CIRCULAR DATAIMAGES FOR DIRECTIONAL DATA

By William Morphet

Dissertation Director

Dr. Juergen Symanzik
Department of Mathematics and Statistics
Utah State University,
Logan, Utah

Key Words

Circular Choropleth/ Circular Data/ Circular Dataimage/ Circular-Spatial Exploratory Data Analysis (CSEDA)/ Circular Statistics/ Dataimage/ Directional Data/ Exploratory Data Analysis (EDA)/ Graphical Statistics/ Hydromagnetics/ Interactive Graphics/ Magnetohydrodyamics/ Paleomagnetism/ Spatial Statistics/ Visual Data Mining (VDM)/

Abstract

Directional data includes vectors, points on a unit sphere, axis orientation, angular direction, and circular or periodic data. The theoretical statistics for circular data (random points on a unit circle) or spherical data (random points on a unit sphere) are a recent development. An overview of existing graphical methods for the display of directional data is given. "Cross-over" occurs when periodic data are measured on a scale for the measurement of linear variables. For example, if angle is represented by a linear color gradient changing uniformly from dark blue at -180 degrees to bright red at +180 degrees, the color image will be discontinuous at +180 degrees and -180 degrees, which are the same location. The resultant color would depend on the direction of approach to the cross-over point. A new graphical method for imaging directional data is described, which affords high resolution without color discontinuity from "cross-over". It is called the circular dataimage. The circular dataimage uses a circular color scale in which colors repeat periodically. Some examples of the circular dataimage include direction of earth winds on a global scale, rocket motor internal flow, earth global magnetic field direction, and rocket motor nozzle vector direction vs. time.

INTRODUCTION

Directional data includes random variates of type circular, axial, and spherical. Circular random variables have the total probability of all possible directions represented as a unit area between a circle of radius one and an enclosing curve. The radial distance between circle and curve is the probability density at an angle. Axial random variables describe random axis direction of axes where there is no reason to distinguish one end of the axis from the other, e.g. crystal axis orientation. Spherical random variables are random points on a sphere of radius one. Random vectors in a plane or 3D space have, in addition to random direction, random magnitude. Recent books on circular statistics include (Mardia, Jupp, 2000) and (Jammalamadaka, Sengupta, 2001).

Applications

Some Applications of Circular Statistics/Probability include: creature departure/vanishing direction at point of release (Biology), direction of fault lines or crystal axis orientation (Geology), magnetic field direction (Geophysics), response to treatment relative to time (Chronotherapeutics), wind direction (Meteorology), and ocean currents (Oceanography).

Historical Summary (Fisher, 1993)

Most of the theoretical advances in the statistics of circular/axial random variables are relatively recent. In 1767, John Mitchell, FRS tested the hypothesis that the distribution of angular separations of stars was uniform. He determined that the number of close stars were too many to support this hypothesis. In 1802, John Playfair noted that directional data should be analyzed differently from linear data, and recommended that average direction be the vector resultant direction. In 1858, Florence Nightingale, chief nurse in the British Army during the Crimean War, created the rose diagram to display the effect of sanitation vs. month of year to save thousands of lives in military hospitals. In 1880, Lord Rayleigh created a statistical test of the uniform circular distribution vs. the unimodal alternative. In 1918, Von Mises defined the circular normal, or Von Mises distribution. This distribution is a basis of parametric statistical inference for circular data. In 1939, Krumbein introduced the transformation of axial to vectorial data for analysis, and back transformation to axial results. In 1955, Potter smoothed directions on a regular grid on a map by averaging direction in four adjacent grid cells. The paper by G. Watson and E. Williams about statistical inference for the mean and dispersion of a sample from the Von Mises distribution began significant theoretical development. In 1976, S. Jammalamadaka created tests for uniformity based on runs or spacing.

This paper is organized as follows. First, existing methods for display of circular data will be summarized, including: The circular data plot for raw data; the rose plot and circular histogram to summarize frequency by in intervals on a circular scale; and non-parametric methods of representing frequency vs. location on a circular scale.

The cross-over problem from applying scaling for linear variables to directional data will be briefly described, a proposed solution detailed, and examples of the new paradigm for the display of high resolution directional data will be given including a circular choropleth (circular data on a map), flow image for CFD flow analysis, and circular time series for periodic and other processes. The new method may also be applied to vectors on contour or perspective plots.

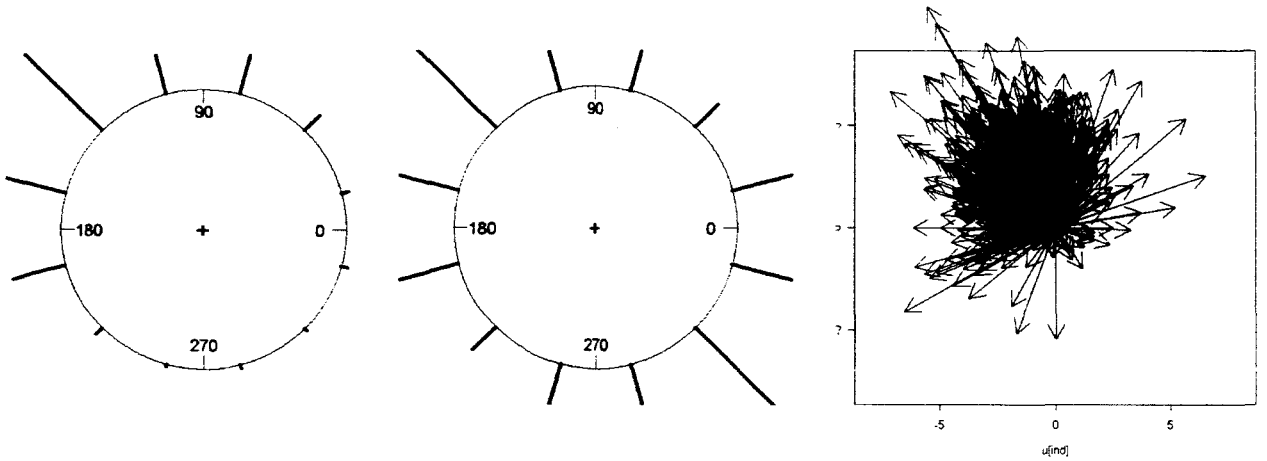
DATA USED IN FIGURES 1-4

The Comprehensive Ocean-Atmosphere Data Set (COADS) was created to provide a consistent and easily used historical record of surface marine data for 1854-1979. COADS began in 1981 as a cooperative project of National Climatic Data Center (NCDC), Environmental Research Laboratories, the cooperative Institute for Research in Environmental Sciences, and the National Center for Atmospheric Research.

70 million unique reports from ships-of-opportunity (primarily) and ocean buoys contain 28 elements of weather, position, etc. This data has been organized, edited removing outliers and data over land, and statistically summarized for each month using 2 x 2 degree latitude-longitude boxes. The trimmed monthly summaries give mean and median statistics for observed air and sea surface temperatures, wind, pressure, humidity, cloudiness, and derived variables.

The wind data used in figures 1-4 was extracted from the COADS dataset, following the tutorial on the COADS webpage (IRI Data Library, 2004), for the period from 1980 to 1990, January to April, and for area latitude 3S to 3N and longitude 93W to 87W. Wind direction is in degrees from East, positive sense CCW, and magnitude in m/s.

OVERVIEW OF EXISTING METHODS - GRAPHS FOR RAW DIRECTIONAL DATA



a) Circular Data b) Axial Data c) Vector Data

Figure 1, Plots for Raw Data, Using Wind Data Subset From COADS Dataset

How These Charts For Raw Data Work

Figures 1 a-b were generated using R (R Development Core Team, 2004) and contributor library CircStats (Lund, 2001). In figures 1 a-b, a bin width is specified as for a histogram. The number of data in each bin is determined. The data points are either overplotted, or stacked on the outside or inside of a circle in their bin. In figure 1a, circular data from the COADS subset have been stacked in 12 bins on the outside of a unit circle. Figure 1b is a graph of artificial axial data derived from the wind data by subtracting pi from any angle greater than pi, then constructing an axis thru each point. Figure 1c is simply an overplot of vectors showing magnitude together with direction. Figure 1c was generated using R and Fields (Nychka, 2004).

Interpretation

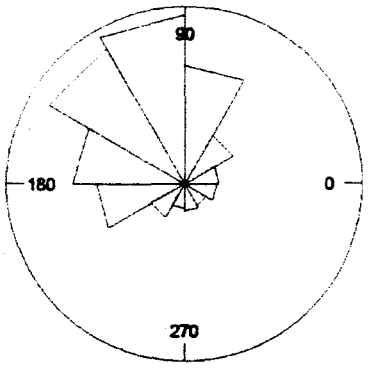
Most of this wind data are in a direction 135 degrees or wind blows from SE to NW. The distribution is unimodal and approximately symmetric. Note that this estimate depends on bin width.

OVERVIEW OF EXISTING METHODS - HISTOGRAMS FOR CIRCULAR DATA

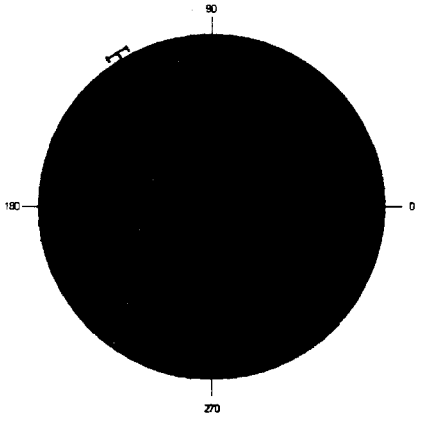
Another common method of representing data is to plot a histogram. The data are grouped in intervals and summarised by count. A rectangular bar is constructed of area proportional ($\propto$) to the count in the corresponding interval. The bars are centered above the midpoints of the intevals. The vertical axis of the histogram shows the count in each interval on the horizontal axis. With the histogram, we can see the frequency with which data occurs relative to value, whether frequency is consistent over the range, or data is concentrated at some value.

A histogram for angular data may be formed by first constructing linear histogram of the angular data as it were linear data. We select some starting point (e.g. 0 degrees if the data are recorded in the range [0,360), or -180 if the data are in the range [-180,180), and select some binwidth or grouping interval (e.g. 5, 10, or 20 degrees). The next step is to wrap the horizontal axis of linear histogram around a unit circle. The bars are aligned with the circle center, centered on the interval midpoint, and have length or area $\propto$ relative frequency. The bars can be replaced by whiskers of length $\propto$ relative frequency, or wedges or sectors with radius or area $\propto$ relative frequency.

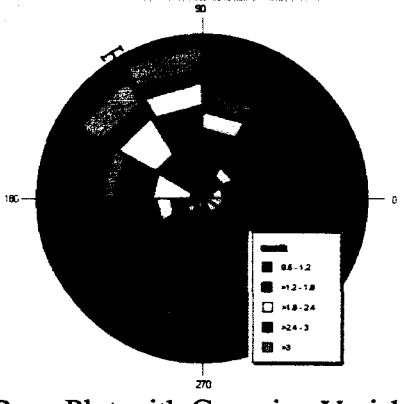
The wind data plotted in figure 2 a-d is the same wind data plotted in figures 1 a-c. Figure 2a was generated using R CircStats (Lund, 2001), and figures 2 b-d were generated by Oriana 2 (Kovach Computing, 2004).



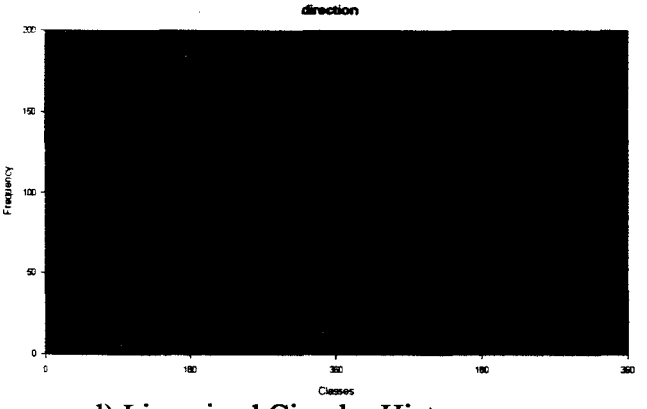
a) Rose Plot



b) Circular Histogram



c) Rose Plot with Grouping Variable



d) Linearized Circular Histogram

Figure 2, Histograms for Summarizing Frequency Relative to Direction
Using Wind Data Subset From COADS Dataset

How These Charts Work/Interpretation

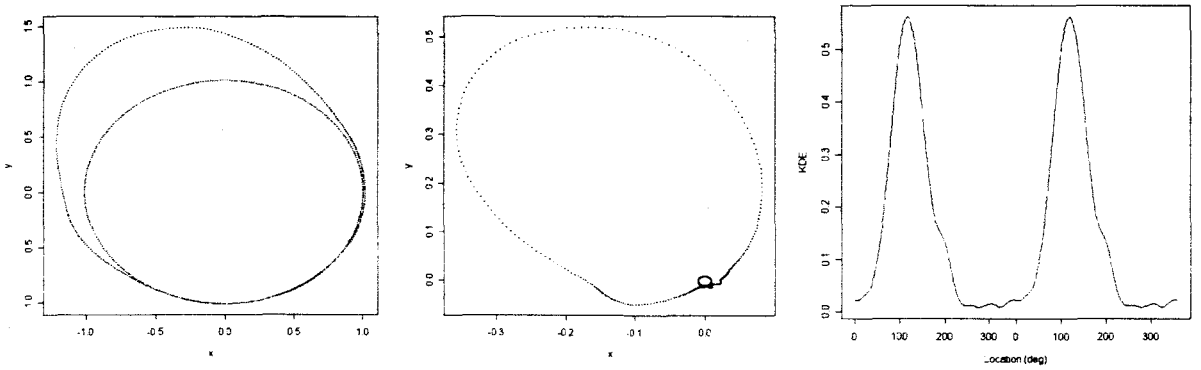
Figure 2a is a rose plot (Nightingale, 1858) generated by R CircStats. The angle of the wedge is the counting bin width. The radius or area of the wedge is set equal to the corresponding bin count. However, area is preferred to radius because "radius results in perceptual bias in favour of larger frequencies" (Fisher, 1993). Fisher's book is a handbook of methods for circular random variables.

Figure 2b is a circular histogram. A radial bar has area equal to frequency. Figure 2c partitions the rose plot of figure 2a by a sub-grouping variable (1=Jan, 2=Feb, 3=Mar, 4=April). Figure 2d unwraps the circular histogram of figure 2b onto a linear scale. This linear histogram repeats one period, which is an advantage if a feature of interest occurs near 0 degrees, and in visually extracting the periodicity. In figures 2 b-c, the mean direction is indicated by a thick black radial line with confidence interval indicated by a thick black arc. In figure 2d, the mean is indicated by a thick black vertical line with confidence limits indicated by the enclosing vertical dashed lines. In the Oriana 2 (Kovach Computing, 2004) plots (figures 2 b-d), the confidence interval changes to red when the estimate of the confidence interval is unreliable. The confidence interval for the mean is about 122 +/- 20 deg and reliable. In figure 2c, the mean direction appears to vary with month.

OVERVIEW OF EXISTING METHODS - NON-PARAMETRIC KERNEL DENSITY ESTIMATES (KDE)

Circular histograms of figure 2, like histograms for linear variables, can have significant distortion of the information in the sample about the number, sizes, and locations of modal groups thru the choice of bin origin and width. The non-parametric kernel density estimate replaces the bin edge and origin decision with an easier band width decision. Alternatively, for a method that is easier to explain, the ASH may be implemented for rose plots or circular histograms (Scott, 1992).

The wind data plotted in figure 3 is the same data plotted in figures 1-2.



a) KDE Plotted on Radius=1 b) KDE Plotted on Min Radius c) KDE Unwrapped & Plotted on Repeating Linear Scale

Figure 3, Non-Parametrically Smoothed Histograms of Frequency vs. Direction Using Wind Data Subset From COADS Dataset

Construction of a Kernel Density Estimate

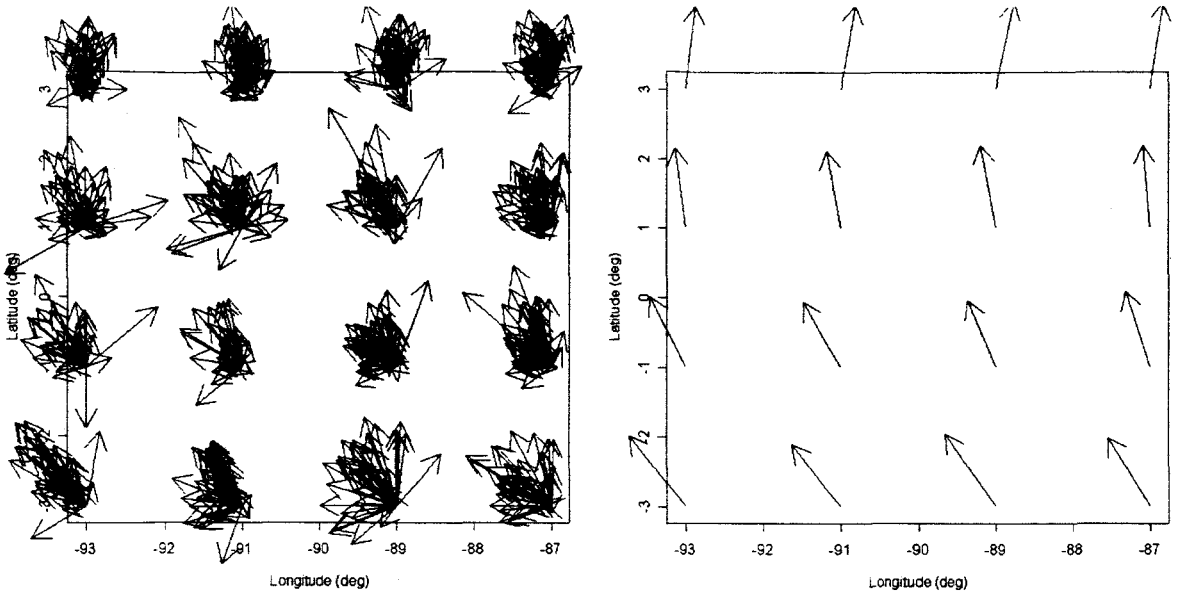
These charts were generated using non-parametric smoothing (N. I. Fisher, 1993, p. 26). In non-parametric smoothing, a symmetric function, e.g. quartic kernel, is placed on each observation. It spreads out the mass of an observation with maximum amplitude at the location of the observation. This is motivated by considering that the observation is an instance of random direction and could have occurred at other locations in neighborhood of the observed direction. Density at an angle of estimation is sum of kernel contributions from observations around the angle of estimation.

How these Charts Work/Interpretation

Different effects result from plotting the estimated density on different radii of one, smallest density, and zero. In figure 3a, the probability density is plotted on a unit circle. This emphases regions of high probability density (Fisher, 1993, p.24). In Figure 3b, the probability density is plotted on a circle of radius = minimum density emphasizing the variation in probability density with direction. In figure 3c, the density is unwrapped onto a linear scale with an extra period. Note that in the middle of the horizontal axis, the scale values repeat beginning at 0 degrees. This method removes the subjectivity associated with choice of plotting radius, and the repeated plot makes it easier to interpret structure, e.g., periodicity and count modes. The extra plot also eliminates breakup of interesting features at the crossover point of zero degrees. Axis (Pisces Conservation Ltd, 2004) provides smoothing using a fast fourier transform. The linearized histogram in figure 3c indicates there is 1 mode at about 118 deg.

OVERVIEW OF EXISTING METHODS - ARROW PLOTS OF VECTOR-SPATIAL DATA

The wind data plotted in figure 4 is the same data plotted in figures 1-3.



a) Raw Data b) Average (Vector Resultants) of Raw Data

Figure 4, Vector-Spatial Plot, Using Wind Data Subset From COADS Dataset

Vectors of magnitude and direction are represented by an arrow superimposed on a surface, contour, or perspective. See also figure 12c. In figure 4b, the regional structure of the mean wind vector is continuous and regular. Direction is changing in the S-N direction and relatively constant in the W-E direction.

INTRODUCTION TO DATAIMAGES FOR CIRCULAR DATA
DATAIMAGES FOR LINEAR DATA

The new dataimage for circular variables (Morphet, 2004) evolved from the dataimage for linear variables (Minnotte, 1998, 1999), which in turn, evolved from the color histogram (Wegman, 1990).

In 1990, Wegman created a relatively simple plot for visualizing hyper-variate (dimension greater than three) data structure. With variables plotted along one axis and observations along the second axis, up to several hundred variables and several hundred observations can easily be observed. Pixels were colorized according to a binned color gradient, and he called the result a color histogram. As suggested by Wegman, positive and negative association among variables is easily seen by sorting by each variable.

The dataimage (Minnotte, 1998, 1999), like the color histogram, arrayed observations on the horizontal axis and colorized variables on the vertical axis forming a hypervariate image. The capability to image hypervariate structure was enhanced by standardizing variables, implementing a continuous linear color gradient, adding distance metrics, computing n-dimensional distance between observations, and adding sorting methods for both variables and observations according to n-dimensional distance.

In the continuous red-cyan gradient (figure 5), as the level of a variable increases, the red, green, and blue levels of a pixel change linearly with red decreasing, and green and blue increasing uniformly. Pure red indicates the minimum of a variable and cyan the maximum.

Distance metrics implemented include: euclidean - L2; manhattan - L1; and maximum – L_Infinity. Sort methods implemented include: none - leave in original ordering; complete - hierarchical clustering, where cluster distances are measured as max of point distances; single - hierarchical clustering, where cluster distances are measured as min of point distances; average - hierarchical clustering, where cluster distances are measured as average of point distances; farthest - farthest insertion spanning tour; nearest - nearest insertion spanning tour; prcomp - sort on first principal component; mds - sort by 1-D multidimensional scaling; and k - sort on variable k.

Figure 5 is a dataimage of weight (wt), displacement (disp), horse power (hp), and miles per gallon (mpg) of 32 1973-74 model cars extracted from the 1974 *Motor Trend* US magazine (Henderson, Velleman, 1981). The image shows an inverse correlation between mpg and {hp, disp, wt}. The vertical bands indicate two major clusters (subgroups) of cars: {low weight, low displacement, low hp, high mpg} at the left,, and {high weight, high displacement, high hp, low mpg} at the right. The Maserati Bora has the greatest hp, the Toyota Corolla has the greatest mpg, and the Chrysler Imperial, Lincoln Continental, and Cadillac Fleetwood have the largest weight, displacement, and lowest mpg..

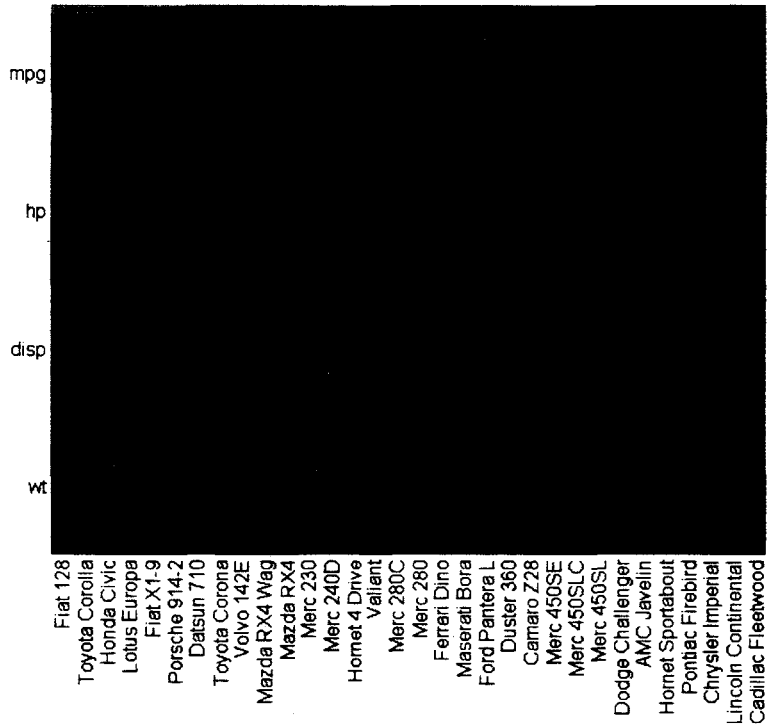


Figure 5, Dataimage, a MultiVariate Plot of Motor Trend Car Data

When observations are ordered by time, dataimages reveal discrete events as pattern perturbations, and multivariate trends as color gradients. Observations and/or variables also indicate clusters, outliers, and correlation as visual structure.

THE "CROSS-OVER PROBLEM"

On a compass, the average direction of 1° and 359° is not the arithmetic mean of 1 and 359 =180. It is 0° or 360°. This has been called the "cross-over problem". Cross-over also occurs when plotting angle on a linear scale vs. time. The plotted angle can vanish at the bottom of the scale at 0° and instantly reappear at the top at 360° resulting in a discontinuous plot.

A NEW GRAPHICAL METHOD - DISPLAY OF DIRECTION DATA USING A COLOR WHEEL, A CIRCULAR/PERIODIC SCALE

The dataimage for linear variables is extended to directional data by displaying the value of direction as a color on a periodic color scale, or "color wheel" such as in figure 6. Displaying direction as a color from the corresponding position of the color wheel, the representation of any cross-over point, e.g. 0 deg, is continuous through the crossover point, whether approached from one side or the other of the crossover point. This eliminates the cross-over problem.

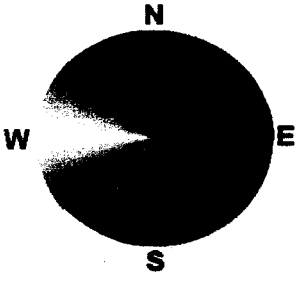


Figure 6, A GBYR (Green Blue Yellow Red) Periodic Color Scale, or Color Wheel

Flow analysis and circular statistics graphics from commercial and free software (see References) were examined to determine if the proposed method had been previously implemented. No examples of this method have been found. In Fluent, a CFD flow analysis system, vector angle is imaged by coloring pixels corresponding to -180° a deep blue. The linear color gradient (Gouraud shading) increases uniformly to bright red at +180°. The resulting image is discontinuous about the 180°. Consequently, users may examine partial images oriented so the direction of interest lies at 0° to avoid directions around 180°.

Following some details of color wheel construction, examples will be given including wind direction, internal flow within a rocket motor chamber, earth magnetic field, and circular time series. The first circular dataimage was of Space Shuttle solid rocket motor nozzle vector direction (Morphet, 2003-4).

COLOR WHEEL CONSTRUCTION

In figure 6, the color wheel has 4 main colors: green, blue, yellow, and red going CCW from East or 0° in this application, so it is called a GBYR (green blue yellow red) color wheel. The selection of this color scheme is a reflection on green forests to the east, cold blue glaciers to the north (for another 30 years?), yellow deserts to the west, and red hot summers to the south. These four colors seem to be more intuitive than three colors in a geographic context.

The color wheel in Figure 6 was implemented in R 1.09 (CRAN, 2004). Like a smoothing spline, it contains 4 continuous linear color gradients (Gouraud shading) with color continuity at the knots. In a GBYR color wheel going 0° to 90°, the color green is decreased and the color blue is increased. The R rgb command creates "colors" corresponding to the given intensities (between 0 and max) of the red, green and blue primaries, respectively:

```
if(Angle < 90) # green to blue gradient
{
        dist = 255*Angle/90
        color = rgb(0, 255-dist, dist, maxColorValue=255)
}
if(Angle >= 90 & Angle < 180) # blue to yellow (red + green) gradient
{
        dist = 255*(Angle - 90)/90
        color = rgb(dist, dist, 255-dist, maxColorValue=255)
}

if(Angle >= 180 & Angle < 270) # yellow to red gradient
{
        dist = 255*(Angle - 180)/90
        color = rgb(255, 255-dist, 0, maxColorValue=255)
}

if(Angle >= 270) # red to green gradient
{
        dist = 255*(Angle - 270)/90
        color = rgb(255-dist, dist, 0, maxColorValue=255)
}
```

green

red

blue

COLOR CONSIDERATIONS

In figure 6, the color saturation at a particular angle is constant. However, vector resultant magnitude, number of observation, or other statistic have been used to modify color saturation by reducing the levels of red, green, and blue proportionately. Weighting saturation by the amount of data and/or resultant magnitude draws attention to significant areas. The colors of all directions fully desaturate to black, which may be the color of missing data. The use of color in computer graphics is described in (Foley, Van Dam, 1992). About 99% of color blindness is the inability to distinguish red and green. Substitution of magenta for red may reduce low saturation red-green color impairment. Colors for the color impaired are described in http://www.toledo-bend.com/colorblind/index.html (9/19/04).

EXAMPLE - CIRCULAR CHOROPLETH

Data used in figure 7: Comprehensive Ocean-Atmosphere Data Set (COADS) wind direction in degrees from East (positive sense CCW) and magnitude in m/s, Longitude -151° to +51°, Latitude -45° to +45°, 1972-3, 1975-6 (IRI Data Library, 2004).

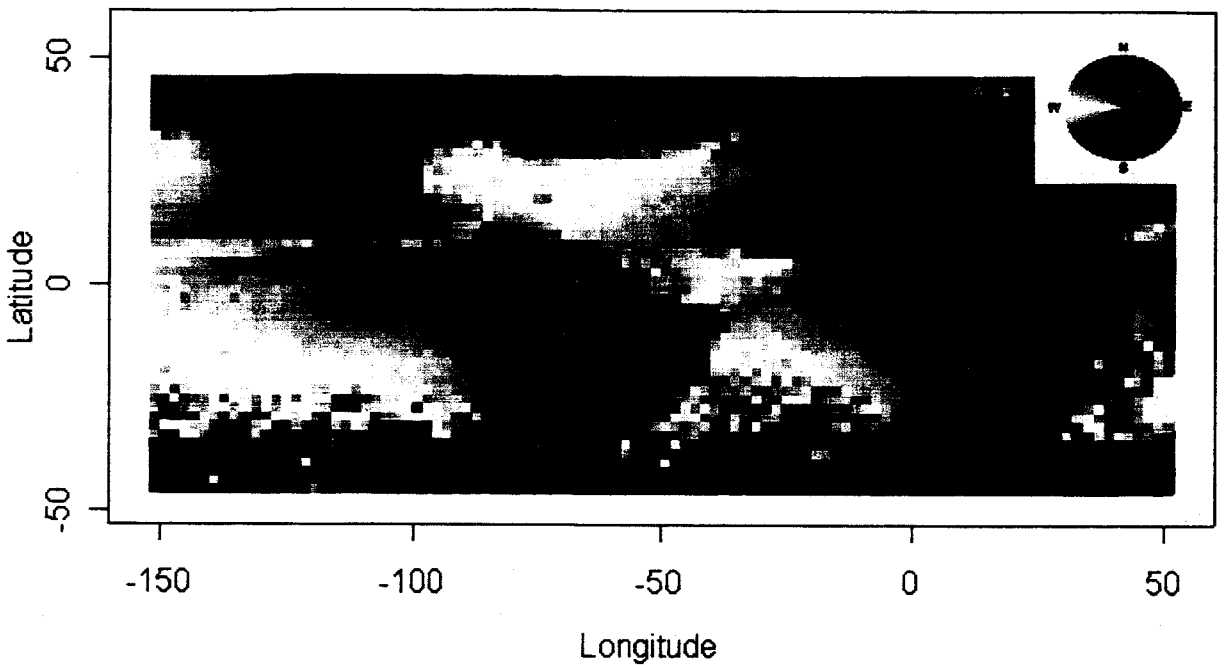


Figure 7, Circular Choropleth of Average Ocean Wind Direction
Using Wind Data Subset From COADS Dataset,
Green:West-to-East, Blue:South-to-North, Yellow:East-to-West, Red: North-to-South

In figure 7, each pixel is horizontally 2 degrees of longitude and vertically 2 degrees of latitude. Direction is indicated by a GBYR (green blue yellow red) color wheel. Black indicates missing data or land mass. Winds were averaged over the years 1972-3 and 1975-6 taking into account vector magnitude wind speed:

$$\text{atan}^{-1}\left(\frac{1}{N_{IJ}}\sum_{K=1}^{N_{IJ}} M_{IJK}\cdot\cos(\theta_{IJK}),\frac{1}{N_{IJ}}\sum_{K=1}^{N_{IJ}} M_{IJK}\cdot\sin(\theta_{IJK})\right)$$

$I \in \{\text{minimum longitude, maximum longitude}\}$

$J \in \{\text{minimum latitude, maximum latitude}\}$

N_{IJ} is the number of observations in grid cell IJ

$K \in 1{:}N_{IJ}$ is the observation index

M_{IJK} is magnitude of vectorial observation K in grid cell IJ

θ_{IJK} is direction in radians of vectorial observation K in grid cell IJ

atan^{-1} is the quadrant specific inverse tangent

Interpretation

West of south end of north America, the green to red to yellow color gradient indicates a turning of wind from east to south to west. On both sides of the equator in this global scale, average wind flows toward the equator, then goes west. The color shape of wind direction on the west side of the Americas is similar to the color shape on west side of Africa. Apparently, the continental masses have similar effect on global wind. The speckled area at lower latitude and in the upper right is due to low sample size resulting in higher variability in the mean (vector resultant) direction.

EXAMPLE - FLOW IMAGE

The solid rocket motor flow data in figures 8-9 is simulated (ATK Thiokol Propulsion, 2004) using arbitrary geometry and fictitious motor properties. Features of the rocket motor chamber shown in figure 8 include the forward end at X=0 in., experimental cavities at about X = 6 in., and nozzle from X = 13 to 20 in.. Combustion products entering the nozzle are compressed with maximum compression occuring at about X = 14 in. where nozzle diameter is minimum. This is called the throat. Aft of the throat, nozzle diameter increases. This causes combustion products to expand and increase in velocity, producing thrust. The magnitude component of flow velocity is in units of inches/second.

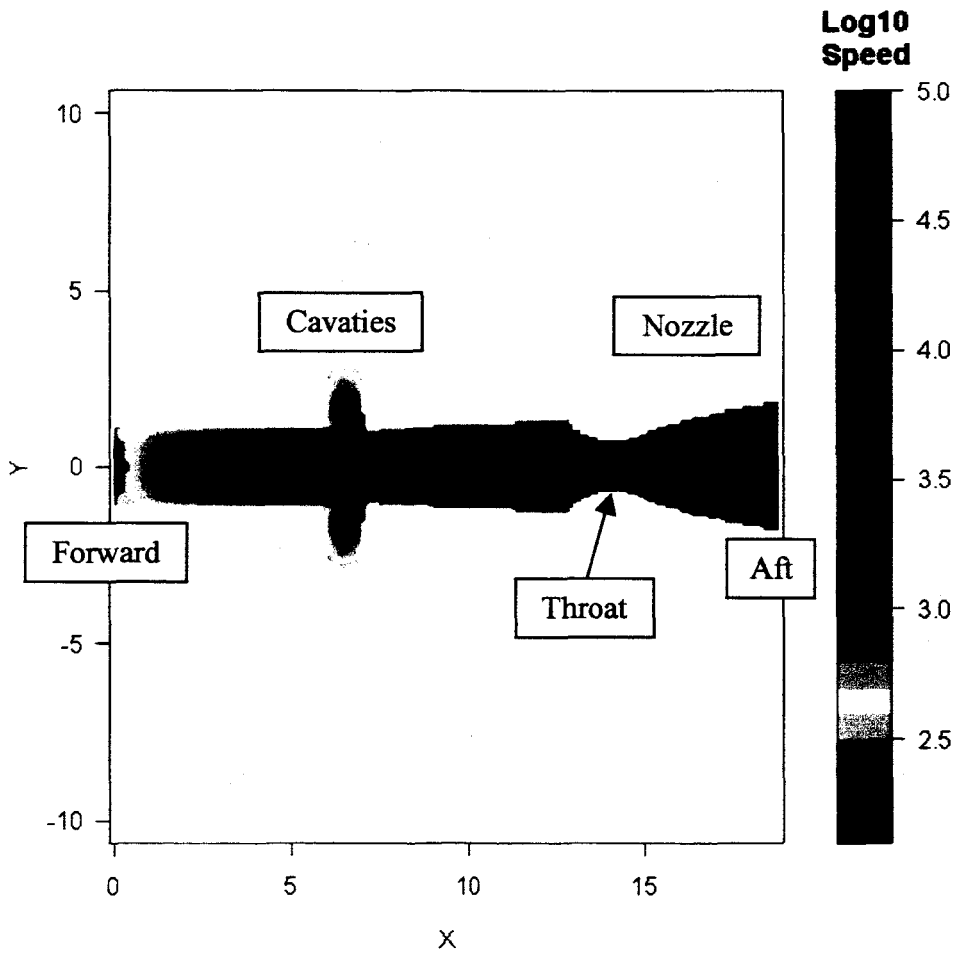


Figure 8, Filled Contour Plot of Log Speed of Simulated Combustion Product Flow Inside Solid Rocket Motor of Arbitrary Geometry and Fictitious Properties

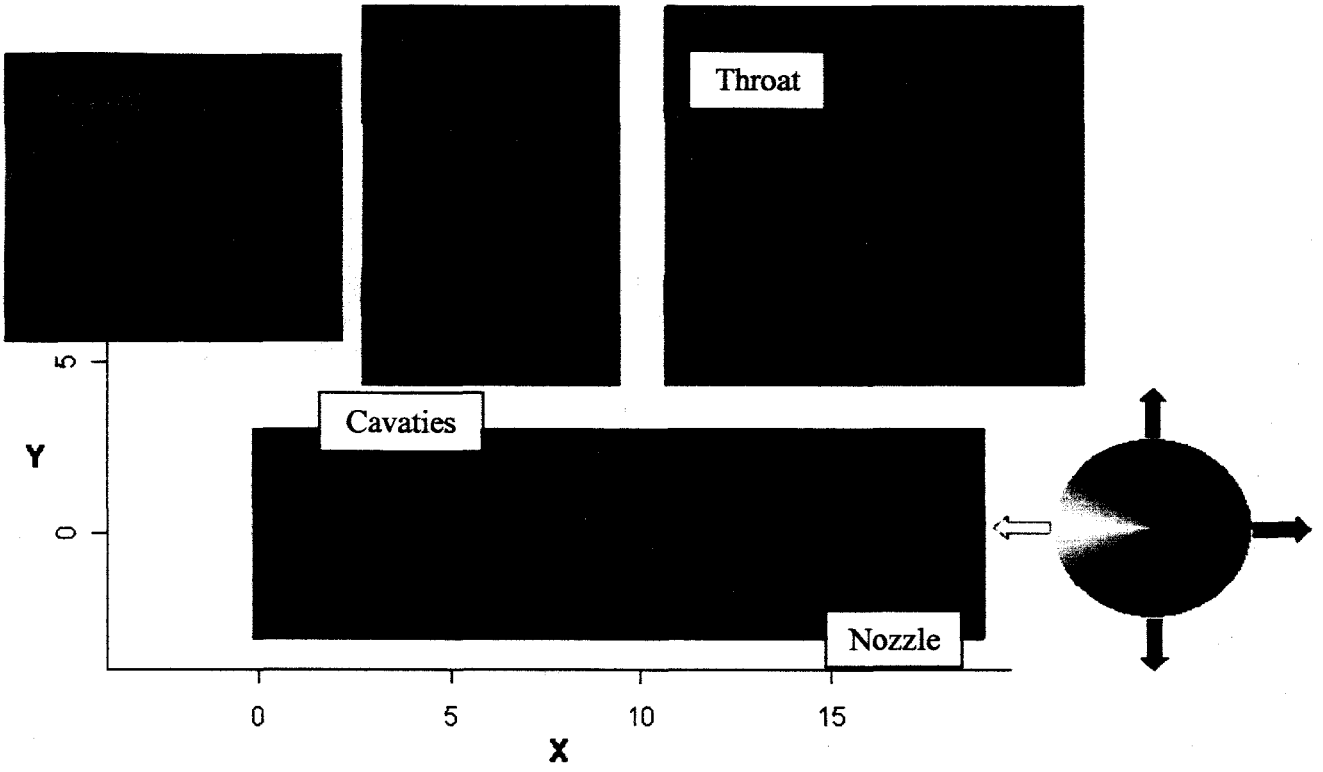


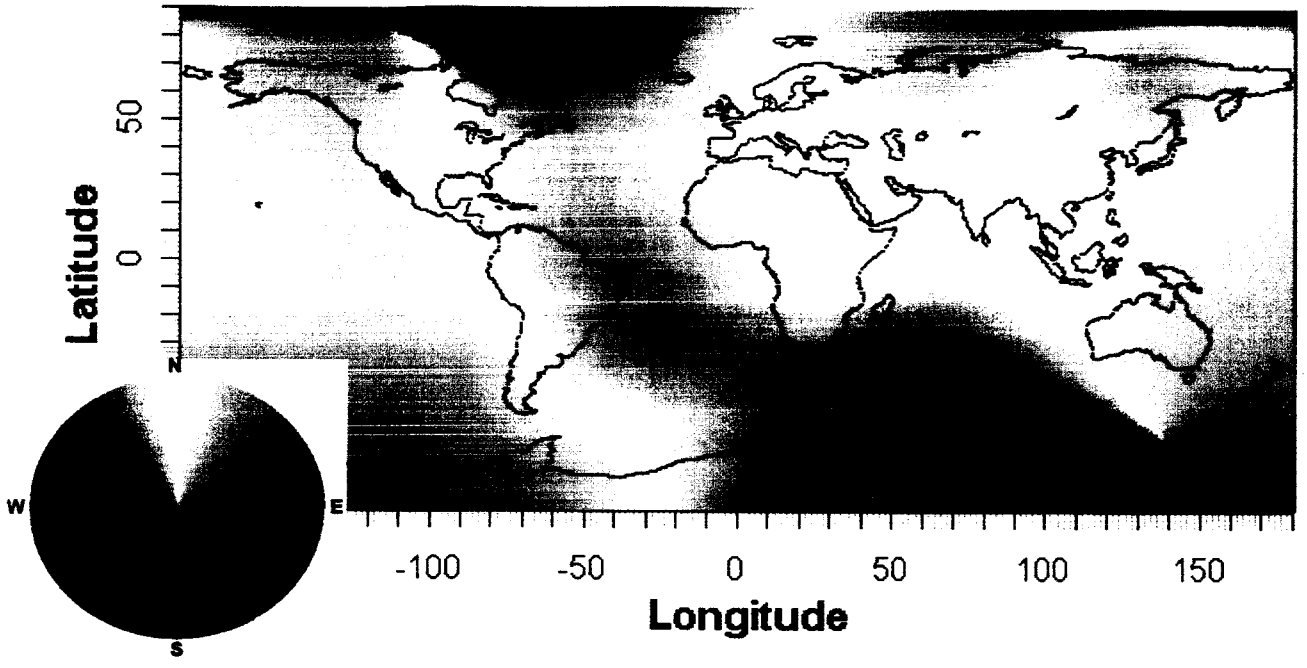
Figure 9, Flow Image of Direction of Simulated Combustion Product Flow Inside Solid Rocket Motor of Arbitrary Geometry and Fictitious Properties Using GBYR Color Wheel

Interpretation of Circular Dataimage

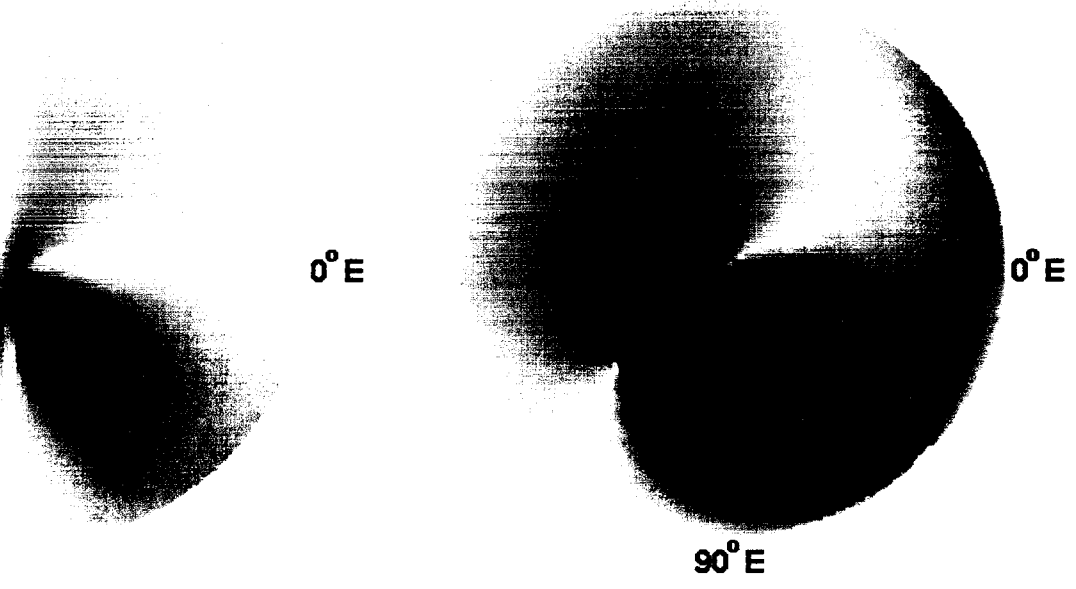
In the forward (left) end of the chamber, combustion products from the burning rocket motor propellant flow toward the center of the chamber. Going aft (to the right), the flow direction turns toward motor axis with the volume of radial flow (toward the center of the chamber) decreases going aft (see upper left view). The radial flow coming from the cavities near X=6 in. into the bore turns toward the axial direction going aft.

In flow analysis, displays like figures 8-9 can be linked, or semi-transparent isoclines of constant pressure or other variable can be overplotted on flow images of direction.

EXAMPLE – DATAIMAGE OF EARTH MAGNETIC FIELD



a) View With Longitude x Latitude Areal Elements As Rectangles Of Constant Area



b) Perspective View from North Geographic Pole Looking Toward Equator

c) Perspective View from South Geographic Pole Looking Toward Equator

Figure 10, Earth International Geomagnetic Reference Field, 9/15/2004, Elevation 0 km, Resultant Direction Displayed by a GYRB Color Wheel for North and East Vectors in Plane Tangent to Earth at (Longitude, Latitude),

Figures 10 a-c plot the direction of the component of earth's main magnetic field which drives a compass using a GYRB (green=0°, yellow=90°, etc) color wheel. The dominant direction (north) has green and red color variations in shading around yellow. Yellow was chosen for north because it seems easier to see variations against yellow than a dark color, e.g. blue.

The geomagnetic field at any point on the Earth's surface is a combination of several magnetic fields generated by various sources. More than 90% of the field is generated in the Earth's outer core. This is referred to as the main field. The main field was estimated for date 9/15/2004, elevation 0 km., and area -180° E to +180° E by -90° N to +90° N in 1 degree x 1 degree areas using the web implementation of the current International Geomagnetic Reference Field (IGRF) model (International Association of Geomagnetism and Aeronomy 8th Generation IGRF, 2004). Computed results include the seven field parameters and the current rates of change: Declination (the angle between magnetic north and true north), positive east in degrees and minutes; Inclination (the angle between the local horizontal plane and the total field vector), positive down in degrees and decimal minutes; Horizontal Intensity (H), in nanoTesla (nT); North Component of H, positive north in nT; East Component of H, positive east in nT; Vertical intensity, positive down in nT; and Total Field = sqrt(north^2 + east^2 + vertical^2) in nT. When used within a valid date range, estimates are generally accurate to within 30 minutes of arc for declination and inclination, and 100-250 nT for the force vector components. This precision is remarkable considering earth's total magnetic field intensity varies from about 23,000 to 67,000 nT. Solar storms can cause significant short-term differences between the estimated and actual field values.

NGDC

The scientific domain of the National Oceanic & Atmospheric Administration (NOAA) spans the distance from the surface of the sun to the bottom of the sea. The strategic plan of the NOAA is "To understand and predict changes in the Earth's environment and conserve and manage coastal and marine resources to meet the Nation's economic, social, and environmental needs." The National Geophysical Data Center (NGDC), is a part of NOAA. The NGDC provides scientific stewardship, products and services for geophysical data describing the solid earth, marine, and solar-terrestrial environment, as well as earth observations from space The NGDC provides the interface to 8th Generation IGRF.

Interpretation

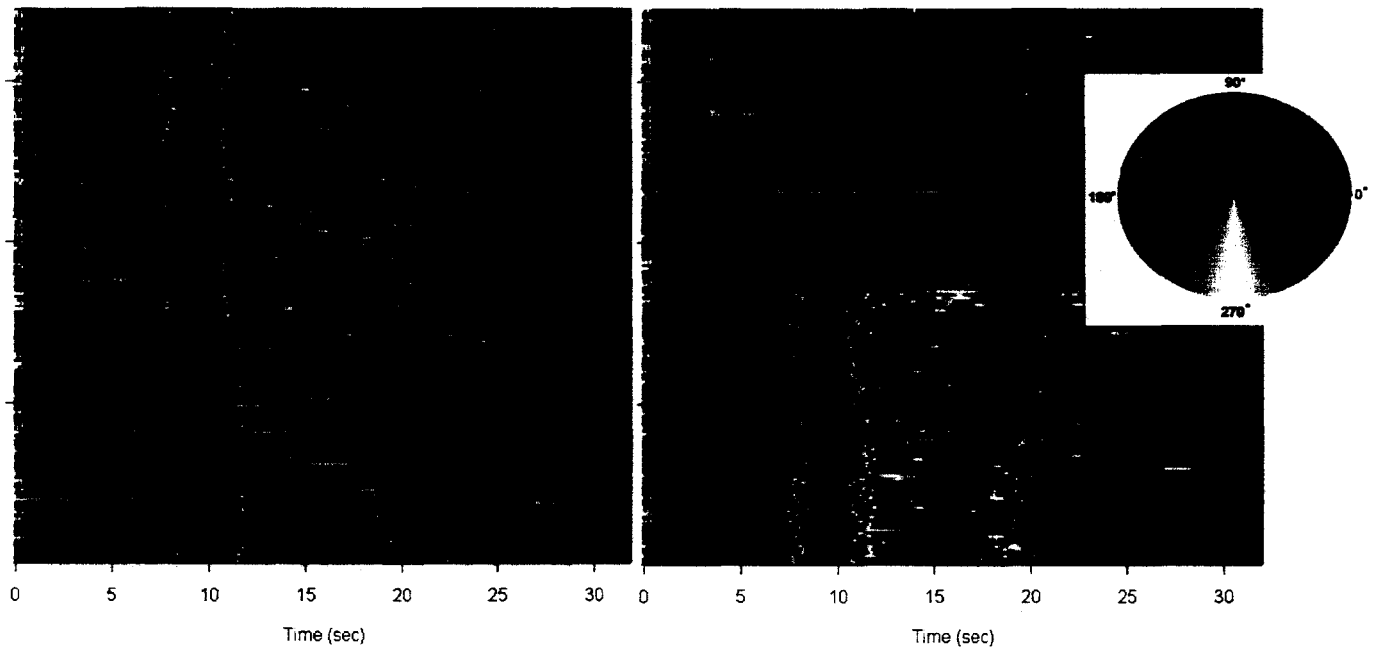
In figure 10a, in the central region and going west to east, magnetic field direction varies from north-northeast, rotates counterclockwise, has a narrow S-shaped band crossing North America and South America in which the direction is nearly north (yellow), continues rotation CCW to north-northwest, and then reverses direction of rotation CW back to north. The rectangular representation in figure 10a distorts the shape of direction. Figures 10b-c corrects this distortion by mapping the color of the magnetic field direction onto a sphere. Starting at the center of Figure 10a and going outward, latitude increases from -90° N to zero at the equator. Continuing, starting at the outside of figure 10c and going toward the center, latitude increases from 0° N to +90° N. The variation around the geographic poles shows why a compass is not reliable within about 10° and must be used with care in the adjacent region.

The circular dataimage of magnetic fields direction provides a new way to compare internal structure of planets.

EXAMPLE - DATAIMAGE OF CIRCULAR TIMES SERIES

Data used in figure 11 is Space Shuttle solid rocket motor nozzle vectoring direction angle (direction nozzle is pointing in plane perpendicular to its motor).

In figures 11a-b, time increases left-to-right. Each row of pixels represents a circular time series of nozzle direction angle vs. time. In each row, a pixel represents nozzle direction angle by color at a corresponding angle on the RGBY color wheel at the right of figure 11. Like the dataimage for linear variables of motor trend cars, the observations and/or variables can be sorted. In figure 11a, the nozzle direction data are ordered in the natural order in which the motors are produced. In figure 11b , the data are first sorted bottom-to-top by side (left side then right side motor), then by inclination (angle of Space Shuttle to plane through equator), and last by orbital altitude.



a) Unsorted data

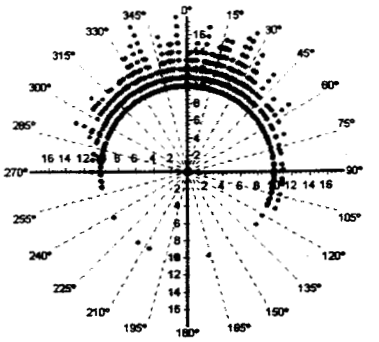
b) Sorted data

Figure 11, Space Shuttle Solid Rocket Motor Nozzle Vectoring Direction Angle vs. Time

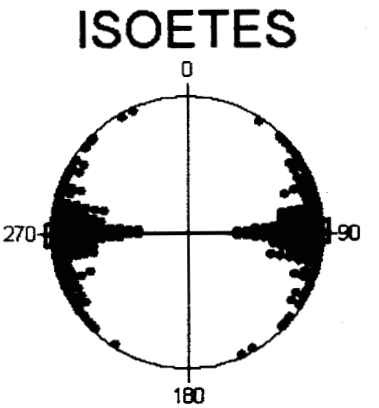
Interpretation

The broad horizontal bands in figure 11b bottom and top half indicate that the left side (bottom half) nozzle is mostly pointed at its 0 degree location, and the right side nozzle (top half) at 180 degree location. The blue vertical bands in figures 11 a-b at about time = 10 seconds is related to the roll maneuver of the Space Shuttle, which occurs early in flight. The diagonal structures in figure 11b indicate effects of Space Shuttle inclination and orbital altitude on nozzle direction angle.

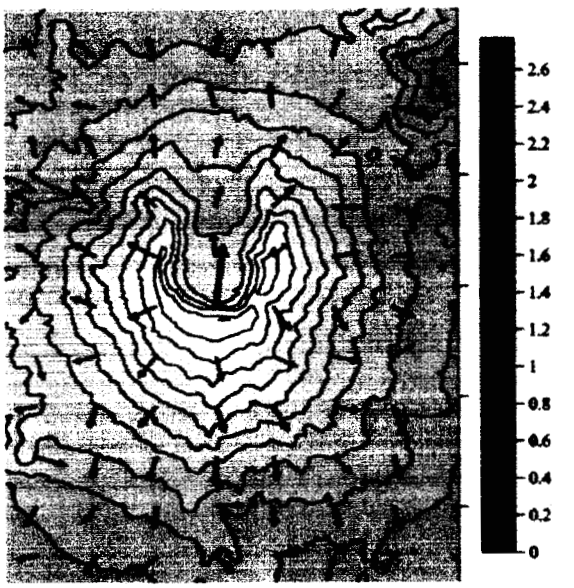
EXAMPLES - EXISTING METHODS OF DISPLAYING DIRECTIONAL DATA



a) Stacked Circular Data Graph
(Pisces Conservation Ltd , Axis, 2004)



b) Stacked Axial Data Graph
Simulated Data Megaspore Element Orientation



c) Vector Plot on Elevation Contour of Mount St. Helens
(Golden Software, Surfer 8, 2004)

Figure12, Examples of Graphics Generated by Commercial Software of Directional Data

Figure 12b is a raw data plot for axial data showing orientation of elements of modern Isoetes megaspore relative to a base line drawn perpendicular to the spore surface (Kovach, 1989).

Figure 12c is an elevation contour of Mount St. Helens generated by Surfer 8 Golden Software (2004). On a regular grid, vectors point in the direction of the steepest descent. The magnitude of the vector is the rate of descent. Mount st. Helens, southwestern Washington, erupted on May 18, 1980, killing 57 people and damaged life in an area of some 70 sq mi.

REFERENCES

ATK Thiokol Propulsion (2004). FEM Builder software, version 1.4.

Fisher, N. (1993). Statistical analysis of circular data. Cambridge University Press, Cambridge.

Fluent (2004). Software for flow analysis and visualization at http://www.fluent.com/ (8/14/04).

Foley, J., Van Dam., A., Feiner, S., Huges, J. (1992). Computer graphics : principles and practice. Addison-Wesley Pub. Co., Reading Mass.

Golden Software (2004). Surfer 8. 2d vector plot of Mt. St. Helens reproduced with permission. www.goldensoftware.com (7/22/04).

Henderson, H., Velleman, P. (1981). Building multiple regression models interactively. *Biometrics*, **37**, 391–411.

International Association of Geomagnetism and Aeronomy 8th Generation IGRF (2004).
IGRF web site at
http://www.ngdc.noaa.gov/IAGA/vmod/
Frequently Asked Questions about earth magnetic field at
http://www.ngdc.noaa.gov/seg/geomag/faqgeom.shtml (9/15/04).
On-line magnetic field calculator at
http://www.ngdc.noaa.gov/seg/geomag/jsp/IGRFGrid.jsp (9/15/04).

IRI Data Library (2004). Comprehensive Ocean-Atmosphere Data Set (COADS), Trimmed Monthly Summaries. Tutorial at http://iridl.ldeo.columbia.edu/dochelp/Tutorial/ (9/12/04).
Data at http://iridl.ldeo.columbia.edu/SOURCES/.COADS/.mean/ (9/12/04).

Jammalamadaka, S., Sengupta, A. (2001). Topics in circular statistics. World Scientific, Singapore.

Kovach, W. (1989). Oriana 2 graphic of orientation of elements of modern isoetes megaspore reproduced with permission.

Kovach Computing (2004). Oriana 2. http://www.kovcomp.co.uk/oriana/newver2.html (7/22/04)

Lund, U. (2001)-S version, Agostinelli, C. (2002)- R port. CircStats, an R contributor package at http://www.cran.r-project.org/ (8/1/04).

Mardia, K., Jupp, P. (2000). Statistics Of Directional Data. John Wiley & Sons, New York.

Minnotte, M. (1998). 1998 Proceedings of the ASA Section on Statistical Graphics.

Minnotte, M., West, R. (1999). The data image: A tool for exploring high dimensional data sets. http://math.usu.edu/~minnotte/research/pubs.html (7/22/04).

Morphet, W. (2004). Space Shuttle Solid Rocket Motor Nozzle Vectoring Direction Angle vs. Time, ATK Thiokol Propulsion. Clearance for publication applied for.

Nychka, D. (2004). Fields, an R contributor package at http://www.cran.r-project.org/ (8/1/04).

Pisces Conservation Ltd (2004). Axis software. Graphic reproduced with permission. http://www.pisces-conservation.com/indexsoftaxis.html (7/22/04).

R Development Core Team (2004). R: A language and environment for statistical computing. Organization is R Foundation for Statistical Computing, address Vienna, Austria. http://www.R-project.org (7/22/04).

Scott, D., Härdle W., (1992). Multivariate density estimation, John Wiley, New York.

Wegman, E. (1990). Color histogram for linear variables. http://www.galaxy.gmu.edu/stats/faculty/wegman.html (7/22/04.