Karl Solchenbach*

Comparing old maps with cartometric methods –
A case study with Bayesian data analysis

Summary:
The planimetric accuracy of old maps can be calculated using cartometric methods like Helmert transformations. As the choice of control points for these transformations is somewhat arbitrary, this may cause stochastic variations in the results. This paper addresses these stochastic variations and proposes to apply Bayesian data analysis methods to quantify the uncertainty of the cartometric calculations. The paper shows that the Bayesian data analysis is consistent with the deterministic calculations, and it provides safe statistical bounds for the results.

Introduction

Cartometric analysis of old maps is a standard tool in historical cartography research. Researchers use it to calculate the scale of an old map, to estimate its accuracy, or to detect dependencies between maps. Since many old maps were copied from previous maps, map historians use quantitative methods to detect how "errors" in maps were generated and propagated. Such methods help to understand how and when historical information, like the change of borders, was reflected in maps. Moreover, they can build map "genealogies", i.e. sequences of maps depending on each other, by using cartometric methods. This is valuable for historians who want to understand how certain historical developments or new geographical information are reflected in old maps for the first time.

"Errors" on old maps can have various sources (Blakemore und Harley, 1980). Three sources of error, in particular, can be distinguished. Firstly, topographical errors which refer to what is shown on the map, secondly geodetic errors related to the projection of the 3D earth surface to the 2D paper. Finally, planimetric errors describe how exact the position of an object is on the map. In this paper, we focus on the latter, ignoring the geodetic effects which are typically small on regional scales. Furthermore, we will not consider methodological errors e.g. how exact can we measure a position of a location, although we acknowledge that measurement errors may contribute to the variability of the data.

The most common method for comparing two maps is a bidimensional linear transformation: For this method control points (CPs) are defined, these can be town centres, bridges, river estuaries, etc. Then the positions of the CPs on both maps are measured. In a last step, the coordinates of the CPs on one map are transformed to the second map while minimizing the resulting "gaps" or residuals. The average residual is a measure for the "distance" of the two maps.

Typically, the residuals are not homogeneous: some CPs may align well, while other outlier CPs show large gaps. The average residual as a single number hides such inhomogeneities. Also, the value of the residual may depend on the selection of the CPs which adds a stochastic component to the process.

The goal of this paper is to investigate how Bayesian data analysis methods can be used to capture the stochastic effects inherent in the transformation between two maps. It is designed as a small case study with a focus on the methods, not on the results of the map comparisons. The paper describes an initial attempt to apply Bayesian data analysis to the field of cartometry of old maps,

---
* University of Luxemburg, Institute for History [karl.solchenbach@t-online.de]

it is not meant to be conclusive. Rather than calculating a single number with no statistical validation, Bayesian data analysis provides a distribution for the distance value with safe statistical bounds. Finally, this will increase the trust in cartometric calculations.

State-of-the-art

Cartometric methods are frequently used in accuracy analyses of old maps, for an overview see (Livieratos, 2006). Conformal transformations are the most commonly applied method, like the 4-parameter Helmert transformation. A good description of the various algorithms can be found in Beineke's dissertation (Beineke, 2001). In order to calculate the Helmert transformation and to visualize the results, many map historians use the public domain tool MapAnalyst (Jenny und Hurni, 2011). Depending on the nature and quality of data, outliers are possible and require special treatment, since they can "spoil" the results. Researchers are exploring different methods to control outliers by applying statistical approaches (Jongepier et al., 2016), alternative transformations based on kernel methods (Herrault et al., 2013), or evolutionary algorithms (Manzano-Agugliaro et al., 2013). To the author's knowledge, the use of Bayesian methods as evaluated in this paper has not been published in the literature so far.

Test cases

As test cases we look at two old maps of the region between the Meuse and Rhine rivers, including the Grand-Duchy of Luxemburg and the greater region, with parts of Belgium, France and Germany.
- The Luxemburg map of Gerhard Mercator, published in his atlas from 1585 on.
- The Luxemburg published in Ortelius' *Theatrum Orbis Terrarum* from 1579 on.

Both maps are well-known iconic examples for maps in the region (van der Vekene, 1980, 1.02 and 1.05), they are based on previous land surveys and have been copied many times.

Both maps will be compared to a modern map and as well among themselves using standard cartometric methods. The focus of the paper will be to understand and validate the statistical aspects of these methods.
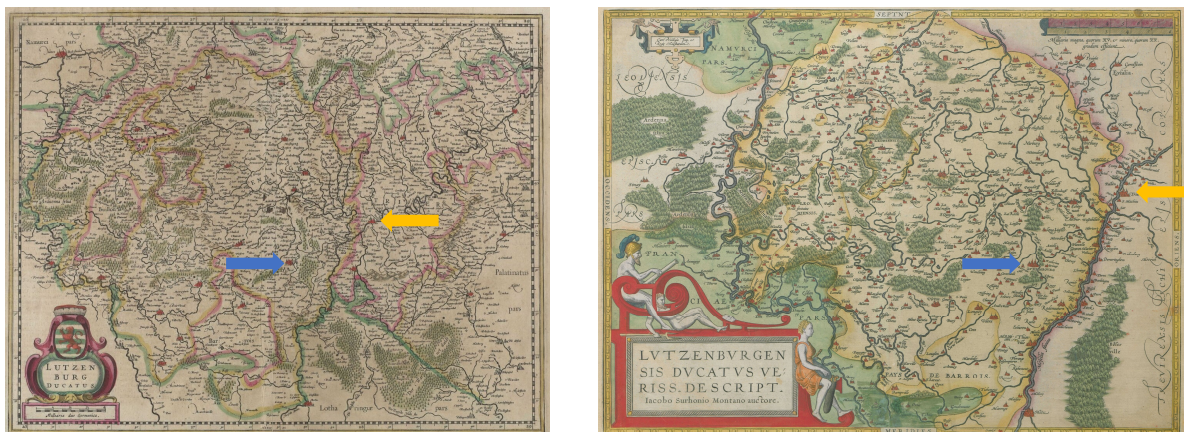

Figure 1: Mercator map of Luxemburg and Trier (left) and the Luxemburg map in Ortelius' Theatrum Orbis Terrarum. The cities of Trier and Luxemburg are marked by orange and blue arrows.

<div align="center">

Methods

The Helmert Transformation

</div>

The classical 4-parameter Helmert transformation consists of the following steps:

1. Define n topographic CPs (town centers, other landmarks) which are present on both maps and have not changed their geographic position over time.
2. Measure coordinates (x_i, y_i) and (X_i, Y_i) of the CPs on both maps $(i=1, ..., n)$. This can be done by a digitizer (unit: pixels), a ruler for paper maps (unit: cm), or direct reading of Gauss-Krueger or UTM coordinates (unit: km).
3. Determine 4 parameters $a_j, j=1, ..., 4$ for the transformation of the (x_i, y_i) coordinates to the (X_i, Y_i) coordinates. The parameters of this affine transformation describe the x- and y-offset, the scaling factor, and the rotation angle. In order to compute the best fit for these parameters an overdetermined system of equations has to be solved, this is described in more detail below.
4. Transform $(x_i, y_i) \rightarrow (\hat{x}_i, \hat{y}_i)$ using the parameters a_j
5. Calculate the residual vector r with $r_i = ((\hat{x}_i - X_i)^2 + (\hat{y}_i - Y_i)^2)^{1/2}$ as the local point errors. The mean value of r_i $(i=1, ..., n)$ is a measure for *distance* between the two maps.

The equations for this transformation in steps 2. and 3. are (following Beineke, 2001, p. 12–16)

$$X = X_0 + m\cdot\cos(\alpha)\cdot x \;-\; m\cdot\sin(\alpha)\cdot y$$
$$Y = Y_0 + m\cdot\cos(\alpha)\cdot y \;+\; m\cdot\sin(\alpha)\cdot x$$

x,y denote the coordinates for the CPs on the old map, X, Y the coordinates of the CPs on the target map (modern map or another old map). X_0 and Y_0 are the offsets in x- and y-direction, m is the relative scale between the maps, and α the angle between the orientation of the maps (old maps are not necessarily aligned to the north).

With $a_1 = X_0$, $a_2 = Y_0$, $a_3 = m\cdot\cos(\alpha)$, $a_4 = m\cdot\sin(\alpha)$ the equations become linear for the parameters a_j

$$X_i = a_1 + a_3 x_i - a_4 y_i \tag{1}$$
$$Y_i = a_2 + a_3 y_i + a_4 x_i$$

or in matrix notation with the Helmert matrix **H**, and the right hand side **l**.

<div align="center">

Ha = l

</div>

with

$$
\mathbf{l} = \begin{Bmatrix} X_1 \\ \vdots \\ X_n \\ Y_1 \\ \vdots \\ Y_n \end{Bmatrix}
\qquad
\mathbf{H} = \begin{Bmatrix} 1 & 0 & x_1 & -y_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_n & -y_n \\ 0 & 1 & y_1 & x_1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & y_n & x_n \end{Bmatrix}
\qquad
\mathbf{a} = \begin{Bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{Bmatrix}
$$

Since the system has *2n* equations for 4 unknowns it is overdetermined, and an approximate solution **â** minimizing the remaining residuals is calculated from

$$\mathbf{\hat{a} = (H^T H)^{-1} H^T \, l}$$

With the back substitution $X_0 = a_1$, $Y_0 = a_2$, $\alpha = \arctan(a_4 / a_3)$, $m = (a_3^2 + a_4^2)^{1/2}$ we can compute the offsets, the scale and the rotation angle between the two maps. Applying transformation (1) to the CPs (x_i, y_i) results in $(\hat{x}_i, \hat{y}_i)$, these points are close to the target CPs (X_i, Y_i). The remaining difference is the local point error, the arithmetic mean over all CPs is the *distance* between the two maps. The Helmert transformation was implemented in R, to have full control over the algorithms and the statistical effects. These transformations can also be performed with the program MapAnalyst.

The Mercator Map of the Duchy of Luxemburg and the Electorate of Trier

For the comparison of the Mercator map and the modern map we selected 55 CPs, mostly towns in Luxemburg, Southern Belgium, Eastern France and Western Germany. The point errors are given in the unit of the target map (here in km), the distance (i.e. the mean value of the point errors) is 5.77 km, see Table 1. For the comparison of the Ortelius map with the modern map 34 CPs were used, the distance is 7.67 km. The *relative distance* is the distance in relation to the longest side of the rectangle represented by the map. Both maps are facing north with a small deviation, the computed scales are consistent with data in the literature: 1:525.000 for the Mercator map (Hellwig, 1985) and 1:400.000 for the Ortelius map (van der Vekene, 1980).

Map 1	Map 2	Control points (CPs)	Distance		Angle	Scale
			km	relative		
Mercator	Modern	55	5.77	2.9%	-1,05°	1:482,000
Ortelius	Modern	34	7.67	5.1%	-4,95°	1:415,500

Table 1: Comparison of old vs modern maps

Figure 2 shows the point errors for each CP. Obviously, there is a large variation in the point errors between the CPs, for the Mercator map ranging from 0.2 km (at CP #4) to 12.9 km (at CP #11), for the Ortelius map ranging from 1.3 km (at CP#20) to 19.0 km (at CP#27).
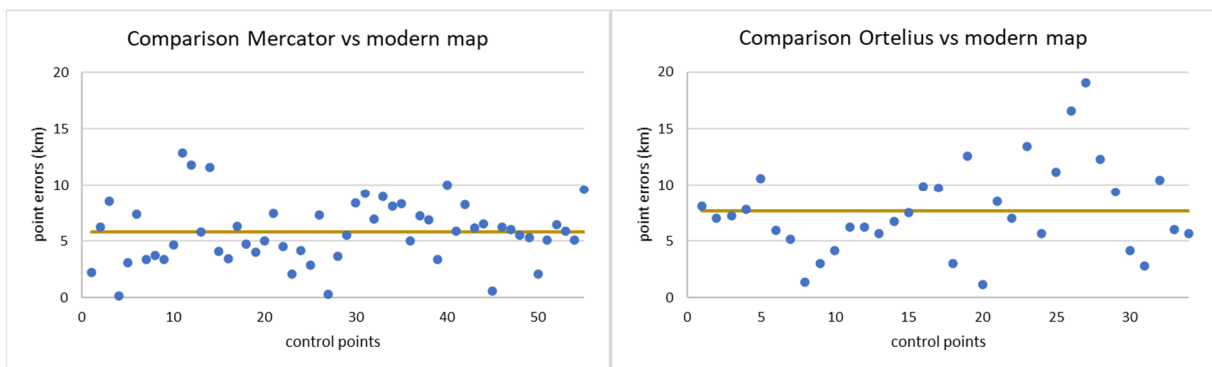

Figure 2: Distribution of point errors, (a) left for the Mercator map, (b) right for the Ortelius map

Given this large variation in the point errors, the question arises how sensitive the cartometric calculation reacts on the selection of the CPs. Had we chosen other CPs, would the distance have been different, and by how much?

In order to understand this better, we cyclically remove one CP from the set of points and run the Helmert transformation for the remaining *n-1* CPs.
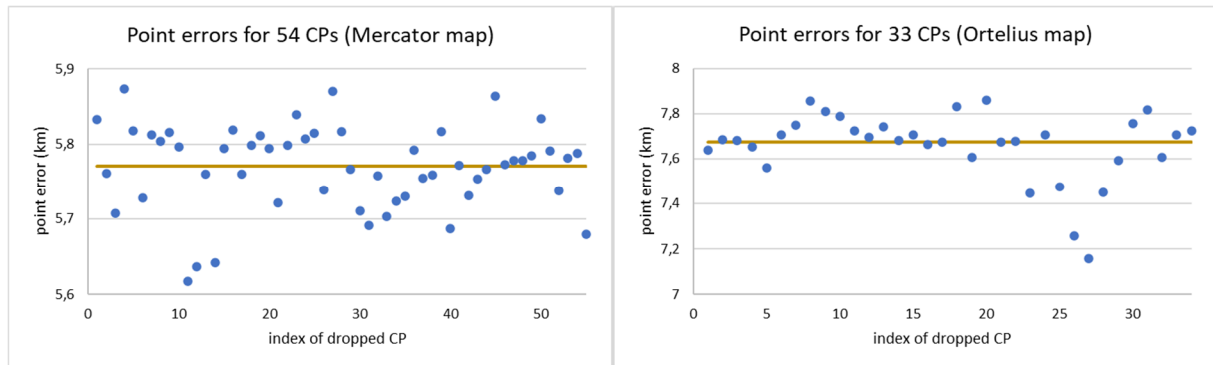

Figure 3: Variation of point errors if one CP is removed from the calculation

We observe a considerable variation of the calculated distance, depending which CP is removed. In particular, when the outlier CPs are not included in the analysis, the distance decreases from 5.77 to 5.62 km (Mercator map) and from 7.67 to 7.16 km (Ortelius map). Not unexpected, the variation of the distance measure is even larger when we remove 2 CPs.

Map 1	Map 2	Control points (CPs)	Distance in km when 1 CP is removed (circular over all CPs)			Distance in km when 2 CPs are removed (randomly over all CPs)		
			min	max	mean	min	max	mean
Mercator	Modern	55	5.62	5.87	5.77	5.48	5.98	5.77
Ortelius	Modern	34	7.16	7,86	7.66	6.70	8.06	7.64
Mercator	Ortelius	33	5.34	5.78	5.65	5.02	5.89	5.64

Table 2: Sensitivity of results when one or two CPs are removed from the calculation

For example, if we had selected only 32 CPs (instead of 34) in the Ortelius map the calculated distance to the modern map could have varied between 6.70 and 8.06 km. The direct comparison of the two old maps (last row) shows a similar interval of uncertainty (between 5.02 and 5.89 km). Although the Helmert transformation is a deterministic algorithm, the choice of CPs is somewhat arbitrary and the data in Table 2 indicate that the distance calculation is inherently stochastic and uncertain. Apparently, a different choice of a few CPs may lead to different results. This observation is the motivation to investigate the possibilities of Bayesian data analysis to capture the stochastic effects.

Bayesian data analysis

Bayesian inference is the process of fitting a probability model to a set of data and summarizing the result by a probability distribution on the parameters of the model and on unobserved quantities (Gelman et al., 2014). If the model parameter space is too large, the computation of the posterior distribution uses a simulation method, typically a Markov Chain Monte Carlo (MCMC) iteration algorithm. Details of Bayesian data analysis and the MCMC method are given in (Gelman et al., 2014) and (Kruschke, 2015).

For our analysis the Helmert transformation algorithm is mapped to a Bayesian model. Assuming a statistical distribution of the CP coordinates, Bayesian regression delivers a distribution for the parameters $a_1, ..., a_4$. These are then used to calculate the transformed coordinates $(\hat{x}_i, \hat{y}_i)$, also as distributions. Finally, we obtain the distance, the angle and the scale between the two maps as distributions, not as single numbers.

We make only weak assumptions about the initial distributions ("prior knowledge" in Bayesian terms), we let them follow a normal distribution, with a uniformly distributed standard deviation.

The model was implemented in R as Markov Chain Monte Carlo (MCMC) iteration process, using the JAGS package and a modified version of the module *Jags-Ymet-XmetMulti-Mrobust* (Kruschke, 2015, p. 509–525).

The result of the simulation is expressed as a 95% Highest Density Interval (HDI) which contains the most credible values of the posterior distribution of the distance, the angle and the scale. For both maps, the data calculated directly with the deterministic Helmert transformation lie in the interval and are close to the modes of the Bayesian statistical model (see Table 3).
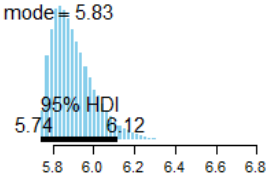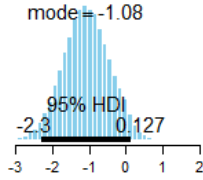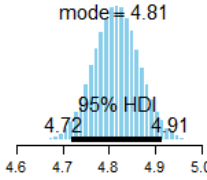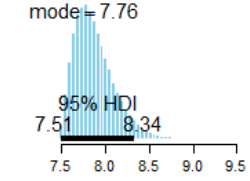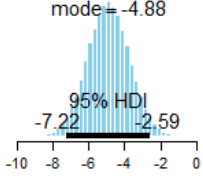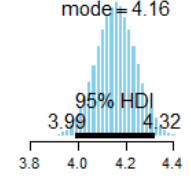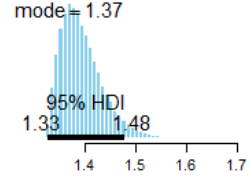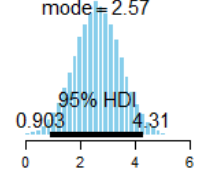
Map 1	Map 2	Algorithm	Distance (km)	Angle (°)	Scale (x 100,000)
Mercator	Modern	Deterministic	5.77	-1.05	4.82
		Bayesian	mode = 5.83; 95% HDI 5.74 – 6.12	mode = -1.08; 95% HDI -2.3 – 0.127	mode = 4.81; 95% HDI 4.72 – 4.91
Ortelius	Modern	Deterministic	7.67	-4.95	4.16
		Bayesian	mode = 7.76; 95% HDI 7.51 – 8.34	mode = -4.88; 95% HDI -7.22 – -2.59	mode = 4.16; 95% HDI 3.99 – 4.32
Mercator	Ortelius	Deterministic	1.36 cm ≙ 5.65 km	2.68	n/a
		Bayesian	mode = 1.37; 95% HDI 1.33 – 1.48	mode = 2.57; 95% HDI 0.903 – 4.31	

Table 3: Comparison of direct (deterministic) and Bayesian calculation of distance, angle and scale. The histograms of the Bayesian distributions show the minimum and maximum of the 95% HDI and the mode (most frequent value).

Conclusion

To compare maps with cartometric methods, the coordinates of control points are measured on both maps, followed by a 4-parameter Helmert transformation. The parameters of the transformation (*x*-offset, *y*-offset, angle and scale) are chosen in order to minimize the resulting mean residual (point errors) which defines what we call *distance* between the two maps. The results of this process may

depend significantly on the choice of control points. To capture the variability in the data, statistical approaches are required.

We developed a Bayesian model for the Helmert transformation and implemented it as a Markov Chain Monte Carlo (MCMC) iteration using the JAGS package. Two old maps (Mercator "Trier and Luxemburg" from 1585 and Ortelius "Luxemburg" from 1579) were selected as examples. They were compared with a modern map and between each other, using the usual deterministic calculation and the new statistical Bayesian model.

The results of the Bayesian analysis are consistent with the data obtained with the deterministic method. Rather than a single number, the Bayesian simulation provides safe (95%) statistical bounds for the distance, the angle and the scale.

The present study shows that the Bayesian analysis is useful to estimate and to control stochastic effects in cartometric map analysis and may contribute to future research.

References

Beineke Dieter, 2001. Verfahren zur Genauigkeitsanalyse von Altkarten. Schriftenreihe Studiengang Geodäsie und Geoinformation. München.

Blakemore M. J., Harley J. B., 1980. Concepts in the History of Cartography - A Review and Perspective: University of Toronto Press (Cartographica, Monograph 26).

Gelman Andrew, Carlin John B., Stern Hal S., Dunson David B., Vehtari Aki, Rubin Donald B., 2014. Bayesian data analysis. Third edition. Boca Raton, London, New York: CRC Press Taylor and Francis Group (Texts in statistical science series).

Hellwig Fritz, 1985. Mittelrhein und Moselland im Bild alter Karten. Katalog zur Ausstellung mit einem Beitrag zur Geschichte der älteren Kartographie vom Mittelrhein und Moselland. Koblenz: Landesarchivverwaltung Rheinland-Pfalz.

Herrault Pierre-Alexis, Sheeren David, Fauvel Mathieu, Monteil Claude, Paegelow Martin, 2013. A comparative study of geometric transformation models for the historical 'Map of France' registration. In: *Geographica Technica* 1, p. 34–46.

Jenny Bernhard, Hurni Lorenz, 2011. Studying cartographic heritage: Analysis and visualization of geometric distortions. In: *Computers & Graphics* 35, p. 402–411.

Jongepier Iason, Soens Tim, Temmerman Stijn, Missiaen Tine, 2016. Assessing the Planimetric Accuracy of Historical Maps (Sixteenth to Nineteenth Centuries): New Methods and Potential for Coastal Landscape Reconstruction. In: *The Cartographic Journal* 53 (2), p. 114–132.

Kruschke John K., 2015. Doing Bayesian data analysis. A tutorial with R, JAGS, and Stan. 2. ed. Amsterdam: AP Academic Press/Elsevier. Online availabe as http://www.contentreserve.com/TitleInfo.asp?ID={38F45CF6-6B5C-433C-85F8-A3568420927D}&Format=50.

Livieratos Evangelis, 2006. On the Study of the Geometric Properties of Historical Cartographic Representations. In: *Cartographica* 41 (2), p. 165–175.

Manzano-Agugliaro F., San-Antonio-Gomez C., Lopez S., Montoya F. G., Gil C., 2013. Pareto-based evolutionary algorithms for the calculation of transformation parameters and accuracy assessment of historical maps. In: *Computers & Geosciences* (57), p. 124–132.

van der Vekene Emile, 1980. Les cartes géographique du Duché Luxembourg. Éd. aux 16,17. et 18. siècles. Catalogue descriptif et ill. 2. éd., corr. et augm. Luxembourg: Krippler-Muller.