Request for comments

RFC20070514NH: New distance formula

1st draft: N. Harvey, May 14, 2007
2nd draft: N. Harvey, September 4, 2007

- added testing info covering Lambert, Albers, Chamberlin projections.

Applies to: Model description v1.1.0

Type of change: Change, for next minor version (3.2)

Summary: This RFC proposes adopting a better formula for distance calculations.

Justification: The formula proposed here is more accurate than the current one and will make the program code simpler, faster, and more readable for maintainers. From the user's point of view, nothing will change – there are no changes to parameters or inputs.

Change: This change applies to Section 4.1 (Direct contact spread). This new text replaces the old paragraph about distance calculations:

> The distance between two units is approximated by a straight-line distance on a map. The Albers Equal Area Conic map projection is used, with the two standard parallels placed at one-sixth and five-sixths of the range of latitudes in the study area, and using the WGS84 model of the Earth[1].

Change: This footnote is added for the text above:

> 1. WGS84 is a set of measures of the Earth, such as its radius at the equator. It is used in many modern mapping applications, including GPS systems.

End of changes

Notes: The main goal of this change is to make the code simpler and easier to read and debug. Currently we need to adjust for latitudes and do a degrees/kilometers conversion every time we do anything related to distance (*e.g.*, finding distance between herds, drawing a zone circle). With this change we would adjust for latitudes once – when doing the map projection – and from then on, work in simple x-y coordinates on a km grid.

Details on why the Albers map projection was chosen, and how much more accurate it is, follow.

Any method of drawing the curved Earth on a flat surface introduces some distortion. We want a method that generally does a decent job of preserving distances, angles, and regular shapes like circles. Four common map projections were considered:

- the Mercator projection, created in 1569, still seen in classrooms everywhere.
- the Lambert Conformal Conic projection, created in 1772, favored by the National Atlas of Canada.
- the Albers Equal Area Conic projection, created in 1805, favored by the US Geological Survey.
- the Chamberlin projection, created in 1950, favored by the National Geographic Society.

For more information and sample pictures of map projections, see:
http://www.radicalcartography.net/?projectionref

The accuracy of distance calculations was tested as follows:

a) Start with a set of locations, in latitude-longitude format.
b) Project the locations onto a flat map.
c) Choose two locations at random.
d) Calculate the straight-line distance between them on the flat map.
e) Record the error in the calculated distance as compared to the true distance.
f) Repeat steps c-e many times.

Five data sets provided the locations mentioned in step (a). Four of them have been used by this group in modeling work before: they are the Ontario, New Zealand, United Kingdom (North Cumbria), and Switzerland data sets. A fifth data set was created containing the locations of U.S. state capitals (contiguous 48 states). All together, the five data sets include both large and small areas, and both Northern and Southern hemispheres.

The "true" distance used to determine the error in step (e) was the Great Circle distance on the WGS84 ellipsoid.

Projection parameters

To construct a Mercator projection, we must choose one "standard parallel". This was placed halfway between the minimum and maximum latitudes in the data set.

The Lambert and Albers projections require two standard parallels. These were placed at one-sixth and five-sixths of the range of latitudes, a rule of thumb given in the manual for the GIS software *Manifold*.

The Chamberlin projection requires three control points. Sometimes it is obvious where to place the points; an example found at http://www.warnercnr.colostate.edu/class_info/nr502/lg2/projection_descriptions/chamberlin.html places the points in Alaska, Labrador and Mexico for a map of North America. But we need an automated method that does not require a human to "eyeball" the problem.

Since the bounding box of the data set is easily found, a simple way to create a triangle is to choose one side of the bounding box, plus the midpoint of the opposite side. A test with the 48 states data set showed no great advantage to any orientation of the triangle (figure 1).

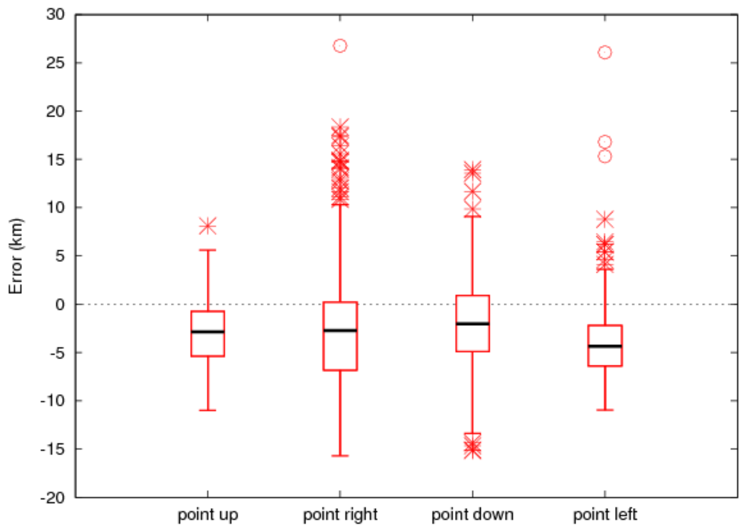

Figure 1. Accuracy when the Chamberlin projection control points form a triangle pointing in various directions. 48 states data set.

For the remainder of the tests, the Chamberlin projection used a triangle pointing up (down if the data set was in the Southern hemisphere).

A further question is whether the triangle should be created entirely inside the bounding box. That triangle will cover only half the area of the bounding box; perhaps a larger triangle would be better. A test with larger triangles, in which the triangle was expanded to 1¼ times its width and height, then 1½ times its width and height, showed no advantage to making the triangle larger (figure 2).

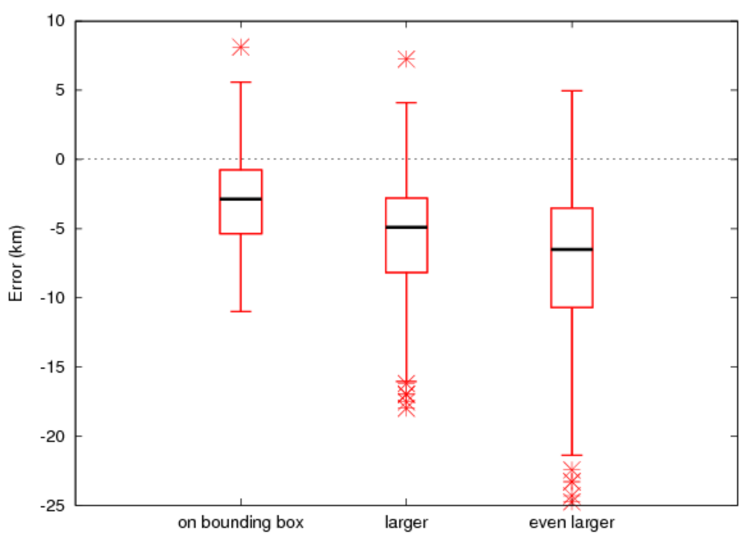

Figure 2. Accuracy when the triangle formed by the Chamberlin projection control points was enlarged. 48 states data set.

Now, with heuristics for choosing parameters for all four projections, we can compare their accuracy on the various data sets. Figure 3 shows the results.

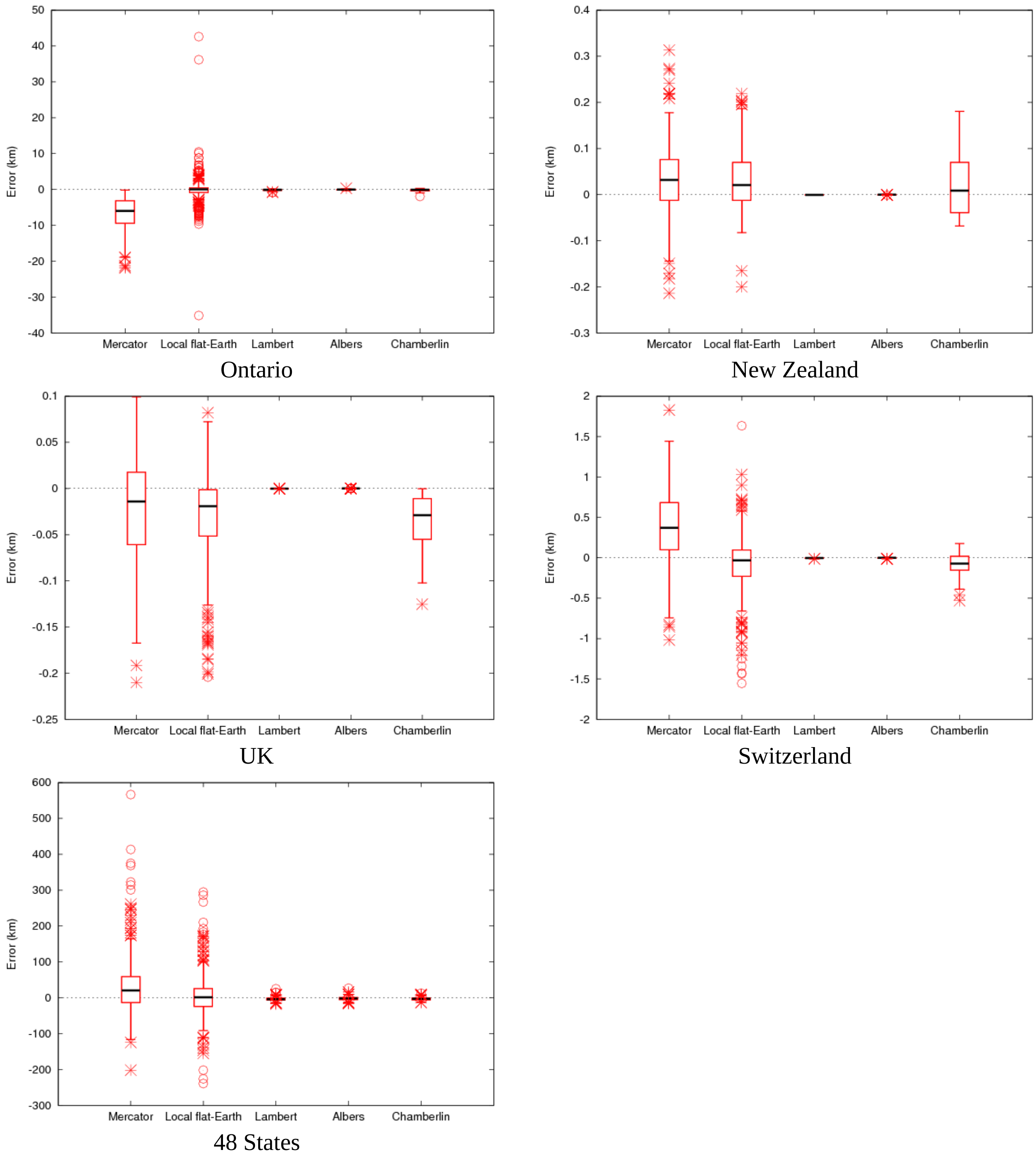

Figure 3. Accuracy of the distance calculations on 5 data sets. The second from the left ("local flat-Earth") is the calculation currently used in NAADSM.

The basic Mercator projection is (unsurprisingly) a poor choice, and so is the calculation currently used in NAADSM. The Lambert and Albers projections do well. The Chamberlin projection is inconsistent, sometimes doing well and other times not.

Figure 4 provides a zoomed-in look at the calculations that did well.

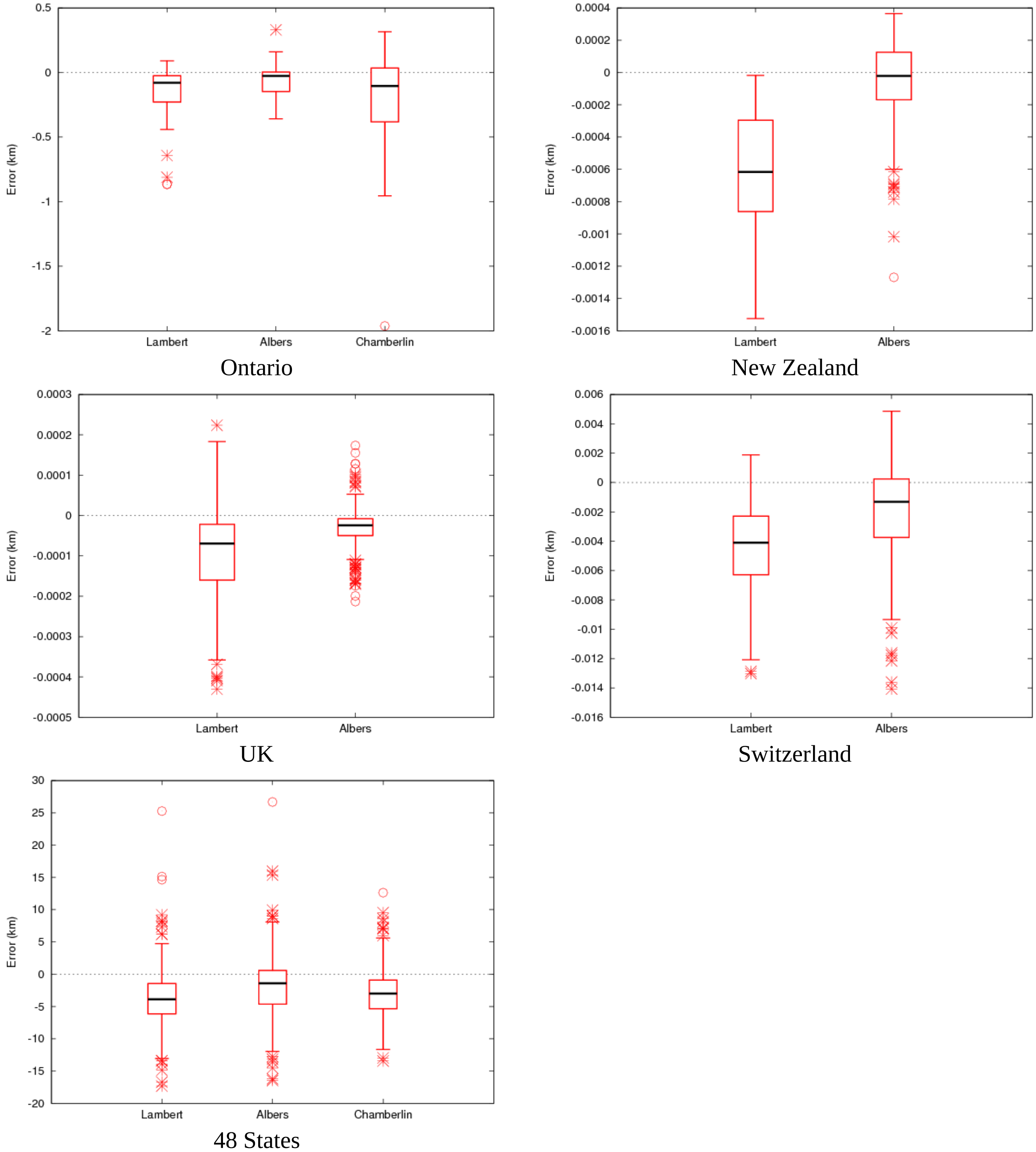

Figure 4. Accuracy of the distance calculations on 5 data sets. Zoomed-in view to compare the most accurate calculations.

The Albers projection was the overall best performer, with its median error closest to zero.