Supplementary material: Influence of regional tectonics and pre-existing structures on elliptical caldera formation throughout the Kenyan Rift

Standard Deviational Ellipse function

The Standard Deviational Ellipse function calculates the best-fitting ellipse from digitised points by first determining the mean centre of all points $\{\overline{X}, \overline{Y}\}$. The coordinate system is then transposed with the mean centre becoming the origin (figure 1A).

Simultaneously for both x- and y-axes (only x-axis example is shown in figure 1), the deviation of each point is calculated and the standard deviation determined using the equations,

$$SDE_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{X})}{n}} \qquad \text{and} \qquad SDE_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \overline{Y})}{n}}, \tag{1}$$

where n is the total number of digitised points (figure 1C).

The axes of the coordinate system are then rotated by angle θ, and the standard deviation recalculated (figure 1B). To recalculate the standard deviation, equation 1 is first recalculated using polar co-ordinates i.e. $y_i = y_i cos\theta - x_i sin\theta$ (figure 2B). SDE_y from equation 1 then becomes

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^{n}(\overline{y}_i cos\theta - \overline{x}_i sin\theta)^2}{n}}, \tag{2}$$

where $\overline{x}$ and $\overline{y}$ are the deviations of each point from the mean centre, $\{\overline{X}, \overline{Y}\}$. The best-fitting ellipse is defined by the angle of rotation, θ, and lengths of the major and minor axes, both of which are derived from 2. Arithmetically, the maximum and minimum deviations are calculated by first expanding equation 2 to

$$\sigma_y = \frac{\sqrt{\sum_{i=1}^{n}\overline{y}_i^2\cos^2\theta - 2\sum_{i=1}^{n}\overline{x}_i\overline{y}_i\sin\theta\cos\theta + \sum_{i=1}^{n}\overline{x}_i^2\sin^2\theta}}{n}, \tag{3}$$

and taking the first derivative and solved for 0 (equation 4) to determine the maximum and minimum values, i.e.

$$\frac{d\sigma_y}{d\theta} = \frac{-\sum_{i=1}^{n}\overline{y}_i^2\cos\theta\sin\theta - \sum_{i=1}^{n}\overline{x}_i\overline{y}_i(\cos^2\theta - \sin^2\theta) + \sum_{i=1}^{n}\overline{x}_i^2\cos\theta\sin\theta}{n\sigma_{\overline{y}^i}} = 0. \tag{4}$$

The two solutions of equation 4 are therefore the angle of maximum deviation and the other, that of the minimum deviation and can be shown in the form of a quadratic equation,

$$\tan\theta = \frac{(\sum_{i=1}^{n}\overline{x}_i^2 - \sum_{i=1}^{n}\overline{y}_i^2) \pm (\sqrt{(\sum_{i=1}^{n}\overline{x}_i^2 - \sum_{i=1}^{n}\overline{y}_i^2)^2 + 4(\sum_{i=1}^{n}\overline{x}_i\overline{y}_i)^2})}{2\sum_{i=1}^{n}\overline{x}_i\overline{y}_i}. \tag{5}$$

The above equation is used to determine the angle of rotation, θ, of the ellipse. The standard deviations of the newly-rotated axes are then recalculated as σ_x and σ_y (figure 1C) using the equations,

$$\sigma_x = \sqrt{2\frac{\sum_{i=1}^{n}(\overline{x}_i\cos\theta - \overline{y}_i\sin\theta)^2}{n}} \quad \text{and} \quad \sigma_y = \sqrt{2\frac{\sum_{i=1}^{n}(\overline{x}_i\sin\theta - \overline{y}_i\cos\theta)^2}{n}}. \tag{6}$$

and represent the length of the major an minor axes. We calculated the caldera eccentricity, e, by dividing the length of the short axis by the long axis whereby values close to 1 are near-circular,

$$e = \frac{\sigma_{short}}{\sigma_{long}}. \tag{7}$$

Fault population statistics

The circular mean, circular variance and confidence intervals were calculated using the CircStat toolbox in Matlab (Berens et al., 2009). The 2θ mean method for lines (Davis et al., 2002) was used prior to statistical calculations which is a technique for analysis of azimuthal data that has an orientation (e.g. E-W) rather than a direction (i.e. values that range between 0-360°). This technique doubles the angle of orientation for each fault, calculates the circular mean of the doubled angles and finally divides the final value by 2 to determine the mean fault orientation.

The mean azimuth of the faults cannot be simply be calculated by averaging the azimuthal data. To illustrate this, figure 2A shows a scenario with two vectors with azimuths of 345° and 15°. The simple arithmetic mean would be calculated as (345 + 15)/2 = 180°, clearly in a direction opposite to intuitive mean azimuth of 0°. Instead we calculate the circular mean by first converting orientations into radians and transformed into unit vectors (figure 2B) in the two-dimensional plane by

$$r_i = \frac{cos\alpha_i}{sin\alpha_i}, \tag{8}$$

these unit vectors r_i are then averaged using,

$$\overline{r} = \frac{1}{n}\sum_i r_i. \tag{9}$$

Graphically the circular mean is calculated by vector addition of all observations (figure 2C). The result in radians is then transformed into the mean angular azimuth $\overline{\alpha}$ using the four quadrant inverse tangent function (Palm et al., 2005).

We calculate the circular variance, a common measure of dispersion in angular data, using the equation $S = 1 - R$, where R is the mean resultant vector length determined by $R = \| \overline{r} \|$. The variance is bounded by the interval [0,1] where a values approaching 1 indicates tighter clustering of the data (Fisher et al., 1995).

The 95% confidence interval for mean orientation is calculated by determining the upper and lower bounds individually. The lower bound is calculated by $L_1 = \overline{\alpha} - d$ and the upper limit by $L_2 = \overline{\alpha} + d$ where d is determined by:

for $R \leq 0.9$ and $R > \chi^2_{\delta,1}/2N$,

$$d = arccos\sqrt{\frac{2N(2R_n^2 - N\chi^2_{\delta,1})}{\frac{4N-\chi^2_{\delta,1}}{R_n}}}, \tag{10}$$

and where $R_n = R.N.$ and for $R > 0.9$:

$$d = \frac{\sqrt{N - (N^2 - R_n^2 exp(\chi^2_{\delta,1}/N)}}{R_n}. \tag{11}$$

Plate motion azimuthal uncertainty

The modern plate-kinematic model of Stamps et al. (2008) is used to calculate the relative plate motion vector at each caldera. Here we propagate the uncertainties associated with the published Victoria-Somalia Euler pole in order to assess the azimuthal velocity uncertainty. The Euler pole is defined in terms of longitude (λ), latitude (ϕ) and rotation rate ($\dot{\omega}$), and the associated uncertainties are defined by an error ellipse with a semi-major axis (a), semi-minor axis (b) and azimuth (α), clockwise from north to the semi-major axis. The rotation rate uncertainty is given as $\sigma_{\dot{\omega}}$.

In order to calculate the azimuthal velocity uncertainty, first we express the Euler pole error ellipse as a rotation rate covariance matrix in geodetic coordinates. The principal axes of the Euler poles error ellipse may be written in terms of a diagonalized partial covariance matrix,

$$\mathbf{\Sigma}'_{\dot{\boldsymbol{\omega}}'} = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix},$$

which is rotated from a regular geodetic coordinate system (λ,ϕ,r) using the equation

$$\begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix} = \mathbf{A} \begin{pmatrix} \sigma_\lambda^2 & \sigma_{\lambda\phi} \\ \sigma_{\phi\lambda} & \sigma_\phi^2 \end{pmatrix} \mathbf{A^T}$$

where A is a rotation matrix given by

$$\mathbf{A} = \begin{pmatrix} \sin\alpha & \cos\alpha \\ -\cos\alpha & \sin\alpha \end{pmatrix}$$

Given that $\mathbf{A^{-1}} = \mathbf{A^T}$ and covariances between the Euler pole location and rotation rate are unknown (so assumed to be zero), the three-dimensional rotation rate covariance matrix in geodetic coordinates is expanded to

$$\mathbf{\Sigma}_{\dot{\boldsymbol{\omega}}'} = \begin{pmatrix} \sigma_\lambda^2 & \sigma_{\lambda\phi} & 0 \\ \sigma_{\phi\lambda} & \sigma_\phi^2 & 0 \\ 0 & 0 & \sigma_{\dot{\omega}}^2 \end{pmatrix} = \begin{pmatrix} \mathbf{A^T \Sigma'_{\dot{\boldsymbol{\omega}}'} A} & \mathbf{0} \\ \mathbf{0} & \sigma_{\dot{\omega}}^2 \end{pmatrix}$$

Next, we propagate this covariance matrix through the plate rotation equation, $\mathbf{v} = \dot{\boldsymbol{\omega}} \times \mathbf{p}$, where $\mathbf{v}$ is the velocity at a point defined by position vector $\mathbf{p}$ and $\dot{\boldsymbol{\omega}}$ is the rotation rate vector. This step is calculated after rotating the covariance matrix in geodetic coordinates ($\mathbf{\Sigma}_{\dot{\boldsymbol{\omega}}'}$) to Earth-centered, Earth-fixed (ECEF) Cartesian (x,y,z) coordinates ($\mathbf{\Sigma}_{\dot{\boldsymbol{\omega}}}$) by

$$\mathbf{\Sigma}_{\dot{\boldsymbol{\omega}}} = \mathbf{J \Sigma_{\dot{\boldsymbol{\omega}}'} J^T}$$

where $\mathbf{J}$ is a Jacobian matrix given by all first-order partial derivatives of the Cartesian components of the rotation rate vector with respect to the geodetic components. Expanding the vector product of the plate rotation equation and rewriting in matrix-multiplication form, $\mathbf{v} = \mathbf{P}\dot{\boldsymbol{\omega}}$, provides a means by which to use the law of covariance propagation, i.e.

$$\mathbf{\Sigma_v} = \mathbf{R \Sigma_{\dot{\boldsymbol{\omega}}} R^T}$$

This relative plate motion vector ($\mathbf{v}$) and associated covariance matrix ($\mathbf{\Sigma_v}$) at a given caldera location may then expressed in local Cartesian (e,n,u) coordinate system by applying the rotation matrix

$$\mathbf{R} = \begin{pmatrix} -\sin\lambda_\mathbf{p} & \cos\lambda_\mathbf{p} & 0 \\ -\sin\phi_\mathbf{p}\cos\lambda_\mathbf{p} & -\sin\phi_\mathbf{p}\sin\lambda_\mathbf{p} & \cos\phi_\mathbf{p} \\ \cos\phi_\mathbf{p}\cos\lambda_\mathbf{p} & \cos\phi_\mathbf{p}\sin\lambda_\mathbf{p} & \sin\phi_\mathbf{p} \end{pmatrix}$$

where $\lambda_\mathbf{p}$ and $\phi_\mathbf{p}$ are the longitude and latitude, respectively, of the position vector for the point at which the relative velocity and uncertainties are being calculated. That is, $\mathbf{v'} = \mathbf{Rv}$ and $\mathbf{\Sigma_{v'}} = \mathbf{R\Sigma_v R^T}$ where $\mathbf{v'}$ and $\mathbf{\Sigma_{v'}}$ are expressed in the local Cartesian coordinate system.

Finally, we calculate the point error in the direction perpendicular to the plate motion velocity vector to obtain the azimuthal velocity uncertainty. We express the velocity vector and associated covariance matrix in a local cylindrical coordinate system (r,θ,u) by applying another Jacobian matrix ($\mathbf{J'}$) relating the components of the two coordinate systems. The equation for the velocity vector in local cylindrical coordinates is then

$$\mathbf{v''} = \mathbf{J'v'J'^T} = \mathbf{J'}\left(\mathbf{Rv}\right)\mathbf{J'^T}$$

and the associated covariance matrix is

$$\mathbf{\Sigma_{v''}} = \mathbf{J'\Sigma_{v'}J'^T} = \mathbf{J'}\left(\mathbf{R\Sigma_v R^T}\right)\mathbf{J'^T}$$

The azimuthal velocity uncertainty (σ_{v_θ}) is the square root of the variance in the θ component.

References

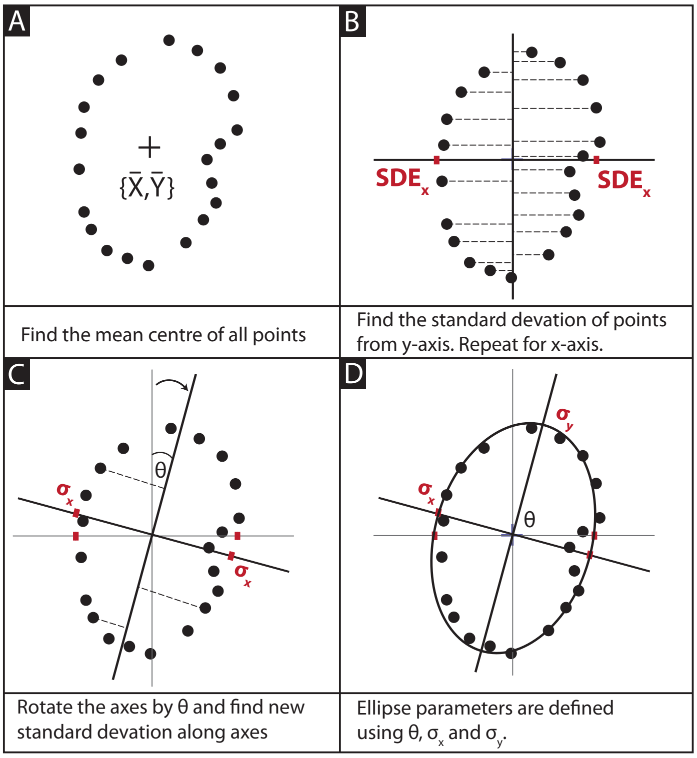

Figure 1: Methodology of ESRI Standard Deviational Ellipse tool for ArcGIS. Firstly, calderas are digitised as a polyline shape layer and converted into point-data at polyline vertices. **A** The mean centre, $\{\overline{X}, \overline{Y}\}$, of all points are calculated and the origin of a coordinate system is transposed to the mean centre. **B** The standard deviation of points from the y-axis is determined (SDE$_x$). **C** The axes are rotated by θ (see text for calculation) and the standard deviation of all points from the newly rotated y-axis is recalculated and plotted (σ_y and σ_x). **D** The best-fitting ellipse is defined using σ_y, σ_x as lengths of the major and minor axes and the rotation angle θ.

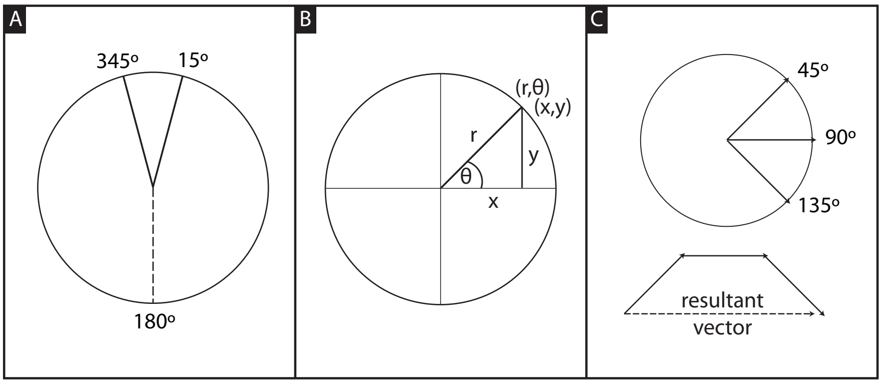

Figure 2: Visualisation of circular statistics used to analyse fault orientation. **A** shows two vectors at 345° and 15° and the associated (incorrect) arithmetic mean at 180° (dashed line) where the correct mean lies at 0°. **B** describes the relationship between Cartesian and Polar coordinates that in used in coordinate conversion. Calculation of the circular mean, **C**, is approached by treating each observation as a unit vector and using vector addition to calculate the resultant vector. The mean azimuth is calculated by the direction of the resultant vector.