Task-based Color Scale Design

Penny Rheingans
Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County, Baltimore MD 21250

Abstract

Color is commonly used in data visualization in order to convey a wide variety of types of information: metric values, pattern, extrema, emphasis, and others. A distressingly large percentage of these visualizations appear to use the default color scale for the visualization package used to create them. While sometimes this is an appropriate choice, careful consideration of the implications of color scale selection can often result in a more effective visualization. Factors which should be considered include the characteristics of the data, the questions of interest about the data, and the expected viewers of the representation.

Keywords: Color-mapping; Pseudo-color; Visualization Design

The color scale selected for a visualization, that is, the sequence of colors used to represent the values of the data range, can have a substantial impact on the effectiveness of the visualization. Unfortunately, there is no one best color scale. The most appropriate color scale for a particular representation is influenced by the characteristics of the data, the questions of interest about the data, and the expected viewers of the representation. This paper first surveys some common, and not very common, color scales used for univariate and bivariate data. It then discusses some of the principles and issues that drive appropriate color scale selection.

1. A Survey of Single-Variable Color Scales

Single-variable color scales map the value of a single scalar variable at each point in the image to a color representing that value. Continuous scales, those in which adjacent colors are similar to one another, necessarily form a continuous path through some color space. Alternatively, color scales could contain discontinuities, i.e. places where adjacent colors were not at all similar. This survey describes primarily continuous color scales.

1.1. Color Model Components

Grey scale. Perhaps the simplest color scale maps the value of a single scalar variable to brightness. See Figure 1. Usually, black represents the lowest value, white represents the highest value, and shades of grey represent the intermediate values. Viewers who are more used to print media, however, may prefer a scale that represents increasing value by the appearance of increasing amounts of ink, mapping the lowest value to white and the highest value to black. The biggest advantage of a grey scale for representing a single scalar variable is the effectiveness of the human visual system at making judgements about shape from lightness variation. This makes a grey scale a particularly good choice for tasks emphasizing the understanding of qualitative shape and pattern information. Another advantage is that there is both an inherent perceived order to the brightness levels and a visual zero value (usually black). The main disadvantages are a limited number of distinguishable display values (approximately 100) and a limited contrast between different levels. These disadvantages make a grey scale a suboptimal choice for tasks involving metric lookup of quantitative values.

Saturation scale. A saturation scale maps a scalar value to colors of increasing saturation, holding hue constant. See Figure 2. Brightness of the color is not generally explicitly controlled, but many color models contain substantial interactions between saturation and actual or perceived brightness. High values are emphasized over low values. The saturation scale has the basic advantages of being simple intuitive. Its basic weakness is a limited number of distinguishable display levels.

Spectrum scale. A spectrum scale is formed by holding saturation and brightness constant and letting hue vary through its entire range. See Figure 3. The scale usually follows the scale of colors in the spectrum, first red, then orange, yellow, green, blue, and finally violet. In general, the problem with the spectrum scale is that many untrained observers (and some trained observers) see no intuitive ordering in the hues, so the scale requires the viewer to impose a learned order [Pizer 83].

By convention, spectrum scales usually start at red with the higher wavelength colors following in order. This convention has the advantage that the resulting scale is intuitive for viewers who have a mental model of the progression of wavelengths of light. It has the potential disadvantage that the colors at the start and end of the scale, red and violet respectively, are very similar. Figure 4 shows a spectrum scale limited to the blue to red range, better differentiating the extremes. Another potential disadvantage is that the yellow in the middle of the scale is very striking. This tends to draw the eye to the places with values represented in yellow. This can be a disadvantage if extreme values are the primary interest. Since the range of hues is circular, a spectrum scale can be started at a place that positions striking colors over the values of the most interest. Another characteristic of a spectrum scale are the perceptual discontinuities that occur at color name boundaries, causing viewers to judge the color difference of two colors with different names as greater than that between two colors of the same basic name. Essentially, this characteristic can lead to the perception of boundaries where none are present in the data.

1.2. Redundant Color Scales

Visualizations that represent data values using multiple display parameters have the potential to portray the data more effectively than visualizations that map each data variable to a single display parameter. There are a number of compelling reasons why this should be true:

1. Different display parameters convey different types of information most efficiently. For example, brightness conveys shape more effectively than hue, but hue provides more accurately distinguishable display levels than does brightness.

2. Multiple display parameters can overcome visual deficiencies. If one display parameter is ambiguous or ineffective because of the visual deficiency, another may compensate. For example, a person with color deficiencies would likely find it easier to unambiguously judge the value represented by a color using a redundant hue and brightness color scale than one using a standard spectrum scale that varies only in hue.

2. Multiple display parameters reinforce each other. In this way, areas with differing values have greater visual difference from one another.

The utility of redundant color scales has been empirically confirmed. Ware [88] conducted three experiments comparing a linear grey scale, a perceptual grey scale, a saturation scale, a spectrum scale, a red-to-green scale, and an experimental redundant scale combining hue and lightness variation for univariate data representation. When subjects were asked to judge the metric value of a colored patch surrounded by a contrasting area, the spectrum scale produced significantly more accurate metric value readings. When subjects were asked to judge information about the surface properties of simulated surfaces, the grey scales were more effective. When the experimental scale was used for a metric query task, subject accuracy was similar to that of the spectrum scale (which had no monotonic lightness variation) and significantly better than the others. This suggests that a color scale that varies in both luminance and hue can be used to accurately represent both metric and surface properties.

Redundant model components. In the most straightforward of redundant scales, data values can be mapped to both hue and lightness. Figure 5 shows such a visualization. Compared to an image using only hue (Figure 3), this representation shows the location and swirling structure of high areas more clearly. This redundant representation also has the advantage that someone with a dichromatic color deficiency can unambiguously interpreted it. For example, a protanope (commonly called red-green colorblind) viewing Figure 3 would find it difficult to distinguish the values near 100 (green) from the values near 170 (red). The same person viewing Figure 5 would be able to distinguish the two values based on lightness; the higher values would appear brighter. Because a legend is included, the value encoded by each color is unambiguous (to the limits of discrimination). Other model components, such as hue and saturation, could also be used redundantly

Heated-object scale. The heated-object scale represents a compromise between the grey scale and the spectrum scale. It increases monotonically with luminance, but not with any of the other opponent color channels. The heated-object scale goes from black through red, orange, and yellow to white. See Figure 6. The resulting color path forms an upward-curving spiral in the HLS color space. The colors of the heated-object scale follow the same scale as those of a black body when heated. The heated-object scale has more distinguishable display values and more contrast between different levels than a grey scale [Pizer 83]. The heated-object scale has a stronger perceived natural ordering than the rainbow scale because of the monotonic increase in brightness and because there exists a basis for remembering the color order that is based in experience.

Optimal color scales. Levkowitz [92] introduces the term *optimal color scale* to describe a scale that maximizes the total number of JNDs (just noticeable differences) while preserving a natural order. Such a color scale goes from black to white, increasing monotonically in red, green, and blue. Levkowitz's optimal color scales increased monotonically in both brightness and RGB components while being linearized with respect to just noticeable differences (JNDs). Levkowitz experimentally compared the optimal color scale to a linearized grey scale and a linearized heated-object scale. He found the grey scale to result in significantly more accurate identifications of simulated lesions in medical images. While this result does not exactly validate the use of redundant color scales, the expected result of the experimental task (shape perception) was that the grey scale would excel.

1.3. Double-Ended Scales

Conceptually, a double-ended color scale is created when two monotonically increasing scales are pasted together at a shared end point. For instance, a scale from grey to red and a scale from grey to blue can be stitched together to form a single scale from red to grey to blue. Figure 7 shows such a scale. Such color scales have three distinct groups of colors, representing the high, low, and middle values. The colors in a double-ended scheme could represent a portion of the hue circle (such as a scale from red to yellow to green), a straight line through a color space (such as a scale from green to grey to purple), or some sort of curved path through color space (such as a scale from purple to grey to brown). The basic advantage of a double-ended scale is the clear visual classification of values as high, low, or middle.

2. A Survey of Multivariate Color Scales

Multivariate color scales map the values of two or more data variables at each point in the image to a single color representing both values. A continuous two-variable color scale forms a curved parametric sheet through a color space. Each of the two sheet parameters corresponds to one of the data variables. These display parameters can defined by components (or combinations of components) of a color space. The figures in this section show color sheets that have been flattened into rectangles. This section discusses primarily continuous, two-variable color scales.

2.1. Color Model Primaries

As with univariate color scales, mapping data values to color model components is an obvious choice. Such scales are simple and reasonably intuitive.

Display Primaries. One obvious two-variable color scheme maps each variable into one component of the RGB color model of the display device. Researchers working with data from remote sensing devices frequently use scales with all display primaries. Landsat "false color" images are commonly produced by representing multispectral scanner (MSS) bands 4, 5, and 7 with levels of blue, green, and red, respectively [Robertson 88]. If the bands displayed are highly correlated, most of the image will be shades of grey because the red, green, and blue components will be roughly equal. This scheme has the advantage that the colors representing the extremes of the variable ranges (black, red, green, and blue) are clearly distinguishable. The disadvantage is that some observers have difficulties decomposing the displayed colors into their component parts. This can result in difficulties perceiving similarities between areas that differ in the values of one variable, but not the other.

Hue-Other Scales. An analogous color scheme in the HLS color model would map two variables to two color model components, generally hue and lightness or saturation. For example, a color scheme could map education level to hue and average hours worked to saturation. Areas with low education levels would be reds, unsaturated when few hours are worked and saturated when many hours are worked. Areas with a relatively average education level would be blues; unsaturated when few hours are worked and saturated when many hours are worked. See Figure 8. Each location in the data set corresponds to a set of characteristics of a person (age, education, gender, etc). The area inside the black contour at the right corresponds to types of people who are likely to make high incomes. The two display parameters of this scheme (hue and lightness) have different characteristics in many of the same ways that the grey scale and spectrum scale have different characteristics. For example, the lightness parameter conveys order and magnitude more intuitively over the whole range of values because we perceive them as having a natural order. Specifically, it is easy to tell that a light grey represents a larger value than a dark grey. Conversely, without a legend, it is difficult to know whether yellow represents values less than or greater than blue. It is also easier to judge the relative magnitude of two lightness values than of two hues. Areas with similar hue components, but differing lightness components are somewhat easier to perceive as related than areas with similar litness, but different hues. Examples of Hue-Other scales include the colorwheel [Johannsen 95] and hue ball [Kindlmann 99] techniques.

2.2. Census Bureau Two-Variable Color Map

The Two-Variable Color Map developed by the Census Bureau represents bivariate information by mapping each variable to a four-level color scale and then taking the Cartesian product of ("crossing") the two scales to produce a sixteen level bivariate scale [Fienberg 79]. One scale uses yellow to represent low values, dark blue to represent high values, and lighter blues to represent intermediate values. The other scale also maps low values to yellow, but maps higher values to reds. The product of the two scales produces a bivariate scale where areas low values of both variables are yellow, areas with high values of both variables are purple, areas where one variable is larger than the other are either predominantly blue or red. Critics of the scheme have noted the lack of an intuitive progression in colors along the rows and columns of the gamut and the great similarity among the nine colors in the upper right of the gamut.

2.3. Complementary Display Parameters

The concept of double-ended color scales extends naturally to bivariate color scales, i.e., mappings from two scalar values to a color [Eyton 84]. One such scale might map the value of one variable to brightnesses of a hue, such as green. The value of the other variable would be mapped to brightnesses of the complementary hue, in this case purple. See Figure 9. The contributions of the two variables are summed additively to give the display color. The resulting color scale contains three clearly discernible classes of colors. Greys represent places where the values of the two variables are comparable because equal chromatic contributions of the two hues cancel each other, producing grey. Dark greys are formed when both values are low, while light greys result when both values are high. Places where one variable is significantly larger are colored green, while places where the other predominates are colored purple. Such a color scale is useful in situations where the two variables are expected to be correlated. Places where this relationship does not hold are immediately apparent.

3. Creating and Evaluating Color Scales

All commonly available visualization packages have some facility for selecting a color scale for a visualization. Completely user-controlled methods include selecting a scale from a supplied list of alternatives, specifying color scales in terms of functions of color components, sketching in curves for color component functions, and simply setting scale tables entry by entry. Getting an effective color scale by any of these methods require the user to understand the characteristics of a good color scale. A few methods have been proposed for more automatic design of color scales. The PRAVDAColor module to IBM's Data Explorer provided the user with a selection of colormaps chosen based on the data type, spatial frequency characteristics, user task, and guided by principals of the human perceptual system [Bergman 95]. An alternative approach proposed generating color and opacity transfer functions for volume rendering using a stochastic generation of alternatives, which are pruned by human evaluation [He 96]. The Design Galleries approach attempts to generate a selection of transfer function alternatives, which are well dispersed in the design space of possible transfer functions [Marks 97].

Identifying the characteristics and requirements of a good color scale is another fruitful area of research. Trumbo [81] presents four basic principles important in the selection of colors for the representation of quantitative information. Trumbo limits his attention to the display of discrete data value levels (classed data), but the ideas generalize to the display of continuous information. The first two principles apply to the representation of both univariate and bivariate information. The Order principle requires that if data value levels are ordered then the colors chosen to represent them should be perceived as ordered. A spectrum scale would violate the Order principle if the viewer did not perceive the hues to be ordered. The Separation principle requires that significantly different levels of variables be represented by distinguishable colors. A grey scale would violate the Separation principle if data variable values with an important difference were mapped to colors with an imperceptible difference. The heated-object scale, optimal color scale, and truncated spectrum scale appear to satisfy both principles.

Trumbo's last two principles apply only to the display of bivariate information. The Rows and Columns principle states that if preservation of univariate information is important, then the display parameters should not obscure one another. This condition is satisfied if rows or columns with a constant value of one variable have constant hue, saturation, or brightness. Using two display primaries (such as red and green) violates the Rows and Columns principle. The Diagonal principle states that if detection of positive association of variables is a goal, then the displayed colors should be easily identified as belonging to one of three classes: those near the minor diagonal, those above it, and those below. This condition could be satisfied by a scheme with the major diagonal made up of greys, elements of maximum saturation, or a constant hue. A hue and lightness scheme violates the Diagonal principle. The Census scheme violates both the Rows and Columns and Diagonal principles. Using complementary display parameters satisfies both principles. Violating one or both of these principles does not necessarily mean that a color scheme is not useful, only that it might not be appropriate for some representation tasks. For example, a hue and lightness scheme would not be the best choice for a representation primarily designed to show positive association between variables because it violates the Diagonal principle. On the other hand, it would be a reasonable choice for a representation where the goal is perception as a class of colors representing similar values of one variable across differing values of the other variable.

4. Design Considerations

Key considerations in the design of effective visualizations include the major goals, the nature of the data, the intended audience, the rest of the visualization, and the cultural connotations of color.

4.1. Consider the goals.

Decisions about the color scale for a visualization should not be made without consideration of the goals of the visualization. Tasks which require the judgement of metric quantities in the data tend to work best with color scales which do not vary monotonically in the opponent color channels (brightness, red-green, yellow-blue). Tasks involving qualitative judgements about value distribution shape are more naturally suited to color scales which vary systematically in brightness, allowing our visual systems to employ familiar shape-from-shading mechanisms.

Some visual experiments have suggested that the color of an object can influence the perceived size of that object. Tedford, Berguist, and Flynn [77] conducted experiments under precisely controlled conditions and found a significant color-size effect. Specifically, rectangles of the same size, saturation, and brightness appeared to have different sizes when colored red-purple, yellow-red, purple-blue, or green (in order of decreasing apparent size). At high saturations, this effect was statistically significant for all color pairs except yellow-red and purple-blue. At low saturations, only the difference between yellow-red and green rectangles was significant. In trials where hue was held constant and saturation varied, rectangles with higher saturations were consistently judged to be smaller than less saturated rectangles. Generalizing from the surveyed studies, they observed that warm colors (red, orange, yellow) appear larger than cool colors (blue, green). Cleveland and McGill [83] investigated the implications of the color-size illusion for statistical maps. Subjects were shown a map of Nevada in which counties where colored red or green with the total area of red and green nearly equal. Subjects were asked to judge which color, if any, represented the larger land area. On the average, subjects judged that the red areas were larger more often than they judged the areas the same or the green areas larger. When the experiment was repeated using low-saturation tones of red and green (formed by adding yellow), no such bias was observed. Their results suggest that the color of a region influences the perceived size of the region, and the effect is strongest for very saturated colors. These findings suggest that highly saturated colors might not be appropriate for tasks where judging size is a goal.

4.2. Consider the data.

Designers of visualizations should take care that the most striking features of the image are also the most important. Representations that draw the viewer's eye to unimportant features may cause more interesting features to be overlooked. Features likely to catch the eye are those that are brightly colored, moving or changing, defined by sharp boundaries, or highly saturated. The common spectrum color scale maps the middle values to yellow, a particularly striking color. In applications where the location of middle values is of particular interest, this is appropriate. Such applications are not very common, however. More often, the high or low values are of greatest interest, and middle values are of least interest. Data sets with both positive and negative values can have a zero point representing no change, average, or expected value. In such data, deviation from zero (and the pattern of such deviation) is what is interesting. For such a data set, a double-ended color scale is often a good choice.

4.3. Consider the audience.

Information about the intended audience of a visualization can also be used profitably in color scale design. For instance, conventions within one application area might generally place the blue/violet colors of a spectrum scale at the low end (in order of increasing wavelength) while another placed them at the high end (in order of increasing frequency). Application area conventions need not be slavishly followed, but an awareness of them can minimize unintentional breaks with viewer expectations.

4.4. Consider the whole visualization.

Overall visualization design choices obviously influence the color scale design of individual elements. For example, three-dimensional visualizations impose different constraints from two-dimensional visualizations. Specifically, viewers use shading cues to judge the 3D shape of a representation object such as an isosurface. A color scale that included brightness variations could interfere with the brightness resulting from shading calculations. A brightness varying scale might still be appropriate for planar objects in a 3D scene, however. Alternatively, the desire to display multiple variables in the same visualization can influence color scale design. Color scales for different objects should not generally overlap. Ideally, the representation for each variable should interfere with the others as little as possible.

4.5. Consider cultural connotations.

Colors tend to have strong cultural connotations (which unfortunately vary from culture to culture). Following these conventions can reduce the cognitive load on the viewer. In some situations, connotations may suggest natural linkings between a variable and the color used to represent it. For example, a variable describing income might naturally be represented by a scale of increasing saturation of green. For an American audience, at least, the color green is reminiscent of the color of money. In other situations, connotations may suggest linkings between variable values and the colors used to represent them. For instance, a visualization of temperature might show high temperatures in red and low temperatures in blue. Like application conventions, cultural connotations need not always be followed.

5. Conclusions

There are no hard and fast rules in the design of color scales. This paper raises questions to be asked about the color scale chosen for a visualization and suggests some general guidelines. The actual answers must come from the visualization designer after consideration of relevant factors, and perhaps with a bit of divine inspiration. In the end, the true test of the value of a color scale is simply "Does it work?"

Acknowledgements

This work was supported in part by a grant from the National Science Foundations (CAREER Grant No. ACIR 9996043).

References

Larry Bergman, Bernice Rogowitz, and Lloyd Treinish, "A rule-based tool for assisting colormap selection," *IEEE: Visualization '95,* IEEE Computer Society Press, Los Alamitos, CA, pp. 118-125, 1995.

William S. Cleveland and Robert McGill, "A Color-Caused Optical Illusion on a Statistical Graph," *The American Statistician*, vol. 37, no. 2, pp. 101-105, 1983.

J. Ronald Eyton, "Complementary-Color Two-Variable Maps," *Annals of the Association of American Geographers,* vol. 74, no. 3, pp. 477-490, 1984.

Stephen E. Fienberg, "Graphical Methods in Statistics," *The American Statistician,* vol. 33, no. 4, pp. 165-178, 1979.

Taosang He, Lichan Hong, Arie Kaufman, and Hanspeter Pfister, "Generation of Transfer Functions with Stochastic Search Techniques," *IEEE Visualization '96*, IEEE Computer Society Press, Los Alamitos CA, pp. 227-234, 1996.

A. Johannsen and R. Moorhead, "AGP: Ocean Model Flow Visualization," *IEEE Computer Graphics and Applications,* Vol. 15, No. 4, July 1995.

Gordon Kindlemann and David Weinstein, "Hue-Balls and Lit Tensors for Direct Volume Rendering of Diffusion Tensor Fields," Proceedings of IEEE Visualization '99, IEEE Computer Society Press, Los Alamitos CA, pp. 183-190, 1999.

H. Levkowitz and G. T. Herman. "Color scales for image data. " *IEEE Computer Graphics and Applications*, 12(1):72-80, January 1992.

J. Marks, B Andalman, P.A. Beardsley, W. Freeman, S. Gibson, J. Hodgins, T. Kang, B. Mirtich, H. Pfister, W. Ruml, K. Ryall, J.Seims, and S. Shieber, "Design Galleries: A General Approach to Setting Parameters for Computer Graphics and Animation," *SIGGRAPH 97 Conference Proceedings*, pp.389 –400, 1997

Stephen M. Pizer, and John B. Zimmerman, "Color Display in Ultrasonography," *Ultrasound in Medicine and Biology*, vol. 9, no. 4, pp. 331-345, 1983.

Penny Rheingans, "Color, Change, and Control for Quantitative Data Display," *Proceedings of Visualization '92*, IEEE Computer Society Press, Los Alamitos CA, pp. 252-259, 1992, 1992.

Penny Rheingans and Chris Landreth, "Perceptual Principles for Effective Visualizations," *Perceptual Issues in Visualization*, G. Grinstein and H. Levkowitz eds., Springer, pp. 59-73, 1995.

P. K. Robertson and J. F. O'Callaghan. "The generation of color sequences for univariate and bivariate mapping." *IEEE Computer Graphics and Applications* , 6(2):24-32, February 1986.

Philip K. Robertson and John F. O'Callaghan, "The Application of Perceptual Color Spaces to the Display of Remotely Sensed Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 49-59, 1988.

W. H. Tedford, Jr, S. L. Berquist, and W. E. Flynn, "The Size-Color Illusion," *The Journal of General Psychology*, vol. 97, 1977, pp. 145-140.

Lloyd Treinish, "Task-Specific Visualization Design," *IEEE Computer Graphics and Applications*, vol. 19, no. 5, September 1999, pp. 72-77.

Bruce E. Trumbo, "Theory for Coloring Bivariate Statistical Maps," *The American Statistician*, vol. 35, no. 4, pp. 220-226, 1981.

Colin Ware," Color Sequences for Univariate Maps:" Theory, Experiments and Principles, *IEEE Computer Graphics and Applications*, Sept. 1988, pp. 41-49, 1988.

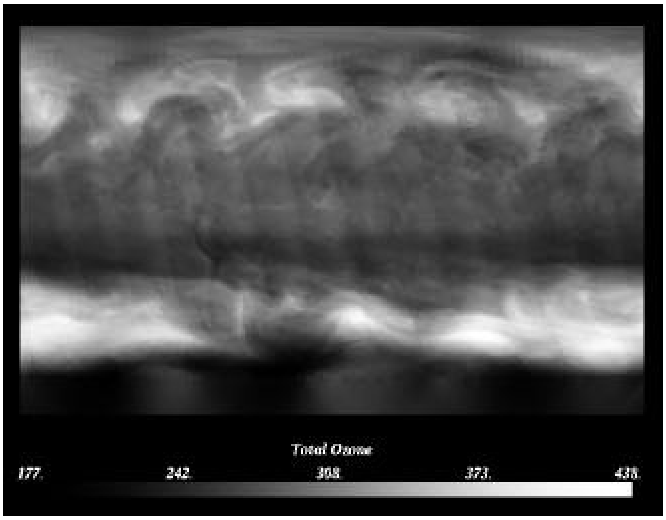


Figure 1. Grey scale.

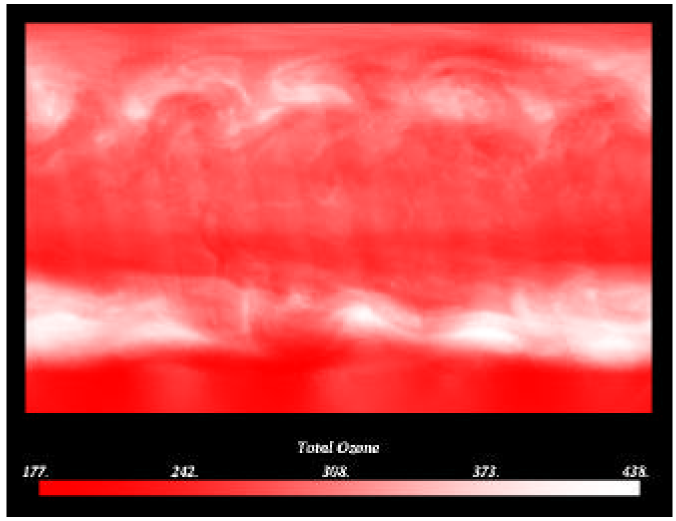


Figure 2. Saturation scale.

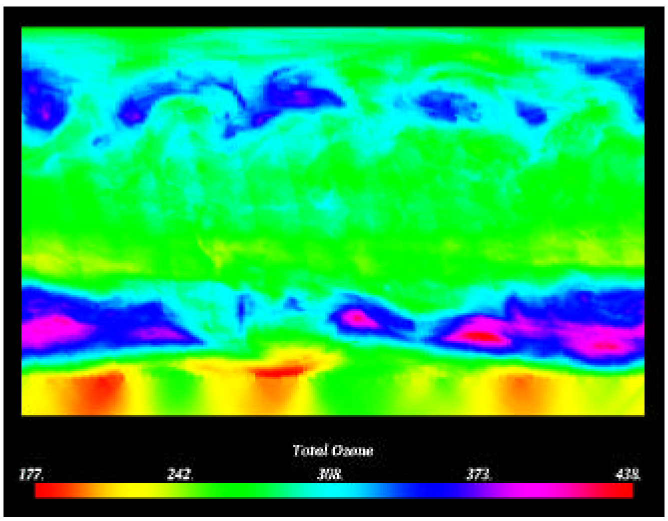


Figure 3. Spectrum scale.

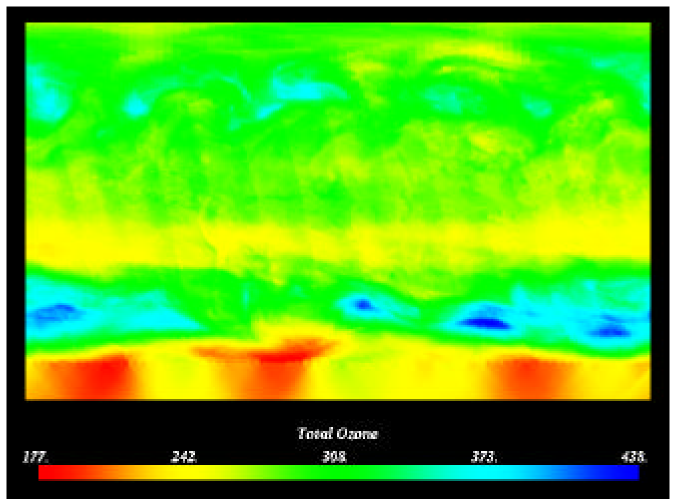


Figure 4. Limited spectrum scale.

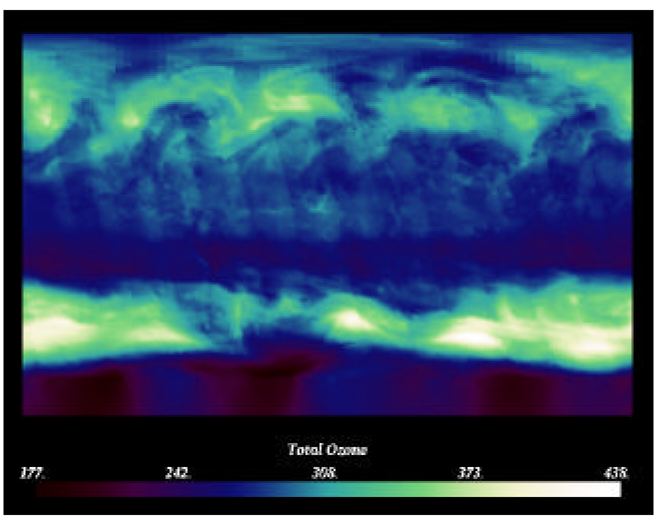


Figure 5. Redundant hue/lightness scale.

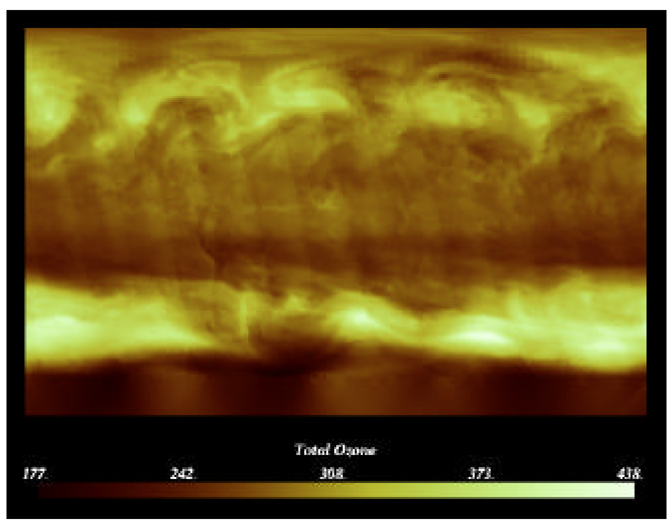
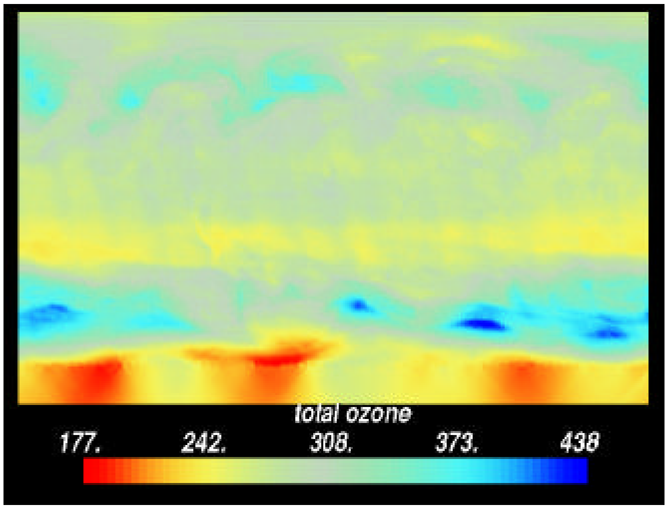


Figure 6. Heated-object scale.



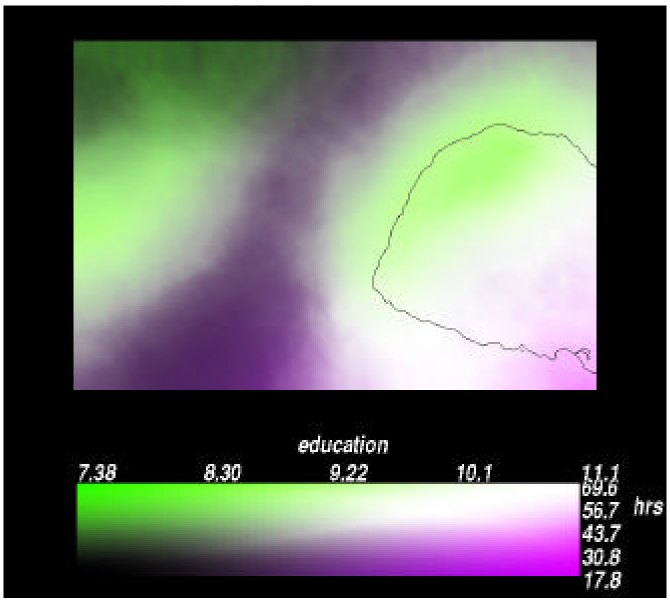


Figure 9. Bivariate complementary color scale.

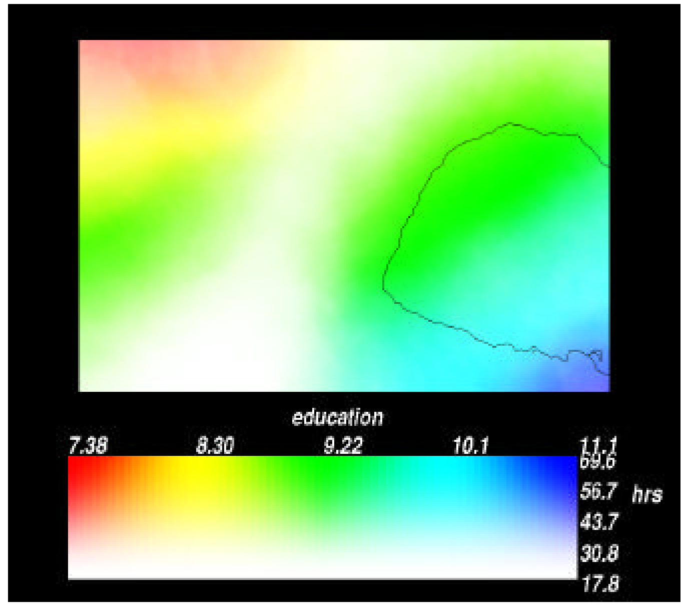


Figure 8. Hue-Saturation scale.