SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 10, No. 2, (2024), pp. 193-211
Open Access. Online & Print

DOI: 10.5281/zenodo.23435355

CERTIFIED INTEGER PROGRAMMING SEGMENTS PTOLEMY'S GEOGRAPHY INTO SIMILARITY UNITS THAT PREDICT SEALED PLACES LESS WELL THAN A SMOOTH DISTORTION FIELD

Luis Eduardo Muñoz Guerrero[1*]

[1]Universidad Tecnológica de Pereira, Pereira, Colombia; Email: lemunozg@utp.edu.co; Orcid ID: 0000-0002-9414-6187

Received: 28/03/2024
Accepted: 19/05/2024

Corresponding Author: Luis Eduardo Muñoz Guerrero (lemunozg@utp.edu.co)

ABSTRACT

The distortion of Ptolemy's coordinates has been explained by transformation units, groups of places sharing a scale and a shift, and by gradual distortion across a map. This study tests whether blocks with their own similarity transformation predict held-out places better than a smooth field. The catalogue was linked to the Pleiades gazetteer without Ptolemy's coordinates, and half the entries were sealed behind a hash commitment and a frozen plan. Segmenting a map into similarity units, a strongly NP-hard minimum-cost cover with outliers, was solved by column generation and integer programming with a certified interval for the optimum. Units, a moving-least-squares field and one frame per map were fitted on the exploration half of 16 maps and compared on sealed places by an exact randomization test. On 897 sealed places, the smooth field predicted better (pooled median difference −6.1 km; 95% interval −10.3 to −3.3; $p = 0.006$), and units beat one frame per map by 16.9 km. Optimality was certified in 8 maps. For identifiable places, a smooth field describes the distortion better than similarity units, by a margin close to the rounding of the coordinates; the published units, with shared scales, were not tested.

KEYWORDS: History Of Cartography, Historical Geodesy, Record Linkage, Column Generation, NP-Hardness, Robust Multi-Model Fitting, Randomization Inference, Pleiades Gazetteer.

Copyright: © 2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License. (https://cre-ativecommons.org/licenses/by/4.0/).

1. INTRODUCTION

This study asks whether the distortion of the coordinates in Ptolemy's Geography is better described by a few blocks, each with its own similarity transformation (one scale, rotation and shift), here called similarity units, or by a deformation that changes gradually from place to place. The question is answered by prediction: both descriptions are fitted on half of the linked places of each regional map and judged on the other half, which stayed sealed until the analysis plan was frozen. A second objective is methodological: the block description is turned into an explicit optimization problem, and every segmentation used in the comparison comes with a certified interval for its optimal cost.

The geography gives longitudes and latitudes for thousands of places of the known world, and the digital catalogue used here covers the two main textual recensions of the work, Ξ and Ω (Graßhoff and Defaux, 2019), which often give different numbers for the same place. Its errors are large and uneven, and two families of explanations have been proposed. The geodetic group of the Technical University of Berlin found groups of places with homogeneous scales and shifts, which they called transformation units, by a multistage combinatorial search for sets of places with homogeneous distortion (Marx, 2012). In their adjustment model there are joint scale parameters for an investigation area and a shift for each group of places (Marx, 2012), and the method was applied chapter by chapter, with the regions of the text as investigation areas (Marx, 2012). Applied to Germania and the neighbouring provinces, it divided Germania Magna into twelve areas of homogeneous displacement, Gallia Belgica into eleven and Raetia and Noricum into three each (Kleineberg et al., 2011). For Ptolemy's Scotland the same group added a rotation, modelling the rotation of Scotland to the east in three dimensions (Marx, 2014). Reviewers noted that when manuscripts disagreed the authors adopted the variant which produced a coherent result (Burri, 2014), and that no calculations were published in the book that presents the rectified coordinates of Germania (Pápay, 2011).

Other studies describe the distortion as a continuous field. Tsorlini (2009) fitted a second-order polynomial transformation, involving a projection and a model of the earth, to the identified toponyms of the northern Mediterranean, and Jenny and Hurni (2011) contrast a global Helmert transformation with multiquadric interpolation for studying old maps. Out-of-sample checks exist for interpolation: Abshire et al. (2016) ran leave-one-out validation on their two strongest models for India, triangulation and flocking. Historical arguments bear on the choice as well. Marx (2012) argues that, since the ancient positions have local shifts in groups, the transformation function cannot be continuous. Shcheglov (2019) describes Ptolemy's space as discrete and hierarchically organized, with most places probably located relative to a few distant reference points, and argues that the displacements cannot be adequately described by a single continuous function, while Tupikova et al. (2014) conclude that places within one region must always be linked to at least one reference point. The clusters of displacement vectors studied by Graßhoff et al. (2017) show that the coordinates were not obtained from astronomical observations at each place. In the works reviewed for this study, however, a segmentation into units and a smooth field have not been compared with each other on places that neither had seen.

The comparison is only fair if the block model is as good as it can be. Previous segmentations were obtained by a multistage combinatorial search followed by manual improvement (Marx, 2012), so a poor predictive result could always be blamed on a poor segmentation. Here the segmentation is a well-defined problem: choose similarity transformations and an assignment of places that minimise a weighted count of units and outliers. The problem is a robust multi-model fit with a cost per model, equivalent to uncapacitated facility location with a continuum of candidate facilities (Delong et al., 2012), and deciding it exactly is strongly NP-hard, by reduction from covering points by lines, which is strongly NP-hard (Megiddo and Tamir, 1982). The algorithm used here therefore does not promise the optimum; it returns a solution together with a lower bound, so that the distance to the optimum is known for every map.

The contributions are four. First, the segmentation problem, its hardness and a bicriteria bound, which proved loose in practice (Section 2). Second, an algorithm that combines pricing by branch and bound, column generation and a covering integer program and certifies an interval for the optimum (Section 2). Third, a sealed comparison of units, a smooth field and one frame per map, with an exact randomization test and validity checks fixed in advance (Section 2). Fourth, a reproducible pipeline in which the sealed half is read only through a logged function that checks the frozen plan (Figure 1).

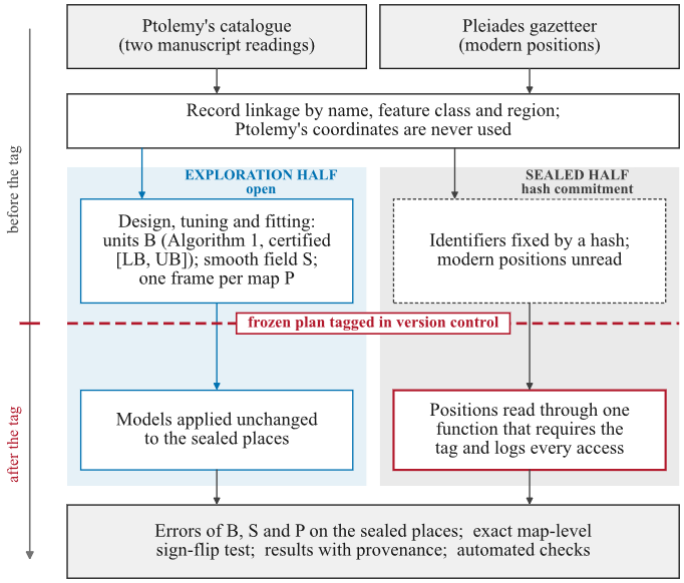


Figure 1: Pipeline Of the Study, Ordered in Time from Top to Bottom. The Identifiers of the Sealed Half Are Fixed by a Hash Commitment, And the Pipeline Reads Their Modern Positions Only Through a Function That Requires the Version-Control Tag of the Frozen Plan and Logs Every Access. The Dashed Line Marks That Tag and the Red Box That Function, A Convention of the Software Rather Than a Cryptographic Guarantee (Section 2).

2. METHODS

Data and record linkage. The catalogue has 10,049 entries, of which 6,338 carry coordinates in the Ω recension and 4,440 in Ξ; the two readings differ for 1,373 entries. Point features with coordinates (settlements, capes, ports and similar) number 5,168. Modern positions come from a numbered release of the Pleiades gazetteer of ancient places (Elliott, 2023), with 32,822 located places.

Linking catalogue entries to gazetteer places is a problem of record linkage (Fellegi and Sunter, 1969); here it is solved with deterministic rules. Only point features (settlements, ports, sanctuaries and capes) are linked, not rivers, mountains or river mouths, each to at most one Pleiades place that has coordinates, a compatible type and at least one reference other than Ptolemy. Three rules produce links: a seed is an exact match of a name root that is unique on both sides; a structural link joins an entry to a Pleiades place that cites the same section of Ptolemy's text and has a similar name; and the remaining entries are matched by the similarity of their names to places inside the modern extent of the map and near seeds of the same chapter. Ptolemy's coordinates never enter these rules, so the positions that are to be predicted cannot leak into the matching; a test replaces them with noise and obtains the same links. The linkage produced 1,854 links (890 seeds, 867 by name and region and 97 structural). A stratified sample of 120 links was labelled twice, against the same written criteria and blind to the other labelling. Counting doubtful links as errors, the precision of the more conservative of the two labels was 0.944 (95% bootstrap interval 0.897 to 0.981), and the two labellings agreed with Cohen's kappa 0.69.

Half of the 10,049 entries, stratified by map, were sealed with a fixed seed before any model was fitted. The identifiers of the sealed half are fixed by the commitment of Equation (1)

$$c = \text{SHA-256}\big(i_{(1)} \,\|\, i_{(2)} \,\|\, \cdots \,\|\, i_{(N)}\big), \qquad (1)$$

where i(1),…,i(N) are the identifiers of the sealed entries in sorted order, each followed by a line break, and c is published in the Declarations. The only function of the pipeline that returns their modern positions refuses to run unless the version-control tag of the frozen plan is an ancestor of the running commit and the working tree is clean; each call is logged. The seal is a convention of the software, not a cryptographic guarantee: the modern positions are public in Pleiades, the full linkage is kept by the author, the commitment was not timestamped by a third party, and the samples drawn to verify the linkage before the seal showed the modern positions of some entries that later fell in the sealed half. The exploration half kept 929 links and the sealed half 925. A map enters the analysis if it has at least 15

exploration links; 16 maps qualify, with 890 exploration places and 897 sealed places; the other 28 sealed links lie in maps that do not qualify.

Maps keep the codes of the catalogue: EU, AF and AS for the maps of Europe, Africa and Asia, followed by the number of the map in Ptolemy's order; the regions they cover are listed in the Supplementary Material.

Local planes. Each map is analysed in two local planes. In the Ptolemaic plane a reading r of place i, with Ptolemy's longitude lon and latitude φ, becomes the complex number of Equation (2)

$$q_{ir} = (\mathrm{lon} - \mathrm{lon}_0)\cos\varphi_0 + i\,(\varphi - \varphi_0), \tag{2}$$

in degrees, with lon0 and φ0 the medians of the map's exploration places. The Pleiades position yi is projected with the spherical azimuthal equidistant projection centred at the median exploration position, in kilometres. A similarity is the map of Equation (3)

$$T(q) = mq + t, \quad m = a + ib = s\,e^{i\omega}, \tag{3}$$

with scale s=|m| in kilometres per Ptolemaic degree, rotation ω=argm and shift t. The least-squares similarity transformation parameters for a set of pairs have a closed-form solution (Horn, 1987; Umeyama, 1991). The primary analysis uses the Ω reading where it exists and Ξ otherwise, the default of the Berlin group, which used the Ω reading by default (Marx, 2012); the multiverse also reverses that order. Table I lists the symbols used below.

Table I: Notation Used in the Methods, With the Unit of Each Quantity.

Symbol	Meaning	Unit
□□□, □□	Ptolemaic coordinates of reading r of place i, and its modern position	degrees; km
□, □, □	similarity □(□) = □□ + □, with complex factor m and shift t	–
□, □	scale |□| and rotation arg □ of a similarity	km per degree; degrees
ε, θ	tolerance of a unit; generic tolerance of a consensus set	km
λ	cost of one unit, measured in outliers	outliers
Θ	box of admissible similarities	–
□(□, □)	consensus set of T at tolerance θ	places
K, ℓ(□)	number of units; unit of place i, zero for an outlier	–
□□, □□	binary choice of a consensus set; outlier indicator of place i	–
□□, □(□)	price of place i; value of the best consensus set at prices π	–
LB, UB, g	lower and upper bound of the optimum and their relative gap	cost; share
□□,□	great-circle error of model M at sealed place i	km
□□, □̂1	paired difference of errors; its pooled median	km
G, s	number of maps; vector of one sign per map	–
δ	relevance threshold of the comparison	km
α	level of the tests; in Algorithm 1, weight of the stabilised duals	–

The segmentation problem. The consensus set of a similarity T with tolerance θ is the set of places that T carries to within θ of their modern positions, as in Equation (4)

$$S(T,\theta) = \left\{ i : \min_r \|y_i - T(q_{ir})\|_\infty \le \theta \right\}, \tag{4}$$

where the norm is the larger of the absolute east and north residuals. The segmentation problem with tolerance ε and unit cost λ chooses K similarities from a box Θ and labels each place with a unit or as an outlier, so that every labelled place lies within ε of its unit, minimising the cost in Equation (5)

$$f = \lambda K + |\{\, i : \ell(i) = 0 \,\}|, \tag{5}$$

where l(i)=0 marks an outlier. The unit cost λ is measured in outliers: a unit pays off only if it absorbs more than λ places that would otherwise be outliers, and a unit with exactly λ members costs as much as leaving them unexplained, so the optimum need not be unique; Figure 2 draws a schematic instance. The box Θ bounds |a| and |b| by 3 times the scale of the map's least-squares similarity and admits any shift that such a similarity can need. Equivalently, with a binary variable xS for each feasible consensus set and oi for each outlier, it is the covering program of Equation (6)

$$\begin{aligned} &\min\ \lambda \sum_S x_S + \sum_i o_i \\ &\text{s.t.} \sum_{S \ni i} x_S + o_i \ge 1, \quad x_S, o_i \in \{0,1\}. \end{aligned} \tag{6}$$

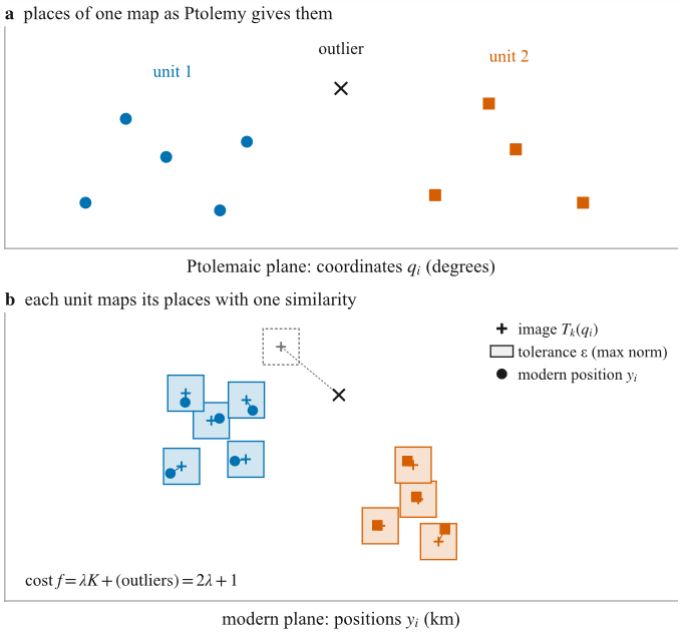


Figure 2: Schematic Instance of the Segmentation Problem, With Invented Coordinates. (A) Places Of One Map in the Ptolemaic Plane. (B) Each Unit Maps Its Places with One Similarity Tk (Plus Signs); A Place Belongs to the Unit When Its Modern Position (Dot or Square) Lies in the Square of Half-Side E Around Its Image, The Tolerance of the Maximum Norm. The Place Marked with A Cross Fit Neither Unit and Is an Outlier, So This Segmentation Costs Λk + 1 With K = 2 (Equation (5)).

Deciding whether an instance of the segmentation problem admits K units and a given number of outliers is NP-complete in the strong sense, even with ε = 0 and one reading per place: a rotation with rational sine and cosine maps an instance of covering points by lines, which is strongly NP-hard (Megiddo and Tamir, 1982), to a segmentation instance in which a unit with two distinct points is exactly a non-vertical line (the proof is in the Supplementary Material). Maximum consensus with a fixed number of models can be solved in polynomial time for a fixed number of models (Ask et al., 2013), but the exponent grows with K, and maximum consensus is APX-hard when the dimension is part of the input (Chin et al., 2018).

Two bounds make the solution checkable. If a unit is feasible with tolerance ε, the similarity that maps a diametral pair of its members exactly onto their positions keeps every member within (2+6) ε; the Supplementary Material gives the proof and a slightly tighter constant. Hence, with ALG(θ) the optimum of the covering program over the consensus sets of all pairwise and single-point candidates at tolerance θ, keeping only candidates in Θ for ALG(ε) and all candidates, solved to optimality, for ALG(cε), the bounds of Equation (7) hold with c=2+6

$$\mathrm{ALG}(c\varepsilon) \le \mathrm{OPT}(\varepsilon) \le \mathrm{ALG}(\varepsilon). \qquad (7)$$

The linear relaxation of Equation (6) has the dual program of Equation (8), with one price πi per place and S ranging over the feasible consensus sets

$$\begin{aligned} \max \ & \sum_i \pi_i \\ \text{s.t.} \ & \sum_{i \in S} \pi_i \le \lambda \quad \forall S, \\ & 0 \le \pi_i \le 1. \end{aligned} \qquad (8)$$

A unit improves the relaxation when its places are worth more than its cost λ. The pricing problem of Equation (9) finds the most valuable consensus set that a similarity in Θ can produce at tolerance ε

$$v(\pi) = \max_{T \in \Theta} \sum_{i \in S(T,\varepsilon)} \pi_i, \qquad (9)$$

and a column enters when v(π)>λ. The second bound adapts Farley's lower bound for linear programs with many columns (Farley, 1990) to the covering program; in the form proved in the Supplementary Material, for any vector π in [0,1]n, feasible for the dual or not, and any U≥v(π), the bound of Equation (10) holds

$$\mathrm{LB}(\pi, U) = \min\left(1, \frac{\lambda}{U}\right) \sum_i \pi_i \leq \mathrm{OPT}(\varepsilon), \quad (10)$$

and it can be rounded up when λ is an integer, because every feasible value is then an integer.

Algorithm 1. Algorithm 1 enumerates pairwise and single-point candidates, removes duplicate and dominated consensus sets, and runs column generation on the linear relaxation of the covering program: the relaxation gives every place a dual price, and each iteration searches for the unit whose places are worth more than its cost λ (the pricing problem). A single-point candidate uses the scale and rotation of the map's least-squares similarity, translated onto its point. Pricing is a best-first search over boxes of (a, b), because for fixed (a, b) the best shift lies at a corner of the residual boxes; within its budget of evaluations, it returns the best consensus set, and otherwise a valid upper bound U on its value. To stabilise the duals, pricing uses a weighted average of the current duals and those of the best Farley bound (Wentges, 1997), with weight 0.7. The final integer program over all generated columns is solved with HiGHS, whose dual simplex implementation is described by Huangfu and Hall (2018), and gives the upper bound; the lower bound is the larger of the rounded Farley bound and ALG(cε). The certificate is summarised by the relative gap of Equation (11)

$$g = \frac{\mathrm{UB} - \mathrm{LB}}{\mathrm{UB}}, \quad (11)$$

which is zero when the optimum is proved. With the plan's budget of 40 iterations, 25 columns per iteration and at most 150,000 evaluations per pricing call, the procedure is deterministic for the library versions of the lockfile as long as the final integer program ends before its time limit of 600 s, which it did in every map (Section 3). Seven invariants (feasibility of every column, exactness of the pair candidates, validity of every bound, and the lower bound never above the optimum on small instances solved by enumeration and by a mixed-integer program) are checked by property-based tests.

Algorithm 1 Certified similarity units	
1:	**Input**: readings q_{ir}, positions y_i, box Θ, tolerance ε, unit cost λ
2:	**Output**: units, outliers and a certificate [LB, UB]
3:	□ ← consensus sets of pair and anchor candidates at tolerance ε
4:	remove duplicates and sets contained in another set
5:	**repeat**
6:	solve the restricted linear master problem; duals π_{LP}
7:	$\pi \leftarrow \alpha\hat{\pi} + (1-\alpha)\,\pi_{LP}$ (stabilised duals)
8:	price with π; update the Farley bound and its duals $\hat{\pi}$
9:	add columns with $\sum_{i\in S} \pi_{LP,i} > \lambda$ to □
10:	**until** no column improves or the iteration budget is spent
11:	UB ← optimum of the covering program over □
12:	LB ← max([Farley], ALG(cε))
13:	**return** the units of the upper bound, refitted by least squares, and [LB, UB]

Models compared. All models are fitted on the exploration places of a map and predict the sealed places of the same map from their Ptolemaic coordinates; predictions return to latitude and longitude through the inverse projection.

Similarity units (B). Algorithm 1 is run with ε = 15 km and λ = 3, values chosen by five-fold cross-validation on the exploration half before the plan was frozen. The cross-validation used a fast variant of Algorithm 1 (pairwise candidates with refinement, local search and the integer program, without column generation or certificate), which is also the variant of B in the calibration and the multiverse, because column generation takes minutes in the largest maps. The chosen tolerance was the smallest of the grid, and with λ = 3 the cross-validated median error of the units grew with ε: 49.24, 51.66 and 53.58 km for ε = 15, 20 and 30 km. Each unit k is refitted by least squares on its members, giving a similarity Tk, and a place with Ptolemaic coordinates q takes the unit of its nearest non-outlier exploration place, as in Equation (12)

$$\hat{y}_B(q) = T_{\ell(j^*)}(q), \quad j^* = \underset{j \in E,\, \ell(j) \neq 0}{\arg\min} |q - q_j|, \tag{12}$$

where E is the set of exploration places of the map and l(j) the unit of place j. Exact ties between equidistant neighbours, which occur when places of different units share their Ptolemaic coordinates, are resolved by the search order of a k-d tree, a spatial search structure, built on those places in identifier order.

Smooth field (S). A moving least squares fit of similarities (Schaefer et al., 2006) predicts each place with its own similarity, fitted to all exploration places with weights that decay with the distance to the place, as in Equation (13)

$$\hat{y}_S(q) = T_q(q), \quad T_q = \arg\min_T \sum_{j \in E} w_j(q) \times |y_j - T(q_j)|^2, \quad w_j(q) = e^{-|q-q_j|^2/h^2}, \tag{13}$$

with h = 2 times the median nearest-neighbour distance of the map. The weights are not normalised: when the weighted spread of the exploration places around a query fall below a fixed absolute threshold, which happens only far from every exploration place, the local fit keeps only the translation, both in prediction and in the leave-one-out step (three sealed and three exploration places, all in one map). Places whose leave-one-out residual exceeds the median plus 3 robust standard deviations are excluded from the fit.

One frame per map (P). A single least-squares similarity per map, as in Equation (14)

$$\hat{y}_P(q) = T_P(q), \quad T_P = \arg\min_T \sum_{j \in E'} |\, y_j - T(q_j)|^2, \tag{14}$$

is iteratively refitted on the set E' of exploration places whose residual does not exceed the same threshold, computed each time from the residuals of all exploration places; it is the regional frame that both richer models should beat. A thin-plate spline (Bookstein, 1989) (T) enters only the multiverse and the exploratory analyses.

The kernel of S, its translation-only guard, the trimming rule of P, the rule for ties in B and the anchors of ALG(cε) were fixed by the frozen code but not by the text of the plan; Amendment 1, written after the confirmatory result was known, records them.

Outcome, estimand and test. The error eM,i of model M at sealed place i is the great-circle distance between its predicted latitude and longitude (φ,lon) and the Pleiades position (φ,lon), in Equation (15)

$$e_{M,i} = 2R \arcsin\sqrt{\eta}, \quad \eta = \sin^2\frac{\hat{\varphi} - \varphi}{2} + \cos\hat{\varphi}\,\cos\varphi\,\sin^2\frac{\widehat{\text{lon}} - \text{lon}}{2}, \tag{15}$$

with R = 6,371 km. The primary contrast is defined in Equation (16)

$$d_i = e_{S,i} - e_{B,i}, \quad \hat{\theta}_1 = \text{median}_i\, d_i, \tag{16}$$

so that θ1<0 favours the smooth field (H1, two-sided). H2 uses di=eP,i−eB,i and asks whether units beat one frame per map (one-sided). Places of one map share their models, so the map is the unit of inference. Under the null hypothesis the vector of differences of each map is symmetric about zero, a stronger condition than a zero median, and the reference distribution flips the sign of all differences of a map at once, as in Equation (17)

$$\theta^*(s) = \text{median}_i\; s_{g(i)}\, d_i, \quad s \in \{-1,1\}^G, \tag{17}$$

over all 65,536 sign vectors of the 16 maps, without random numbers. The two-sided p-value is the share of sign vectors at least as extreme as the observed estimate, as in Equation (18)

$$p(\theta_0) = 2^{-G} \sum_{s\in\{-1,1\}^G} \mathbf{1}\{|\theta^*_{\theta_0}(s)| \geq |\hat{\theta}_1 - \theta_0|\}, \quad (18)$$

where θθ0∗(s) is Equation (17) applied to di−θ0 and the reported p-value is p(0); the one-sided version of H2 counts only the upper tail. Intervals invert the test, as in Equation (19)

$$\mathrm{CI}_{1-\alpha} = \{\theta_0 : p(\theta_0) > \alpha\}, \quad (19)$$

which assumes that the differences of every map are symmetric about the same shift. The pooled differences are skewed (mean −14.44 km against a median of −6.11 km), so the interval is best read as a statement about a location shift under that assumption. H2 is tested one-sided and reported with the same two-sided interval. The two hypotheses form one family corrected with Holm's sequentially rejective multiple test procedure (Holm, 1979). Asymptotic cluster-robust tests can over-reject when there are few clusters (Cameron et al., 2008), and the plan rejected a cluster bootstrap of the median on similar grounds.

Rerun after the result with fixed seeds, in 500 null simulations with the real map sizes, the percentile bootstrap rejected a true null in 8.6% of them at the 5% level and the sign-flip test in 3.6%; when the median was zero but the differences were as skewed as the observed ones, the sign-flip test rejected in 4.8% (Supplementary Material).

The relevance threshold is δ = 5 km, half the finest step of the coordinates, and equivalence is judged with two one-sided tests (Lakens, 2017), which declare the estimate smaller than δ when the interval of Equation (19) at the 90% level satisfies Equation (20)

$$\mathrm{CI}_{1-2\alpha} \subset (-\delta, \delta). \quad (20)$$

Validity checks, robustness and secondary analyses. The plan required four checks before a positive result could be called valid.

(V1) Calibration with known structure: synthetic positions were generated for all linked places, once from the smooth field and once from the units fitted to the exploration data, each with the noise that its structure leaves in the real data; the text of the plan says the same noise for both, the frozen code follows a decision recorded before the plan was frozen, and Amendment 2 corrects the text. Each scenario had 5.6% gross errors, and the comparison was repeated 40 times per scenario with the fast variant of B; with an observed advantage of the smooth field, the rate of significant 'smooth better' conclusions when the truth is blocks had to stay at or below 10%. (V2) No leakage: the linkage and the seal were tested as described.

(V3) Minimal robustness: at least 80% of the specifications of a multiverse analysis (Steegen et al., 2016) had to keep the sign of the estimate.

The multiverse crosses three tolerances, three-unit costs, the two smooth models and the two reading orders, and adds six one-at-a-time variants (majority of three neighbours, four subsets of the linkage and a stricter map threshold). (V4) With seed and structural links only, the estimate had to keep its sign.

Secondary analyses change the inference (a weighted and an unweighted mean of map medians, and signs flipped by chapter instead of by map) and repeat the comparison on an independent identification: the modern positions that Shcheglov (2020) assigned to Ptolemy's coastal points of the Pontus Euxinus, the Black Sea.

That study also recalls that the Ξ recension is recognised as the earlier and more authentic, although it survives in a single manuscript (Shcheglov, 2020).

Exploratory analyses, corrected for the false discovery rate (Benjamini and Hochberg, 1995), compare units with the thin-plate spline; two others were planned, on the choice of reading and on whether the links labelled wrong in the verification samples are outliers of the segmentation, and were not run (Section 3).

Exact baseline, benchmark and reproducibility. As an exact baseline, the segmentation problem was written as a mixed-integer linear program (MILP) with big-M constraints (continuous similarity parameters per unit and binary assignments), whose behaviour with big-M is known to be unpredictable (Li, 2009). The MILP was solved with HiGHS once on the real instance of every map and on nested subsamples of the largest map; a stronger exact

method, such as branch and price on the covering program, was not tried. Wall-clock time, processor time and peak resident memory were measured in fresh processes with the numerical libraries limited to one thread, after a discarded warm-up, over 5 repetitions. The solver's processor time exceeded its wall-clock time in 4 of the 16 MILP runs on the real instances, so the solver itself used more than one thread there, and other computations of this study kept the processor load of the system between 46.6% and 98.5% during those runs. The machine has a desktop processor with 8 cores and 15.6 GB of memory available to the system; its graphics card was not used.

Every result is written to a structured file with its provenance (commit, plan tag, data and partition hashes, seed, library versions, detected hardware and thread settings), and the text reads every number from those files. Before the sealed half was opened, the confirmatory code was rehearsed on one fold of the exploration half (184 places; pooled median −3.05 km, p = 0.167). The sealed half was then opened 5 times for the primary analysis, each time with the analysis code of the tag: once at the tagged commit, twice after the provenance records were completed, and twice after the numerical libraries were limited to one thread. Those last two complete runs on the same commit gave identical result files apart from time stamps and measured run times, and they are the ones reported (Supplementary Material). An independent re-implementation, written from the text of the plan without access to the pipeline code, recomputed the primary analysis (Section 3).

Deviations from the plan. For lack of time, the limit of the exact program was lowered from 1,800 s to 300 s, the scaling curve repeated its measurements only up to 80 places, the table of stages covered the maps with at most 60 places, and the benchmark ran while the validity checks used other cores of the same machine; two planned exploratory analyses were also not run (Section 3). Amendment 1 records the details of the models that the frozen code fixed and the text of the plan did not, and Amendment 2 corrects the description of the calibration noise and records that the diagnostics of the primary analysis and the simulations behind the choice of test were produced after the result. Neither amendment changes an analysis.

3. RESULTS

Certified segmentations. Algorithm 1 closed the certificate, LB=UB, in 8 of the 16 maps, and the largest relative gap was 25% (Table II). In 5 maps column generation reached its iteration budget before proving the linear relaxation optimal; the Farley bound is valid for any duals, and the gaps of those maps reflect it. The lower bound was always the Farley bound: the bicriteria bound of Equation (7) was never the tighter one (Figure 3), and in the largest map it gave 20 against 51. The segmentations use 123 units and leave 101 of the 890 exploration places as outliers (11.3%), a fraction that reaches 54%, 47% and 37% in AS10, AS06 and AS03.

Units with exactly λ members cost as much as leaving their members as outliers; the segmentations contain 15 such units in 9 maps, so those segmentations are not the only optimal ones, and the predictions of B depend on which of them the integer program returns.

The unit scales (Figure 4) range from 11.5 to 173.5 km per degree (median 83.5); 14 units have a scale below 0.5 or above 1.5 times the median of their map, and 7 are rotated by more than 45° from the median rotation of their map, so some units fit a few places with transformations far from those of their neighbours. The 16 maps took 1,254 s in total, and the largest map, with 156 places, took 530 s for all its steps, so no final integer program can have reached its own limit of 600 s.

Table II: Certified Segmentation of Each Analysable Map: Exploration Places N, Sealed Places Evaluated, Units K And Outliers of the Solution, Lower Bound (LB) And Upper Bound (UB) Of the Certificate, Relative Gap (UB − LB)/UB And Wall-Clock Time of Algorithm 1 In Seconds (One Run).

Map	n	Sealed	K	Outliers	LB	UB	Gap	Time
EU01	15	18	3	4	13	13	0%	0.2
EU02	104	123	15	8	43	53	19%	58.5
EU03	49	45	10	5	34	35	3%	12.4
EU05	33	45	7	2	22	23	4%	13.0
EU06	118	121	13	5	33	44	25%	419.8
EU07	24	31	4	2	14	14	0%	13.3
EU09	28	39	5	1	16	16	0%	3.3
EU10	108	106	10	7	31	37	16%	70.4
AF01	22	16	4	3	15	15	0%	4.2
AF02	36	28	6	4	22	22	0%	6.2
AF03	49	40	6	6	23	24	4%	53.3

AS01	156	138	19	7	51	64	20%	530.3
AS03	19	15	3	7	16	16	0%	0.5
AS04	64	65	9	6	31	33	6%	66.8
AS06	19	14	3	9	18	18	0%	0.9
AS10	46	53	6	25	43	43	0%	0.9

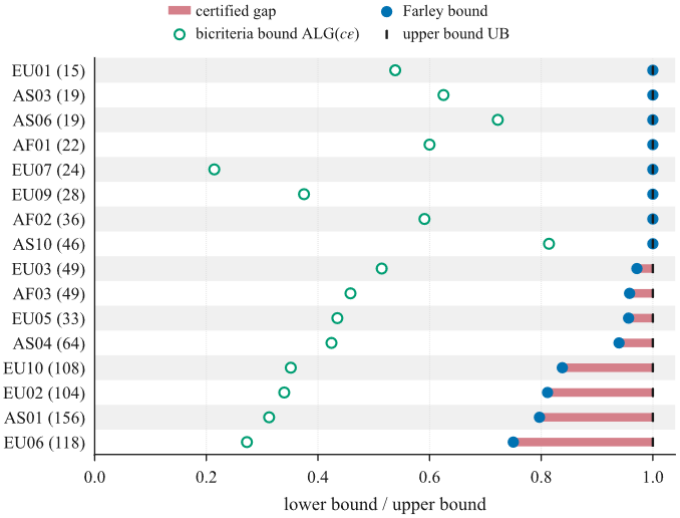


Figure 3: Bounds Of the Certificate of Each Map, Divided by Its Upper Bound UB: The Bicriteria Bound ALG(Cε) Of Equation (7) (Open Circles), The Rounded Farley Bound of Equation (10) (Dots) And the Certified Relative Gap of Equation (11) (Bars). Rows Give the Map and Its Number of Exploration Places.

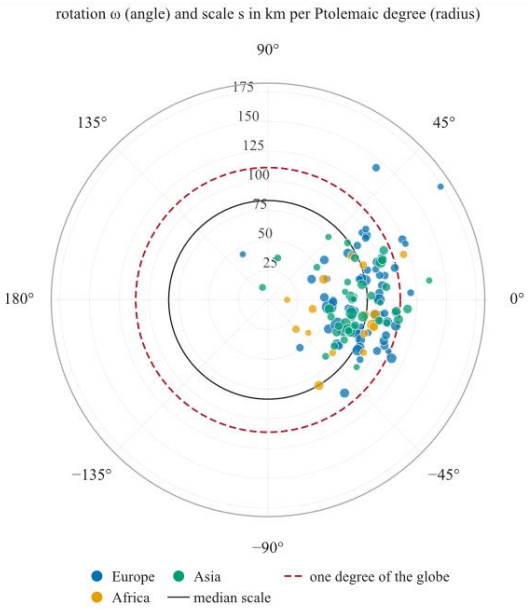


Figure 4: Scale S And Rotation Ω (Equation (3)) Of The 123 Certified Units, With Area Proportional to the Number of Members and Colour by Continent. The Solid Circle Marks the Median Scale and the Dashed Circle the Length of One Degree of a Great Circle on the Sphere of Radius R Used for the Errors (Equation (15)).

Figure 5 draws the certified units of two maps in Ptolemy's plane and, beside them, the local rotation of the smooth field with respect to the single frame of the map, together with the residuals that the single frame leaves at the exploration places. Both use the exploration half only.

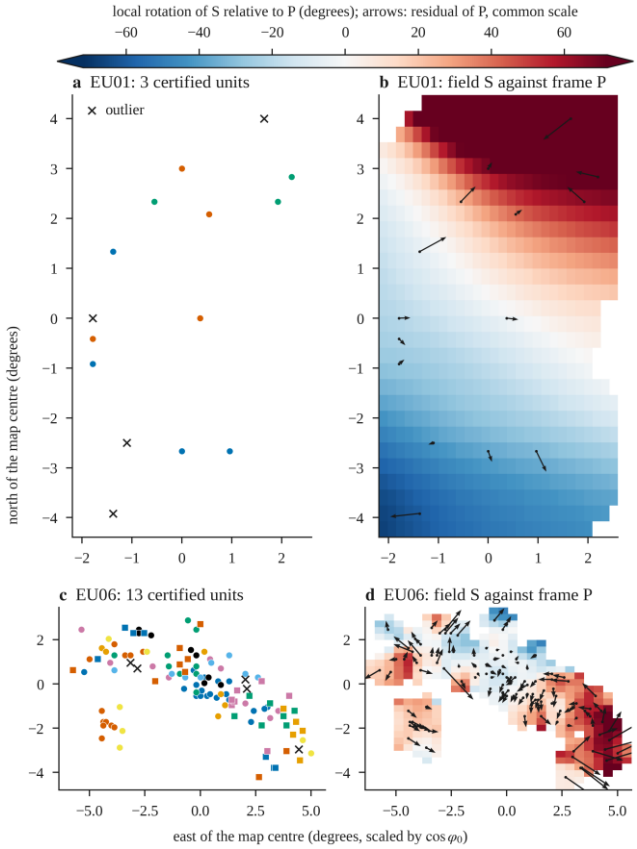


Figure 5: Certified Units and Smooth Field in Two Maps, Exploration Places Only. (A, C) Units of the Certified Segmentation in the Ptolemaic Plane (Equation (2)), One Colour and Marker Per Unit; Crosses Are Outliers. (B, D) Local Rotation Of The Smooth Field S Relative to the Frame P Of the Map, From the Conformal Part of the Derivative of Its Prediction (Colour, Shown Only Within Two Nearest-Neighbour Distances of an Exploration Place, And Saturated Beyond the Colour Bar), And Residual of P At Each Exploration Place (Arrows, One Scale for All Panels).

Primary result. On the 897 sealed places of the 16 maps, the smooth field predicted better than the similarity units. The pooled median of eS−eB was −6.11 km, with a 95% interval from −10.32 to −3.26 km and an exact two-sided p = 0.0061 (Holm-adjusted 0.0061). Units predicted better than one frame per map: the pooled median of eP−eB was 16.90 km (two-sided 95% interval 8.17 to 29.02; one-sided p = 0.0013, Holm-adjusted 0.0026). The median prediction errors were 44.4 km for the smooth field, 49.9 km for the units and 73.5 km for one frame per map.

By the rule of the plan the result is positive and favours the smooth field. Its size is modest: the point estimate exceeds the relevance threshold of 5 km, but the interval reaches values below it, so the data cannot say whether the advantage is larger than the rounding of Ptolemy's coordinates. The equivalence test is not met either, because the 90% interval, −9.48 to −3.71 km, leaves the band. The smooth field had the lower map median in 13 of the 16 maps (Figure 6), and the 3 maps where the units did better are all European. Leaving out one map at a time, the pooled median stayed between −7.02 and −5.61 km, and the largest p-value was 0.0097. Without AS10, AS06 and AS03, the three maps where the segmentation leaves most exploration places as outliers and where the smooth field wins by the widest margins, it was −5.72 km (p = 0.018). Map-level signs are not stable, however: 5 of the 16 maps changed sign between the cross-validation on the exploration half and the sealed places. Table III sets these results beside the decision rules of the plan and the validity checks reported below. The absolute errors behind these differences (Figure 7) show the same order over most of their range: the distributions of the smooth field and of the thin-plate spline lie above that of the units, which lies above that of one frame per map.

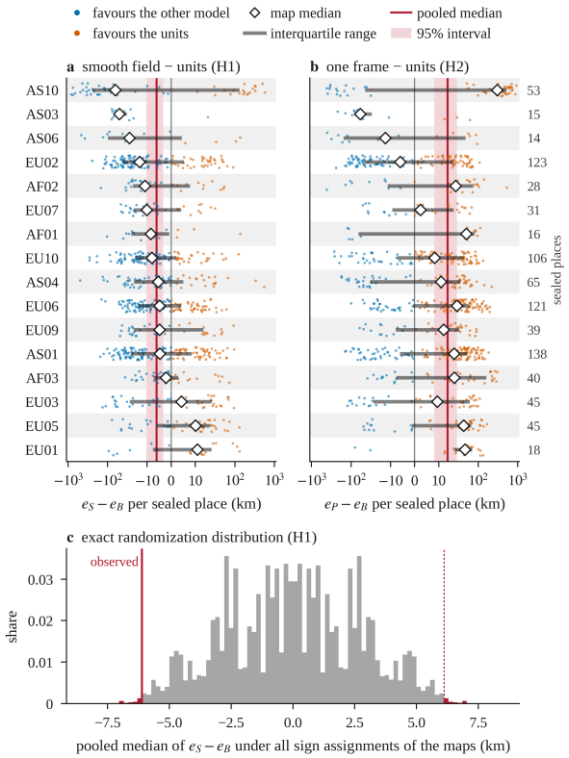


Figure 6: Paired Differences of Prediction Error on the Sealed Places. (A) Smooth Field Minus Units; Negative Values Favour the Smooth Field. (B) One Frame Per Map Minus Units; Positive Values Favour the Units. Dots Are Sealed Places, Diamonds Map Medians and Grey Bars Interquartile Ranges; The Red Line Is the Pooled Median and the Band Its 95% Interval from the Inverted Sign-Flip Test. The Horizontal Axis Is Linear Within 10 Km of Zero and Logarithmic Beyond. (C) Exact Randomization Distribution of the Pooled Median of Panel (A) Over The 65,536 Assignments of Signs to the Maps; The Red Tails, As Extreme as the Observed Value, Give the Two-Sided P-Value.

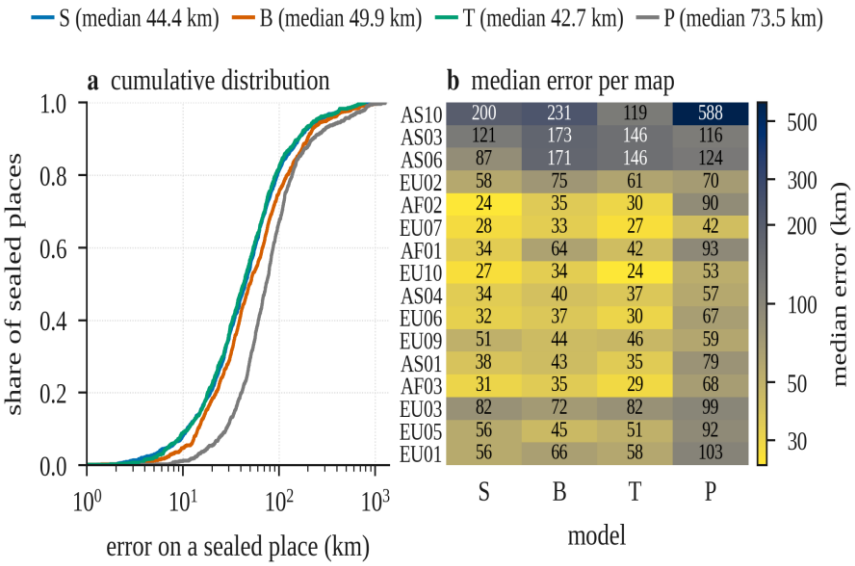


Figure 7: Absolute Prediction Errors on the Sealed Places. (A) Cumulative Distribution of the Error of Each Model, With Its Median in the Legend; The Horizontal Axis Is Logarithmic. (B) Median Error Per Map and Model, In Kilometres, With Maps in the Order of Figure 6.

Table III: Hypotheses And Validity Checks of the Frozen Plan: What Each Compares, The Decision Rule Fixed Before the Sealed Half Was Opened, The Result (Estimates With 95% Intervals) And Whether the Rule Was

Met.

	Comparison or check	Rule of the plan	Result	Rule
H1	smooth field against units, two-sided	Holm-adjusted p below 0.05	−6.11 km (−10.32 to −3.26); adjusted p = 0.0061	met
H2	units against one frame per map, one-sided	Holm-adjusted p below 0.05	16.90 km (8.17 to 29.02); adjusted p = 0.0026	met
V1	calibration with similarity units as the truth	significant 'smooth better' in at most 10% of replicates	0% of 40 replicates	met
V2	leakage from the seal or from Ptolemy's coordinates	every leakage test passes	all passed	met
V3	multiverse of specifications	at least 80% keep the sign	100% of 42 specifications	met
V4	seed and structural links only	the estimate keeps its sign	−6.22 km; p = 0.0059	met

Independent re-implementation. A re-implementation of the primary analysis, written from the text of the plan without access to the pipeline code, reproduced the objective and the units of every certificate, the bicriteria bounds, every per-place error, the pooled and map medians, the exact p-values and the intervals, once the details that the plan text had left open were read as the frozen code reads them (Amendment 1).

Those readings were given to the re-implementation after the confirmatory result was known, so the agreement checks the computation, not the independence of the choices. With the readings it had first taken from the plan text, all 897 predictions of the smooth field and 281 of the frames per map differed from those of the pipeline, and the pooled median was −6.47 km (95% interval −11.40 to −1.90; p = 0.012), with the same decision. With the kernel and the trimming rule of the pipeline but its own rules for ties, for the translation-only guard and for the anchors, 4 predictions of the units and 4 of the smooth field differed, and the pooled median was −6.10 km (p = 0.005).

Validity checks. In the calibration with known structure, the test declared the smooth field better in none of the 40 replicates whose truth was made of units, against a limit of 10%. When the truth was smooth, it declared the smooth field better in 100% of the replicates, and when the truth was made of units it declared the units better in 100% (Figure 8). The leakage tests passed. Of the 42 specifications of the multiverse, 100% kept the sign of the primary estimate and 100% also had p below 0.05 (Figure 9). With seed and structural links only, the estimate was −6.22 km (p = 0.006). All four checks passed, so the result is valid in the sense of the plan.

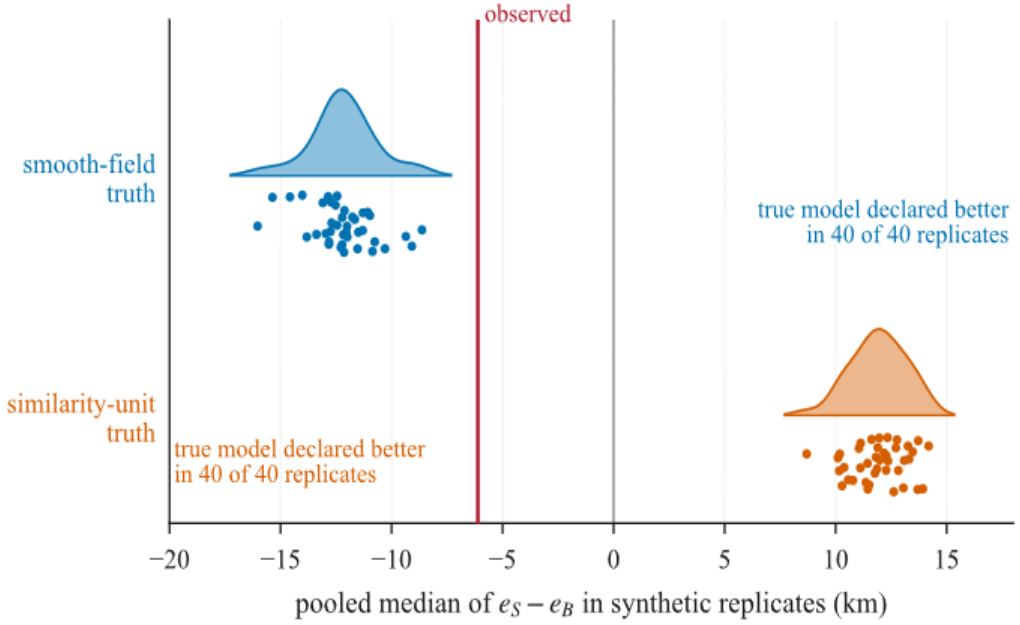


Figure 8: Calibration With Known Structure: Pooled Median of Es−Eb in Each Synthetic Replicate (Dots, Filled When the Exact Test Is Significant; Curves, Kernel Density Estimates) For Data Generated from the Smooth Field and from The Similarity Units Fitted to the Exploration Half. The Notes Count the Replicates in Which the Test Declared the Generating Model Better; The Red Line Marks the Observed Estimate.

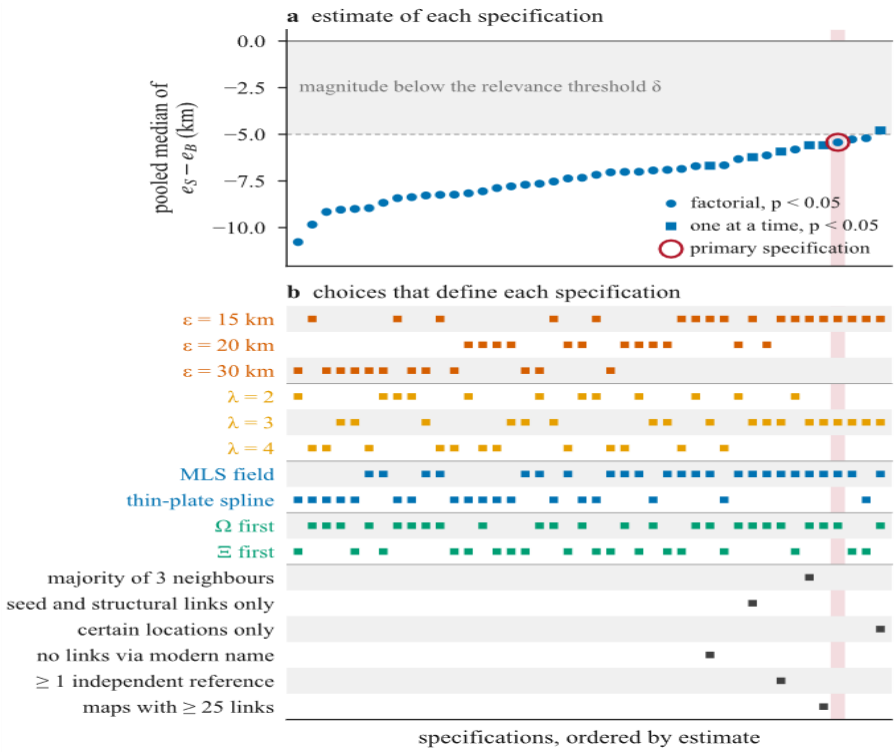


Figure 9: Multiverse Of the Comparison Between the Smooth Field and the Similarity Units. (A) Estimate Of Each Specification (Circles, Factorial Grid; Squares, One-At-A-Time Variants; Filled When P Is Below the Level of the Test; Grey Band, Magnitude Below the Relevance Threshold Δ; Red Circle and Pink Column, Primary Specification). (B) Choices That Define Each Specification. The Units Use the Fast Variant of Algorithm 1, So the Primary Specification Gives −5.43 Km Here, Against −6.11 Km with the Certified Algorithm.

Sensitivity of the inference. The weighted and unweighted means of the map medians were −10.19 km (p = 0.005) and −12.09 km (p = 0.013), and flipping signs by chapter instead of by map, with 53 clusters, gave p < 0.001. The two means weight maps instead of places and estimate a different quantity, which is why they lie further from zero; the chapter version uses a random sample of sign vectors instead of all of them and ignores that chapters of one map share their models, so it is anticonservative.

External replication. The positions of Shcheglov (2020) for points of the Pontus Euxinus that the models had not seen gave 51 comparisons in 3 maps. The pooled median of eS−eB was −4.03 km (p = 1.000), and with so few maps the smallest attainable p-value is 0.250, so no interval can be formed. The map medians had different signs (AS01 −8.82 km with 30 points, EU09 19.82 km with 20 points, AS03 51.55 km with a single point). Restricted to the 12 points whose modern position was not taken from Pleiades, the pooled median was 17.11 km in 2 maps. This check therefore neither supports nor contradicts the primary result.

Exact baseline. With a limit of 300 s per instance, the MILP proved the optimum of 1 of the 16 maps, none with more than 15 places. In 12 maps its best solution at the limit was worse than the upper bound of Algorithm 1, and in 15 its dual bound stayed below the lower bound of Algorithm 1, so the decomposition gave tighter intervals on both sides (Table IV). No instance contradicted a certificate.

Table IV: Exact Baseline on the Real Instance of Each Map: Places N, Certificate of Algorithm 1 (Lower and Upper Bound) And Dual Bound, Best Value and Wall-Clock Seconds of the MILP Under Its Time Limit (One Run Per Map).

Map	n	Alg. LB	Alg. UB	MILP bound	MILP value	MILP s
EU01	15	13	13	13	13	21
EU02	104	43	53	6	95	300
EU03	49	34	35	8	43	300
EU05	33	22	23	10	27	300
EU06	118	33	44	6	77	300

EU07	24	14	14	12	15	300
EU09	28	16	16	10	17	300
EU10	108	31	37	7	91	300
AF01	22	15	15	13	15	300
AF02	36	22	22	10	26	300
AF03	49	23	24	8	30	300
AS01	156	51	64	4	155	300
AS03	19	16	16	12	16	300
AS04	64	31	33	7	52	300
AS06	19	18	18	11	18	300
AS10	46	43	43	12	46	300

Cost of each stage. Table V gives the wall-clock time, processor time and peak resident memory of each stage of the pipeline, as medians over the repetitions, on the maps with at most 60 places (record linkage on the whole catalogue).

Table V: Cost Of Each Stage of the Pipeline: Median Wall-Clock Time and Processor Time in Seconds and Median Peak Resident Memory in Megabytes, Over the Repetitions After a Discarded Warm-Up.

Stage	Wall-clock	Processor	Memory
Record linkage	84.85	81.86	286
Pair and anchor candidates	0.11	0.12	107
Column generation	131.60	128.95	127
Final integer program	0.61	0.75	125
Bicriteria bound	0.60	0.80	119
Prediction and exact test	112.41	110.09	227

Computational cost. On nested subsamples of the largest map, the median wall-clock time t of Algorithm 1 was fitted by least squares in the log-log form of Equation (21)

$$\log t\,(n) = a + b \log n. \qquad (21)$$

The slope was b = 3.26 (95% bootstrap interval 2.96 to 3.64), compatible with the cubic cost of enumerating consensus sets (Figure 10). The MILP proved the optimum of the subsamples up to 14 places and reached its time limit beyond them.

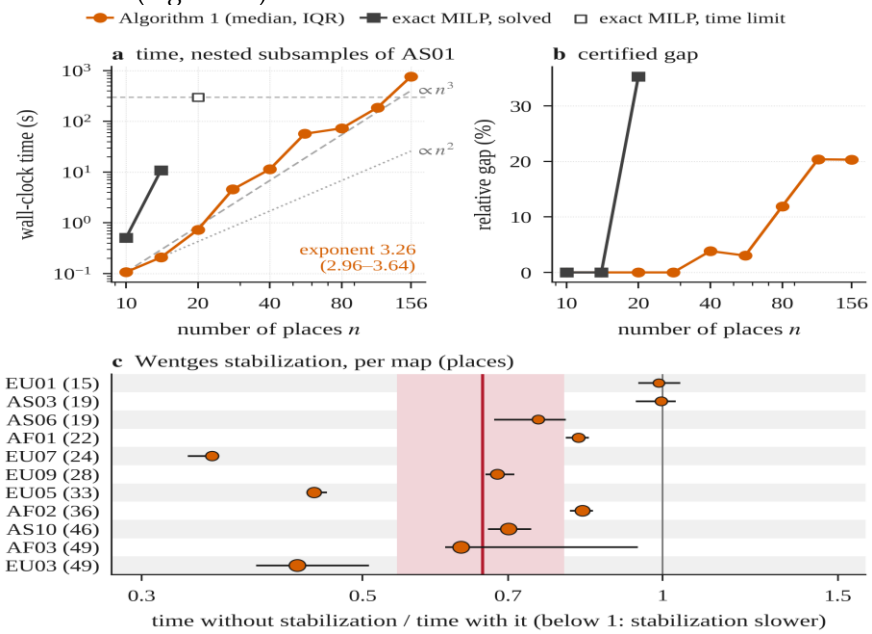


Figure 10: Cost Of Algorithm 1. (A) Median Wall-Clock Time with Interquartile Range on Nested Subsamples of the Largest Map, Against the Exact Program (Open Squares Reached the Time Limit); Dotted and Dashed Lines Grow as N2 And N3. (B) Relative Gap Between the Bounds. (C) Ratio Of the Median Time Without to the Median Time with the Stabilization of the Duals, Per Map, with 95% Bootstrap Intervals; The Red Line and Band Are the Geometric Mean and Its Interval.

Stabilisation. On each of the M maps with at most 60 places, the median wall-clock time of column generation was measured without and with the weighted duals, tm− and tm+, and the effect is summarised by the geometric mean of Equation (22)

$$\rho = \exp\left(\frac{1}{M}\sum_{m=1}^{M}\log\frac{t_m^-}{t_m^+}\right). \qquad (22)$$

Column generation with the weighted duals was slower than without them: the geometric mean of the ratios of median times, without over with, was 0.66 (bootstrap interval 0.54 to 0.80). On these small maps the weighted duals did not pay for themselves, so the stabilisation is not an optimisation that this study can claim (Figure 10c).

Figure 11 follows column generation iteration by iteration in two maps of the same size, rerun on the exploration half with and without the weighted duals: the value of the restricted master problem falls, the best Farley bound rises, and the run with the weighted duals reproduces the Farley bound of its certificate. Without the weighted duals, column generation stopped after fewer iterations in both maps, as the timings above suggest.

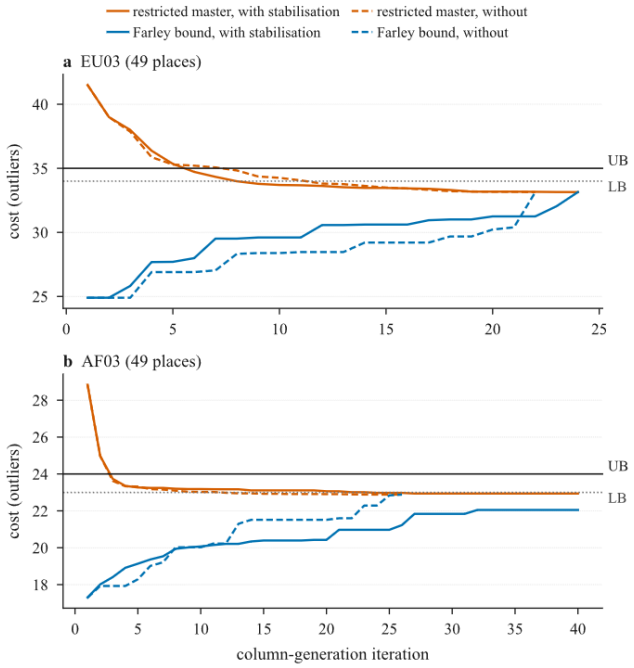


Figure 11. Column Generation in Two Maps: Value of the Restricted Linear Master Problem and Best Farley Bound (Equation (10)) At Each Iteration, With the Stabilised Duals of Algorithm 1 (Solid) And Without Them (Dashed); Horizontal Lines Mark the Upper Bound UB Of the Final Integer Program and the Rounded Lower Bound LB.

Exploratory analyses. These results were not planned as tests of the hypotheses. The thin-plate spline also predicted better than the units (pooled median of eT−eB −6.74 km, false-discovery-adjusted p = 0.001). Two planned exploratory analyses were not run, because each requires segmenting every linked place of every map again: the choice between the Ω and Ξ readings, and whether links judged wrong in the verification samples become outliers more often than correct ones. The false-discovery adjustment therefore covers the comparison with the thin-plate spline alone.

4. DISCUSSION

The units describe the exploration places of most maps with a few blocks, but they predict unseen places less well than a field that lets the scale, rotation and shift change gradually. They still beat a single frame per map by a wide margin, so the distortion is far from uniform within a map. What the sealed places disfavour is one way of defining blocks, a similarity per unit found by minimising a weighted count of units and outliers at a tolerance of 15 km, and not every block model: the units of the Berlin group share the scales of their area and differ in their shifts, and the interval of the advantage reaches values below the relevance threshold. The calibration with known structure declared the smooth field

better in none of the replicates whose truth was made of units, so the verdict is not simply an artefact of the smooth model's extra flexibility, although the calibration used the fast variant of the units and synthetic noise whose block scenario keeps every member near its unit; with no false declaration in 40 replicates, the upper limit of a 95% interval for the rate is 8.8%.

These results bear on the historical literature only indirectly, because a model that predicts well is not an account of how the coordinates were compiled. Marx (2012) attributes the shifts of groups of places to reference places on which the localization of neighbouring places was based, and considers an interpolation of the remaining errors inappropriate; on the sealed places, a local interpolation of similarities predicted better than units found by a search for homogeneous groups. Shcheglov (2019) describes Ptolemy's space as discrete and hierarchically organized and argues that the displacements cannot be adequately described by a single continuous function.

The smooth field is continuous, and its median error on the sealed places was still 44.4 km, so the result does not contradict that argument: neither description reproduces the displacements closely, and the smooth field is only less wrong. The advantage is small, of the order of the rounding of Ptolemy's coordinates, and 3 of the 16 maps favour the units.

The certificate changes what a segmentation can be used for. In 8 of the 16 maps the covering program proved its segmentation optimal for the chosen tolerance and unit cost, and in the others the interval bounds the distance to the optimum, with gaps of up to 25%.

The certificate does not make the segmentation unique, because units of exactly λ members can be dropped at no cost, but it lets a reader who disputes a unit separate two questions that earlier segmentations mixed: whether the data support a unit at that tolerance, and whether the search found it.

The certificate concerns the segmentation itself; the predictions of B also depend on the least-squares refit of each unit and on the nearest-neighbour assignment, which it does not cover.

The MILP proved the optimum of only 1 of the 16 maps within its time limit, none with more than 15 places, which is the reason for the decomposition; a big-M formulation is a weak baseline, however, and a stronger exact method was not tried.

Several limitations bound the conclusion. The comparison is only as good as the linkage and the modern positions, and a systematic identification error could favour either model.

The linkage never reads Ptolemy's coordinates, but some of the modern identifications that Pleiades and the catalogue record may have been made with them; the multiverse drops, separately, the 71 links made through the catalogue's modern identification and the 34 without a reference independent of Ptolemy.

Only the 16 maps with at least 15 exploration links of Ptolemy's 26 regional maps, enter the test, and linked places are the identifiable ones, so the conclusion applies to identifiable places of those maps and says little about poorly documented regions.

The tolerance of the units was chosen with the fast variant of Algorithm 1 at the edge of its grid, where the cross-validated error was still falling, so the units may be under-tuned; the multiverse varies it only upwards.

The two models treat outliers differently: the units leave 11.3% of the exploration places unexplained, while the smooth field excludes only places with an extreme leave-one-out residual, and the smooth field wins by most in the maps where the units leave most places out, although the pooled result holds without them.

The units tested here give each block its own scale and rotation, whereas the published transformation units share the scales of their area and add a rotation only for Scotland; that model was not tested, and a segmentation constrained in the same way might predict differently. Similarities also have one scale for both axes, whereas the Berlin model allows separate scales in longitude and latitude, and the cross-validation did not try tolerances below the grid, so units with a finer tolerance or anisotropic scales remain untested alternatives. The test compares predictive structures, not historical procedures. Finally, the benchmark was shortened for lack of time, the exact baseline had a shorter time limit than planned, and the timings were taken while other computations of the study shared the machine.

5. CONCLUSIONS

Judged by the prediction of places sealed before the analysis, the distortion of the identifiable places in 16 of Ptolemy's 26 regional maps is better described by a smooth field than by units with one similarity each, and those units are better than a single frame per map. The advantage of the smooth field is statistically clear but close to the resolution of the coordinates, and it concerns blocks that each carry their own similarity, not the transformation

units of the Berlin group, whose scales are shared by an area and which were not tested here. The segmentation into units is a strongly NP-hard covering problem that column generation and integer programming solve with a certified interval for every map of the study; the interval says how far each segmentation can be from the optimum, although it does not make the optimum unique. The same design, a sealed half, a frozen plan and certified models, can be applied to other historical coordinate lists.

Declarations

Conflicts of Interest. The author declares that there are no conflicts of interest.

Funding. This research received no specific grant from any funding agency.

Data Availability. The catalogue of Ptolemy's Geography and the Pleiades datasets are openly licensed releases that the pipeline downloads and verifies by hash; the modern positions of the Black Sea points come from the supplement of the cited study. The SHA-256 commitment of the identifiers of the sealed half is bd2c57ddcff242a8eee1140dc6ad34b6ca5c12d5a9688a4f70b4e635e69832b1.

Ethics. The study uses published historical and geographical data only and involves no human participants, personal data or cultural objects.

REFERENCES

Abshire, C., Gusev, D., Papapanagiotou, I. and Stafeyev, S. (2016) A mathematical method for visualizing Ptolemy's India in modern GIS tools. e-Perimetron, Vol. 11(1), pp. 13–34. http://www.e-perimetron.org/Vol_11_1/Abshire%20et%20al.pdf

Ask, E., Enqvist, O. and Kahl, F. (2013) Optimal Geometric Fitting under the Truncated L2-Norm. 2013 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1722-1729. https://doi.org/10.1109/cvpr.2013.225

Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B: Statistical Methodology, Vol. 57(1), pp. 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bookstein, F. L. (1989) Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11(6), pp. 567–585. https://doi.org/10.1109/34.24792

Burri, R. (2014) Review of: Europa in der Geographie des Ptolemaios. Die Entschlüsselung des "Atlas der Oikumene": Zwischen Orkney, Gibraltar und den Dinariden. Bryn Mawr Classical Review. https://bmcr.brynmawr.edu/2014/2014.01.19

Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2008) Bootstrap-Based Improvements for Inference with Clustered Errors. Review of Economics and Statistics, Vol. 90(3), pp. 414–427. https://doi.org/10.1162/rest.90.3.414

Chin, T.-J., Cai, Z. and Neumann, F. (2018) Robust Fitting in Computer Vision: Easy or Hard?. Computer Vision – ECCV 2018, pp. 715-730. https://doi.org/10.1007/978-3-030-01258-8_43

Delong, A., Osokin, A., Isack, H. N. and Boykov, Y. (2012) Fast Approximate Energy Minimization with Label Costs. International Journal of Computer Vision, Vol. 96(1), pp. 1-27. https://doi.org/10.1007/s11263-011-0437-z

Elliott, T. (2023) isawnyu/pleiades.datasets: Pleiades Datasets 3.2. Zenodo. https://doi.org/10.5281/zenodo.10070421

Farley, A. A. (1990) A Note on Bounding a Class of Linear Programming Problems, Including Cutting Stock Problems. Operations Research, Vol. 38(5), pp. 922–923. https://doi.org/10.1287/opre.38.5.922

Fellegi, I. P. and Sunter, A. B. (1969) A Theory for Record Linkage. Journal of the American Statistical Association, Vol. 64(328), pp. 1183–1210. https://doi.org/10.1080/01621459.1969.10501049

Graßhoff, G., Mittenhuber, F. and Rinner, E. (2017) Of paths and places: the origin of Ptolemy's Geography. Archive for History of Exact Sciences, Vol. 71(6), pp. 483–508. https://doi.org/10.1007/s00407-

017-0194-7
Graßhoff, G. and Defaux, O. (2019) Catalog of Ptolemy's Geography, main books. Zenodo. https://doi.org/10.5281/zenodo.3585758
Holm, S. (1979) A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics, Vol. 6, pp. 65–70. https://api.openalex.org/works/W2121044470
Horn, B. K. P. (1987) Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A, Vol. 4(4), pp. 629–642. https://doi.org/10.1364/josaa.4.000629
Huangfu, Q. and Hall, J. A. J. (2018) Parallelizing the dual revised simplex method. Mathematical Programming Computation, Vol. 10(1), pp. 119–142. https://doi.org/10.1007/s12532-017-0130-5
Jenny, B. and Hurni, L. (2011) Studying cartographic heritage: Analysis and visualization of geometric distortions. Computers & Graphics, Vol. 35(2), pp. 402-411. https://doi.org/10.1016/j.cag.2011.01.005
Kleineberg, A., Marx, C., Knobloch, E. and Lelgemann, D. (2011) Die antike Karte von Germania des Klaudios Ptolemaios. zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement, Vol. 136(2), pp. 105–112. https://geodaesie.info/images/zfv/136-jahrgang-2011/downloads/zfv_2011_2_Kleineberg_et-al.pdf
Lakens, D. (2017) Equivalence Tests. Social Psychological and Personality Science, Vol. 8(4), pp. 355–362. https://doi.org/10.1177/1948550617697177
Li, H. (2009) Consensus set maximization with guaranteed global optimality for robust geometry estimation. 2009 IEEE 12th International Conference on Computer Vision, pp. 1074-1080. https://doi.org/10.1109/iccv.2009.5459398
Marx, C. (2012) Rectification of the ancient geographic coordinates in Ptolemy's Geographike Hyphegesis. History of Geo- and Space Sciences, Vol. 3(1), pp. 99-112. https://doi.org/10.5194/hgss-3-99-2012
Marx, C. (2014) Rectification of position data of Scotland in Ptolemy's Geographike Hyphegesis. Survey Review, Vol. 46(337), pp. 231–244. https://doi.org/10.1179/1752270613y.0000000085
Megiddo, N. and Tamir, A. (1982) On the complexity of locating linear facilities in the plane. Operations Research Letters, Vol. 1(5), pp. 194-197. https://doi.org/10.1016/0167-6377(82)90039-6
Pápay, G. (2011) Review of: Germania und die Insel Thule: Die Entschlüsselung von Ptolemaios' "Atlas der Oikumene". H-Soz-u-Kult (H-Net Reviews). http://www.h-net.org/reviews/showrev.php?id=34081
Schaefer, S., McPhail, T. and Warren, J. (2006) Image deformation using moving least squares. ACM Transactions on Graphics, Vol. 25(3), pp. 533–540. https://doi.org/10.1145/1141911.1141920
Shcheglov, D. A. (2019) The structure of space in Ptolemy's Geography: methodological challenges in georeferencing ancient toponyms. ΣΧΟΛΗ. Ancient Philosophy and the Classical Tradition, Vol. 13(1), pp. 115–136. https://doi.org/10.25205/1995-4328-2019-13-1-115-136
Shcheglov, D. A. (2020) The configuration of the Pontus Euxinus in Ptolemy's Geography. History of Geo- and Space Sciences, Vol. 11(1), pp. 31–51. https://doi.org/10.5194/hgss-11-31-2020
Steegen, S., Tuerlinckx, F., Gelman, A. and Vanpaemel, W. (2016) Increasing Transparency Through a Multiverse Analysis. Perspectives on Psychological Science, Vol. 11(5), pp. 702–712. https://doi.org/10.1177/1745691616658637
Tsorlini, A. (2009) Spatial distribution of Ptolemy's Geographia coordinate differences in North Mediterranean eliminating systematic effects. e-Perimetron, Vol. 4(4), pp. 247-266. http://www.e-perimetron.org/Vol_4_4/Tsorlini.pdf
Tupikova, I., Schemmel, M. and Geus, K. (2014) Travelling along the Silk Road: A new interpretation of Ptolemy's coordinates. Max-Planck-Institut für Wissenschaftsgeschichte (Preprint 465). https://www.mpiwg-berlin.mpg.de/sites/default/files/Preprints/P465.pdf
Umeyama, S. (1991) Least-squares estimation of transformation parameters between two point patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13(4), pp. 376-380. https://doi.org/10.1109/34.88573
Wentges, P. (1997) Weighted Dantzig-Wolfe decomposition for linear mixed-integer programming. International Transactions in Operational Research, Vol. 4(2), pp. 151–162. https://doi.org/10.1016/s0969-6016(97)00001-4