Visualizing of Berkeley Earth, NASA GISS, and Hadley CRU averaging techniques

Robert Rohde
Lead Scientist, Berkeley Earth Surface Temperature

1/15/2013

Abstract

This document will provide a simple illustration of how different climate research groups combine discrete observations to construct large-scale views of Earth's land areas. This illustration is meant as a simple way to summarize the different methodologies, though it has no specific research value beyond that.

Introduction

There are four major efforts to synthesize the Earth's disparate temperature observations into a coherent picture of our planet's climate history. These efforts are led respectively by NOAA's National Climate Data Center (NOAA NCDC), NASA's Goddard Institute of Space Sciences (NASA GISS), a collaboration between the University of East Anglia's Climatic Research Unit and the UK Met Office's Hadley Centre (CRU[1]), and the Berkeley Earth Surface Temperature group. Each group uses different averaging techniques, quality control procedures, homogenization techniques, and datasets.

The current discussion will briefly look at the land-surface averaging methods applied by Berkeley Earth, CRU, and NASA GISS. NOAA's method uses information from multiple times to construct an empirical orthogonal function representation of the Earth. Since the illustration performed here will only look at a single time slice, it isn't possible to consider NOAA's method, and hence their averaging technique will not be discussed here.

Visualizing the Methods

We shall provide an illustrative example of how the various averaging methods view the world. Rather than using temperature data, we will perform this illustration using a visual image of Earth's land surface. Since most people are familiar with the appearance of the Earth this approach provides an easy

[1] The land surface temperature analysis published by the Climatic Research Unit / Hadley Centre collaboration is known as CRUTEMP, in recognition that the land component was originally developed by Climatic Research Unit alone. Following their naming convention, we will use the acronym CRU in describing this methodology. The collaboration is sometimes referred to HadCRU, though we have decided not to use that abbreviation here in order to avoid confusion with the HADCRUT land plus ocean data set that they develop. The present discussion focuses only on the CRUTEMP land-surface analysis methods.

way to examine the effects of the different averaging techniques with respect to discreteness, smoothing, and extrapolation.

Figures 1 shows an image of the Earth as seen from space. This image is taken from NASA's Next Generation Blue Marble project and the ocean has been replaced with a lighter blue to aid in the recognition of land-ocean contrasts. In Figure 2 we "sample" this image at the 5883 sites where weather monitoring stations existed in the monthly Global Historical Climatology Network (GHCN-M) in 1970. We chose 1970 as this is the time period with the largest number of simultaneously active stations in the GHCN-M. Hence this is a way of looking at the best case scenario for GHCN-M. Each circle represents the color of terrain at a station's location, as seen from space, and has an area of approximately 50,000 km^2. The underlying land is indicated in red to help show the effects of sparse sampling. As shown, Antarctica, Africa and part of South America are sparsely sampled at this time compared to the rest of the world. By examining Figure 2 it is possible to get a sense of how completely (or incompletely) regions of the world were sampled by the GHCN-M network. From this sparse sampling, we aim to apply the various averaging approaches to help illustrate their characteristics and how well they reproduce the underlying field.

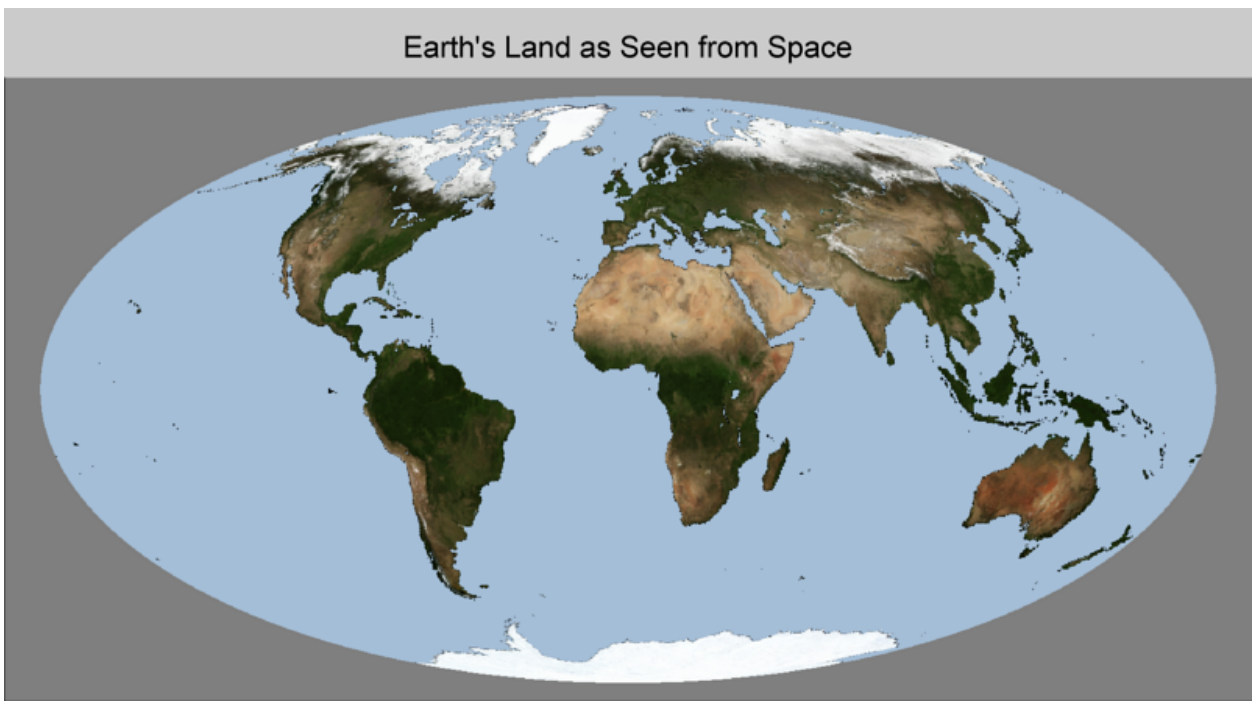

Figure 1: NASA image showing the colors of the of Earth's land areas with an equal-area projection. Ocean areas have been replaced with a light blue to aid in recognition of land-ocean contrasts.

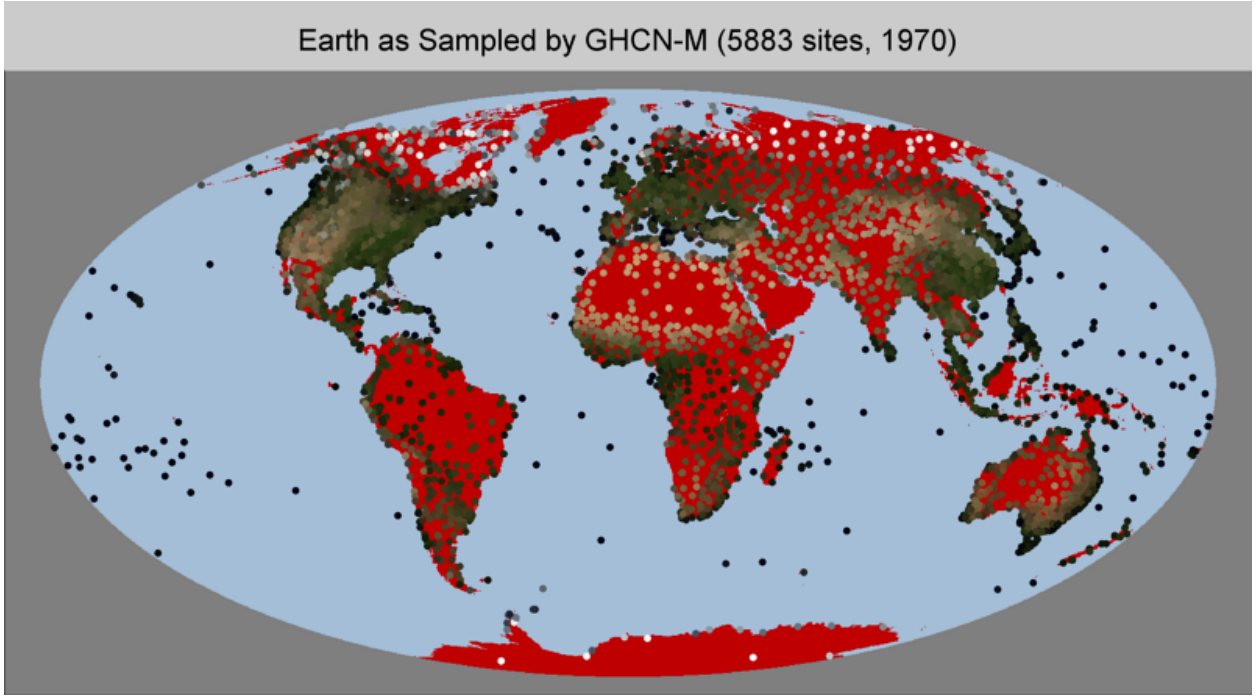

Figure 2: The visual field of Earth's land areas as "sampled" at each of the locations where a weather station existed in the GHCN-M as of its peak station count in 1970. Each circle represents about 50000 km^2 at the surface. A red land mask is added to help identify regions of the Earth that are poorly sampled at this time.

CRU

CRU relies on a simple 5°x5° latitude-longitude gridding and only considers the local data within each grid cell when determining the grid cell average. Applying this methodology to the visual data results in the somewhat pixilated and incomplete field shown in Figure 3.

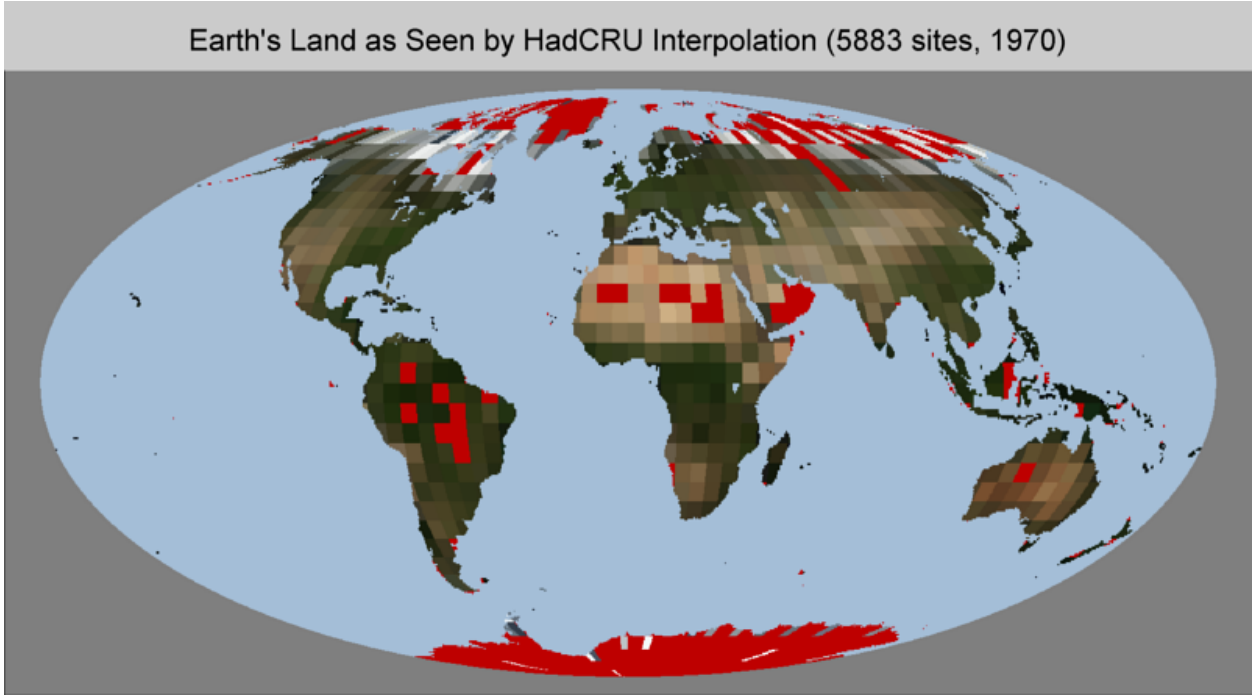

Figure 3: Interpolation from the sampled visual data using CRU style gridding. The 5°x5° latitude-longitude gridding is a defining characteristic of this method, as is the presence of many empty grid cells where no estimate is made due to the lack of local data. Such missing cells are indicated in red.

As indicated by the image, the relatively large grid boxes result in a map that has a blocky texture. More importantly, the reliance on a flat latitude-longitude grid results in polar regions that will be mostly incomplete. This gridding structure compounds the already poor sampling of Antarctica to produce a field with almost no content near the Southern pole. The CRU gridding approach is the only methodology of the major groups that does not produce a complete map from the sampled field. To a degree, CRU's focus on local data does ensure that each grid cell is a good representation of local features. However, the incompleteness of their reconstruction is ultimately the biggest limiting factor in their ability to reconstruct global means (as will be shown in a separate discussion).

NASA GISS

GISS uses an equal area gridding, which at the equator has a resolution of 2.25°, significantly more detailed than CRU. However, GISS combines this with a distance weighted average that incorporates data out to 1200 km from each grid cell. This distance weighted averaging has the effect of smoothing the resulting field and blurring any small scale details. The effects of applying this methodology to the sampled field are shown in Figure 4.

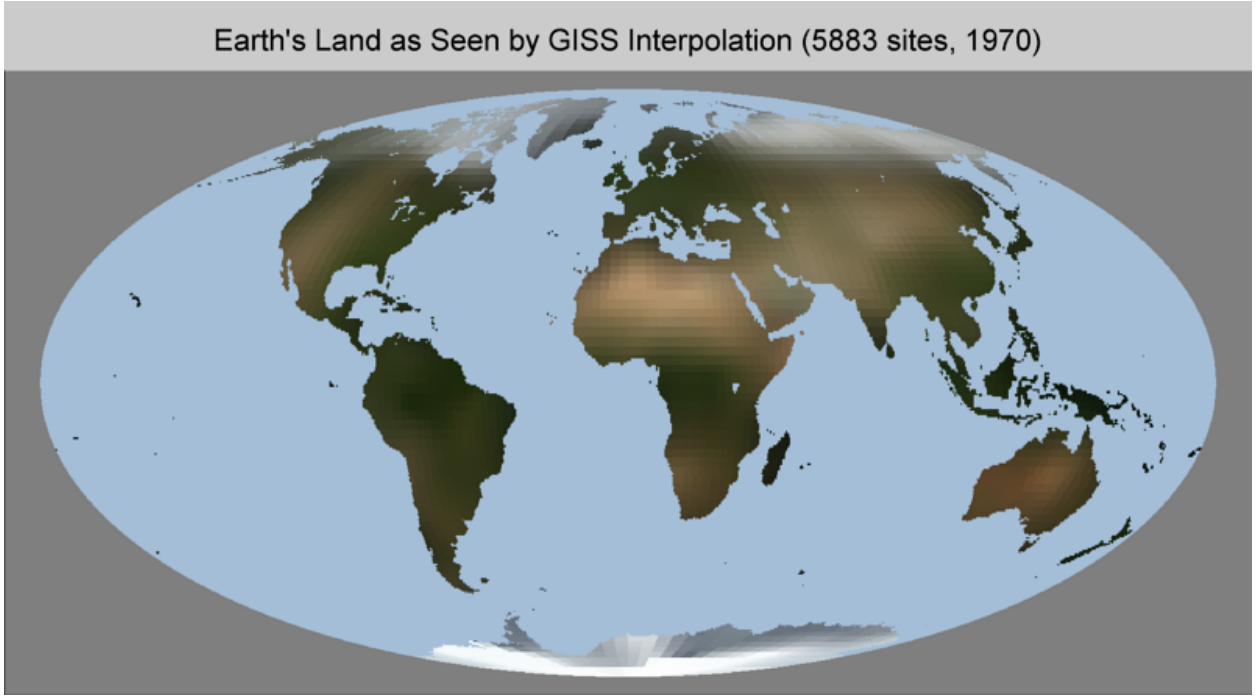

Figure 4: Interpolation from the sampled visual data using NASA GISS style averaging. The relative smoothness (i.e. blurring) of the field is a characteristic of the GISS field. A globally complete field is generated through the use of regional averaging, but any features that may exist below the scale of the averaging cannot be represented.

The NASA GISS method has the advantage of generating a globally complete field and the use of an equal area grid helps produces reasonable approximations at high latitude. However, the NASA GISS method will generally do a poor job of representing any structures at scales less than a few million km^2. Not coincidentally, the natural scale for weather anomalies in the monthly average temperature field is about 2-3 million km^2. That is to say, NASA GISS chose their blurring scale to roughly match the typical feature size for weather anomalies. This is sensible from the point of view of needing to extrapolate over regions where no coverage is available. However, it also means that the method will never capture finer structure even if densely sampled data were available. As will be discussed in a future memo, the globally complete field allows GISS to have more accurate global averages than CRU, even though the blurring can make the GISS method less accurate than CRU at many individual sites.

Berkeley Earth

The Berkeley Earth method applies a Kriging-based interpolation using an inferred correlation vs. distance structure between observations. This allows extrapolations over regions with sparse data to be computed smoothly without sacrificing the ability to capture detail in regions that happen to be densely sampled. By its nature, Kriging allows for a continuously defined temperature field, though due to the limits of practical computation, the field is generally sampled on an equal area grid having a resolution of 1.6° at the equator (approximately twice as many total elements as the GISS grid). Figure 5 shows the effect of applying the Berkeley Earth method to the sampled visual field.

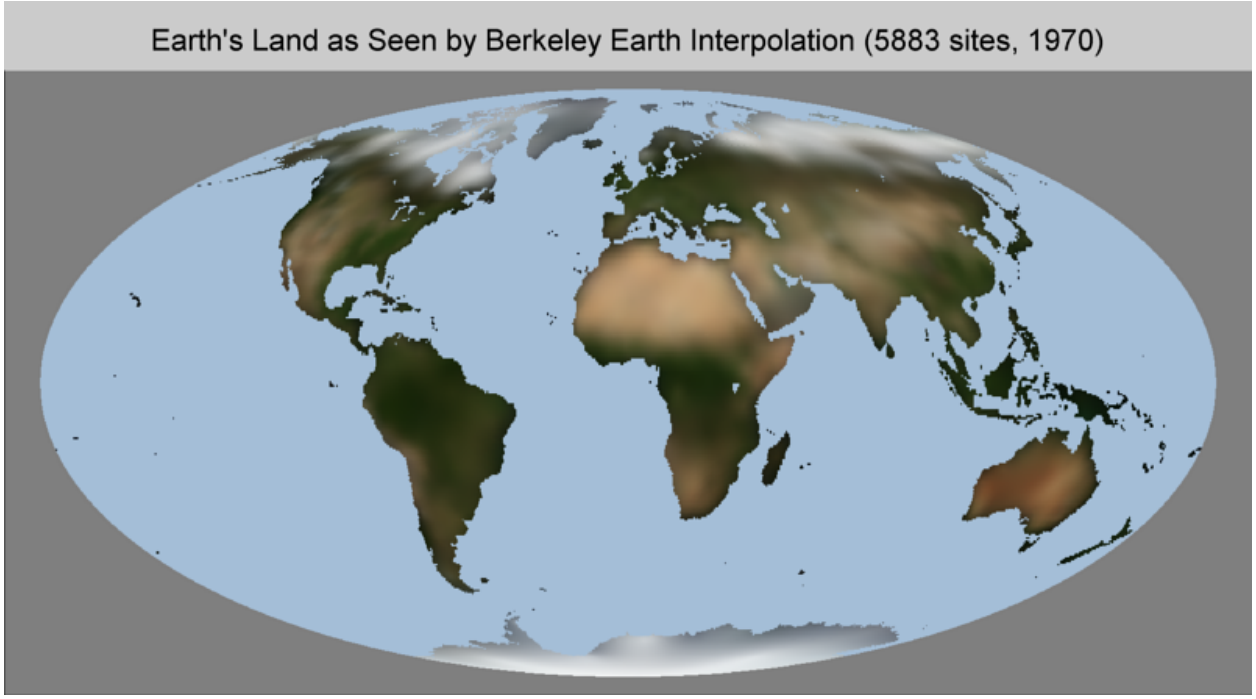

Figure 5: Interpolation from the sampled visual data using the Berkeley Earth methods. Like GISS, the resulting field is globally complete on this data; however, the Kriging approach makes it possible for more fine structure to be preserved when dense data is available.

Kriging, allows for a better set of trade-offs between local accuracy and global completeness than either the CRU or GISS methods. Extrapolations can be made when necessary, while local structure can also be captured when densely sampled data exists to make that possible. Nevertheless, there is still blurring in the global field since ultimately it is still based on only 5883 observations.

This figure provides an idea of the level of detail with which the global temperature field can be reconstructed from historical weather data. As previously stated, at monthly and longer time scales, it is generally observed that most weather variations have a feature size of more than 1 million km^2, so the remaining blur present in the Berkeley average is not expected to be a major limitation.

It is known mathematically that if the underlying data is normally distributed, and the correlation structure is accurately known, then Kriging (also known as Gaussian Process Regression) provides the best unbiased linear estimate of the underlying field. As discussed in the accompanying memo, if accurate data is available then the Berkeley Earth method will produce more accurate estimates of both local and global features than the GISS and CRU techniques.