Supplement to: GC Insights: Choosing the right map projection, and what it means for scale bars and north arrows

Dominic Royé — Biological Mission of Galicia (MBG), CSIC, Santiago de Compostela, Spain

S1 Why a survey, and what it can and cannot show

The main text makes a normative argument: the projection should be chosen for the purpose the map serves, and the validity of the scale bar and north arrow follows from that choice. What follows does not test the argument; it measures something narrower: how often a reader of the published literature is in a position to determine which projection was used. The unit of observation is the published map, not the author who drew it.

The design is a cross-sectional content analysis of published figures, coded against a protocol fixed before coding began (Sect. S3). It is not a systematic review. The aim is not to locate every relevant map or to summarise what the sampled papers concluded, but to estimate how often given features occur. The sample is accordingly drawn from a defined frame, and the results are reported as prevalences with confidence intervals rather than as an attempt at completeness.

The full dataset, including the DOI of every sampled article, is published so that the coding can be verified rather than taken on trust (Sect. S6). What is reported here is nonetheless aggregate: no article, author or journal is identified in the text, tables or figures, and no published map is reproduced as an example of poor practice. That is a deliberate choice about how to report, not a restriction of the data itself: the argument concerns an inherited convention and a set of software defaults, not any individual's work.

S2 Sampling

S2.1 Frame

The frame is research articles published between 2016 and 2025 by Copernicus Publications (Crossref member 3145) in 27 peer-reviewed geoscience journals, about 36 300 articles in all. Crossref also types conference proceedings and abstract volumes as journal articles, so the frame is defined by an explicit allow-list of journals rather than by the type field. That list, with the reason each excluded container was dropped, is released with the data.

The frame was restricted to open access for practical as well as principled reasons. Reviewers can verify the sample without a paywall, figures can be reproduced under CC BY, and retrieval is automatic and exactly repeatable. This is a subset of the geoscience literature, EGU-affiliated, Earth and space sciences, weighted toward Europe, and it is not claimed to represent the field as a whole.

S2.2 Two-stage design

Stage 1 drew 600 articles by simple random sampling within strata defined by journal group (five domains) crossed with three periods (2016-2018, 2019-2021, 2022-2025), 40 articles per stratum, with the random seed fixed and reported in the code. Stage 2 shuffled the figures within each sampled article and coded the first one that turned out to be a map. Taking the first eligible item from a random permutation is the same as drawing uniformly at random among the eligible items, which gives a random sample of maps without needing to know in advance which figures qualify. Coding one map per article keeps the observations independent of each other.

Coding proceeded in random article order within each group until it reached a quota of 60 maps. Four groups reached 60 and one reached 61; biogeosciences and soil stopped at 58, once every article drawn for that group had been exhausted. That leaves 299 maps in all. The 600 drawn articles span 26 of the 27 journals, and the 299 with a coded map span 24. Keeping the groups roughly balanced matters, because the frame weights used in Sect. S4 lose precision quickly once one group dominates the others.

Stage	Count
Records harvested from Crossref	60 010
After the journal allow-list	36 275
Articles drawn (stage 1)	600
Articles with a retrievable PDF	596
Figures located and rendered	5 069
Figures inspected by the coder	2 534
Articles reaching a coded map	299
Articles found to contain no eligible map	204
Maps coded (analysis set)	299
Second maps coded (within-article substudy)	24

Table S1. Sampling and screening cascade.

S3 Coding

Coding was limited to what is visible on the figure itself: whether a coordinate system can be identified and what geometry it shows, which map furniture is drawn, and how much of the world the map covers. What a given projection does to scale and to north follows mathematically from its geometry, not from this sample; the survey shows only which projections are actually in use.

Variable	Values
crs_kind	none_apparent / degrees_rect / degrees_curved / utm_metres
proj_name	the projection as named in the paper, verbatim, if it names one
scalebar	present / absent. A companion field scalebar_multiple records more than one bar; no sampled map drew one
north_arrow	none / indicator (an arrow or a compass rose). Rows coded before the two were merged carry compass_rose, which is equivalent
graticule	none / edge_ticks (frame only) / graticule (lines across the map face)
extent_class	local (<100 km) / regional / continental / global
lat_band	tropical / mid / polar, or a range such as mid-polar or tropical-mid; empty for global maps

Table S2. Coded variables.

One coding rule needs spelling out, because the results depend on it. Degree labels only indicate what the axis values are, not how the map was projected: a Lambert conformal conic and a polar stereographic both label their graticules in degrees. A map was read as unprojected geographic coordinates only when the grid is rectangular and evenly spaced; where meridians converge, parallels curve, or the spacing along a straight frame is uneven, the map is projected, but its projection cannot be identified. Of the 153 maps read as a rectangular degree grid, 48 draw no coordinate lines or frame ticks at all, so that classification rests on the axis labels and on the shape of the coastlines shown. The distinction matters because a local map at mid-latitude is off by 17.6 % if it is genuinely unprojected, and by roughly 0 % if it is equal-area or Transverse Mercator.

Extent is coded as a category, not as typed coordinates. What matters for the results is how much of the world a map covers, and the category captures that directly; entering four coordinates per map would have taken far longer than the extra precision was worth. Continental maps can carry two latitude bands, since a map spanning 40° of latitude routinely straddles them.

S4 Results

Prevalences are weighted by how many articles each journal group contributes to the frame, because the sampling quota made the groups equal-sized while the underlying population is not: atmosphere and climate alone account for 15 350 of the 36 275 articles, biogeosciences and soil for 3 982. Pooling the raw counts unweighted would answer a different question; the average across five domains rather than the published literature as a whole. Confidence intervals are Wilson intervals evaluated at Kish's effective sample size, since unequal weights mean the raw count overstates how much information the sample actually carries.

Outcome	n	Unweighted	Weighted	95 % CI
Names its projection	299	2.3 %	1.7 %	0.6-4.3
No coordinate system visible	299	33.8 %	33.2 %	27.4-39.6
Projected but unidentifiable	299	10.0 %	10.7 %	7.3-15.4
Projection not determinable (either)	299	43.8 %	43.9 %	37.6-50.4
Scale bar drawn	299	32.4 %	26.2 %	20.9-32.3
North indicator drawn	299	23.1 %	20.1 %	15.4-25.8
Graticule across the map	299	14.7 %	16.1 %	11.9-21.5

Table S3. Main prevalences.

Extent	n	Projection not determinable	Scale bar	Grat.	North ind.
Local (<100 km)	97	55 %	68 %	8 %	48 %
Regional	91	36 %	31 %	19 %	23 %
Continental	53	49 %	2 %	23 %	2 %
Global	58	33 %	3 %	12 %	2 %

Table S4. By mapped extent, unweighted. Scale bars are overwhelmingly a convention of large-scale maps. That is why the failure the literature usually warns about, a bar drawn across a world map, is rare in practice, while the quieter failure, an unstated projection under a plausible-looking local bar, is common.

The projections actually in routine use here are cylindrical or close to it. Three-quarters of the world maps that draw any coordinate lines draw them as a rectangular grid, and a rectangular grid means straight meridians. The north-arrow problem discussed in the cartographic literature, where north tilts away from vertical, therefore barely arises in this corpus. That is a property of what authors actually publish, not a limitation of the sample, and coding more maps would not change it.

S4.1 Within-article agreement

A second map was coded in 24 articles, solely to estimate how similar two maps from the same paper tend to be; these pairs are never pooled into the prevalences above. Thirty articles had been earmarked for this in advance, but only 14 of them offered two eligible maps; the remaining 10 pairs arose incidentally, where a second map was coded regardless. The pairs are therefore not a pre-designated random subsample, and rho is reported as a descriptive quantity rather than a design-based estimate. Rho is the agreement in excess of what the marginal frequencies alone would predict, equivalent to Cohen's kappa applied to the two members of each pair. A simulation at this number of pairs puts its standard deviation at about 0.09.

Variable	Agreement	rho
Scale bar drawn	95.8 %	0.92
North indicator drawn	87.5 %	0.59
Projection not determinable	75.0 %	0.49
Extent class	83.3 %	0.76

Table S5. Within-article agreement. A rho of 0.92 for the scale bar means the practice is essentially fixed within a paper: it belongs to the research group rather than to the individual map, which is why the main text addresses the remedy to journal guidelines rather than to individual authors. It also justifies, after the fact, coding only one map per article: a second map would have added just 1 - rho of a genuinely new observation.

S5 Limitations

The frame is EGU-affiliated, Earth and space sciences, and weighted toward Europe; it is not claimed to represent the geoscience literature as a whole. Coding was carried out by a single coder. The protocol called for a 20 % double-coded subsample to assess inter-coder reliability, and this was not achieved, so no such figure is reported. Nor was a systematic second pass made over every decision already logged. Extent is coded as a category rather than measured as a bounding box, which is adequate for the comparisons made here but not exact.

Two limits of the sampling are worth keeping in mind alongside the results. The target was 60 maps per journal group; Biogeosciences and soil reached only 58, because every article drawn for that group had already been checked and only 58 of them contained a map. The analysis set is 299 maps rather than the intended 300. Because prevalences are weighted by frame size using the counts actually achieved, a shortfall of two maps in the smallest group does not change any estimate materially.

Second, the figures excluded as not being maps were not re-inspected. A figure wrongly excluded, a small inset or a map-like panel folded into a composite figure, would remove its article from the sample, and no estimate of how often that occurred is reported here.

S6 Data and code

The archive holds everything needed to reproduce the survey: the journal allow-list that defines the frame, the DOI list of the 600 sampled articles, the figure worklist with captions and the randomised order used in stage 2, the coding protocol with its deviations log, both append-only decision logs, and the analysis scripts. Only the downloaded PDFs and the rendered page images are left out, purely for size; the scripts rebuild them from the DOI list. The code is released under the MIT licence and the data under CC BY 4.0.

Releasing the DOI list at all is only possible because of what got sampled. Crossref lists an open licence for every article drawn: 465 CC BY 4.0, 134 CC BY 3.0, and one 2016 article with no open licence, which was consequently never downloaded and never coded. Redistribution is therefore permitted provided attribution is given, and the DOI list serves as that attribution.

The per-article rows deserve one caveat. This is a random draw from 36 275 articles: no article is in it because something was found wrong with it, and none is held up as an example of poor practice. The coded fields describe what is drawn on the page, not the quality of the work behind it. The subject here is an ordinary convention of the field that the sampled authors happened to follow. The coding is by one person and unadjudicated, with no inter-coder reliability reported (Sect. S5), so the study speaks to aggregates: a single row is not evidence about a single paper.

Figure S1

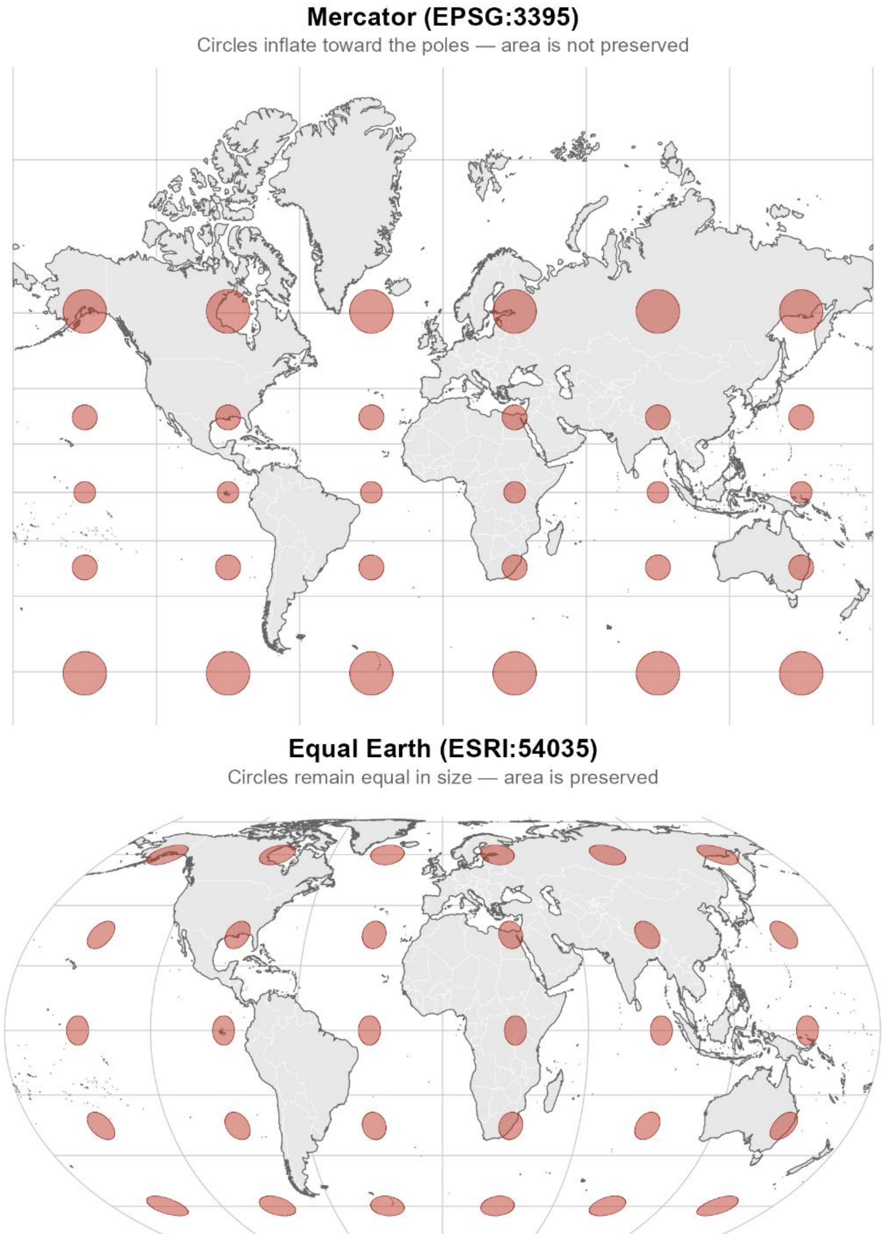

Figure S1. Tissot's indicatrices for the projection families discussed in Sect. "Why projections matter" of the main text. Circles of identical size on the globe are shown as each projection renders them: equal-area projections keep their areas correct while distorting shape, conformal projections keep them circular while distorting size, and compromise projections give a little on both.