6-1-2008

Components of information for multiple resolution comparison between maps that share a real variable

Robert Gilmore Pontius
Clark University, rpontius@clarku.edu

Olufunmilayo Thontteh
Regional Centre for Training in Aerospace Surveys

Hao Chen
Clark University

Components of information for multiple resolution comparison between maps that share a real variable

Authors

Robert Gilmore Pontius Jr, Olufunmilayo Thontteh, and Hao Chen

Clark University

School of Geography

Department of International Development, Community, and Environment

950 Main Street, Worcester MA 01610-1477, USA

PHONE 508 793 7761

FAX 508 793 8881

EMAIL rpontius@clarku.edu

Keywords

accuracy, error, MAE, raster, scale, RMSE.

Abstract

This paper presents quantitative methods that allow scientists to compare the patterns in two maps that show a shared real variable. Specifically, this paper shows how to budget various components of agreement and disagreement between maps. The components are based on the separation of a map's information of quantity from its information of location. The technique also examines how variation in resolution influences the measurement of the components of information. The manner in which the measurements change as a function of spatial resolution can be more important and interesting than the results at any single particular resolution, because the results at a single particular resolution may indicate more about the format of the data than about the overall pattern in the landscape. An example illustrates the mathematical concepts, and an application to compare mapped vegetation indices in Africa illustrates the usefulness of the proposed approach vis-à-vis a conventional approach. The results are presented visually in the form of stacked bar graphs that show separable components of information. The entire analysis is performed twice, each time with a different mathematical measurement of deviation: 1) Root Mean Square Error, and 2) Mean Absolute Error. This paper compares these two approaches and discusses their relative advantages and disadvantages. Hopefully, this approach of budgeting components of information at multiple resolutions will become adopted as standard practice in the measurement of patterns.

1. Introduction

1.1 The need to measure at multiple resolutions

Scientists face fundamental problems when measuring landscapes, because there is no natural or obvious spatial unit of analysis for most landscapes. Landscapes typically contain many types of patterns that numerous factors create by operating at multiple scales. Nevertheless, there is a need to format data concerning a landscape in units that facilitate the measurement and analysis of landscapes.

In many cases, data concerning landscapes are expressed in the form of raster maps, which consist of rows and columns of square units called pixels. The size of the pixel dictates the resolution of the digital information. The pixel is usually a function of the technology that generates the raster, whether the technology is a satellite or Geographic Information Science (GIS) software. The pixel is a natural unit of a digital map, but the pixel is not a natural unit of the landscape, meaning that real landscapes are not organized in terms of pixels. Square pixels do not dictate natural processes, and humans do not manage landscapes according to square pixels that are oriented along the flight paths of satellites.

In spite of this, there are a few related reasons why there is tremendous temptation on the part of applied scientists to treat the pixel as a unit of observation and to adopt the resolution of the raw data as the resolution of the applied analysis. Many applied scientists use data that are organized in terms of pixels, because these types of data are readily available and statistical techniques with accompanying software packages are designed to perform pixel-level analysis relatively easily. Some scientists are reluctant to reformat the data in a subjective manner, because any reorganization of the data may have a large influence on the conclusions drawn from such data, as illustrated by the well known modifiable areal unit problem (Openshaw 1984). However, scientists must appreciate that the available data are already formatted in some manner for some reason. Usually the reason is convenience of data collection or storage. Consequently, the unit of analysis and its resolution are frequently selected by default, as a function of issues such as the precision of the satellite or the capacity of the computer. Ultimately, this adoption by default can be worse than subjective selection. If there is a mismatch between the format of the available data and the resolution of the substantive question, then adoption of the format of the available data can be more dangerous than subjective modification of the format of the data, because a subjective decision about reformatting the data could be based on at least some knowledge about the phenomenon of interest and its relevant resolution.

There are insufficiently developed methods to guide scientists in how to rescale the data, if at all. Nevertheless, scientists are aware that there is a need to examine the influence of resolution in map comparison (Veldkamp et al. 2001).This issue is so important that the University Consortium on Geographic Information Science has articulated consistently that research priorities should include Scale and Representation (McMaster and Usery 2004).

This paper addresses this need directly. It offers quantitative methods that allow scientists to examine how resolution and representation influence statistical measurements. The basic approach is to examine how measurements change as a function of the resolution of the pixels. The strategy is to examine the data at many resolutions, and not to focus on any one particular resolution. The manner in which the measurements change as a function of resolution can be more important and interesting than the results at any single resolution, because the results at any single particular resolution may indicate more about the format of the data, than about the overall pattern in the landscape.

1.2 The need to budget components of information

This paper has a second purpose, which is to introduce a statistical approach that focuses on comparing maps in terms of components of information. Specifically, this paper shows how to budget various components of agreement and disagreement between maps. Applied scientists should find this approach helpful because it allows them to visualize important types of information that can explain the patterns in the data. For example, a map producer needs to know which types of errors are relatively more important in order to improve the process of map production. Therefore, it would be helpful to have a method to budget the sources of error for any particular mapping exercise. This paper’s proposed approach compliments other thoughtful techniques for visualization of spatial data, which are becoming increasingly possible, useful, and popular (Bailey and Gatrell 1995).

Over the last few years, Pontius (2000) has been developing an approach to statistical analysis that focuses on budgeting components of information during the comparison of two maps that share the same categorical variable. Pontius (2002) extends the approach to include multiple resolutions, while Pontius and Suedmeyer (2004) extend the approach to allow for spatial stratification. The present paper describes analogous methods for a real variable. One of the goals of the present paper is to establish a philosophy of map comparison that unifies the methods for a categorical variable and for a real variable.

1.3 The need for accessibility

If this new approach is to be adopted successfully, it must be accessible conceptually and mathematically to applied scientists who are thoughtful non-experts in statistics. This paper offers such scientists an approach for which the most complicated mathematical operation is a square root. Ultimately, this paper recommends a method for which the most complicated mathematical operation is an absolute value. This paper's proposed approach involves no calculus, no probability density functions, and no p-values. The approach is designed specifically to allow the results to be presented graphically. A graphical display is essential to facilitate interpretation, while it must be founded on sound mathematical principles. If this paper is successful, it will transform how scientists approach statistical analysis and how we teach statistical concepts to students.

This paper proposes a statistical approach that is fundamentally different than the approach of hypothesis testing that continues to be taught to millions of statistics students. Hypothesis tests rely on integral calculus and/or combinatorics to examine whether randomness can explain patterns in data. For many applications, comparison to randomness is not an interesting question, because the scientist already knows that the patterns in the data are not random, while the scientist still has important questions concerning the patterns. For example, even if a hypothesis test shows that randomness can not explain the patterns in a map, a map producer would still want to understand the patterns in the map in such a way that would be useful to improve the map production process. Furthermore, any approach that relies on a clearly defined unit of observation, such as hypothesis testing, is questionable for an application where the phenomenon of interest has no natural unit of observation. A hypothesis test's p-value can be extremely sensitive to the number of observations, but for most GIS applications, the number of so-called observations (e.g. number of pixels) indicates more about the format of the data than about the character of the landscape. Therefore, p-values can be worse than unhelpful; p-values can be misleading when there is no natural unit of analysis, because they can indicate artifacts due to the data storage mechanism, rather than the patterns in the real world. A portion of the community of applied scientists has been weary of the hypothesis testing paradigm for quite some time (Gaile and Willmott 1984). This paper offers a statistical approach that gives scientists an alternative technique of quantitative analysis, in order to complement or replace more conventional statistical approaches, thus is answers directly the calls for new methods that are designed specifically for spatial analysis and GIS (Unwin 1996). The sections below explain the methods by using both an example that the reader can compute by hand and an application that illustrates the usefulness for a practical problem in environmental science.

2 Methods

2.1 Data for example

It is easiest to grasp the concepts and subsequent equations in the context of example maps, such as the ones shown in figures 1 and 2. The purpose of the figures is to illustrate the logic of the method to compare any two maps that show a single real variable.

[Insert figure 1 here.]

[Insert figure 2 here.]

Figure 1 gives the raw example data. The top of figure 1 gives two maps called X and Y. Both maps consist of sixteen pixels arranged in 4 rows and 4 columns. The pixels of map X consists of the sequence of integers -8, -7, … , -1 in the west and the sequence 1, 2, …, 8 in the east, thus the average of all pixels in map X is zero. The pixels of map Y consist of various even integers in the interval [-4, 8] such that the average of the pixels in map Y is 1. The maps are organized into two strata, as defined by the thick dashed line that vertically bisects the maps. Stratum 1 is in the west and stratum 2 is in the east. The bottom of figure 1 shows the membership of each pixel to each of the strata, where a membership of 1 means the pixel belongs completely to the stratum and a membership of 0 means that the pixel does not belong at all to the stratum. These memberships are weights that dictate the influence of each pixel on the analysis, so the weight could be any non-negative real number for other applications.

Figure 2 introduces the notation and equations that show how to convert the raw data to coarser resolutions, whereas subsection 2.3 gives the details of the notation and equations. The left side of figure 2 shows the notation for the pixels in map X at each of three resolutions. The top-left map is the notation for the raw data at the fine resolution. The middle-left map gives the equations for the middle resolution, which is generated by taking a weighted average of each cluster of four neighboring pixels, using the weights illustrated in the bottom of figure 1. The bottom-left map shows the equation for the information about map X when it is converted to one large pixel that contains the entire study area. The values for map Y are computed for multiple resolutions in an identical manner. The maps on the right side of figure 2 show the difference between maps Y and X. The top-right map is Y-X at the fine resolution of the raw data. The middle-right map is Y-X at the middle resolution, where all of the pixels are completely nested in exactly one of the two strata. The bottom-right map is Y-X at the coarsest resolution where the entire analysis is contained within one coarse pixel, while the two strata are maintained. If there were no stratification (i.e., if the entire analysis were to consist of exactly one stratum), then the value in the coarsest pixel of figure 2 would be 1, because the average of the raw resolution pixels in map Y is 1 and the average of the raw resolution pixels in map X is 0.

2.2 Logic of analysis

It is easiest to explain the logic of the analysis if the reader envisions map X as reference data, meaning that X is assumed to have high accuracy. In practice, map X could be the ground information or the truth data to which another map is compared. It is helpful to envision that map X shows the mass of a substance, in which case map X contains the perfectly accurate overall quantity of the mass and shows the location of that mass distributed accurately in space at the precision of the resolution of the raw data. In this respect, map X has two types of perfect information: 1) perfect information concerning the quantity of the mass and 2) perfect information concerning the location of the mass.

Map Y is any other map that is compared to map X, with the condition that map Y expresses the same real variable that map X shows. For example, map Y could be information from a satellite, a prediction from a simulation model, a map from some previous point in time, or a map produced from an alternative cartographic technique. It is helpful to envision that map Y shows a prediction of the mass of the same substance that map X shows.

This paper focuses on two important respects in which map Y can differ from map X. These two are: 1) information of quantity, and 2) information of location. If map Y has the same total quantity of mass as map X, then the information of quantity in map Y is perfect by definition. In addition, if the spatial distribution of the mass within map Y is identical to the spatial distribution within map X, then the information of location in map Y is perfect by definition. In general, the total quantity of mass in map Y can be different than in map X, and the mass' spatial distribution within map Y can also be different than it is within map X.

We use the word “medium” to describe the types of information that map Y actually displays. Thus the quantity of mass in map Y is the medium quantity, and the manner in which the mass is distributed spatially within map Y is called the medium location. A more naïve version of map Y would have a “null” level of information of the quantity of mass, which would be the mass that the map maker would have assumed in the absence of better information. This null information of quantity is not necessarily found within map Y.

Figures 3 through 6 illustrate the logic of the analysis in terms that describe two components of information: 1) information of quantity, and 2) information of location. All four figures have the same organization in columns and rows that are aligned along orthogonal axes that show the accuracy of the information. The three columns are aligned along the horizontal axis, which shows perfect information of quantity on the left and worse information concerning quantity toward the right. The five rows are aligned along the vertical axis, which shows perfect information of location at the bottom and worse information towards the top. Thus perfect information exists at the origin of the space, and information becomes less perfect farther from the origin. Figure 3 shows seven maps in this space to illustrate seven important components of information. Each map shows how the pattern would appear, if it were to have the combination of information designated by its position in the space. This subsection describes the logic of the sequence of the seven maps as they emanate from the origin to the upper right corner of the information space.

[Insert figure 3 here.]

[Insert figure 4 here.]

[Insert figure 5 here.]

[Insert figure 6 here.]

The map closest to the origin at the lower left of the information space in figure 3 shows a map in the perfect information of quantity column and the perfect global information of location row. This map is map X given in figure 1, where darker pixels show larger numbers. Map X resides at the origin of the information space, because map X has perfect information of both quantity and location by definition. It is important to compare the other six maps in figure 3 to map X at the origin.

The next map to the right of map X within figure 3 is in the medium information of quantity column and the perfect global information of location row. It has the same spatial pattern as map X, which is why it is in the bottom row according to the information of location axis. It is equal to map X plus a constant where the constant is the overall average quantity in map Y minus the overall average quantity in map X. The effect of adding the constant is to modify map X so that the result adopts the same overall average quantity as map Y. All the maps in the middle column have the same overall average quantity as contained in map Y, which is why the middle column is called medium on the information of quantity axis.

Next, we begin to climb up the middle column of the sequence of maps in figure 3. As we ascend, each subsequent map has less accurate information of location with respect to map X. Let us examine the map in the medium information of quantity column and the perfect stratum information of location row. It has the same visual pattern as map

X within each stratum, because it is a modified version of map X, whereby all the pixels within each stratum are shifted by a constant. In our example, a constant of 2.75 is added to the pixels in the western stratum, and 0.75 is subtracted from the pixels in the eastern stratum. The constants are selected to make the average in each stratum equal to the corresponding average in each stratum of map Y. Consequently, the modified map has the same overall quantity of mass as map Y, thus it is in the medium information of quantity column.

Map Y resides at the center of figure 3. It is in the medium information of quantity column and the medium pixel information of location row. The visual pattern within each stratum of map Y does not match map X at the pixel level, nevertheless there is some identifiable positive correlation between the spatial pattern in map Y and map X for our example.

As we continue to climb up the middle column of the sequence of maps in figure 3, the next map above map Y is in the medium information of quantity column and the uniform stratum information of location row. This map is created by modifying map Y in a manner that redistributes the mass within each stratum of map Y uniformly within the stratum. This eliminates pixel-level details concerning the information of location within the strata, but it maintains stratum-level information of location between the strata. Consequently, each pixel has a value of -1.75 in the western stratum and a value of 3.75 in the eastern stratum for our example.

The next map at the top center of figure 3 is in the medium information of quantity column and the uniform global information of location row. This map is created by redistributing the mass of map Y uniformly over the entire space. This homogenizes the information of location within the map, but maintains the medium information of quantity that map Y displays. For our example, every pixel in this map has a value of 1, which reflects the difference between the overall average of map Y and the overall average of map X.

The final map in the upper right within figure 3 is in the null information of quantity column and the uniform global information of location row. This map is created by distributing the null quantity uniformly over the entire map. Every pixel in this map has a value of 6, which reflects the assumed null information of quantity for the example.

Figure 4 shows seven scatter plots that correspond to the seven maps in figure 3. Each of the seven plots in figure 4 shows 16 points, which relate to the 16 pixels of the fine resolution maps in figure 3. Also, each plot shows the line Y=X for reference, along with the X and Y axes. For all of the scatter plots, the fundamental question is “How close are the plotted points to the line Y=X?” At the origin of the information space, all the points are exactly on the line Y=X because there exist perfect information of both quantity and location at that position in the information space. We see a shift up in the plotted points as we move to the medium column on the information of quantity axis, which reflects the fact that the overall mass in map Y is different than the overall mass of X. The slope of the plotted points is 1 for the plots at the bottom of figure 4, because there is perfect information of pixel-level location in that region of the information space. The points migrate farther from a slope of 1 as we climb the central column in figure 4 because there is less information of location the farther we ascend the column. The central scatter plot compares directly map X to map Y. The slope of the points is zero at the top of the figure where there is uniform information of location.

Measures of goodness-of-fit can be computed for all of the scatter plots in figure 4, in order to address the question "How close are the plotted points to the line Y=X?" Two conventional measures are Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Figure 5 gives the equations to compute RMSE from the raw data for each of the seven scatter plots at the respective positions in the information space. Similarly, figure 6 gives the equations to compute MAE from the raw data for each of the seven scatter plots.

Ultimately, each sequence of seven maps in figure 3 gives a sequence of seven measures of goodness-of-fit, as we step through the sequence from the origin to the upper right position in the information space. Each additional step in the sequence is likely to show a worse fit, because each additional step worsens some type of information. Therefore, we can measure the importance of each component of information by seeing how the measured deviation increases as we step through the sequence. The next subsection gives the details of the calculations to use the mathematical expressions in figures 5 and 6 to compute components of disagreement and agreement between map X and map Y.

2.3 Notation of analysis

This subsection defines the notation in figures 2, 5, 6, and the remainder of this paper. Both maps X and Y are georeferenced to the same raster of pixels. Let r denote the resolution of the information with respect to the raw data in a manner such that the resolution of the raw data is denoted as r = 1, and coarser resolutions can be denoted as r = 2, …, R, where R is the maximum of the raster's number of rows and number of columns. Each increasingly coarse resolution is created by aggregating square clusters of neighboring pixels, as shown in figure 2. At each increasingly coarse resolution, r is a multiple of the side of a pixel of the raw data. The maximum possible resolution is the resolution at which the entire study area is in one pixel, which is denoted as r = R. If stratification is relevant, then the entire analysis can be performed by stratum, where each pixel has some quantifiable membership to each stratum as illustrated in figure 1.

The proposed method relies on the following terms:

r = resolution of the information as a multiple of the side of a pixel of raw data,

R = maximum resolution,

e = index for strata,

E = number of strata,

Nre = number of pixels in stratum e at resolution r,

Wren = weight for pixel n in stratum e at resolution r,

Xren = reference value for pixel n in stratum e at resolution r,

Yren = comparison value for pixel n in stratum e at resolution r.

The weight is constrained such that $0 \leq$ Wren, and usually also Wren ≤ 1.

Equations 1 and 2 compute the average for each stratum e for each variable. Equation 3 computes the global average of all Yren, denoted by $\hat{Y}$. Note that these averages do not change with resolution, therefore the left hand side of equations 1, 2 and 3 have no subscript r.

$$\overline{X}e = \frac{\sum_{n=1}^{Nre}(Wren \times Xren)}{\sum_{n=1}^{Nre} Wren} \qquad \text{equation 1}$$

$$\overline{Y}e = \frac{\sum_{n=1}^{Nre}(Wren \times Yren)}{\sum_{n=1}^{Nre} Wren} \qquad \text{equation 2}$$

$$\hat{Y} = \frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}(Wren \times Yren)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre} Wren} \qquad \text{equation 3}$$

2.4 Equations for components based on RMSE

There are two popular techniques to measure the average deviation in the comparison of X versus Y: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). This paper uses both techniques in order to contrast them. We present the RMSE first, because it tends to be more popular among statisticians. The next subsection presents analogous equations for MAE.

Throughout this subsection, keep in mind that X is considered accurate. Therefore, any deviation between Y and X is attributable to error in Y as measured by the vertical distance between Y and the line Y=X. Figure 5 gives the mathematical expressions to compute RMSE for each of the scatter plots in figure 4, so this subsection does not present those expressions directly. Instead, this subsection presents differences between sequential pairs of those expressions in figure 5 in order to measure the additional deviation that is accumulated at each step through the sequence from the origin to the upper right corner of the information space. Each segment of additional deviation is a component of some type of information, denoted by a three letter abbreviation. The first letter of each component in this subsection is S, because squared residuals form the mathematical foundation to compute each component of information in this subsection. The second letter is either D to denote disagreement or A to denote agreement. The third letter denotes the type of information of location or information of quantity.

The first component of information is the disagreement due to quantity (SDQ). This component does not change at multiple resolutions because information of quantity is independent of information of location, as indicated by the orthogonality of the axes in figures 3-6. SDQ is equivalent to the average of all Y pixels minus the average of all X pixels. Equation 4 presents it in the form of RMSE in order to allow for direct comparison to the other equations.

$$SDQ = \sqrt{\left[\frac{\sum_{e=1}^{E} \sum_{n=1}^{Nre} \left(Wren \times \{ Yren - Xren \} \right)}{\sum_{e=1}^{E} \sum_{n=1}^{Nre} Wren} \right]^2} \qquad \text{equation 4}$$

The second component of information in the sequence is the disagreement due to stratum-level location (SDS). This component does not change at multiple resolutions because the quantity within each stratum is independent of the pixel resolution, and the stratification does not change with resolution. Equation 5 computes this component of additional deviation by subtracting the previous component (SDQ) from the RMSE shown by figure 5 in the perfect global information of location row and the medium information of quantity column.

$$SDS = \sqrt{\sum_{e=1}^{E}\left[\frac{\sum_{n=1}^{Nre}\left(Wren \times \left\{Yren - Xren\right\}\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre} Wren}\right]^{2}} - SDQ \qquad \text{equation 5}$$

The third component of information is disagreement due to pixel-level location (SDPr), where the subscript r indicates the resolution of the pixels. It is necessary to denote the resolution r because SDPr can change with modification of resolution. Equation 6 computes this component at resolution r by computing the RMSE for the direct comparison between map X and map Y, then subtracting the two previous components of disagreement due to stratum-level location (SDS) and disagreement due to quantity (SDQ).

$$SDPr = \sqrt{\frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(Wren \times \left[Yren - Xren\right]^{2}\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre} Wren}} - SDS - SDQ \qquad \text{equation 6}$$

Notice that the square root part of equation 6 is the RMSE that compares the pixel-level X values to the pixel-level Y values. Therefore, we can express this total RMSE as the sum of three separable components of disagreement: SDPr, SDS, and SDQ. As we continue to climb up and to the right through the information space in figures 3-6, the subsequent components of information indicate agreement, as opposed to disagreement.

The fourth component of information is agreement due to pixel-level location (SAPr), where the subscript r indicates the resolution of the pixels, since this component can change with modification of resolution. This component compares the goodness-of-fit of the pixel-level X data with the pixel-level Y data to the goodness-of-fit that one would observe if the mass of the Y variable within each stratum were spread uniformly among the pixels within each stratum of map Y. If the pixel-level values of X and Y are strongly positively associated within each stratum, then this component of agreement due to pixel-level information is positive. However, it is possible that the pixel-level values of X and Y are not positively associated within each stratum, in which case equation 7 defines the component of agreement due to pixel-level location to be zero.

$$\begin{aligned} \mathrm{SAPr} &= \sqrt{\frac{\sum_{e=1}^{E} \sum_{n=1}^{Nre} \left(\mathrm{Wren} \times \left[\overline{\mathrm{Y}}\mathrm{e} - \mathrm{Xren} \right]^2 \right)}{\sum_{e=1}^{E} \sum_{n=1}^{Nre} \mathrm{Wren}}} - \mathrm{SDPr} - \mathrm{SDS} - \mathrm{SDQ} \quad \text{if positive} \\ &= 0 \qquad \text{else} \qquad\qquad \text{equation 7} \end{aligned}$$

The fifth component of information is agreement due to stratum-level location (SASr), where the subscript r indicates the resolution of the pixels. This component is positive if the stratum-level averages for X and Y are strongly positively associated among the strata. If this is not the case, then the agreement due to stratum-level location is zero.

$$\mathrm{SASr} = \sqrt{\frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(\mathrm{Wren} \times \left[\hat{Y} \quad - \quad \mathrm{Xren}\right]^{2}\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\mathrm{Wren}}} - \mathrm{SAPr} - \mathrm{SDPr} - \mathrm{SDS} - \mathrm{SDQ} \quad \text{if positive}$$

$$= \quad 0 \qquad \text{else} \qquad\qquad \text{equation 8}$$

If it exists, $\tilde{Y}$ denotes the null estimate for quantity. $\tilde{Y}$ can be used to compute a sixth component of agreement due to quantity (SAQr), where the subscript r indicates the resolution of the pixels. If the medium information of quantity is more accurate than the null information of quantity, then the component of agreement due to quantity is positive as computed by equation 9. If the opposite is true, then equation 9 defines the component to be zero.

$$\mathrm{SAQr} = \sqrt{\frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(\mathrm{Wren} \times \left[\tilde{Y} \quad - \quad \mathrm{Xren}\right]^{2}\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\mathrm{Wren}}} - \mathrm{SASr} - \mathrm{SAPr} - \mathrm{SDPr} - \mathrm{SDS} - \mathrm{SDQ} \quad \text{if positive}$$

$$= \quad 0 \qquad \text{else} \qquad\qquad \text{equation 9}$$

2.5 Equations for components based on MAE

This subsection follows the same logic as the previous subsection, but this subsection uses the MAE as the measurement of error as opposed to the previous subsection that uses RMSE. Therefore, the first letter in the abbreviation for each component of information is A, which signifies that absolute residuals serve as the mathematical foundation of the calculation. The second and third letters of the abbreviation are the same as in the previous subsections. The procedure computes first the components of disagreement then the components of agreement, by stepping through the sequence of mathematical expressions in figure 6. The typical case is that each subsequent mathematical expression in figure 6 gives a measurement of deviation that increases as we move from the origin to the upper right corner of the information space. Each additional increase in deviation constitutes a component of disagreement or agreement.

The first component of information is the disagreement due to quantity (ADQ). This component does not change with resolution because overall quantity is not a function of resolution. Notice that the component of disagreement due to quantity that is based on MAE is equivalent to the corresponding component that is based on RMSE, since equation 10 is equivalent to equation 4. Disagreement due to quantity is the only component of information for which the measurement based on MAE is mathematically equivalent to the measurement based on RMSE.

$$\mathrm{ADQ} = \left| \frac{\sum_{e=1}^{E} \sum_{n=1}^{Nre} \left(\mathrm{Wren} \times \{ \mathrm{Yren} - \mathrm{Xren} \} \right)}{\sum_{e=1}^{E} \sum_{n=1}^{Nre} \mathrm{Wren}} \right| \qquad \text{equation 10}$$

The second component of information is the disagreement due to stratum-level location (ADS), as equation 11 indicates. This component does not change with resolution.

$$\mathrm{ADS} = \sum_{e=1}^{E} \left| \frac{\sum_{n=1}^{Nre} \left(\mathrm{Wren} \times \{ \mathrm{Yren} - \mathrm{Xren} \} \right)}{\sum_{e=1}^{E} \sum_{n=1}^{Nre} \mathrm{Wren}} \right| - \mathrm{ADQ} \qquad \text{equation 11}$$

The third component of information is the mean absolute disagreement due to pixel-level location at resolution r (ADPr), given by equation 12. It has a subscript of r because it can vary with resolution.

$$ADPr = \frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(Wren \times \left|Yren - Xren\right|\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre} Wren} - ADS - ADQ \qquad \text{equation 12}$$

The fractional part of the right hand side of equation 12 is the total MAE for the comparison between the pixels of maps X and Y. This total MAE is the sum of the three separable components of disagreement: ADPr, ADS, and ADQ.

The fourth component of information is the agreement due to pixel-level location at resolution r (AAPr), given by equation 13. This component can vary as a function of resolution. If the pixels of X and Y are not strongly positively associated within the strata, then equation 13 defines the component of agreement due to pixel-level location to be zero.

$$AAPr = \frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(Wren \times \left|\overline{Y}e - Xren\right|\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre} Wren} - ADPr - ADS - ADQ \quad \text{if positive}$$

$$= 0 \qquad \text{else} \qquad \text{equation 13}$$

The fifth component of information is the average absolute agreement due to stratum-level location (AASr), which can vary with resolution. If the stratum-level averages for maps X and Y are not strongly positively associated, then equation 14 defines the component of agreement due to stratum-level location to be zero.

$$\mathrm{AASr} = \frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(\mathrm{Wren} \times \left|\hat{\mathrm{Y}} - \mathrm{Xren}\right|\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\mathrm{Wren}} - \mathrm{AAPr} - \mathrm{ADPr} - \mathrm{ADS} - \mathrm{ADQ} \quad \text{if positive}$$

$$= 0 \qquad \text{else} \qquad \text{equation 14}$$

If it exists, then the sixth component of information is agreement due to quantity (AAQr), as given by equation 15. Its existence requires a null estimate of the overall quantity of Y, denoted as $\tilde{\mathrm{Y}}$.

$$\mathrm{AAQr} = \frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(\mathrm{Wren} \times \left|\tilde{\mathrm{Y}} - \mathrm{Xren}\right|\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\mathrm{Wren}} - \mathrm{AASr} - \mathrm{AAPr} - \mathrm{ADPr} - \mathrm{ADS} - \mathrm{ADQ} \quad \text{if positive}$$

$$= 0 \qquad \text{else} \qquad \text{equation 15}$$

2.6 Application to environmental science

This subsection presents an application of the methods proposed in the previous subsections to a practical case study in environmental science in order to illustrate how the proposed approach compares to a conventional approach. Figure 7 shows two maps for a section of Southeastern Africa. The white lines show the country borders to help to orient the reader and to delineate the Indian Ocean, which is masked from the analysis. There are 49976 pixels in the study area for each of the maps in figure 7. Each pixel on the land shows the amount of vegetation, where darker shades indicate more vegetation. The underlying data derive from the Advanced Very High Resolution Radiometer (AVHRR), which is a sensor that collects information via satellite in pixels that are 8 kilometers on a side. The reference map (X) is the vegetation as observed by the satellite and the comparison map (Y) is the vegetation predicted by an extrapolation model that is being developed by Eastman (personal communication). The amount of vegetation in each pixel is expressed as a z-score that gives the deviation of the 2003 growing season from the long term average for the Normalized Difference Vegetation Index (NDVI). The long term average consists of the 18 years from 1985 to 2002, and the growing season consists of the months January through May. For each pixel, a positive z-score indicates more vegetation in 2003 compared to the previous 18-year average, and a negative z-score indicates less vegetation in 2003 compared to the previous 18-year average. The extrapolation model is calibrated with NDVI data from 1985 to 2002. A null model would predict no variation from the long term average, hence would predict a z-score of zero for every pixel, thus the null quantity is zero for this application.

[Insert figure 7 here].

The applied substantive question concerns famine, because Southeastern Africa experiences intense droughts that can lead to severe food shortages during some types of El Niño events. For this case study, agencies that manage food security are the decision makers who would like to know a few months before the growing season whether there will be an unusually low level of primary production in Southeastern Africa. If these agencies can trust a model that predicts plant productivity, then they can prepare famine relief supplies with confidence. Therefore, the practical applied question is "How accurately does the extrapolation model predict variation in vegetation for a particular year?" Most importantly, the food security agencies need to know whether or not a particular year will be above or below average in terms of overall quantity of vegetation. In addition, agencies might want to know the likely variation in terms of the general location of vegetation. There does not exist a unique relevant spatial resolution for these questions; however it is clear that that the resolution of 8-kilometer pixels is not particularly important and coarser resolutions may be suitable for this application. Hence this application is perfectly suited to the proposed method, which compares the maps at multiple resolutions in terms of information of quantity and information of location. Ultimately, decision makers would like for scientists to help them decide how to interpret the accuracy of the prediction appropriately.

Figure 8 shows an obvious first step in the analysis, which is to plot the predicted vegetation versus the observed vegetation, where each point in the figure corresponds to a position of a pixel in the maps of figure 7. The points are clustered on the negative side of the horizontal axis which means that 2003 experienced an unusually low amount of vegetation relative to the long term average. The cluster is mostly on the negative side of the vertical axis, which means that the extrapolation model predicted an unusually low amount of vegetation for 2003. Furthermore, the cluster is centered below the one-to-one line, which means that extrapolation model predicted that there would be less vegetation than the amount actually observed in the reference map. Recall that a null model predicts a zero z-score for each pixel, so a null model would produce points that reside exclusively on the horizontal axis of figure 8.

[Insert figure 8 here].

This paper examines the maps of figure 7 at multiple resolutions by assessing a plot similar to figure 8 for each resolution. An averaging algorithm aggregates the fine resolution pixels into coarser resolution pixels in order to transform the data in a manner identical to the example in figure 1.

We contrast a conventional statistical approach with the proposed approach in order to illuminate the important differences. A conventional approach fits a least squares line through the plotted points and computes confidence intervals around the slope of the line. The proposed approach generates a budget of components of agreement and disagreement concerning the information of quantity of vegetation and the information of location of vegetation. The country-level stratification in figure 7 is ignored in order to allow for simple direct comparison between a conventional approach and the proposed approach. The following Results section gives the output for this case study, immediately after the output for the example.

3 Results

3.1 Results for example

Figures 9 and 10 present the results for the comparison between the maps in figure 1, using the RMSE and MAE respectively on the vertical axis. The horizontal axis indicates resolution changing from fine to coarse, thus the bar on the left shows the results at the resolution of the raw data and the bar on the right shows the results when the entire study area is in one coarse pixel.

[Insert figure 9 here].

[Insert figure 10 here].

The total error in the direct comparison of map X to map Y is the sum of the bottom three components of disagreement. The remaining components show agreement attributable to either information of location or information of quantity. Information of location can derive from the pixel-level information or from the stratum-level information, since both pixels and strata express spatial distribution.

Disagreement due to quantity is identical for RMSE and MAE, because the respective mathematical equations for this component are identical, as illustrated by the bottom section of each bar. Working our way up the bar, the next component is disagreement due to stratification, which independent of resolution, as illustrated by the fact that it does not change as the resolution varies from fine to coarse. The next two components of information are disagreement due to pixel-level location and agreement due to pixel-level location, which are both positive at the fine resolution. Disagreement due to pixel-level location is converted into agreement due to pixel-level location as errors of location are resolved in the conversion from the fine to the middle resolution. The overall importance of pixel-level information of location shrinks as resolution becomes coarser, until both agreement and disagreement due pixel-level location are zero at the coarsest resolution. The next component is agreement due to stratum-level location. Both components associated with stratification remain positive as resolution changes, because the information that the stratification expresses is independent of the resolution of the pixels. The final component is agreement due to quantity. Both figures show a positive component of agreement due to quantity, based on an assumption that the null quantity is 6, meaning that each pixel in the null map has a value of 6.

3.2 Results for environmental application

A conventional approach focuses on confidence intervals concerning the slope of the least squares line. Figure 11 gives the slope of the least squares line and 95 percent confidence intervals around it as a function of resolution, which progresses from the fine resolution of the raw data where there are 49976 pixels to a very coarse resolution where four pixels contain the entire study area. The number of pixels decays exponentially as resolution changes from fine to coarse, so the width of the confidence interval grows correspondingly. For nearly all of the resolutions, the slope of the least squares line is positive and significantly different than zero, which means the variation in observed vegetation is more closely associated with the extrapolation model's prediction map than with a map that would be derived from a random rearrangement of the pixels in the prediction map. The positive slope of the least squares line is due mainly to inaccurately predicted pixels that cause outliers that are far from the one-to-one line in figure 8. The slope is statistically significantly different than zero and different than one at fine resolutions because of the large number of pixels. The statistical significance suggests that the relationship is somehow important; however the R-squared is less than 3 percent. The results become less stable and less certain as resolution changes from fine to coarse. Slope and R-squared are zero for the null model at all resolutions, because the null model specifies Y = 0 for all pixels.

[Insert figure 11 here].

[Insert figure 12 here].

[Insert figure 13 here].

The proposed approach quantifies components of information concerning the quantity and the location of the vegetation in the maps. Figure 12 gives the results from the proposed approach using RMSE as the measurement of deviation, whereas figure 13 gives the results using MAE. In these figures, the vertical axis shows the measurement of deviation between the maps and the horizontal axis shows the resolution changing from fine to coarse. The components give crucial information to answer important aspects of the practical question, “How accurately does the extrapolation model predict variation in vegetation for a particular year?” Each component addresses an important aspect of the answer, so we examine each of the components in sequence from bottom to top.

Disagreement due to quantity is the first component at the bottom of the figures. It has a value of 0.2 for both RMSE and MAE at all resolutions, because the average z-score predicted by the extrapolation model is -0.7 and the average z-score observed is -0.5. This first component answers the most important aspect of the practical question concerning the map comparison, because it measures the degree to which the extrapolation model predicts an overall amount of vegetation that is different than the overall amount of vegetation observed in the reference map.

Disagreement due to location is the second component, which is stacked on top of the first component in figures 12 and 13. The sum of these first two components of disagreement is the overall average deviation between the maps. The disagreement due to location indicates the severity of the errors in terms of their spatial distribution. The fact that the disagreement due to location is positive means there is room for improvement for the extrapolation model to predict the spatial distribution of the vegetation more accurately than it did, given the quantity that the extrapolation model predicted. The disagreement due to location is larger than the disagreement due to quantity at the fine resolutions, while the opposite is true at coarse resolutions since disagreement due to location shrinks to zero at the coarsest resolution. Section 4.4 interprets the rate of shrinkage.

The component of agreement due to location is zero according to both RMSE and MAE, therefore this component does not appear in figures 12 and 13. This result answers an aspect of the practical question concerning the extrapolation model, because it indicates that the extrapolation model has gained no accuracy by its attempt to predict the spatial distribution of the vegetation in a non-uniform manner. In other words, if the extrapolation model would have predicted its average value of -0.7 in every pixel, then the prediction would have been more accurate than the prediction in figure 7.

Lastly, the component of agreement due to quantity shows how the extrapolation model compares to a null model that predicts a z-score of zero in every pixel. Agreement due to quantity is positive if and only if the prediction model is more accurate than the null model. Figure 12 shows that the extrapolation model is more accurate than the null model at resolutions coarser than 32 kilometers, since there is a positive component of agreement at resolutions coarser than 32 kilometers. Conversely, the null model is more accurate than the extrapolation model at resolutions finer than 32 kilometers, since the component of agreement is zero at resolutions finer than 32 kilometers. The null resolution is defined as the resolution at which the accuracy of the null model equals the accuracy of a prediction model, therefore 32 kilometers is the null resolution according to RMSE (Pontius et al. 2004). Figure 13 shows that the extrapolation model is more accurate than the null model at all resolutions, therefore the null resolution does not exist according to MAE.

4 Discussion

4.1 Interpretation for environmental application

The overall purpose of the African case study is to advise food security organizations whether to prepare for famine relief. The decision makers need to know whether the general region is likely to experience to famine, thus there is no single natural unit of observation for this application. Nevertheless, data to forecast variation in vegetation are available in the particular format in which the satellite collects the information. This format is a function of the satellite, and has nothing to do necessarily with the scales of El Niño, drought, vegetation, or famine. For our example, the resolution of the raw data is 8 kilometers per pixel side, which is probably finer than the level of detail for the substantive questions. Nevertheless, scientists want to use the most detailed data available to run predictive models, which is understandable and justified. However, if the resolution of the raw data is different than the resolution of the substantive questions, then scientists should assess the performance of the model at other resolutions, in addition to the single resolution of the raw data. This paper's proposed approach offers scientists a useful technique to compare maps in an interpretable manner that frees the analysis from commitment to any one specific unit of analysis or resolution. The proposed approach addresses directly many important aspects of the answer to the substantive question and quantifies the information the human eye can see in figure 7. The predicted year was a low vegetation year, and the model predicted that it would be a low vegetation year; in fact, the model predicted that it would be lower than it actually was. The component of disagreement due to quantity measures this error of overall quantity. The component of disagreement due to location decreases as resolution becomes coarser, which indicates the spatial distribution of the errors, as section 4.4 describes. The zero agreement due to location shows that the prediction would have been more accurate if it were to have allocated the vegetation uniformly in space; by this criterion, the extrapolation model is not reliable in terms of the spatial allocation of the prediction. The positive component of agreement due to quantity at coarse resolutions indicates that the extrapolation model is more accurate than a null model that predicts no change from the long term average.

The conventional analysis shown in figure 11 fails to give any information whether the extrapolation model predicted more or less than the amount of observed vegetation, so it can not answer the most important applied question. This is due in part to the fact that the conventional approach has been designed for situations where X and Y are two completely different real variables, not for this case where both X and Y show different expressions of a single shared variable. The slope of the least squares line is statistically significantly different than zero and different than one at fine resolutions because of the large number of pixels, which is an artifact of the format of the data and has nothing to do with El Niño, drought, vegetation, or famine. For this case study, the pixels are units of convenience, which are not directly related to the phenomenon of interest. However, conventional methods have been developed for situations where the units of analysis have substantive meaning, so it treats the transformation of the pixels from fine resolution to coarse resolution as a process that renders the information less certain as indicated by the growth in the width of the confidence interval. Thus a conventional hypothesis testing paradigm can be misleading for multiple resolution analyses for which the units of analysis are not necessarily related directly to the relevant questions. Furthermore, a conventional approach compares the pattern in the map to a random distribution, which is not necessarily an appropriate or interesting null model.

4.2 Similarities between RMSE and MAE

There are many important similarities and differences between RMSE and MAE especially in the context of comparing observed X values to predicted Y values (Willmott 1981, Willmott 1982, Willmott et al. 1985). The top three rows of table 1 describe three important characteristics that RMSE and MAE have in common.

First, both RMSE and MAE can be used to budget components of disagreement and agreement between two maps. It is important to present results in a visual manner such as figures 9-10, 12-13, where it is easy to see the relative sizes of the components of information. This is essential in order to understand the additional information that would be necessary to improve the accuracy of map Y for cases where the research concerns accuracy assessment for map production. Hopefully, applied scientists will find this graphical format of presentation more immediately useful than the more conventional presentation of tables of regression line coefficients along with their accompanying p-values.

[Insert table 1 here].

The second row of table 1 indicates that the components of disagreement are mathematically identical when one compares X to Y as when one compares Y to X. This commutative property matches intuitive sense, especially because it is not immediately obvious which of the two maps should be selected as X or Y in some cases. This property is apparent in the bottom three rows of mathematical expressions in figures 5 and 6.

The third row of table 1 indicates that it can make a difference which of the two maps is considered X and which is considered Y when computing the components of agreement, because the expressions in the uniform stratum and uniform global rows of figures 5 and 6 are sensitive concerning which map is called X and which is called Y. This lack of a commutative property for components of agreement requires some thought in order to make intuitive sense. An example helps. Consider two maps A and B, each with two pixels and no stratification, such that map A has ordered pixel values of {0,4} and B has ordered pixel values {1,3}. Both maps have an average value of 2, so the corresponding uniform map is {2,2}. If map A is considered X and map B is considered Y, then the agreement due to location is 1, because there is more similarity between map A and map B than between map A and the uniform map. Conversely, if map B is considered X and map A is considered Y, then the component of agreement due to location is 0, because the similarity between map B and map A is equal to the similarity between map B and the uniform map. The spatial pattern in map B is more subtle than it is in map A, so map A's intense spatial pattern does not explain the pattern in map B any better than a uniform distribution explains the pattern in map B.

This is the phenomenon that explains why there is zero agreement due to location in the African case study. The predicted z-scores show more variation than the observed z-scores, so a uniform map gives a better fit than the predicted map to the observed z-scores.

4.3 Advantage of RMSE over MAE

Row 4 of table 1 indicates a major difference between RMSE and MAE. Namely, given map X and a fixed overall quantity of the mass in map Y, there exists a unique spatial pattern for map Y that would minimize RMSE, whereas there could be an infinite number of potential spatial patterns for map Y that would minimize MAE. For illustration, consider map Y in figure 1. If one were to rearrange the mass within map Y, then there would be a unique spatial arrangement that would minimize RMSE, with respect to map X. That spatial rearrangement would cause the residual Y-X in each pixel to be 1. The spatial pattern would be a perfect match visually. In fact, this is the pattern in figure 3 in the medium information of quantity column and the perfect global information of location row. However, there could be an infinite number of spatial rearrangements of the mass in map Y that would give the minimum possible MAE. Any rearrangement that results in all the residuals being non-negative would yield the minimum MAE of 1.

In general, the unique spatial pattern in map Y that minimizes RMSE is the pattern such that all the pixel-level residuals Y-X are identical. When this is the case, the visual appearance of the spatial pattern in map Y matches perfectly the spatial pattern in map X, meaning that map Y has perfect information of location with respect to map X, and the only difference between map Y and map X is a difference in information of quantity. MAE is at a minimum for any arrangement of the mass in map Y such that the residuals Y-X for the pixels are either all non-negative or all non-positive. If map X and map Y have the same overall mass (i.e. if map Y has perfect information of quantity), then there is a unique spatial arrangement that minimizes MAE, which is the same spatial arrangement that minimizes RMSE. If the total quantity of mass in map Y is different than the total quantity of mass in map X, then there can be an infinite number of spatial rearrangements of the mass in map Y that will give the minimum MAE. This property is related to the fact that MAE ≤ RMSE for each of the corresponding expressions in figures 5 and 6. MAE < RMSE when some residuals are larger than other residuals, because RMSE assigns a disproportionately large influence to large residuals. MAE = RMSE when all residuals are equal, which occurs when map Y has perfect information of location.

4.4 Neutral differences between RMSE and MAE

Row 5 of table 1 states that RMSE is more sensitive than MAE to outliers, which is related to the property described in the previous subsection. This is an important difference between RMSE and MAE, but it is not clear whether this is an advantage or disadvantage for applied work in general. In some cases, scientists want statistical methods to help to find outliers, because outliers can alert scientists to potentially important information. In other cases, outliers are a nuisance and do not indicate interesting signals in the data, in which case it is desirable to use a statistical technique that is not sensitive to outliers.

This difference in sensitivity to outliers explains why RMSE is also more sensitive than MAE to changes in resolution, as stated in row 6 of table 1. The components of information of location based on RMSE can shrink faster than those for MAE as resolution becomes coarser because the outliers dissolve at coarser resolutions. As resolution changes from fine to coarse, RMSE shrinks towards MAE and RMSE = MAE at the coarsest resolution where the entire study area is in one pixel.

4.5 Advantages of MAE over RMSE

MAE has at least two important conceptual and practical advantages over RMSE, which table 1 highlights in rows 7-8. The two advantages are related conceptually.

When MAE serves as the measure of deviation, then we can interpret each component of information as an amount of mass in the maps, thus we can interpret the variation in components as a function of resolution in terms of moving the mass in map Y over various distances. It is especially interesting to examine how disagreement due to pixel-level location shrinks as resolution becomes coarser. For example, if all the disagreement due to pixel-level location were attributable to misregistration by a distance of one pixel width in map Y, then all the disagreement due to pixel-level location could be resolved by moving some of the mass in map Y the distance of one pixel width, in which case the disagreement due to pixel-level location would vanish when the resolution doubles. On the other hand, if the mass in map Y were concentrated at a location in the map that is far from where the mass is concentrated in map X, then we would need to move the mass in map Y a large distance in order to rectify the disagreement due to pixel-level location. In this latter case, the disagreement due to pixel-level location would vanish in the multiple resolution analysis only after the resolution becomes very coarse. When MAE serves as the measure of deviation, then the disagreement due to pixel-level location is directly proportional to the amount of mass that would have to be moved in order to rectify the disagreement.

The mass in pixels that contain positive residuals would be moved to pixels that contain negative residuals in order to reduce the disagreement due to pixel-level location. For this paper's example in figures 1 and 2, a mass of 2 in pixel $X_{1\,e\,14}$ would be moved to pixel $X_{1\,e\,16}$ in order to reduce the total absolute deviation by 4, hence reducing the component of disagreement due to pixel-level location by 4/16 as resolution grows from fine to middle (figure 10). In general, if a total mass of H is moved within the map, then the total absolute deviation in the map decreases by 2×H, because the movement reduces the total absolute deviation by H in the pixels that lose the mass and also reduces the total absolute deviation by H in the pixels that gain the mass. Consequently, equation 16 gives the total mass that would need to be moved in order to reduce the component of disagreement due to pixel-level location by U-V, when MAE serves as the basis for the disagreement due to pixel-level location in a map of resolution r (ADPr). Note that U-V ≤ ADPr ≤ 0, and the sum of Wren is usually the number of pixels.

$$\text{Mass to move to reduce ADPr by U - V} = \left(\sum_{e=1}^{E}\sum_{n=1}^{Nre}\text{Wren}\right)\times\left(\frac{\text{U - V}}{2}\right) \qquad \text{equation 16}$$

It is possible to specify mathematically the maximum distance of the necessary movement of the mass. Let F be the distance of the side of a pixel at the fine resolution of the raw data. Let C be an integer greater than 1 that denotes the multiplication factor by which the fine resolution pixels are aggregated to form coarser pixels. Thus the distance of the side of a pixel at the coarser resolution is F times C. Let U be the disagreement due to pixel-level location at resolution F, and let V be the disagreement due to pixel-level location at resolution F×C. It is possible to reduce by U-V the disagreement due to pixel-level location by moving mass in map Y through a distance of less than or equal to the distance specified by equation 17. It is necessary to include the square root of two as a factor in the distance in order to account for the possibility that the mass in map Y may need to be moved across the diagonals of the coarse square pixels.

$$\text{Maximum distance to move mass to reduce ADPr by U - V} = \text{F}\times(\text{C}-1)\times\sqrt{2} \qquad \text{equation 17}$$

For example, in figure 13, the disagreement due to location is 0.26 at the finest resolution of 8 kilometers per pixel side and is 0.20 at a coarser resolution of 32 kilometers per pixel side, which is four times the finest resolution. According to equation 16, the total mass of z-scores that must be moved is 49976×0.06/2 ≈ 15000. According to equation 17, the maximum distance this mass would need to be moved is 8*(4-1)*√2 = 34 kilometers.

We can not use the moving mass analogy to interpret the results based on RMSE because the influence of each residual on RMSE is in proportion to the square of the residual's size, not in proportion to its absolute size. Large residuals have disproportionately more influence on RMSE than small residuals. Consequently, it does not make sense to draw an analogy about moving the mass in map Y when RMSE measures the information of pixel-level location, but the analogy makes sense when MAE measures the information of pixel-level location.

Row 8 of Table 1 gives the final important reason why this paper endorses MAE as the definition of deviation. If scientists are to use a unified general technique of map comparison to apply to both categorical variables and real variables, then scientists should measure deviation in terms of MAE, because MAE is the basis for the formulas to compute components of agreement and disagreement for the comparison between two maps that share a categorical variable (Pontius 2000; Pontius 2002; Pontius and Suedmeyer 2004). There is a good reason why MAE serves as the basis of measurement for the categorical case where each pixel is a complete member of exactly one category. Specifically, let Bren be the error in pixel n of stratum e at resolution r for the categorical case such that each Bren is either zero or one. If the category in the reference map matches the category in the comparison map then Bren is zero, otherwise Bren is one. Let p be the proportion of pixels in the map for which the error is one, and let Wren be the weight for pixel n of stratum e at resolution r, as described section 2.3. Equation 18 shows that p is the average error in the map according to MAE, while the average error according to RMSE is the square root of p.

$$p = \frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(Wren \times Bren\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre} Wren} = \frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(Wren \times \left|Bren\right|\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre} Wren} \le \sqrt{\frac{\sum_{e=1}^{E}\sum_{n=1}^{Nre}\left(Wren \times Bren^{2}\right)}{\sum_{e=1}^{E}\sum_{n=1}^{Nre} Wren}} = \sqrt{p} \qquad \text{equation 18}$$

If either p = 0 or p = 1, then MAE and RMSE give identical results. However, if $0<p<1$, then the measure produced by MAE is strictly less than the measure produced by RMSE. MAE produces the more intuitively satisfying result, because p is the proportion of pixels classified erroneously in the map.

6 Conclusions

This paper offers quantitative methods to budget important components of information that indicate fundamental ways in which patterns in maps compare. The approach is based on an intuition that the human eye can identify. When comparing two maps, humans usually see immediately how the overall quantity compares between the maps, and also how the location of the spatial pattern compares between the maps. Scientists should use statistical methods that both match this intuition and respect mathematical rigor. Hence, this paper's methods separate quantitatively the information of quantity from the information of location during map comparison. The results are presented visually in the form of stacked bar graphs that show separable components of information. The technique is designed specifically to examine how results vary as a function of changes in the resolution, because the resolution of raw data is often irrelevant to the resolution of the substantive questions. This proposed approach reveals information that is potentially more useful than conventional approaches that are based on hypothesis testing. In contrast to a more conventional approach, the proposed approach relies on simpler mathematics and a more flexible interpretation of the unit of observation. Hopefully, the proposed approach of budgeting components of information at multiple resolutions will become adopted as standard practice in the measurement of patterns.

Acknowledgements

The National Science Foundation supported this work via three of its programs: 1) Human-Environment Regional Observatory program via grant 9978052, 2) Long Term Ecological Research via grant OCE-0423565, and 3) Center for Integrated Study of the Human Dimensions of Global Change through a cooperative agreement between Carnegie Mellon University and the National Science Foundation SBR-9521914. Wageningen University's C. T. de Wit Graduate School for Production Ecology & Resource Conservation partially funded Pontius' sabbatical, during which he wrote some of this paper. Ron Eastman supplied the data for the environmental application to African vegetation. We thank the editor and anonymous reviewers for their helpful comments. Clark Labs facilitated this work by creating the GIS software Idrisi®.

Literature

Bailey, T. C. and Gatrell, A. C. (1995) *Interactive Spatial Data Analysis*. Prentice Hall.

Gaile, G. and Willmott, C.J. (eds) (1984) *Spatial Statistics and Models*. D Reidel Publishing Company, Dordrecht.

McMaster, R.B. and Usery, E.L. (eds) (2004) *A research agenda for geographic information science*. CRC Press, Boca Raton FL.

Openshaw, S. (1984) *The modifiable areal unit problem*. GeoBooks, Norwich.

Pontius Jr, R.G. (2000) Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering & Remote Sensing*, **66(8)**, 1011-1016.

Pontius Jr, R.G. (2002) Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering & Remote Sensing*, **68(10)**, 1041-1049.

Pontius Jr, R.G., Huffaker, D. and Denman, K. (2004) Useful techniques of validation for spatially-explicit land change models. *Ecological Modeling*, **179(4)**, 445-461.

Pontius Jr, R.G. and Suedmeyer, B. (2004) Components of agreement in categorical maps at multiple resolutions. In *Remote Sensing and GIS Accuracy Assessment*, R.S. Lunetta and J.G. Lyon (eds), CRC Press, Boca Raton FL. p. 233-251.

Unwin, D. J. (1996) GIS, spatial analysis and spatial statistics. *Progress in Human Geography*, **20(4)**, 540-551.

Veldkamp, A., Verburg, P.H., Kok, K., de Koning, G.H.J., Priess, J. and Bergsma, A.R. (2001) The need for scale sensitive approaches in spatially explicit land use change modeling. *Environmental Modeling and Assessment*, **6**, 111-121.

Willmott, C.J. (1981) On the validation of models. *Physical Geography*, **2(2)**, 184-194.

Willmott, C.J. (1982) Some comments on the evaluation of model performance. *Bulletin American Meteorological Society*, **63(11)**, 1309-1313.

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., and Rowe, C.M. (1985) Statistics for the evaluation and comparison of models. *Journal of Geophysical Research*, **900(C5)**, 8995-9005.

Biographical Sketches

Robert Gilmore Pontius Jr is Associate Professor at Clark University, where he coordinates the Master of Arts program in Geographic Information Sciences for Development and Environment. He earned a Master of Applied Statistics from The Ohio State University and a doctorate from the State University of New York / College of Environmental Science and Forestry. Many of the quantitative methods that he derives become incorporated into the GIS software Idrisi®.

Olufunmilayo Thontteh is Lecturer of Photogrammetry, Geographic Information Systems and Cartography in the Regional Centre for Training in Aerospace Surveys at Obafemi Awolowo University, Nigeria. She earned a Master of Arts in Geographic Information Science for Development and Environment at Clark University.

Hao Chen is a doctoral candidate in the Graduate School of Geography at Clark University. His research concerns simulation models and map comparisons. He has programmed several modules in the GIS software Idrisi®.

Table

Table 1. Characteristics of Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).

Characteristic	RMSE	MAE
1. Ability to budget components of disagreement and agreement	Yes	Yes
2. Commutative property for components of disagreement	Yes	Yes
3. Commutative property for components of agreement	No	No
4. Unique solution for minimum deviation	Yes	No
5. Sensitivity to outliers	More	Less
6. Sensitivity to change of resolution	More	Less
7. Interpretable in terms of moving mass	No	Yes
8. Consistent with categorical case	No	Yes

Figures

-2	-1	7	8
-4	-3	5	6
-6	-5	3	4
-8	-7	1	2

X

2	0	8	8
-2	0	6	6
-4	-4	2	6
-4	-2	-4	-2

Y

1	1	0	0
1	1	0	0
1	1	0	0
1	1	0	0

Weights for stratum 1

0	0	1	1
0	0	1	1
0	0	1	1
0	0	1	1

Weights for stratum 2

Figure 1. Example for maps of X & Y on top and strata 1 & 2 on bottom.

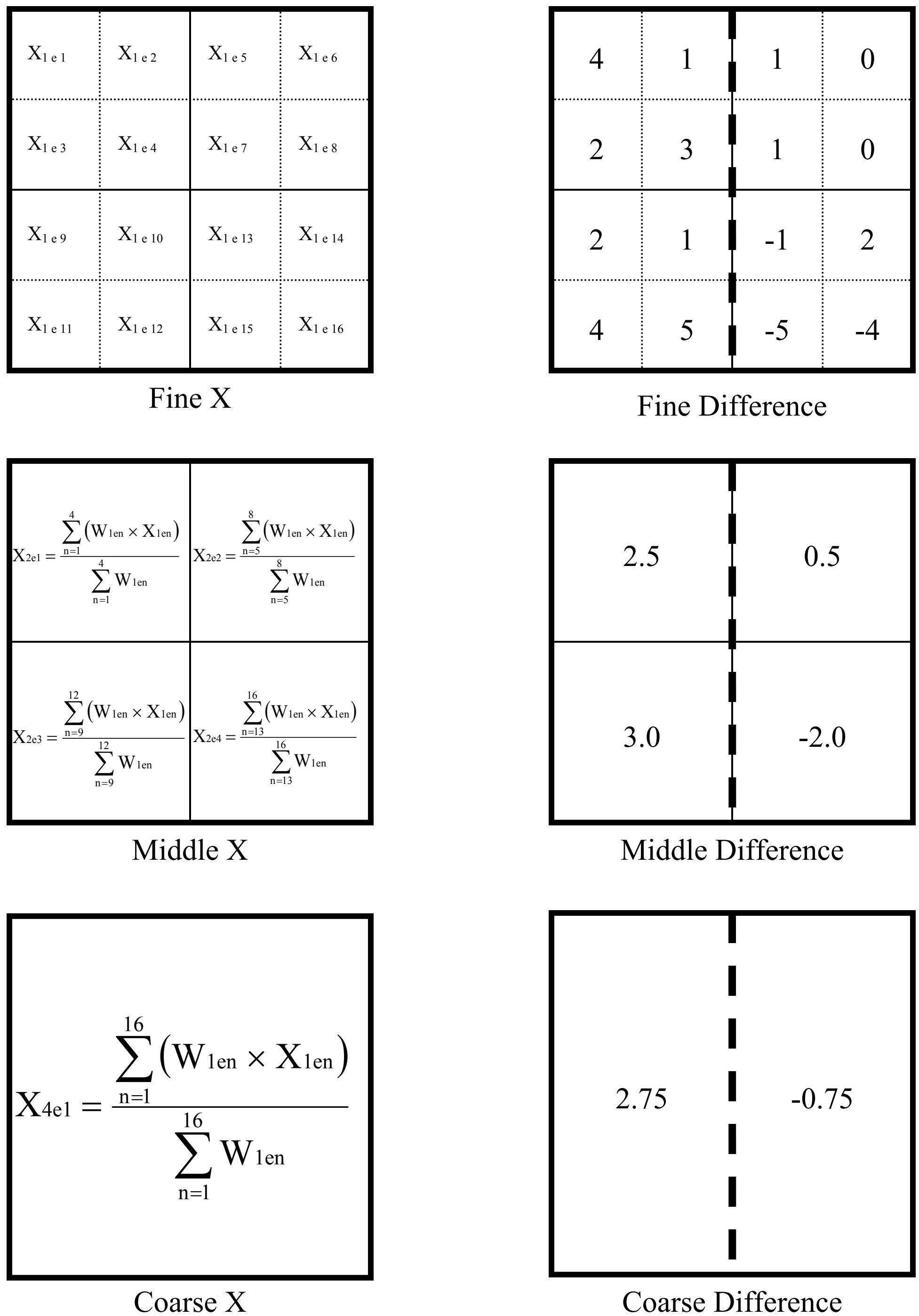


Figure 2. Conversion from fine to coarser resolutions for stratum e on the left and values of Y minus X on the right.

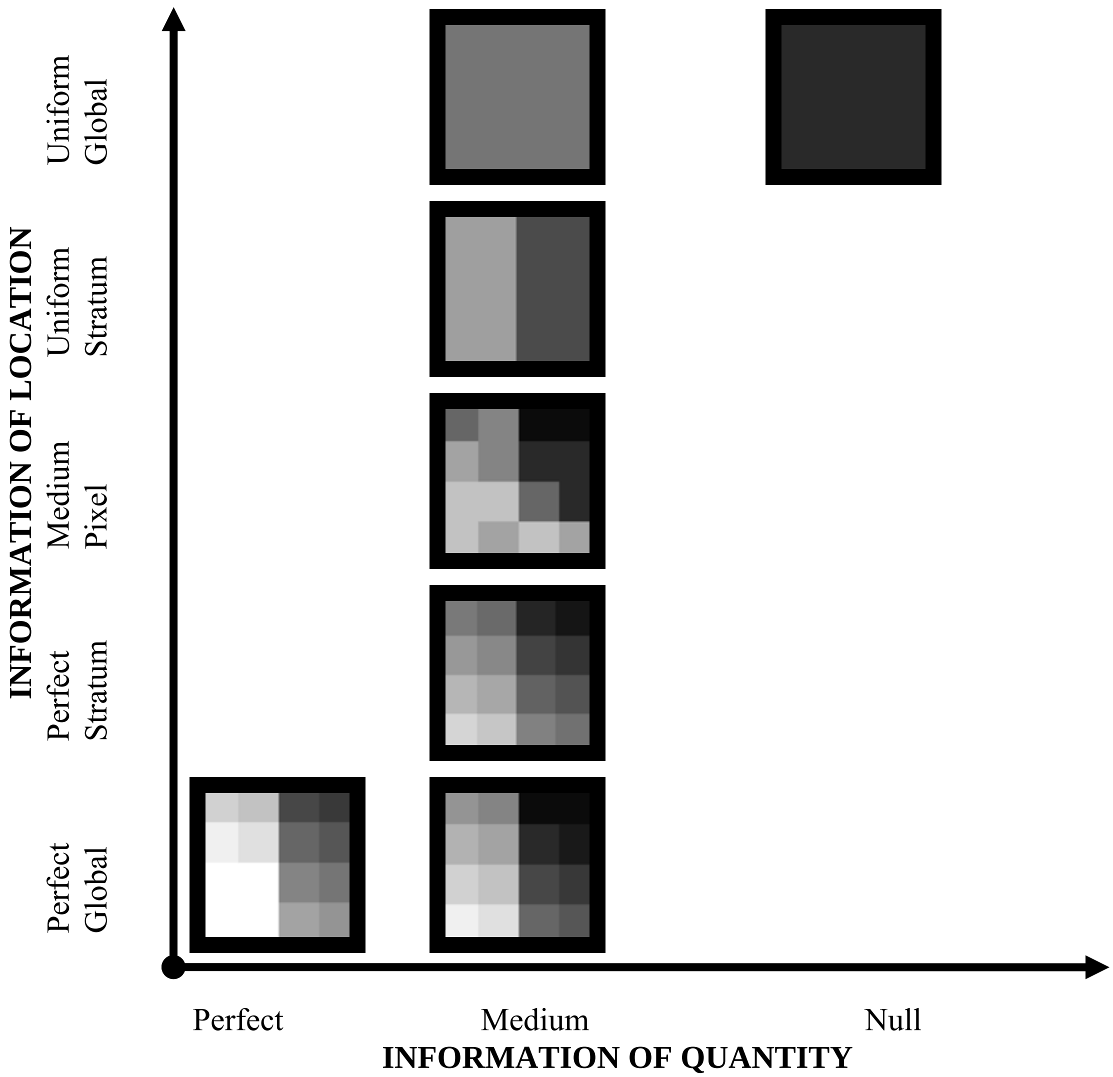


Figure 3. Fine resolution maps that have the combination of information of quantity and information of location as designated by the position in the information space.

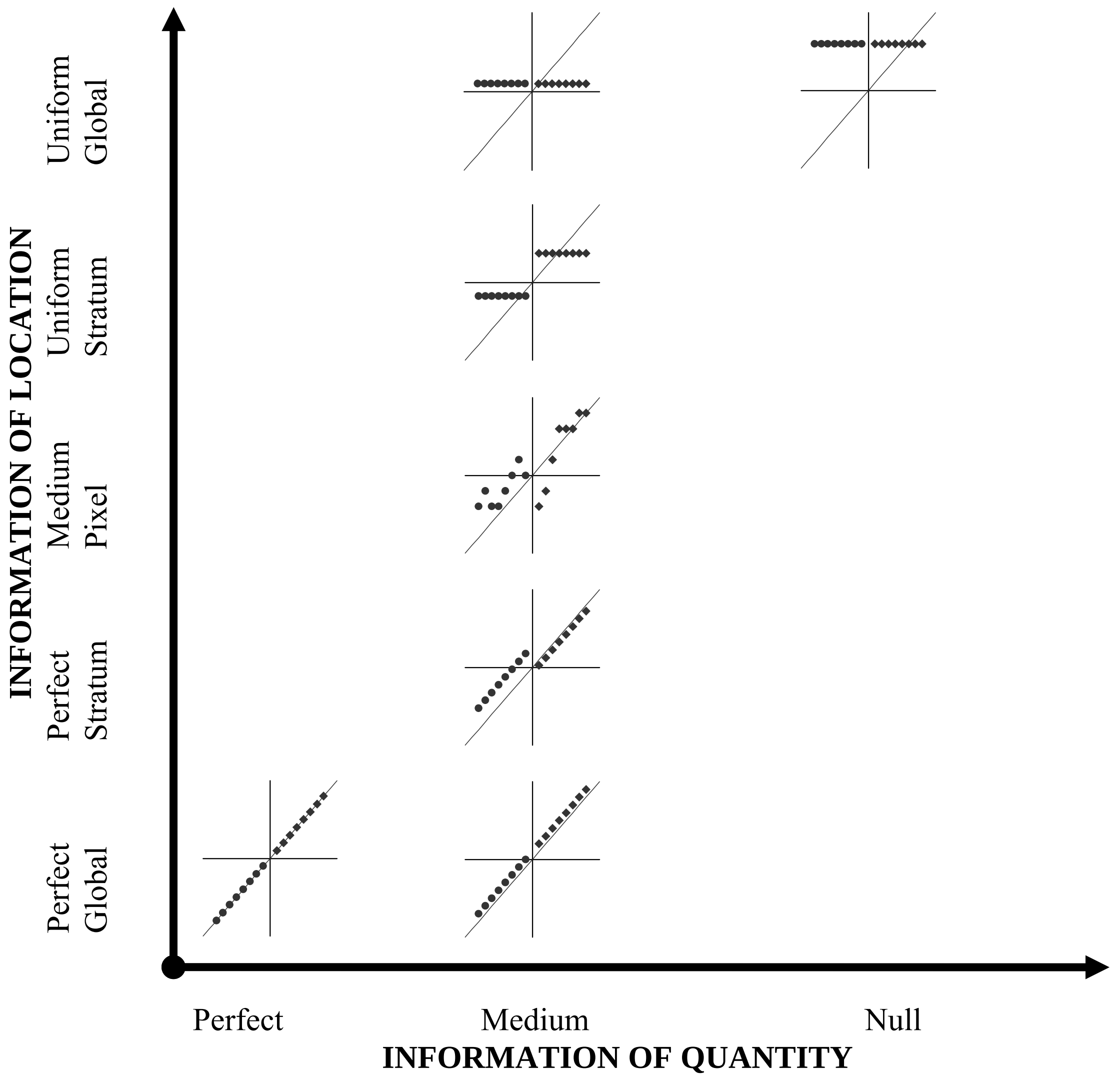


Figure 4. Scatter plots that compare the 16 pixels of map X to the map at the corresponding position in figure 3.

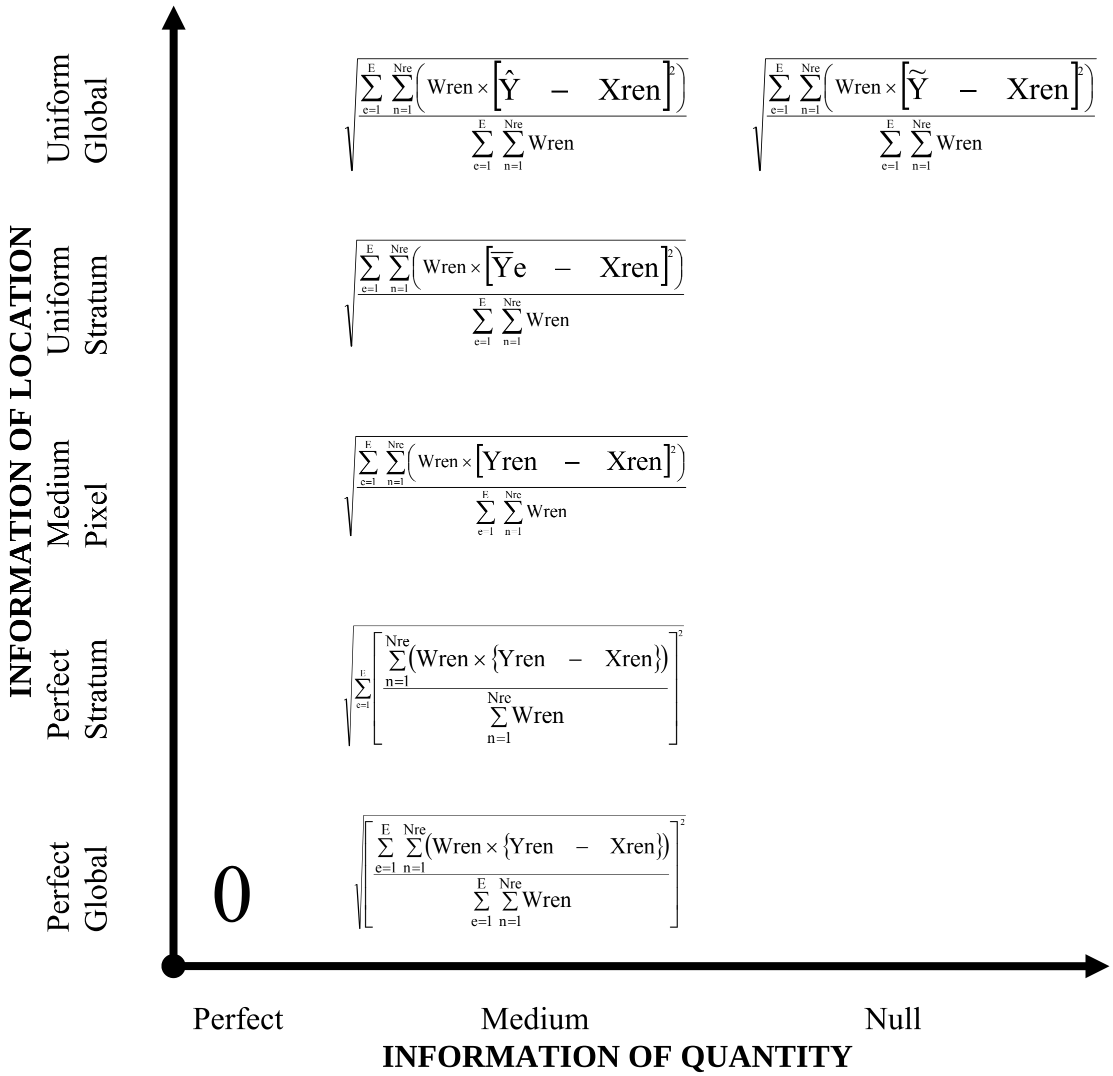


Figure 5. Mathematical expressions based on Root Mean Square Error (RMSE) that measure the deviation between map X and the map at the corresponding position in figure 3.

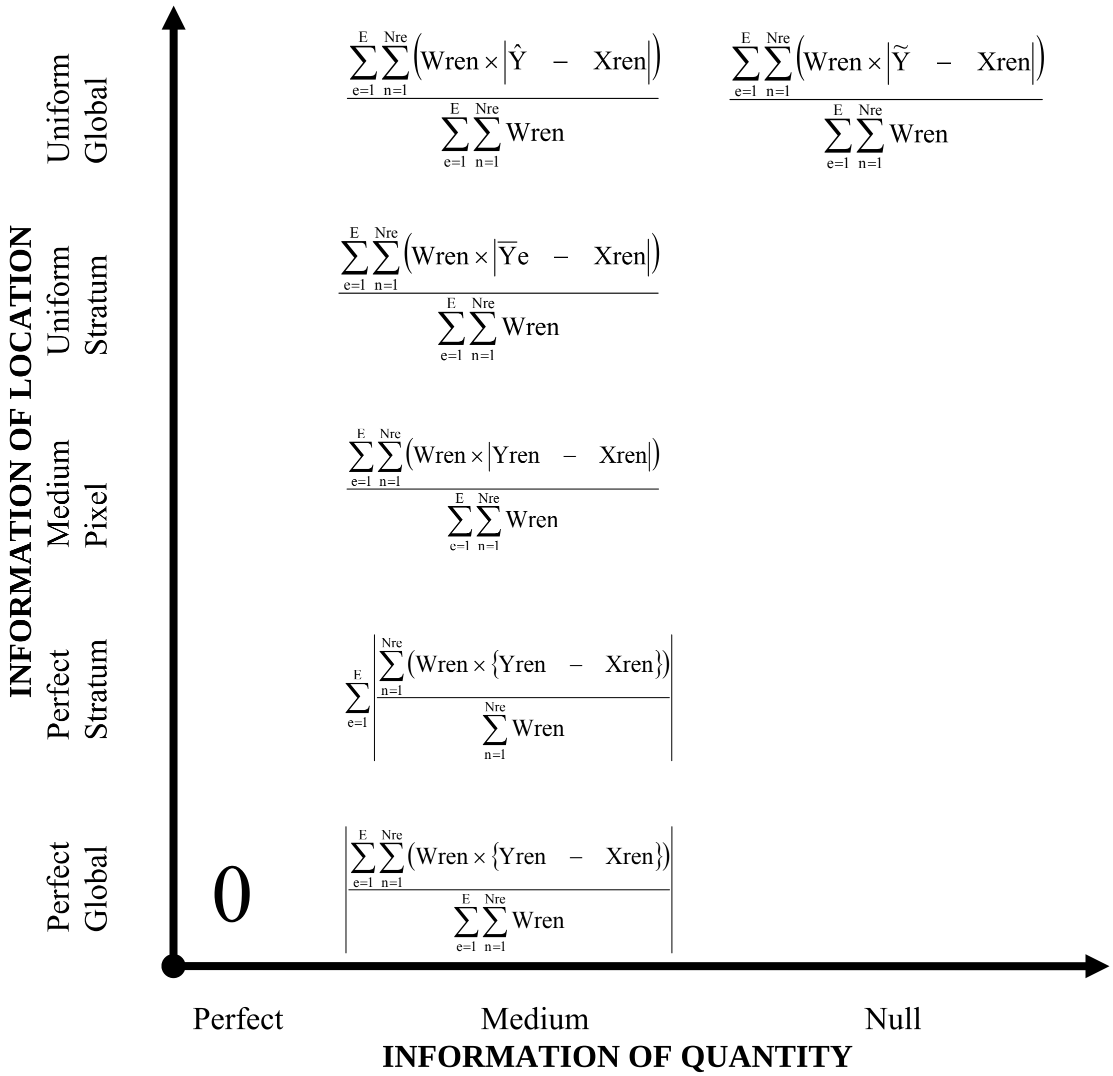
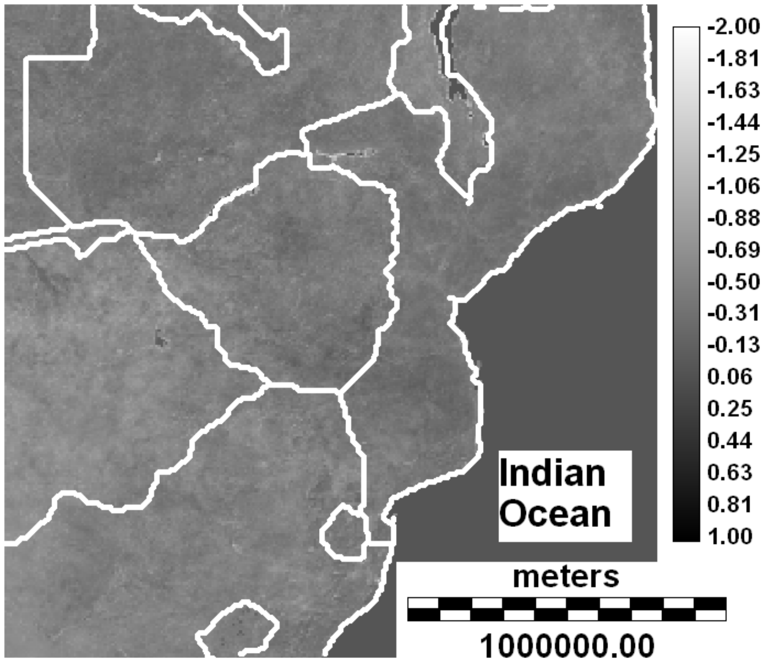


Figure 6. Mathematical expressions based on Mean Absolute Error (MAE) that measure the deviation between map X and the map at the corresponding position in figure 3.



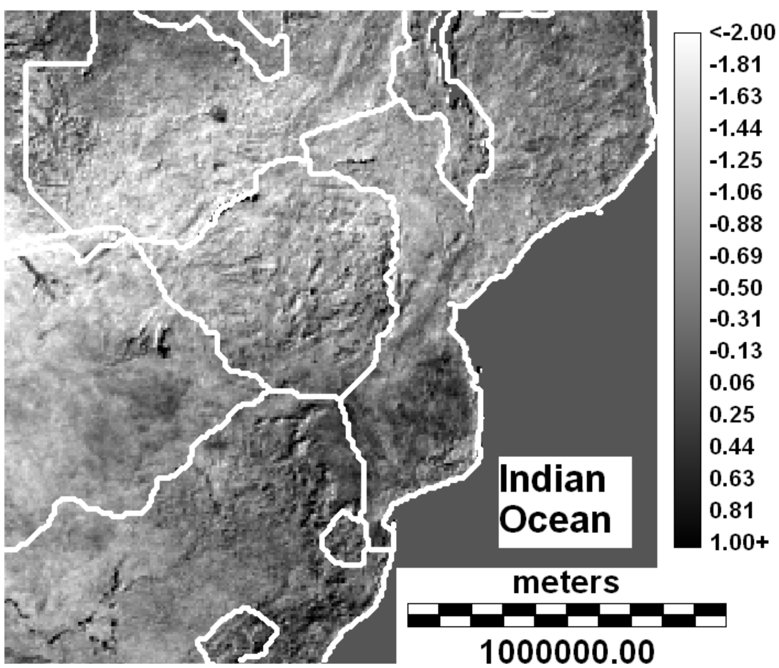


Figure 7. Maps of z-scores for NDVI during 2003 at 8-kilometer by 8-kilometer resolution for the observed (X) on the top and the predicted (Y) on the bottom.

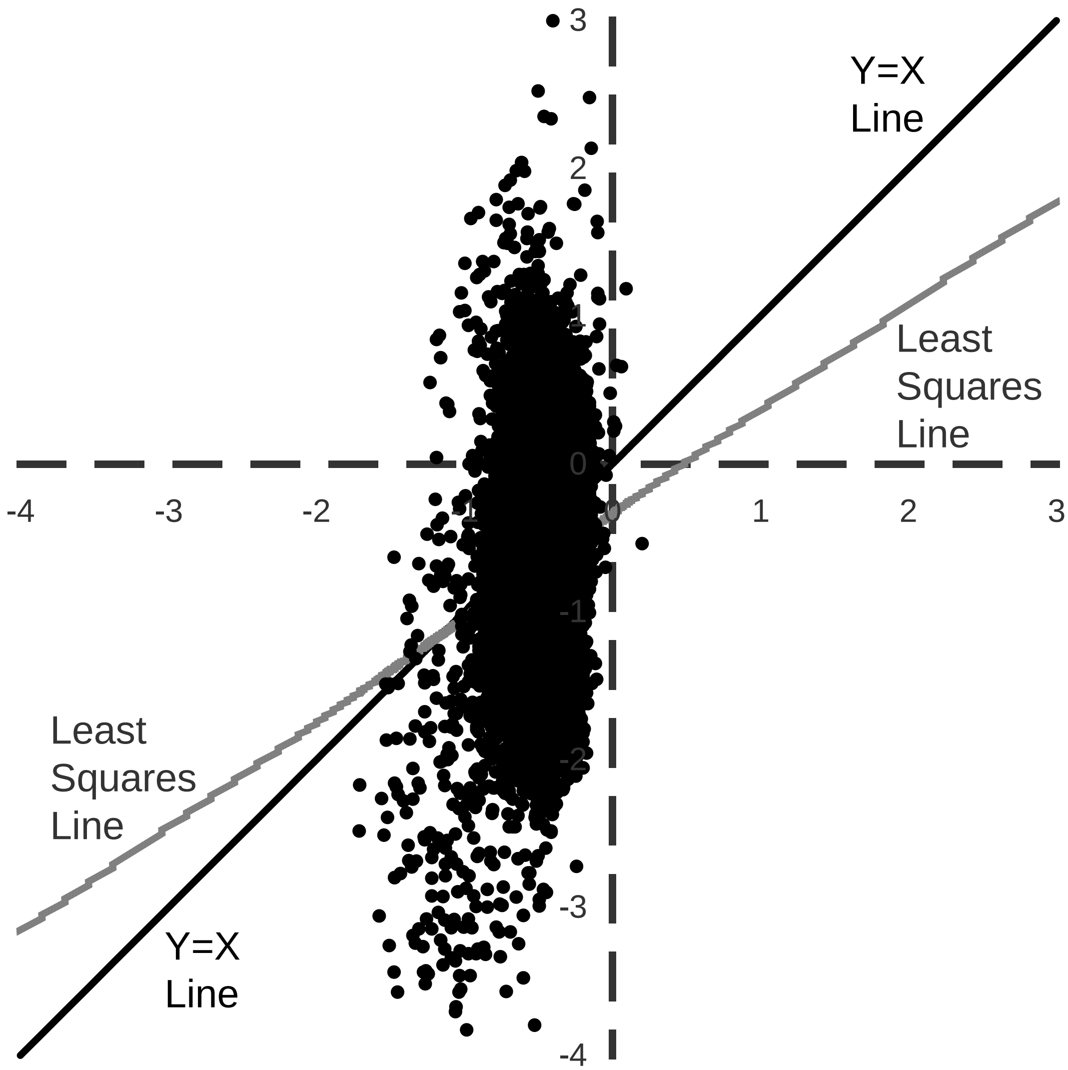


Figure 8. Scatter plot of pixels in figure 7 where the dashed horizontal line is the observed z-score axis and the dashed vertical line is the predicted z-score axis.

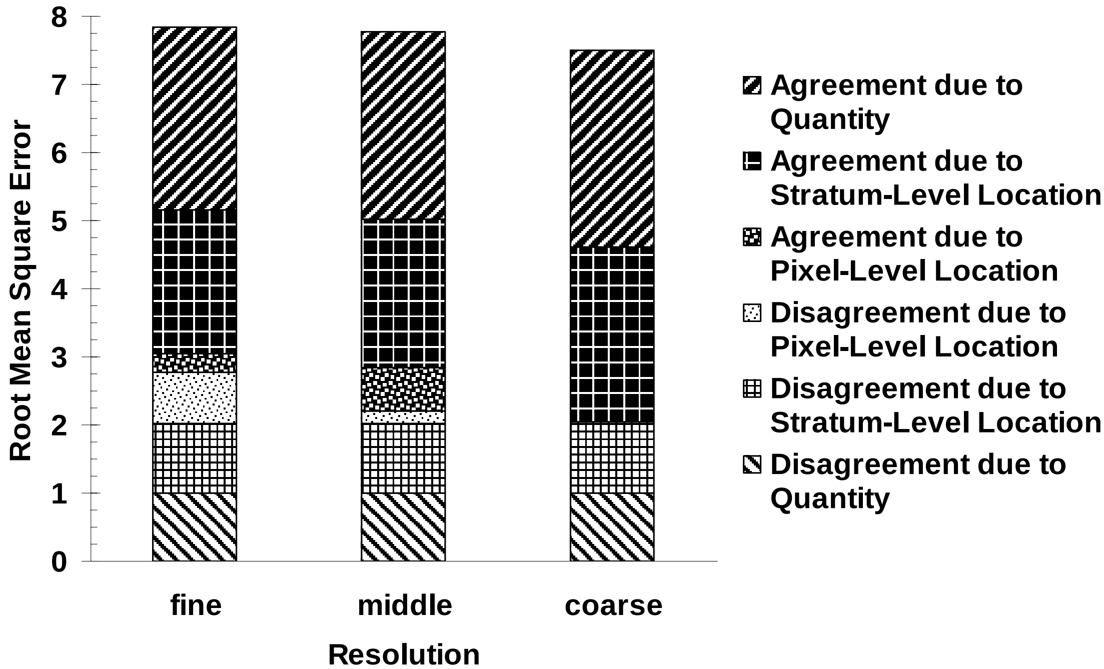


Figure 9. Budget of components of information based on Root Mean Square Error at multiple resolutions for the example in figures 1-4.

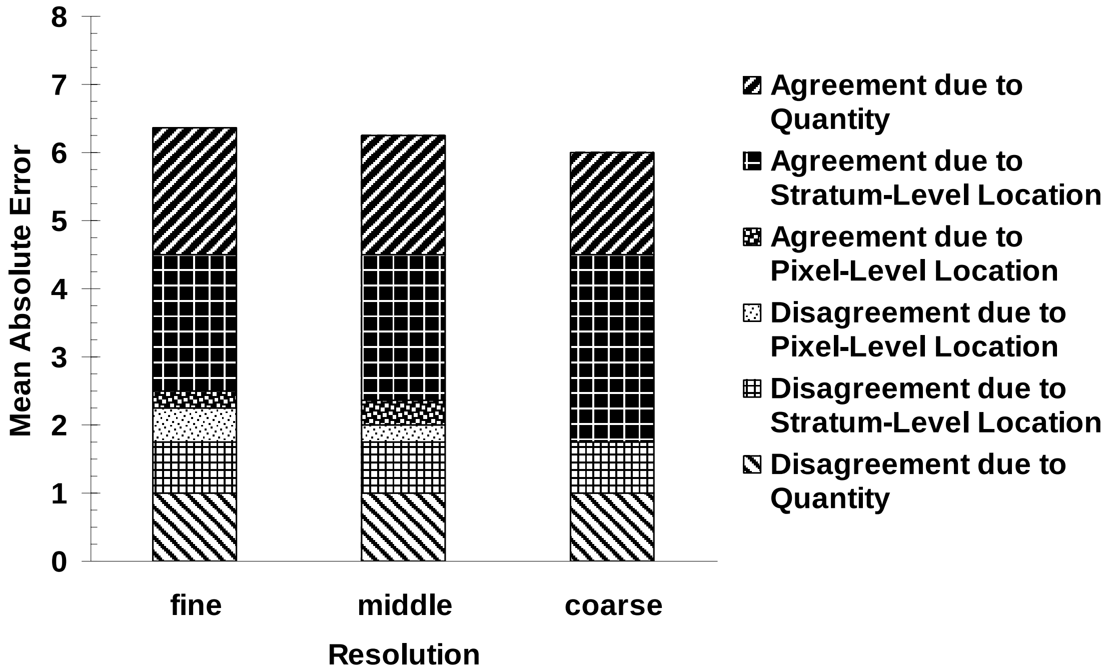


Figure 10. Budget of components of information based on Mean Absolute Error at multiple resolutions for the example in figures 1-4.

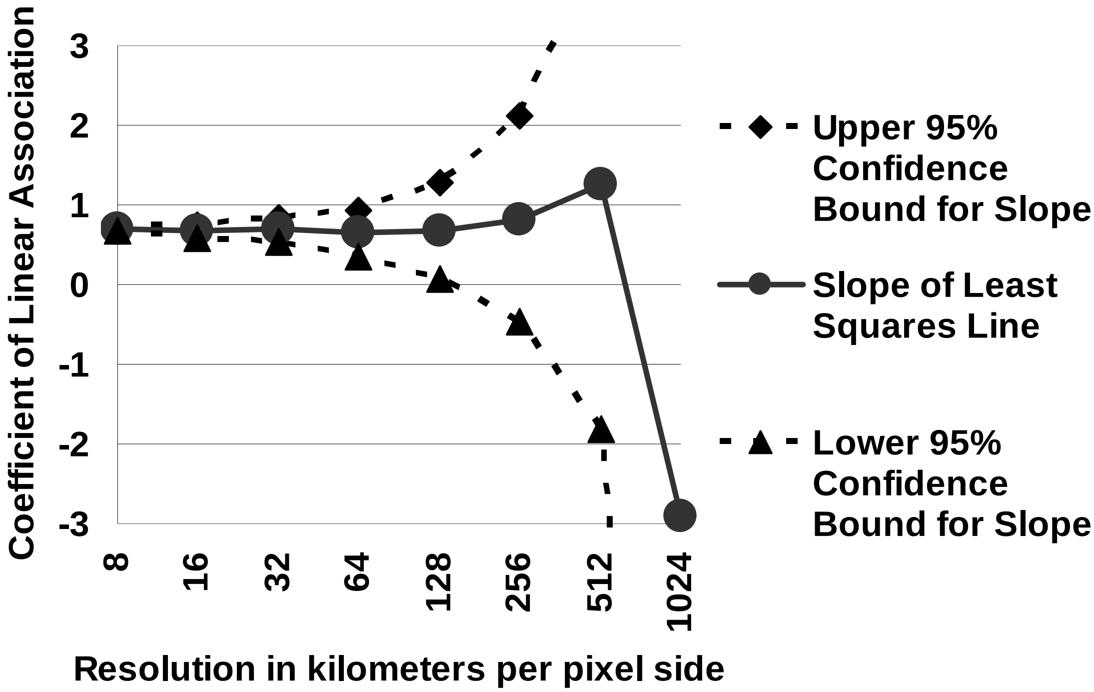


Figure 11. Confidence intervals around slope of regression line at multiple resolutions for the application to vegetation in Southeastern Africa.

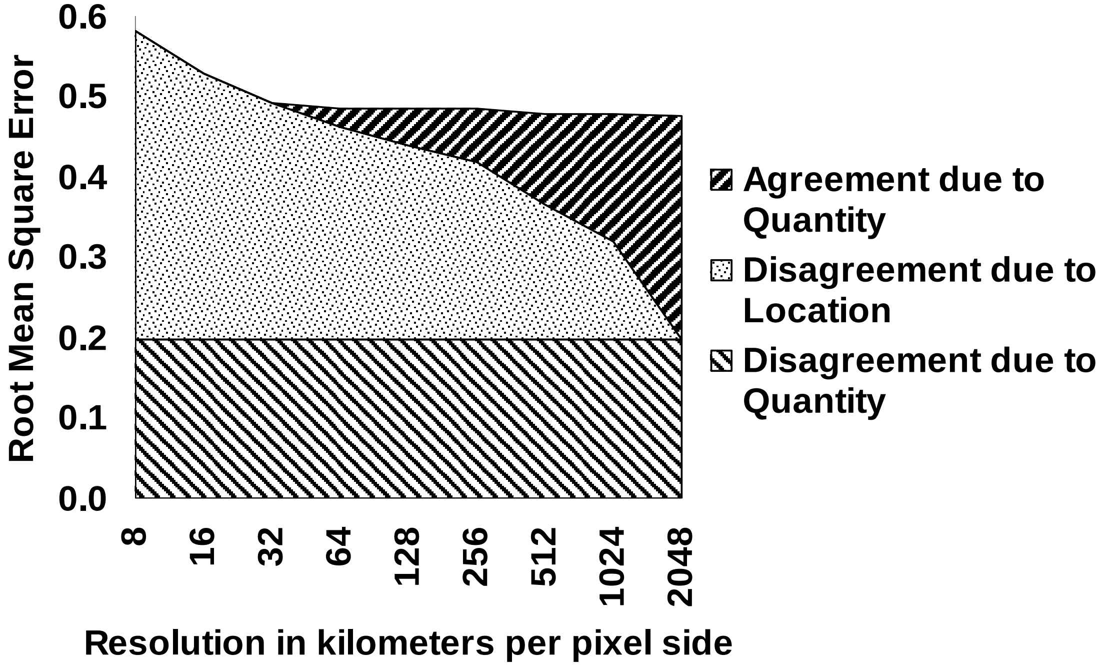


Figure 12. Budget of components of information based on Root Mean Square Error at multiple resolutions for the application to vegetation in Southeastern Africa.

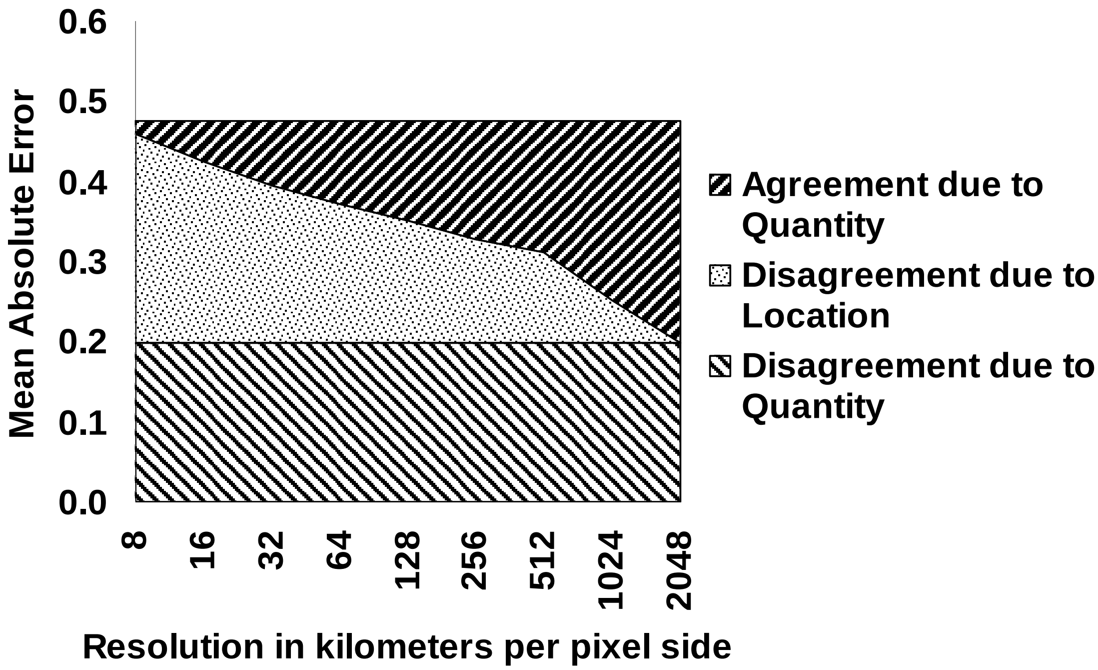


Figure 13. Budget of components of information based on Mean Absolute Error at multiple resolutions for the application to vegetation in Southeastern Africa.