness and have no additional tuning parameters in their standard form. The interpolating surface is available in closed form and the partial derivatives exist everywhere and are easily computable. As a matter of fact, we are ensured of a smooth (continuous and differentiable) surface, with continuous first derivatives. Thin plate splines are often used as the interpolation technique in the 'rubbersheeting' process to align aerial imagery with

maps. Like multiquadratic interpolation, thin plate splines can be seen as a technique based on radial basis functions (Buhmann 2003).

Other custom techniques

Some techniques were designed to yield interpolation functions with special properties. An important class consists of techniques that prevent ‘foldover’ of the interpolation (Tideman et al. 2001, Seo and Cordier 2010, Hongchuan et al. 2012). This can occur for maps with extremely inaccurate control points, and is mathematically expressed by the interpolation’s Jacobian determinant flipping sign. Furthermore, similar to local regression, true interpolation techniques with improved locality have been developed (Fornefett et al. 2001).

Which interpolation technique is best suited for our general-purpose distortion analysis method? True interpolation is generally preferred over regression, since it is more objective, as it does not require one to draw the arbitrary line between what is the global trend and what are local displacements. Regression can, however, still be preferred in some cases, for instance, if one is more interested in the general trend than in the full relation, or if one suspects that the ground control points contain some errors. Hardy (1990) studied the accuracy of general true interpolation schemes, finding multiquadratic interpolation and thin plate spline interpolation to generally be the best and second or third best interpolation technique, respectively, with a relatively small difference between both. Both have a comparable computational time complexity of $\mathcal{O}(n^3)$. Other desirable properties for an interpolation technique are a guaranteed existence of partial derivatives (as we look ahead to the differential distortion analysis), robustness with respect to the number and distribution of control points, and objectivity of the interpolation. These three considerations disfavour multiquadratic interpolation, due to the possible erratic behaviour of its derivatives at the control points and the ambiguous determination of the smoothing factor, which uses heuristics or a user’s opinion and may need to be readjusted if the control points change. Hence, we chose **thin plate spline interpolation** for our method.

2.4. Techniques for distortion computation and visualisation

We have already mentioned that in order to compute and visualise inaccuracies, we need to extend the georeferencing process with an extra step. This step should describe the ‘local distortion’ at different points on the map, in order to obtain distortion information of high spatial resolution. There are two distinct ways in which this can be accomplished.

Regression residuals carrying local distortion information

In case regression is used, one can collect distortion information from the regression’s residuals at the ground control points. The prime example here is the technique of *displacement vectors*, which are most commonly obtained by applying a Helmert transformation and visualising the residuals in the codomain as arrows. Some authors vary this technique by using other regression techniques such as affine transformations or by visualising the displacements using circles (Forstner and Oehrli 1998).

Interpolation function carrying local distortion information

The other possibility is to visualise the local distortion information contained in the true interpolation function or regression function. A popular technique here is the *distortion grid*, introduced by Wagner as early as 1896 (Wagner 1896). Such a grid is drawn by transposing a rectangular grid from the domain to the codomain using the interpolation function. This yields a deformed grid that intelligibly visualises how the modern map should be warped to exactly (for true interpolation) or approximately (for regression) coincide with the old map (in the modern-to-old setting).

One can also use techniques which first compute measures of local distortion from the interpolation function, and consequently visualise these measures on the map. The popular MapAnalyst software follows this procedure, and computes measures of local scale and rotation angle at a number of grid cells as the scale and rotation *parameters of a 'local affine' transformation* – an affine transformation that uses only the control points within a specified radius of each grid cell and additionally weighs each control point as function of its distance to the cell. This is illustrated in Figure 2. In the software, the resulting distortion measures are then visualised using isolines. This focus on the study of local distortion is an important part of the progress contained in Jenny's work. Still, the proposed technique leaves room for improvement.

In Section 3, we introduce an improved technique to compute local distortion metrics from the interpolation function, called *differential distortion analysis*. It is similar to how the distortion of a cartographic projection is computed from its map projection function, and we adapt this analysis to allow for its application to the map-to-map setting. This technique is not only more elegant, but it also computes the exact distortion at every point, without any spatial averaging, from the differential properties of the interpolation function. Of course, the quality of the resulting distortion values is still subjected to the quality of the interpolation function, and hence to the quality and quantity of the ground control points.

As will become apparent in Section 3, the relation between the analytical technique of differential distortion analysis and the approximate technique used in the MapAnalyst software can be seen, to some extent, as the two-dimensional analog of the relation between exact computation of the local slope of a one-dimensional function through analytical derivation of the explicit function, and its approximate computation through some finite difference approach based on a limited number of known points.

Which visualisation technique is suited for our general-purpose method? Displacement vectors are frequently used in literature. They are easy to calculate and seemingly easy to interpret. However, they only visualise distortion information at the ground control points, and an interpretation of the continuous distortion trend is left to the observer. This can be a hard task, since the vectors are often very heterogeneous in length and direction, and since distortions are not necessarily represented by large vectors, but rather by a discontinuity between neighbouring vectors. Also, the fact that their appearance is strongly impacted by the preceding regression (which is influenced by all control points) reduces the local character of this visualisation technique. Distortion grids, on the other hand, are a truly local technique and are easily interpretable. Their disadvantages are that the human perception of the deformation can be influenced by the deformation along the directions of the gridlines, and that they do not provide quantitative distortion information or emphasise the distorted areas. Also, detailed visualisations need many gridlines, which compromises the visibility of the underlying map. Therefore, techniques which compute measures of local distortion are preferred. The resulting measures can

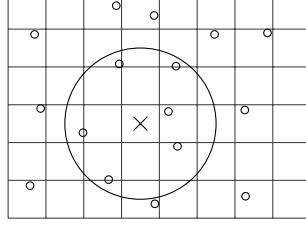


Figure 2.: In the MapAnalyst software, local distortion metrics are computed at a number of grid cells. At every cell, an affine transformation is performed taking into account only those ground control points within a specified radius of the cell's centre – six control points in the above example. These ‘local affine’ transformations yield estimates for the local scale and rotation angle at every cell. The differential distortion analysis technique proposed in this paper provides a more precise and generally applicable alternative.

be plotted over the map to visualise trends and highlight distorted areas. The approach used in MapAnalyst manages to compute a local interpretation of distortion at every grid cell. However, when doing so it inevitably applies some form of spatial averaging, and its result is therefore always approximate. Furthermore it introduces undesirable extra requirements (such as a minimal number of nearby ground control points) and parameters (such as the radius, which must be tuned by the user in order to balance fine resolution (small radius) and sufficient data (large radius)). This paper presents **differential distortion analysis** as an improved version of this approach, and therefore we select it for our general-purpose distortion analysis method.

2.5. Our general-purpose distortion analysis method

Gathering the chosen options, our general-purpose distortion analysis method compares the old map to a modern reference map in the same projection or in a local conformal projection. It analyses how the modern map must be warped in order to fit on the old map. The warping is modelled using a thin plate spline interpolation. Finally, the distortions induced by this warping are computed and visualised using differential distortion analysis.

3. Differential distortion analysis

3.1. Differential distortion analysis in the globe-to-map setting

Cartographic map projections map the curved surface of the earth onto a flat plane. A sphere (or ellipsoid) and a plane have a different Gaussian curvature, and thus, by Gauss’s famous Theorema Egregium, a different local metric. The consequence is that distances, angles or areas are inevitably distorted when the former is mapped onto the latter. In map projection theory, these distortions are expressed and studied using differential distortion analysis. We briefly review this analysis here.

Differential distortion analysis considers two main local error metrics: the *area scale factor* σ indicates how much a local infinitesimal surface element is enlarged on the map (relative to the map’s scale) and the *maximum angular distortion* 2Ω indicates the maximal difference between an angle on the globe and on the map, making it a measure for shearing. Due to the curvature of the earth, it is impossible for both σ and 2Ω to equal

their optimal values 1 and 0 everywhere simultaneously. However, projections where one of both is optimal everywhere exist and are called equivalent and conformal, respectively. All other projections try to find a balance between these extremes. Drawing the contour lines or coloured maps of the quantities σ and 2Ω on a map allows for an intelligible overview of the map's distortions.

The french mathematician Tissot investigated how the projection of infinitesimal circles on the earth's surface resulted in infinitesimal ellipses on the projected map. The orthogonal axes of these ellipses show the orientation and magnitude of the maximal and minimal scale factor at that location on the map. To observe the spatial variation of distortion, small ellipses, called 'Tissot indicatrices', are drawn on the map at the intersection of the main parallels and meridians.

A full mathematical derivation of map projection theory, including the formulas for the area scale factor σ , the maximum angular distortion 2Ω and the magnitude and orientation of the main axes of the Tissot indicatrices, can be found in literature for the spherical and the elliptical earth model as well as for more general Riemannian manifolds (Canters and Decleir 1989, Bugayevskiy and Snyder 1995, Grafarend and Krumm 2006). Here, we briefly lay out the start of this derivation and the main results for the spherical earth model. In addition to distortion analysis of map projections, the equations of this section can also be used for distortion analysis of old maps in the globe-to-map modern-to-old case, by due application of the chain rule of calculus.

The derivation typically starts by expressing the squared length of an infinitesimal line element ds^2 , both on the projected map in cartesian coordinates (x, y) and on the surface of the earth (with *reduced earth radius* R) in longitude-latitude coordinates (λ, ϕ) . The *scale factor* μ is defined as the root of their ratio:

$$\mu^2 = \frac{ds^2}{ds'^2} = \frac{dx^2 + dy^2}{(R d\phi)^2 + (R \cos \phi d\lambda)^2}. \quad (5)$$

After inserting the projection formulas, $x = x(\phi, \lambda)$ and $y = y(\phi, \lambda)$ and applying the chain rule of calculus, the so-called *Gaussian fundamental quantities* E , F and G are defined, allowing to rewrite the squared scale factor as

$$\mu^2 = \frac{ds^2}{ds'^2} = \frac{Ed\phi^2 + 2Fd\phi d\lambda + Gd\lambda^2}{(R d\phi)^2 + (R \cos \phi d\lambda)^2}, \quad (6)$$

with

$$E = \left(\frac{\partial x}{\partial \phi} \right)^2 + \left(\frac{\partial y}{\partial \phi} \right)^2 \quad (7a)$$

$$F = \frac{\partial x}{\partial \phi} \frac{\partial x}{\partial \lambda} + \frac{\partial y}{\partial \phi} \frac{\partial y}{\partial \lambda} \quad (7b)$$

$$G = \left(\frac{\partial x}{\partial \lambda} \right)^2 + \left(\frac{\partial y}{\partial \lambda} \right)^2. \quad (7c)$$

The scale factor can vary with the location and orientation of the line element. The Gaussian fundamental quantities also depend on the location. From this point on, the general formula for the scale factor μ is used to derive equations for the scale factor along the meridians and parallels, the transformed angle between both and the area

scale factor σ . All are function only of the Gaussian fundamental quantities, the latitude and the reduced earth radius. The formula for the area scale factor σ reads

$$\sigma = \frac{\sqrt{EG - F^2}}{R^2 \cos \phi}. \quad (8)$$

The study of the indicatrices of Tissot can also be started from the general formula for the scale factor, and the length of the major axis a and minor axis b of the ellipses can be solved from a system of equations:

$$\begin{cases} a^2 + b^2 = \frac{E}{R^2} + \frac{G}{R^2 \cos^2 \phi} \\ ab = \frac{\sqrt{EG - F^2}}{R^2 \cos \phi} \end{cases}. \quad (9)$$

It can then be shown that the maximum difference between an angle on the globe and the corresponding angle on the map 2Ω is given by

$$2\Omega = 2 \arcsin \frac{a - b}{a + b}. \quad (10)$$

In order to draw a Tissot indicatrix at a certain location, we need to know how its axes are oriented. The angle θ_a between the ellipse's major axis and the cartesian x -axis of the projected map is given by the difference of the angle θ_p between the local parallel and the x -axis and the angle α_p between the parallel and the major axis:

$$\theta_a = \theta_p - \alpha_p. \quad (11)$$

It holds that

$$\theta_p = \arctan \frac{\frac{\partial y}{\partial x}}{\frac{\partial x}{\partial \lambda}} \quad (12)$$

and

$$\alpha_p = \pm \arcsin \sqrt{\frac{1 - \frac{a^2 R^2 \cos^2 \phi}{G}}{1 - \frac{a^2}{b^2}}}. \quad (13)$$

The sign of α_p depends on the angle (on the projected map) between the local parallel and local meridian being obtuse or acute, which is related to the sign of the fundamental quantity F , such that we have

$$\operatorname{sgn} \alpha_p = -\operatorname{sgn} F. \quad (14)$$

All necessary information is now available to draw a Tissot indicatrix at a certain location. We note that Laskowski (1989) discussed an alternative approach to compute indicatrices.

3.2. Differential distortion analysis in the map-to-map setting

We have argued that in a general-purpose distortion analysis it is preferable to directly compare the old map to a modern reference map. Hence, in order to apply the technique of differential distortion analysis within our general-purpose method, we now translate this technique to the map-to-map setting. We again start by expressing the scale factor as the root of the ratio of the squared lengths of a line element. This time both line elements are on a flat surface with cartesian coordinates (x', y') and (x, y) (for example a modern reference map and an old map, in the modern-to-old case), yielding

$$\mu^2 = \frac{ds^2}{ds'^2} = \frac{dx^2 + dy^2}{dx'^2 + dy'^2}. \quad (15)$$

This can again be expressed using the Gaussian fundamental quantities, after inserting the interpolation function, $x = x(x', y')$ and $y = y(x', y')$, resulting in

$$\mu^2 = \frac{ds^2}{ds'^2} = \frac{Edx'^2 + 2Fdx' dy' + Gdy'^2}{dx'^2 + dy'^2}, \quad (16)$$

with

$$E = \left(\frac{\partial x}{\partial x'} \right)^2 + \left(\frac{\partial y}{\partial x'} \right)^2 \quad (17a)$$

$$F = \frac{\partial x}{\partial x'} \frac{\partial x}{\partial y'} + \frac{\partial y}{\partial x'} \frac{\partial y}{\partial y'} \quad (17b)$$

$$G = \left(\frac{\partial x}{\partial y'} \right)^2 + \left(\frac{\partial y}{\partial y'} \right)^2. \quad (17c)$$

One immediately sees that Eqs. (16)–(17) would be equal to Eqs. (6)–(7) if we would replace (λ, ϕ) by (x', y') , if R would equal 1 and if ϕ would be sufficiently small such that $\cos \phi \approx 1$. Indeed, at small enough latitudes, a sphere with radius 1 behaves as a plane with Cartesian coordinates (λ, ϕ) . Instead of deriving the expressions from differential distortion analysis again for the map-to-map setting, we can simply translate them from the globe-to-map setting by putting $R = 1$ and $\cos \phi = 1$. One important note is that by replacing (λ, ϕ) with (x', y') , E and G in the globe-to-map case, respectively, correspond to G and E in the map-to-map case, because in map projection theory E and G are historically defined with partial derivatives with respect to ϕ and λ .

Executing this translation, we find that the area scale factor is given by

$$\sigma = \sqrt{EG - F^2}. \quad (18)$$

The maximum angular distortion is still given by

$$2\Omega = 2 \arcsin \frac{a - b}{a + b}, \quad (19)$$

but the system of equations from which a and b can be solved simplifies, and is given by

$$\begin{cases} a^2 + b^2 = E + G \\ ab = \sqrt{EG - F^2} \end{cases} . \quad (20)$$

Assuming $a \geq b$, this leads to

$$a = \sqrt{\frac{E + G + \sqrt{(E - G)^2 + 4F^2}}{2}} \quad (21a)$$

$$b = \sqrt{\frac{E + G - \sqrt{(E - G)^2 + 4F^2}}{2}} . \quad (21b)$$

The formulas for the orientation of Tissot indicatrices simplify slightly. In the domain, the x' -axis takes the place of the parallel. The angle between the major axis of the indicatrix and the cartesian x -axis in the codomain is given by

$$\theta_a = \theta_{x'} - \alpha_{x'} \quad (22)$$

with

$$\theta_{x'} = \arctan \frac{\frac{\partial y}{\partial x'}}{\frac{\partial x}{\partial x'}} \quad (23)$$

and

$$\alpha_{x'} = -\operatorname{sgn} F \times \arcsin \sqrt{\frac{1 - \frac{a^2}{E}}{1 - \frac{a^2}{b^2}}} . \quad (24)$$

The partial derivatives, the Gaussian fundamental quantities and all variables based on them are functions of the coordinates in the domain (for example the reference map, in the modern-to-old case). One can now evaluate the partial derivatives of the interpolation function and all following variables on a dense regular grid or mesh in the domain. The resulting values are valid at the corresponding points in the codomain. The values for the area scale factor and the maximum angular distortion can be interpolated to scalar fields, which can be visualised using coloured maps or contour plots. Tissot indicatrices can be plotted at desired locations to visualise the deformation in detail.

For practical purposes, the old map and the reference map are often brought to the same scale before starting this procedure, either using their scale bars or using a Helmert transformation. Also note that dimension-wise, for a scaling m obtained through a Helmert transformation, the scale factor μ and area scale factor σ relate as $\sigma \propto \mu^2 \propto m^2$.

Apart from the area scale factor and maximum angular distortion, map projection theory discusses a handful of other metrics that can potentially be useful to analyse the distortion of an old map. Their equivalents in the map-to-map setting (for example in the modern-to-old setting) describe, amongst others, the course of the scale factor on the old map along the horizontal and vertical axis of the reference map, the orientation of both axes on the old map (which can be used to check the North direction) and the angle between them, and the orientation of the axes a and b of maximal and

minimal scale factor. Also, some measures have been proposed that combine the effects of areal and angular distortion. The most popular examples are Airy's first and second distortion measures, Jordan's distortion measure and their modified versions proposed by Kavrayskiy (Canters 2002, Kavrayskiy 1958). The Airy-Kavrayskiy measure is often the most practical and balanced one, and is given by:

$$e_{AK} = \frac{1}{2} ((\ln a)^2 + (\ln b)^2). \quad (25)$$

In the next section, we illustrate differential distortion analysis on a historical map. We note that this technique also finds application in other fields, where some of the equations mentioned in this work can arise in a different context. In the research area of 'map projection transformations', which studies how spatial coordinates transform directly from one map projection to another, it is sometimes investigated how the already present distortion is altered during this transformation (Yang *et al.* 1999). In the research area of 'low-error projection design', the area or angular distortion of map projections is reduced by application of polynomial and complex-polynomial transformations. These can be designed so as to reduce one type of distortion without adding distortion of the other type (Canters 2002). A notable example outside of cartography is deformation-based brain morphometry, where neuroimaging data is mapped to reference data and volumetric changes are examined (Ashburner and Friston 2000).

4. Application example: a 1798 map of the Basel and Frickthal region

As an example, we compute and visualise the distortion of an old map using the general-purpose distortion analysis method proposed in Section 2.5, including thin plate spline interpolation and the differential distortion metrics derived in Section 3.2. Additionally, we also compute the displacement vectors and distortion grid. For comparability, we apply our method to the same map as the one used in the papers on the popular software MapAnalyst (Jenny *et al.* 2007, Jenny and Hurni 2011): a 1798 map of the Basel and Frickthal region (Haas 1798). The dataset can be downloaded from the MapAnalyst website and includes the old map, a modern map with corresponding 'world file' and a list of coordinates of the 343 ground control points on both maps. The Swiss National Map 1:200 000 by 'swisstopo' – the Swiss Federal Office of Topography – is used as modern map (swisstopo 2015). It employs the Swiss oblique Mercator projection CH1903/LV03–EPSG:21781. Our general-purpose distortion analysis method was implemented through a handful of MATLAB/Octave scripts and functions, which are publicly available on Github¹. They make use of the MATLAB Curve Fitting Toolbox functions or Octave Splines Package functions to build a thin plate smoothing spline and obtain its derivatives. The values of σ , 2Ω and e_{AK} were evaluated on a 110×160 grid. The computation took about 2 seconds on a standard desktop computer. The results were visualised in QGIS using simple TIN interpolation. The results of this analysis are displayed in Figures 3a–3f, and some summary statistics are given in Table 1. The last four figures were generated using the differential distortion analysis technique, and these

¹A MATLAB/Octave implementation of the general-purpose distortion analysis method proposed in Section 2.5, including thin plate spline interpolation and differential distortion metrics, is available at: <https://github.com/mclaeysb/distortionAnalysis>

types of visualisations are not included in the MapAnalyst toolset. Figure 4 compares one of our results with a similar visualisation obtained using MapAnalyst.

In Figure 3a the old map is displayed with its control points and displacement vectors with respect to a Helmert transformation. The result is the same as in Jenny *et al.* (2007). The main deviations from the linear trend appear in the lower right corner of the map. In Figure 3b a distortion grid is shown. It is easily interpretable as a warped reference map lying on top of the old map. The result is comparable to Jenny *et al.* (2007), who employs multiquadratic interpolation, but differs slightly. This is mainly due to the default setting of MapAnalyst, which, to speed up computation, computes a relatively coarse distortion grid and subsequently smoothes it using Bézier curves. For fairly distorted maps, like the one under investigation, this yields a bad approximation of the actual distortion grid. If one produces a more detailed distortion grid using MapAnalyst (finer and without smoothing), one observes that it almost perfectly compares to Figure 3b, indicating that, for this map, using multiquadratic interpolation or thin plate spline interpolation has little impact on the result. For completeness, we note that for this map, four zones with a very limited amount of ‘foldover’ are present – a concept introduced earlier in Section 2.3. They occur with both multiquadratic and thin plate spline interpolation, can be identified upon close inspection of a fine distortion grid or by computation of the Jacobian determinant, and are located in the western zone of high distortion, and along three highly stretched straps in the east.

Figure 3c displays how the area is deformed. The old and reference maps were first brought to the same scale, and subsequently the area scale factor σ was computed. The base-2 logarithm of σ is plotted, so that enlarging or shrinking an area with the same factor is displayed as an equally large distortion. A linear colour ramp [blue – transparent – red] indicates that the area is displayed at smaller, equal or larger size on the old map. Figure 3d shows a plot of the maximum angle deformation 2Ω , using a linear colour ramp [transparent – green] for low to high values of 2Ω . Note that the distortion information in Figures 3c and 3d agree with the distortion grid: enlarged and shrunken areas are coloured red and blue, respectively, shearing corresponds to a green colouring, and uncoloured areas have little distortion. In Figure 3e the Airy-Kavrayskiy measure of distortion, combining the effects of areal and angular distortion, is plotted using a linear colour ramp [transparent – yellow] for low to high values of distortion. These plots make it easy to identify distortion trends, which can provide valuable input for historical research. For example, straps of high σ or 2Ω , or zones with a different mean σ , may indicate that the mapper used different older maps or terrain surveys as input source material.

Finally, Figure 3f shows a collection of Tissot indicatrices plotted over the old map. These ellipses show how the scale factor (at their centerpoint) varies with direction. They allow us to visualise, on the same plot, the deformation of area and angle as well as the directions with maximal and minimal scale factor. To facilitate interpretation, a sufficient number of ellipses needs to be plotted, making this technique less suited for overviews of the distortion of highly inaccurate maps.

We now interpret Table 1, which presents some summary statistics on the values of the distortion measures at the 110×160 grid points and 343 ground control points used in this analysis. Displacement vectors reach lengths of more than 5 km. The orientation of such extreme vectors can be observed in Figure 3a. The mean and standard deviation of our three distortion statistics were used during the design of the corresponding colour-bars of Figures 3c–3e. As can be seen from the ‘min’ and ‘max’ columns, some points have values surpassing the corresponding colour-bar’s extent. A noteworthy effect can be

measure		mean	std. dev.	min	max
displacement vector length [meter]		962	846	36	5 170
base-2 logarithm of area scale factor	$\log_2 \sigma$	-0.078	0.542	-8.713	3.667
maximum angular distortion [rad]	2Ω	0.417	0.335	0.001	3.008
Airy-Kavrayskiy measure	e_{AK}	0.121	0.407	0.000	20.682

Table 1.: Summary statistics describing the distortion measures computed at the 110×160 grid points and the 343 ground control points in the analysis of the 1798 map of the Basel and Frickthal region.

observed when comparing the $\log_2 \sigma$ summary statistics and its plot. The mean value is close to zero (and even slightly negative), while the plot seems to indicate that positive values are dominant. To understand this paradoxical observation, first recall that, as explained in Section 3.2, the distortion measures are evaluated on a regular grid in the domain and are valid at the corresponding points in the codomain – in this case, the modern reference map and the old map, respectively. The regular grid on the reference map assures that the set of grid points is a representative sample, and the validity at the points in the deformed old map reminds us that the grid is distorted (see Figure 3b) before the values are interpolated and plotted, implying that areas with positive and negative values of $\log_2 \sigma$ are displayed larger and smaller, respectively, in the codomain. This explains why red is the dominant colour in Figure 3c.

In Figure 4, we compare our visualisation of the area scale factor to the scale isolines generated by MapAnalyst, for a varying radius of that method’s ‘local affine’ transformation. As the radius decreases, less spatial averaging is applied and, as expected, the isolines gradually become curvier and cover a wider range of values. In order to obtain area scaling information of high spatial detail, one is tempted to make the radius ever smaller. However, MapAnalyst’s approach does not allow for arbitrarily small radii, since it requires a minimal number of nearby ground control points for every ‘local affine’ transformation. For this example map, the radius of 5 km is approximately the lower bound, as preluded by the ‘blocky’ behaviour of some isolines in Figure 4c. Interestingly, one observes that with decreasing radius the isolines increasingly agree, both in shape and in value, with the area scaling computed using our method. For example: the difference between the largest and smallest scales increasingly grows, and in Figure 4c these scales attain the values of 1:90 000 and 1:255 000, relating by a factor 2.83 and implying $\log_2 \sigma = 1.50$, which indeed approaches the range of values obtained through differential distortion analysis. This convergent evolution illustrates how our approach can be seen as the analytical limit of MapAnalyst’s approach when the radius becomes infinitesimally small. Our approach thus has multiple clear advantages: it applies no form of spatial averaging, making it more accurate and of the highest possible spatial resolution, and it does not introduce undesirable extra parameters (such as the radius) or requirements (such as the minimal number of nearby ground control points). Finally, note how the isolines quickly clutter the image, making them less favourable than colour ramp visualisations.

In their paper, Jenny *et al.* (2007) conclude that this map “was probably compiled from at least two base maps”, upon observing the patterns of the displacement vectors and their isoline visualisation of scale (equivalent to Figure 4a). If multiple maps with different origins were indeed used, we would expect clearly different behaviour of the local distortion in well-distinguishable areas, or bands of high distortion occurring where

the base maps were patched. Our spatially more detailed analysis method produces distortion visualisations that do not reflect such behaviour, and in this way gives us more information that seems to refute the earlier hypothesis.

5. Conclusion

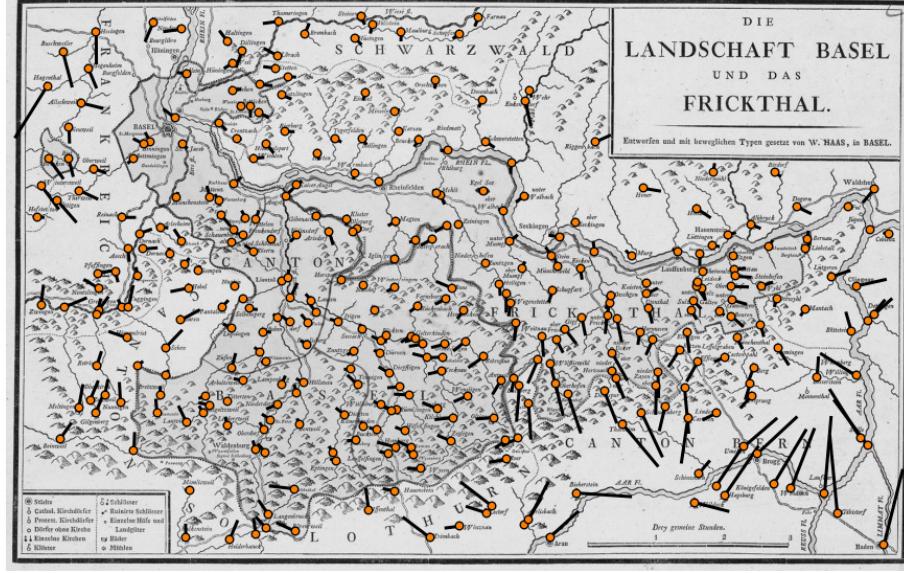
In this paper we discussed the study of distortion of old maps, and proposed a new technique to compute and visualise such distortion. We began by explaining how distortion analysis methods are structured: they extend the popular procedure of georeferencing, which aligns raster data (e.g. areal imagery or, in our case, images of old maps) with a modern reference map using a set of ground control points and an interpolation technique, by adding an extra step where distortions are computed. We discussed the various available options for every step in these methods. First, there are different settings in which the old map can be viewed: one can compare distances and angles map-to-map or globe-to-map, and old-to-modern or modern-to-old. Furthermore, a variety of interpolation techniques are available, including true interpolation techniques and regression techniques. Finally, there are different ways to compute and visualise distortion. In order to correctly embed our proposed technique, we selected the optimal options for a general-purpose study, which enables the development of our technique within a well-defined distortion analysis method.

Most current research uses the ‘MapAnalyst’ software by Jenny (Jenny and Hurni 2011), which computes distortion metrics, such as the area scaling, at a number of grid cells using a ‘local affine’ transformation. This focus on local distortion is an important part of the progress contained in Jenny’s work. Still, this distortion computation technique leaves room for improvement: it applies some form of spatial averaging, resulting in approximate values and coarser spatial resolution, and furthermore introduces undesirable extra parameters and requirements. In contrast, we proposed to compute distortion using differential distortion analysis, the dominant technique used in map projection theory. This analytical technique calculates local distortion metrics by direct analysis of the map projection function and its partial derivatives. It computes the exact distortion at every point without any spatial averaging. To enable the use of this technique we translated the mathematical framework of differential distortion analysis to the setting where an old and reference map can be directly compared (map-to-map), with the interpolation function taking the place of the map projection function. This resulted in formulas that can be readily used for the computation of the area scale factor, maximum angular distortion, Airy-Kavrayskiy distortion and Tissot indicatrices.

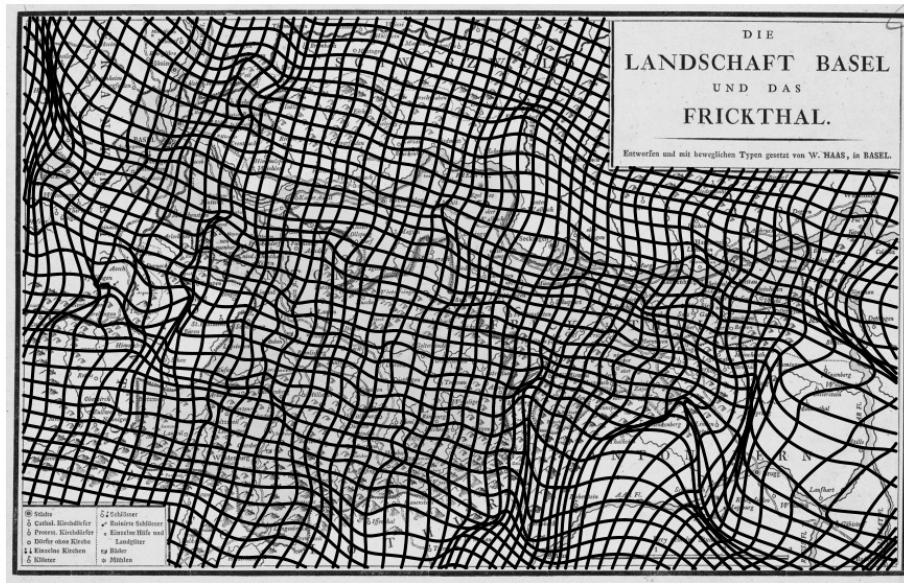
To conclude, we applied our general-purpose method, including the differential distortion analysis technique, to an example map also used in other literature, and discussed the resulting visualisations. Even though distortion has been studied in map projections using differential distortion analysis for multiple centuries, only a handful of authors seem to have drawn from this elegant technique for the study of old maps. The literature section mentions the only three cases known to the authors where similar ideas are mentioned or applied to some degree (Tobler 1976, 1994, Beineke 2007). The authors hope that the comprehensive discussion, readily applicable mathematical expressions and publicly available code of the present paper will allow this technique to find wide usage in future research on the distortion of old maps.

Acknowledgements

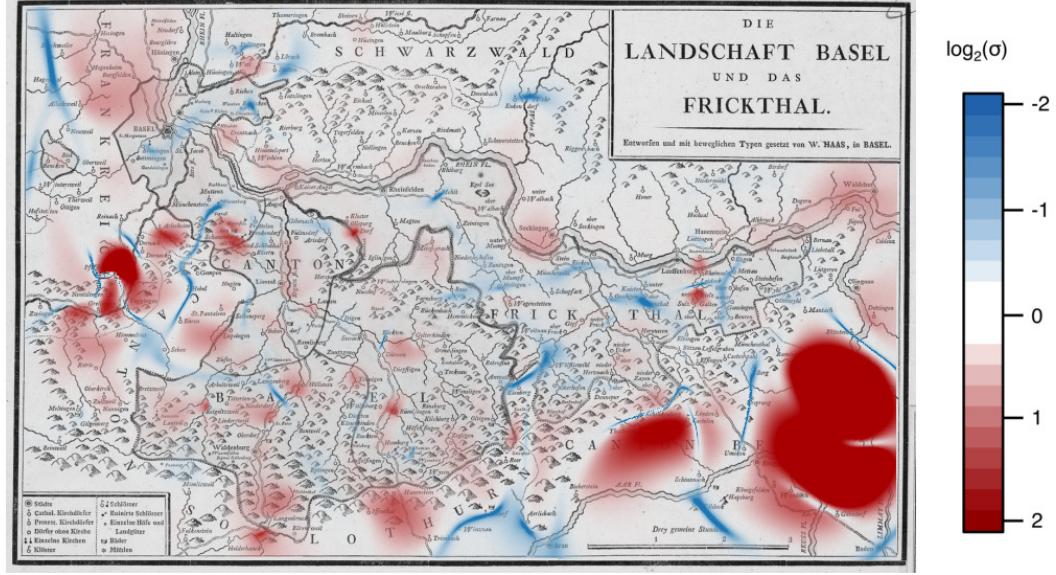
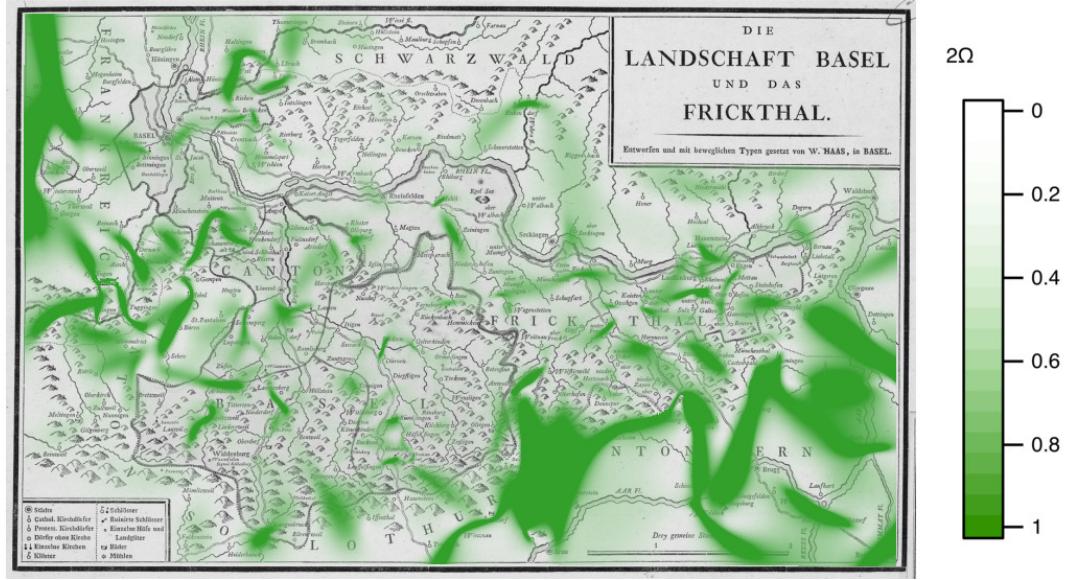
Manuel Claeys Boùùaert is a Fellow of the Fund for Scientific Research–Flanders (FWO–Vlaanderen). The author would like to thank Prof. F. Canters (VUB, Brussels) for his inspiring lectures on map projections. We also thank swisstopo and Martin Rickenbacher for allowing us to use the map of the Basel and Frickthal region and the ground control points, respectively.

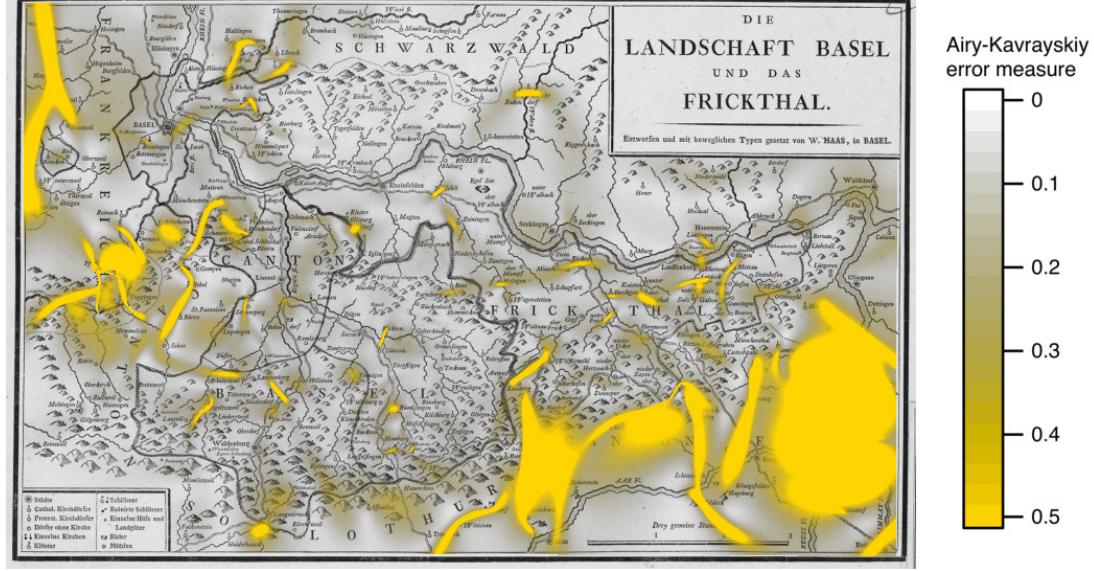
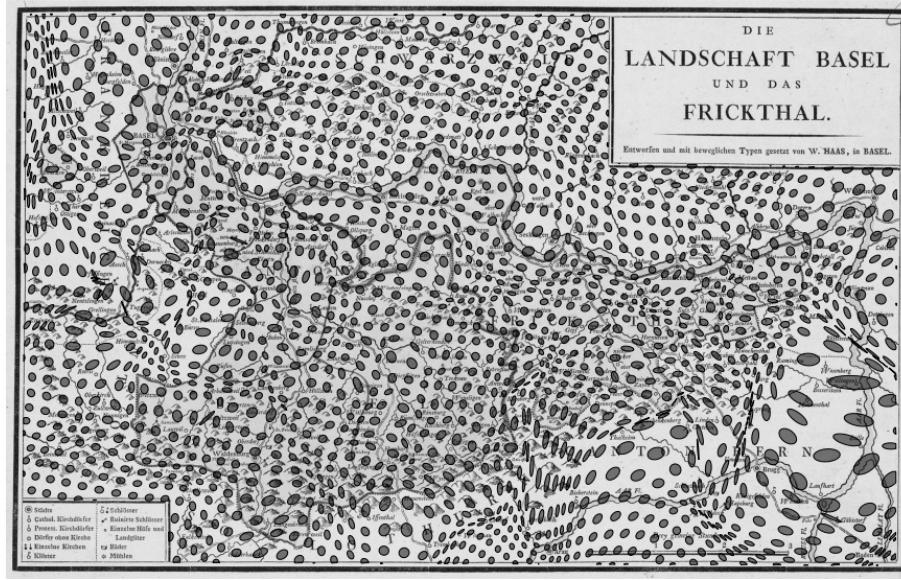


(a) Ground control points, displayed in orange, and displacement vectors, computed with respect to a Helmert transformation.



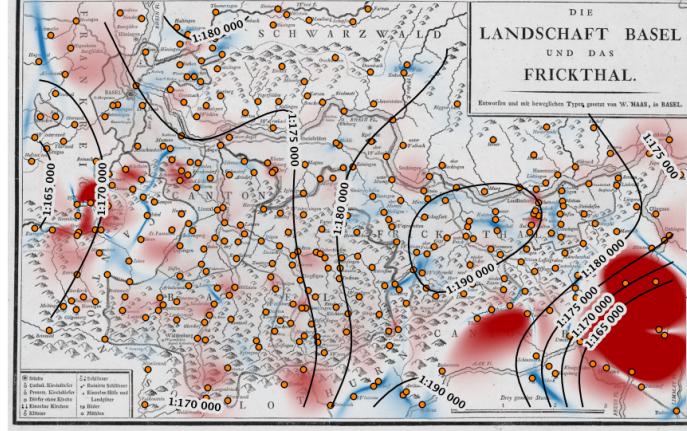
(b) The distortion grid, showing how the modern map must be warped in order to fit on the old map.

(c) Spatial variation of the base-2 logarithm of the area scale factor: $\log_2 \sigma$.(d) Spatial variation of the maximum angular distortion: 2Ω [rad].

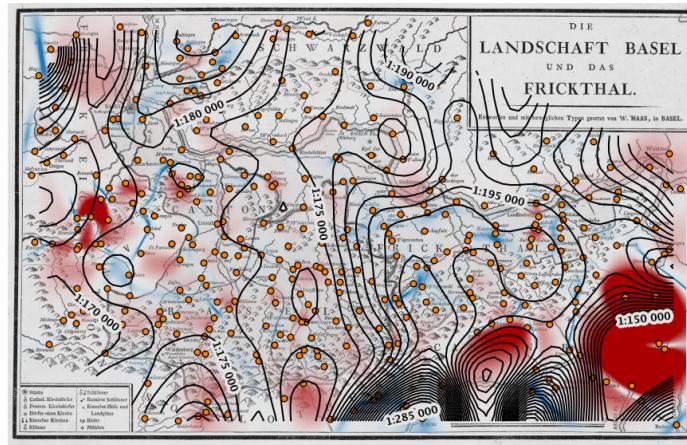
(e) Spatial variation of the Airy-Kavrayskiy measure of distortion: e_{AK} .

(f) Spatial variation of the Tissot indicatrices.

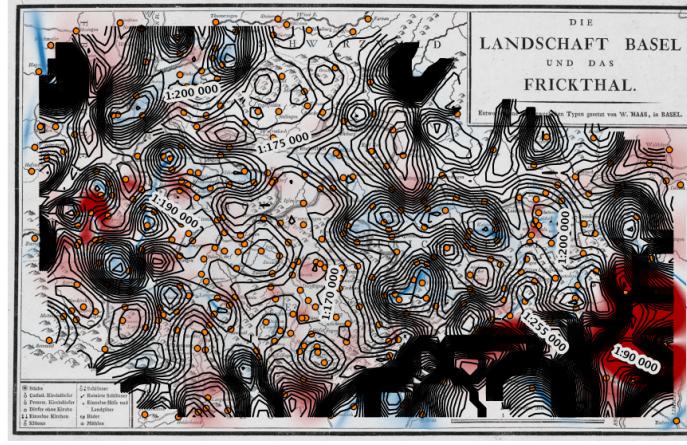
Figure 3.: Various techniques to compute and visualise distortion, applied to a 1798 map of the Basel and Frickthal region. The analysis is performed using the default method proposed in this paper, including the map-to-map and modern-to-old setting, a thin plate spline interpolation and the technique of differential distortion analysis. The latter is proposed and discussed in this paper, and is used to compute the area scale factor, the maximum angular distortion, the Airy-Kavrayskiy distortion measure and the Tissot indicatrices displayed here. The resulting distortion information is of higher accuracy and finer spatial resolution than that of previous studies (Jenny *et al.* 2007, Jenny and Hurni 2011).



(a) Local radius: 20 km.



(b) Local radius: 10 km.



(c) Local radius: 5 km.

Figure 4.: Scale isolines as generated by MapAnalyst (Jenny and Hurni 2011) for a varying radius of the ‘local affine’ transformation, displayed on top of our linear colour ramp visualisation of the area scale factor $\log_2 \sigma$ (as in Figure 3c). Isolines are plotted with an interval of 1:50 000 in all figures. As the radius decreases, the shape and values attained by the isolines converge to our visualisation.

References

- Ashburner, J. and Friston, K.J., 2000. Voxel-based morphometry—the methods. *Neuroimage*, 11 (6), 805–821.
- Battersby, S.E., Finn, M.P., Usery, E.L. and Yamamoto, K.H., 2014. Implications of web Mercator and its use in online mapping. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 49 (2), 85–101.
- Beineke, D., 2007. Zur Bestimmung lokaler Abbildungsverzerrungen in Altkarten mit Hilfe der multiquadratischen Interpolationsmethode. *Allgemeine Vermessungsnachrichten*, 114, 19–27.
- Blakemore, M.J. and Harley, J.B., 1980. Concepts in the history of cartography. A review and perspective. *Cartographica*, 17, 1–120.
- Bugayevskiy, L.M. and Snyder, J., 1995. *Map Projections: A Reference Manual*. CRC Press, London, Bristol.
- Buhmann, M.D., 2003. *Radial Basis Functions: Theory and Implementations*. Cambridge University Press, Cambridge.
- Canters, F., 2002. *Small-scale Map Projection Design*. CRC Press, Boca Raton.
- Canters, F. and Decleir, D., 1989. *The World in Perspective. A Directory of World Map Projections*. John Wiley & Sons Ltd, New York.
- Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74 (368), 829–836.
- Duchon, J., 1977. Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In: W. Schempp and K. Zeller, eds. *Constructive Theory of Functions of Several Variables*. Springer, Berlin, Heidelberg, 85–100.
- Fornefett, M., Rohr, K. and Stiehl, H.S., 2001. Radial basis functions with compact support for elastic registration of medical images. *Image and Vision Computing*, 19 (1–2), 87–96.
- Forstner, G. and Oehrli, M., 1998. Graphische Darstellungen der Untersuchungsergebnisse alter Karten und die Entwicklung der Verzerrungsgitter. *Cartographica Helvetica*, 17, 35–43.
- Franke, R., 1979. A critical comparison of some methods for interpolation of scattered data. Technical Report NPS 53-79-003. Naval Postgraduate School, Monterey, Calif.
- Grafarend, E.W. and Krumm F.W., 2006. *Map Projections*. Springer, Berlin, Heidelberg.
- Haas, W., 1798, *Die Landschaft Basel und das Frickthal*, Basel. The map data is courtesy of swisstopo, map collection, LT K BL 1798. The ground control points are courtesy of Rickenbacher, M., Bern.
- Hardy, R.L., 1971. Multiquadric equations of topography and other irregular surfaces. *Journal of Geophysical Research*, 76, 1905–1915.
- Hardy, R.L., 1990. Theory and applications of the multiquadric-biharmonic method: 20 years of discovery 1968–1988, *Computers & Mathematics with Applications*, 19, 163–208.
- Hill, L.L., 2009. *Georeferencing: The Geographic Associations of Information*. MIT Press, Cambridge.
- Hongchuan, Y. et al., 2012. An rbf-based reparameterization method for constrained texture mapping. *IEEE Transactions on Visualization and Computer Graphics*, 18 (7), 1115–1124.
- Jenny, B., Weber, A. and Hurni, L., 2007. Visualizing the planimetric accuracy of historical maps with MapAnalyst. *Cartographica*, 42, 89–94.
- Jenny, B. and Hurni, L., 2011. Studying cartographic heritage: Analysis and visualization

- of geometric distortions. *Computers & Graphics*, 35, 402–411.
- Jenny, B., 2015a. List of publications on techniques for visualisation of the accuracy of old maps and computation of statistical measures [online]. Available from: <http://mapanalyst.org/literature.html> [Accessed: 27 April 2015].
- Jenny, B., 2015b. List of publications on the accuracy of specific old maps that use MapAnalyst software [online]. Available from: <http://mapanalyst.org/press.html> [Accessed: 27 April 2015].
- Kavrayskiy, V.V., 1958. *Izbrannyye trudy, vol. II: Matematicheskaja kartografija*. Moscow Is. Nedra.
- Laskowski, P.H., 1989. The traditional and modern look at Tissot's indicatrix. *The American Cartographer*, 16 (2), 123–133.
- Seo, H. and Cordier, F., 2010. Constrained texture mapping using image warping. *Computer Graphics Forum*, 29 (1), 160–174.
- Swisstopo, 2015. *Swiss National Map 1:200 000*. [online] Maptiles are provided through geo.admin.ch [Accessed: 18 June 2015].
- Tiddeman, B., Duffy, N. and Rabey, G., 2001. A general method for overlap control in image warping. *Computers and Graphics*, 25 (1), 59–66.
- Tobler, W.R., 1976. The geometry of mental maps. In: R.G. Golledge and G. Rushton eds. *Spatial Choice and Spatial Behavior*. Ohio State University Press, Columbus, 69–81.
- Tobler, W.R., 1994. Bidimensional regression. *Geographical Analysis*, 26, 187–212.
- Wagner, H., 1896. *Das Ratsel der Kompasskarten im Lichte der Gesamtentwicklung der Seerkarten*. Geographische Verlagshandlung Deitrich Reimer, Berlin.
- Wahba, G., 1990. *Spline Models for Observational Data*. Siam, Philadelphia.
- Yang, Q., Snyder, J. and Tobler, W., 1999. *Map Projection Transformation: Principles and Applications*. CRC Press, Boca Raton.