Classification of planetary craters using outline-based morphometrics

Thomas J. Slezak  
*Brigham Young University - Provo*

Jani Radebaugh  
*Brigham Young University - Provo*

Eric H. Christiansen  
*Brigham Young University - Provo*

Mark C. Belk  
*Brigham Young University - Provo, mark_belk@byu.edu*

Follow this and additional works at: https://scholarsarchive.byu.edu/facpub

Part of the Biology Commons

**Original Publication Citation**

**BYU ScholarsArchive Citation**
Slezak, Thomas J.; Radebaugh, Jani; Christiansen, Eric H.; and Belk, Mark C., "Classification of planetary craters using outline-based morphometrics" (2020). *Faculty Publications*. 5389.  
https://scholarsarchive.byu.edu/facpub/5389

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Classification of planetary craters using outline-based morphometrics

Thomas J. Slezak a,⁎, Jani Radebaugh a, Eric H. Christiansen a, Mark C. Belk b

a College of Physical and Mathematical Sciences, Department of Geological Sciences, Brigham Young University, Provo, UT 84602, USA
b College of Life Sciences, Department of Biology, Brigham Young University, Provo, UT 84602, USA

A R T I C L E  I N F O

Article history:
Received 24 June 2020
Received in revised form 12 October 2020
Accepted 27 October 2020
Available online 5 November 2020

A B S T R A C T

The morphologies of craters on planetary surfaces reveal clues about the geologic mechanisms by which they originate and subsequently evolve, as well as the materials and physical variables inherent to the environment in which they formed. We carried out a quantitative multivariate analysis of shape descriptors derived from the outlines of craters formed by volcanic processes on Mars, Io, and Earth and by impact cratering on the Moon using elliptic Fourier analysis (EFA) and the Zahn-Roskies (Z-R) shape function. Canonical variate analysis (CVA) was used to construct a statistical model of differences between the crater groups to classify craters produced by various volcanic and impact processes.

The classification model from canonical variate analysis of EFA shape descriptors yielded a 90% rate of success for the assignment of group membership among 406 examined craters. It correctly classified 138 of 154 (90%) ionian paterae, 154 of 155 (99%) lunar impact craters, 31 of 35 (89%) terrestrial basaltic shield calderas, 32 of 38 (84%) terrestrial ash-flow calderas, and 12 of 24 (50%) martian basaltic shield calderas. The classification model from canonical variate analysis of Z-R shape function descriptors classified 84% of the total population of the examined craters correctly. The analysis correctly classified 96% of ionian paterae, 100% lunar impact craters, 51% terrestrial basaltic shield calderas, and 63% martian calderas, but only 16% of the terrestrial ash-flow calderas were correctly classified.

Canonical variate analysis of EFA and Z-R results shows that the shapes of ash-flow calderas and paterae on Io differ the least of all groups included in this study, and basaltic shield calderas and martian calderas analyzed together also have few differences. The Z-R model successfully classifies more ionian patera and impact craters than the EFA classification model but performs poorly at classifying the other crater groups. This result shows that the descriptors convey different shape information. The Z-R model is robust in its ability to classify end-member differences in complexity while the EFA model is robust in its ability to reliably classify among more groups.

These differences and similarities in shape confirm previously understood commonalities related to the origin of various types of craters. In general, basalt shield calderas on Earth and Mars are morphologically similar and are thought to have similar origins; this study confirms that the 2-D shapes of their craters are quantitatively correlated. Similarities have been noted between terrestrial ash-flow calderas and paterae on Io, principally in their large sizes, shallow magma chambers and complex evolution; this study confirms their shapes are also similar. Impact craters and ionian paterae are most dissimilar, as are their evolutions. This study demonstrates rigorous landform shape analysis can greatly increase our understanding of the diversity in craters and the processes involved in their formation.

1. Introduction

Craterforms, subcircular forms created by impact, volcanic collapse or explosion, tectonism, karst and other processes, are the most common geologic feature on planetary surfaces, and identifying the processes by which they formed is fundamental to interpreting the geologic history (Zimbelman, 2001; Greeley, 2013). Studies of planetary landform shapes have traditionally employed measurement-derived shapes, sizes, and dimensionless ratios (e.g., Michalski and Bleacher, 2013; Watters et al., 2017; Radebaugh et al., 2001). However, the substantial conceptual and mathematical developments in the quantitative analysis of natural, irregular forms by the biological sciences (e.g., Zahn and Roskies, 1972; Kuhl and Giardina, 1982; Lohmann, 1983; Rohlf and Marcus, 1993) remain largely unutilized in contemporary planetary landform studies.
Geomorphic investigations aim to understand how variations in origin and evolution produce differences in morphology, similar to biological studies (Lestrel, 2000; Neal and Russ, 2012). Morphology provides a link to investigate the geologic processes involved in the origin and evolution of geologic forms ranging in scale from sub-millimeter sedimentary particles (Powers, 1953) to planetary landforms (Zimbelman, 2001) hundreds of kilometers in diameter. Existing methods used to classify these craterforms using shape rely on criteria such as qualitative descriptions (e.g., “angular”, “elongate”, “circular”), indices derived from measurements (e.g., width to length ratio), and the fitting of mathematical models (e.g., power laws, quadratic equations) across 1 or more morphologic variables (Evans, 1986). Existing landform classifiers are not robust to complex differences in morphology and assume that the morphology of landforms produced by natural processes is not inherently multivariate (Evans, 1972).

This study adopts outline-based shape analysis from systematic evolutionary biology and paleontology, where it is used to identify biological forms based on morphology alone, and applies this methodology to planetary landform classification. Samples from planetary bodies and inferred formation mechanisms were selected and analyzed to assess the value of this methodology.

2. Background

Landform classification is a fundamental task of planetary geology (Shoemaker, 1963; Greeley, 2011). Heavily impact cratered terrains suggest older surface ages and an inactive interior while the presence of volcanic craters and few impact craters suggests active resurfacing and thus an active interior (e.g., Basilevsky and McGill, 2007). As different geologic processes can produce craterforms similar in shape, improving our understanding of how morphology varies by process and environment will enhance geologic interpretations of planetary surfaces.

The morphological variation across some craterforms is relatively simple (e.g., lunar impact craters) and links to the processes from which morphological differences originate are well-established (Wilhelms et al., 1987; Melosh, 1989). In contrast, the morphological variation of paterae on Io (defined by the International Astronomical Union as “complex, or irregular craters with scalloped edges”), is not as easily identified and fundamental aspects of their formation remain unresolved (Radebaugh et al., 2001; Keszthelyi et al., 2004; Radebaugh, 2005; Slezak et al., 2015; Dundas, 2017). This study compares the morphologic information of paterae on Io with the morphology of craterforms with better-constrained formation processes including martian calderas, lunar impact craters, terrestrial ash-flow calderas, and terrestrial basaltic shield volcanoes.

Traditional morphometric methods can be useful to examine size and shape properties collected from physical measurements of natural forms. Common morphometric measures use length, width, perimeter, and area to derive simple shape descriptors including circularity, aspect ratio, best-fit ellipse, effective diameter, and others (Neal and Russ, 2012; Liu et al., 2015). While these quantities generalize shape information to communicate some visually apparent differences, they are not as effective for complex forms (Evans, 1972; MacLeod, 1999; Saynue et al., 2014). Furthermore, analyses of shape using these measures can yield results that fail to identify visually intuitive differences among the shapes in the analysis.

2.1. Geometric morphometrics

This study uses outline-based methods of shape analysis that is a branch of geometric morphometrics known as boundary morphometrics (Lestrel, 1997) to investigate patterns in the geomorphology of a variety of craterforms. Geometric morphometrics (Rohlf and Marcus, 1993) is a set of shape analysis methods that express shape quantitatively and preserve all geometric information, allowing shapes to be compared invariant to scale, translation, and rotation (Slice, 1996). Geometric morphometrics has been widely employed in biological research (MacLeod, 2017), and its methods can address a wide number of scientific problems.

2.2. Previous shape outline studies in the geological sciences

Pioneering studies in the analysis of quartz grains by Ehrlich and Weinberg (1970) introduced outline-based analysis approaches to the geological sciences community through Fourier analysis as an “exact” method to characterize grain shapes. The Fourier series expansion can be used to quantitatively examine shapes if the outline is treated as closed curve periodic functions by providing an approximation to the curve using any infinite number of harmonics to derive frequencies and amplitudes. The application of traditional, radial Fourier analysis to examine the variability among simple closed curves derived from the outline of geologic forms has been applied to lunar impact craters by Eppler et al. (1977a, 1977b, 1978, 1983) to examine the variability in, and to identify influential factors affecting, morphology of lunar crater populations. Additional applications of Fourier analysis in planetary studies include Kordesh et al. (1982), Kordesh and Basu (1982), Kordesh et al. (1983), Kordesh (1983a); Kordesh (1983b), in which Fourier analysis was used to analyze the differences between lunar soil particle shapes and clasts in meteoritic breccias. Ultimately, the conclusions of these outline-based studies were limited by available computational power. More recently Watters et al. (2017) used radial Fourier analysis to examine the size dependent scaling of impact craters on the Moon and Mars.

Some studies have explored quantitative frameworks to standardize lunar crater classifications using methods such as the application of Chebyshev polynomials (e.g., Mahanti et al., 2014). These methods rely upon high-resolution topographic data which is not common to all observed planetary surfaces, such as Io. The use of polynomials (e.g., Craddock and Howard, 2000) and power laws (e.g., Baldwin, 1963; Pike, 1977) have also been used to describe specific aspects of morphologic relationships for some crater populations.

Using these quantitative studies of craters as a background, we explore the quantification and comparison of crater morphology using planform shape and modern methods of multivariate statistical analysis.

3. Data

We examined the shapes of 406 craterforms consisting of paterae on Io (154), lunar impact craters (155), martian calderas (24), terrestrial basaltic shield calderas (35), and terrestrial ash-flow calderas (38). Before tracing each craterform boundary, a stereographic (conformal) map projection was applied using ArcGIS or JMARS to the central coordinates of each crater to minimize shape distortion in each image. Fig. 2 illustrates an example of problematic distortions in shape and scale that can occur as a result of different map projections. Craters of ambiguous origin (other than paterae on Io) were not included in this study and the highest resolution imagery available was used. Images were exported as 2000 by 2000-pixel jpg files including a scale bar (denoted in km) from the projected map.

3.1. Data sources

3.1.1. Paterae on Io

A sample consisting of 154 paterae on Io with resolved boundaries was examined in this study using observations made by the Galileo Solid State Imaging camera (SSI) (Belton et al., 1992) and the Voyager 1 Narrow Angle Camera (NAC) (Smith et al., 1977). These were analyzed using the Io Galileo SSI / Voyager Color Merged Global Mosaic fully registered base map provided by the U.S. Geological Survey Astrogeology Division (Belton et al., 1992; Geissler et al., 1999; Becker and Geissler,
2005; Barth et al., 2009; Veeder et al., 2009; Williams et al., 2011). The mosaic was loaded into ESRI's ArcMap 10.4 software as a USGS .cub file using the Io 2000 IAU geographic datum. The spatial resolution of the imagery used ranges from ~7 m/px to ~2 km/px, and our sample is only a subset of all paterae able to be resolved in the imagery (>400; Radebaugh et al., 2001; Williams et al., 2011), although many of those have extremely poor resolution and were not used in this study.

3.1.2. Terrestrial volcanoes

Locations of 35 terrestrial basaltic shield volcanoes and 38 terrestrial ash-flow calderas were taken from global databases provided by the ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) Volcano Archive (AVA) and Smithsonian Global Volcanism Program (Venzke et al., 2002). We used the criteria of Pike and Clow (1981) as modified by Radebaugh and Christiansen (1999) to classify the

Fig. 1. Examples of craterforms examined in this study including ionian paterae (A & D), lunar impact craters (B), martian calderas (C), terrestrial ash-flow calderas, (E) and terrestrial basaltic shield volcanoes (F). While the crater morphologies are similar, differences in shape can be quantitatively derived. Image credits: (A) the Voyager mission, (B) the Lunar Reconnaissance Orbiter mission, (C) the Mars Reconnaissance Orbiter mission, (D) the Galileo mission, (E & F) Google Earth.

Fig. 2. A lunar impact crater in two different map projections illustrates potential distortions in shape and scale. (A) Rheita on a colored shaded relief map (100 m/px LROC DEM) using an equidistant cylindrical projection and (B) Rheita using a stereographic map projection of LROC WAC (100 m/px) imagery. Image credit: lroc.sese.asu.edu.
terrestrial volcanoes we examined. Supplementary records and geologic mapping of ash-flow shields are provided by Newhall and Dzurisin (1988), Lipman (1997), Radebaugh and Christiansen (1999), and Hughes and Mahood (2008). The images were studied in ArcGIS using ESRI's Online World Imagery base map layer using a stereographic projection. This base map layer contains imagery ranging in resolution from 15 m/px to >1 m/px depending on location. For cases where ESRI imagery was unable to resolve the craterform bounds due to data anomalies or obscuration from clouds or snow, DigitalGlobe 2016® imagery ranging from ~1–15 m/px was used. In some locations, 30 m/px Landsat Enhanced Thematic Mapper (ETM) and 15 m/px ASTER Visible and Near Infrared (VNIR) images from the USGS and NASA’s Land Processes Distributed Active Archive Center (NASA LP DAAC) were used.

3.1.3 Martian calderas

Outlines of 24 martian calderas documented by Hodges and Moore (1994), Williams et al. (2009), Robbins et al. (2011), and Tanaka et al. (2014) were digitized from the Thermal Emission Imaging System Infrared (THEMIS IR) Day 100 m v12 Global Mosaic (Edwards et al., 2011) and Context Camera (CTX) imagery using the JMARS software (Christensen et al., 2009). Martian calderas were located and the map was re-projected using the JMARS Reproject tool to the central coordinates of each caldera to preserve shape. CTX imagery (~7 m/px) and was used for some calderas where shape could not be traced using the THEMIS IR Day v12 mosaic (100 m/px).

3.1.4 Lunar impact craters

Lunar impact craters were selected from the Lunar and Planetary Institute’s Lunar Impact Crater Database (Losiak et al., 2015). The simple to complex transition causes shape variation as the crater diameter increases (Wilhelms et al., 1987). Simple craters are nearly circular and lack central peaks, while complex craters display more complicated outlines and have central peaks or pits. “Transitional” impact craters lack central peaks but have more complex rim outlines than simple impact craters. The simple to complex transition occurs at ~21 km for craters in the lunar highlands and at ~16 km diameter for craters in the mare.

Fig. 3. Outlines of the rims of ionian Camaxtli patera and lunar impact crater Linne shown following digitization and normalization (scaling is removed).

Fig. 4. Transformation from raw, manually placed Cartesian points (red dots in A) along the rim of the patera to points with equidistant spacing (B) from linear interpolation is illustrated here. Image of Camaxtli patera on Io courtesy of the Galileo mission.
(Wilhelms et al., 1987). A simple random sample of lunar craters was selected from the Lunar and Planetary Institute 2015 impact crater database (Losiak et al., 2015) binned by diameter ranges of 0–12 km, 12–35 km, and 35–220 km to distinguish crater morphology associated with these ranges in size (Wilhelms et al., 1987). While the lunar simple–complex transition is correlated to scale, it does not define the phenomena, and this method of selection allows the global variance of impact crater morphology on the Moon to be sampled. Images of lunar impact craters were taken from the LRO LOLA and Kaguya Terrain Camera DEM merge base map (Barker et al., 2016) and the LRO LROC-WAC 100 m June 2013 Global Mosaic (Wagner et al., 2015).

4. Methodology

The tps series software (Rohlf, 2015) was used to digitize outlines from the reprojected images. A tps file consists of a series of Cartesian coordinates for a shape, the image traced, and other metadata such as ID, image name, and scale factor. Outlines were obtained for each craterform by tracing numerous points along the most continuous and well-defined outer boundary of the observed morphology using tpsDig2. The initial point of each raw outline was placed (arbitrarily) in the northwest part of the crater’s rim with additional points placed in a clockwise fashion.

Fig. 5. (Upper) Plot of Zahn-Roskies shape function values for Camaxtli patera and Linné crater (Fig. 1B) computed from the digitized outlines in Fig. 4. A perfect circle sampled at a sufficiently high resolution would plot with a constant value slope. (Lower) Plot of normalized Z-R shape function values for Camaxtli patera and Linné crater. A perfect circle sampled at a sufficiently high resolution would plot at 0.

Fig. 6. Zahn-Roskies analysis of Camaxtli patera using the \( \phi \)-based approach, where \((\phi_i, \phi_{i+1}, \phi_n, \ldots)\) describes shape as the angular deviation from an ideal circular form.
The number of “raw” points needed to outline each craterform varied with the complexity of its shape and the final point of the outline was placed just before the first point. The set of points comprising each outline was then resampled to 99 points with equidistant spacing using linear interpolation (Fig. 4b). This approximate number of points is conventionally fixed and is consistent with other studies (Lohmann, 1983; Rohlf and Archie, 1984; Ferson et al., 1985). The points were interpolated to positions of equidistant spacing along the outline to compare the collected outlines at “positions of maximum correlation” (Lohmann, 1983). Since craterforms inherently lack traditional, homologous “landmarks” on biological forms, “pseudo-homologous” points of reference (Sneath and Sokal, 1973) were derived for each outline by computing the distance between each point and the shape centroid, and then reordering each collection of coordinates so that the point of maximum distance to the centroid appears first in the series. Examples of outlines following digitization and processing are shown in Fig. 3.

5. Analysis

5.1. Shape descriptors

Once these procedures were applied to each outline, outline-based shape analysis methods, including the Zahn and Roskies (Z-R) shape function (known as the tangent angle approach) and elliptic Fourier analysis (EFA) were applied to produce a set of shape descriptors from the outlines of the craterforms. Statistical classification techniques including principal component analysis (PCA) and canonical variate analysis (quadratic discriminant analysis) were applied to examine group differences in shape and subsequently produce a supervised shape classification model. The resultant model was then used to test the hypothesis that craterforms differing in origin can be classified using outline information alone. These analyses are described in detail below.
5.1.1. The Zahn and Roskies (Z-R) shape function

The Zahn and Roskies (1972), or “Z-R” shape function evaluates the shape of an object by the curvature of its outline, $\phi$. It provides an intrinsic representation of shape as it deviates from that of a circle (Bookstein et al., 1982), making it ideal to quantify impact and volcanic crater shapes.

The Z-R shape function was computed for the outlines using the Wolfram Mathematica software and MacLeod’s (2011) “Z-R Shape Function 1.4” notebook. Fig. 5 shows the Zahn and Roskies (1972) shape function ($\phi$) and normalized shape function ($\phi^*$) (MacLeod, 2011) values calculated for Camaxtli patera and Linné crater. The quantity $\phi$ describes the net angular change around the perimeter of shape as cumulative radians. The quantity $\phi^*$ describes the net angular change around the perimeter of shape with the ramp of circularity removed. The ramp of circularity is removed by subtracting a cumulative-constant term corresponding to the radian-equivalent ($2\pi$) at each step divided by total number of points used to represent the outline.

Fig. 9. The first two canonical variates of the Canonical Variate Analysis Elliptic Fourier Analysis (10 harmonics) classifier of planetary craters. Canonical axis 1 describes 77% of the differences in shape and canonical axis 2 represents 11% of shape differences. This plot illustrates 88% of the overall differences in shape as described by the Canonical Variate Analysis – Elliptic Fourier Analysis 10 classifier.

Fig. 10. Canonical variate analysis of Z-R shape function values for the examined planetary craterforms.
As Linné crater is a nearly circular simple crater, it plots as a nearly straight line while the more irregular Camaxtli patera deviates more significantly from circularity. These values are invariant to scale and record shape information so they can be subjected to multivariate statistical analysis. An analysis of boundary coordinates computed by the Z-R shape function results in a set of angles, \( \phi \), expressed in radians, whose cumulative sum represents the angular change around the perimeter of the shape. The values are output from the function for each craterform and are collected into an array for further analysis. Fig. 6 illustrates these quantities and their derivation along the bounds of Camaxtli patera.

5.1.2. Elliptic Fourier analysis (EFA)

Elliptic Fourier analysis (Kuhl and Giardina, 1982) is a widely-used technique of geometric morphometrics in outline analysis studies that provides a more precise approximation of complex shapes in comparison to traditional Fourier analysis and is well-suited to boundary morphometrics (Rohlf and Archie, 1984; Lestrel, 2000). Elliptic Fourier analysis (Kuhl and Giardina, 1982) uses Fourier decomposition to separate the x (horizontal) and y (vertical) inter-point components of the outline as independent parametric functions of arc length and distance of each point from the starting point along the outline (Marcus et al., 1996; Lestrel, 1997; Kuhl and Giardina, 1982; Ferson et al., 1985; MacLeod, 2012). A recent study used EFA to describe the shapes of lakes on the north polar region of Titan and sort them according to possible formation mechanisms or stages (Dhingra et al., 2019).

The PAST3 (PAleontological STatistics) software package (Hammer et al., 2001) was used to compute the normalized elliptic Fourier analysis coefficients with the implementation of the Ferson et al. (1985) normalization procedure. The procedure produces 4 values per computed harmonic that correspond to \( \cos x \), \( \sin x \), \( \cos y \), and \( \sin y \) coefficients that describe independent interpoint orientation in the horizontal and vertical directions. We used the normalization procedure of Ferson et al. (1985) to ensure the shape was invariant to non-shape variables (scaling, translation, and rotation). As a result of this normalization procedure, three coefficients of the first harmonic, \( \cos x_1 \), \( \sin x_1 \), and \( \cos y_1 \)
(a', b', and c') were degenerated to values of 1, 0, and 0 respectively and thus were not valuable to include in statistical analysis. The fourth coefficient of the 1st harmonic, $\sin y_1$ (d''), however, provided the ellipticity (also known as axial ratio) of the best fit ellipse for each shape, and thus contributed meaningful shape information. Successive harmonics yielded additional coefficients that approximated a shape function to the outline of a shape with increasing precision. As an array of these values was then subjected to statistical analysis, we limited the number of coefficients to the fewest necessary to adequately represent the major shape differences in the craterforms.

The number of harmonics necessary to optimally describe the complexity across all of the craterforms was evaluated by comparing the effect of successive harmonics with the raw data (Fig. 7; Crampton, 1995).

The amount of shape information contributed by successive harmonics was also analyzed using a Fourier power plot (Fig. 8). Similar to other Fourier analyses, the first few consecutive harmonics (following the 1st) provided most of the information needed to approximate the shape. From this a posteriori examination, we selected the first 10 harmonics to represent the complexity and variability of the shapes included in this study. A model using only the first 5 harmonics was also initially selected, but was found to not be as successful (Slezak, 2017). The coefficients for these harmonics were subjected to multivariate statistical analysis in order to compare the ability of the two quantitative descriptors (Z-R shape function and EFA) to differentiate and successfully classify craterforms by shape alone. These coefficients were subsequently subjected to discriminant analysis and classifiers were built to determine which model classified the craters into their groups identified a priori in this study.

5.2. Statistical analysis and classification

Multivariate statistical analyses allow multiple variables to be examined simultaneously for an individual among a larger population. Dimensionality reduction of the quantities representing each shape is critical to enable effective comparisons. We used canonical variate analysis (CVA) and principal components analysis (PCA) to reduce the dimensionality of shape descriptors produced by the Z-R shape function and elliptic Fourier analysis from the digitized outlines.

Canonical variate analysis (multigroup discriminant analysis) (Fisher, 1936) maximizes the differences between multiple groups ($k > 2$) in a population (Lestrel, 2000) and predicts group membership using a classification function (Rencher and Christensen, 2012). The procedure projects the data onto a lower dimensional space where within-group differences are statistically significant. The resulting canonical axes represent maximum between-group differences in the data (Friedman, 1989) and we do not assume the shape quantities from each class to be normally distributed. The classifiers can then be evaluated with scoring metrics that summarize a model by comparing predicted classifications with actual results. As accuracy alone is a misleading metric for classification problems with imbalanced classes, precision (the proportion of correct positive classifications to all positive classifications), recall (the proportion of correct positive classifications to all correct classifications), and F1 score (the harmonic mean of precision and recall) are used to evaluate each classifier.

The application of principal components analysis to normalized values of the Z-R shape function, a procedure known as “eigenshape analysis” (Lohmann, 1983; Lohmann and Schweitzer, 1990), provides a visually intuitive way to interpret shape differences between groups in a population. The objective of this method is to represent maximum shape variability visually using the smallest number of components.

6. Results

6.1. Canonical variate analysis of EFA shape descriptors (10 harmonics)

The first 10 harmonics of EFA shape descriptors were also subjected to canonical variate analysis. The resulting classifier, “CVA-EFA10”, is shown to be statistically significant in Table 1.

The $p$-values in Table 1 are less than 0.05 and thus fail to reject the null hypothesis, that the means of the covariates for the canonical axes are equal across all groups. Table 2 shows that both canonical axis 1 and canonical axis 2 are statistically significant and together describe 88% of the total shape differences between groups. Fig. 9 illustrates the significant canonical axes and depicts the ability of the classifier to discriminate between groups.

The CVA-EFA10 classifier scores an overall accuracy of 90%, precision of 94%, recall of 82%, and F1 score of 88%. The confusion matrix produced by the model built from the first 10 harmonics of EFA coefficients is provided in Table 3 and are shown as percent of total count in Table 4.

This model correctly classifies 99%, or 154 of the 155 total lunar impact craters included in the study; 1 (<1%) is misclassified as an ionian patera. Of the 38 ash-flow, 32 (84%) are successfully classified with 5 (13%) misclassified as ionian paterae, and 1 (3%) ash-flow caldera classified as a lunar impact crater. Of 35 basaltic shield calderas, 31 (89%) are correctly classified with the remaining 4 basaltic shield calderas (11%) misclassified as lunar impact craters. Of 154 total paterae, this model classifies 138 (90%) successfully with 13 (8%) misclassified as lunar impact craters, 2 (~1%) classified as basaltic shield calderas, and 1 (<1%) misclassified as an ash-flow caldera. Of the 24 martian caldera shapes examined in the study, this model correctly classifies only 12 (50%) with 6 (25%) classified as ionian paterae and 6 (25%) misclassified as lunar impact craters.

6.2. Canonical variate analysis of Z-R shape descriptors

The 98 Z-R shape function ($\phi$) values produced by the shape function for each outline can also be examined using canonical variate analysis. The resulting classifier, “CVA-ZR98”, is shown to be statistically significant by Table 5.

The differences between canonical axes are more widely dispersed than those of elliptic Fourier analysis plots due to the larger number of features resulting from the Z-R shape function as angles between the interpolated boundary points.

As shown in Table 6, the first two canonical axes are statistically significant for this analysis and together describe 80% of the shape differences. The confusion matrix for this classifier is shown in Table 7 and cast as accuracy rates in Table 8. The overall model accuracy is 84%, precision is 77%, recall is 65%, and the F1 score is 71%.

Fig. 10 illustrates the significant canonical axes and the ability of the classifier to discriminate between groups. Of paterae on Io, 148 of 154...
Table 2
Summary statistics of canonical variates for 10 harmonics of EFA shape descriptors.

| Axis | Eigenval | % Total | Canonical Corr | p-val
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.262</td>
<td>77</td>
<td>0.747</td>
<td>0.310 &lt;0.0001*</td>
</tr>
<tr>
<td>2</td>
<td>0.177</td>
<td>11</td>
<td>0.387</td>
<td>0.702 0.0221*</td>
</tr>
<tr>
<td>3</td>
<td>0.115</td>
<td>7</td>
<td>0.321</td>
<td>0.826 0.2981</td>
</tr>
<tr>
<td>4</td>
<td>0.086</td>
<td>5</td>
<td>0.282</td>
<td>0.920 0.5231</td>
</tr>
</tbody>
</table>

Table 3
Confusion matrix of CVA-EFA10 classifier results.

<table>
<thead>
<tr>
<th>EFA (10 harmonics)</th>
<th>Group</th>
<th>Predicted count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ash-Flow</td>
<td>Basaltic Shield</td>
</tr>
<tr>
<td>Ash-Flow</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>Basaltic Shield</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>Io Patera</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Lunar Impact</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mars Caldera</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 4
Confusion matrix of CVA-EFA10 classifier results as % of total count. Values in bold are the true positive rates for each group.

<table>
<thead>
<tr>
<th>EFA (10 harmonics)</th>
<th>Group</th>
<th>Predicted count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ash-Flow</td>
<td>Basaltic Shield</td>
</tr>
<tr>
<td>Ash-Flow</td>
<td>84%</td>
<td>0%</td>
</tr>
<tr>
<td>Basaltic Shield</td>
<td>0%</td>
<td>89%</td>
</tr>
<tr>
<td>Io Patera</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Lunar Impact</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Mars Caldera</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 5
Significance tests for CVA of Z-R shape function values.

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>DF</th>
<th>p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.105</td>
<td>392</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>1.609</td>
<td>392</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Hotelling-Lawley</td>
<td>3.472</td>
<td>392</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Roy’s Max Root</td>
<td>2.051</td>
<td>98</td>
<td>&lt;0.0001*</td>
</tr>
</tbody>
</table>

Table 6
Summary statistics for the canonical axis scores of the CVA-ZR98 classifier.

<table>
<thead>
<tr>
<th>Axis</th>
<th>Eigenval</th>
<th>% Total</th>
<th>Canonical Corr</th>
<th>p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.047</td>
<td>59</td>
<td>0.820</td>
<td>0.105 &lt;0.0001*</td>
</tr>
<tr>
<td>2</td>
<td>0.719</td>
<td>21</td>
<td>0.647</td>
<td>0.319 &lt;0.0001*</td>
</tr>
<tr>
<td>3</td>
<td>0.400</td>
<td>12</td>
<td>0.535</td>
<td>0.549 0.1679</td>
</tr>
<tr>
<td>4</td>
<td>0.301</td>
<td>9</td>
<td>0.768</td>
<td>0.768 0.5508</td>
</tr>
</tbody>
</table>

Table 7
Confusion matrix of CVA-ZR98 classifier results.

<table>
<thead>
<tr>
<th>Z-R Results</th>
<th>Predicted Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Ash-Flow</td>
</tr>
<tr>
<td>Ash-Flow</td>
<td>6</td>
</tr>
<tr>
<td>Basaltic Shield</td>
<td>0</td>
</tr>
<tr>
<td>Io Patera</td>
<td>0</td>
</tr>
<tr>
<td>Lunar Impact</td>
<td>0</td>
</tr>
<tr>
<td>Mars Caldera</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 8
Confusion matrix of CVA-ZR98 classifier results as % of total count. Values in bold are the true positive rates for each group.

<table>
<thead>
<tr>
<th>Z-R Results</th>
<th>Predicted Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Ash-Flow</td>
</tr>
<tr>
<td>Ash-Flow</td>
<td>16%</td>
</tr>
<tr>
<td>Basaltic Shield</td>
<td>0%</td>
</tr>
<tr>
<td>Io Patera</td>
<td>0%</td>
</tr>
<tr>
<td>Lunar Impact</td>
<td>0%</td>
</tr>
<tr>
<td>Mars Caldera</td>
<td>0%</td>
</tr>
</tbody>
</table>

(96%) were correctly classified, with 6 (4%) misclassified as lunar impact craters. Of the 38 ash-flow calderas, only 6 (16%) are classified correctly with 23 (61%) misclassified as martian calderas, 5 (13%) misclassified as ionian paterae, and 4 (11%) misclassified as basaltic shield calderas. 18 of 35 (51%) basaltic shield calderas were classified correctly with 12 (34%) misclassified as martian calderas and 5 (14%) misclassified as ionian paterae. Of 24 total martian calderas, 15 (62%) were correctly classified, with 4 (17%) misclassified as basaltic shield calderas, 3 (13%) misclassified as ash-flow calderas, and 5 (14%) misclassified as basaltic shield calderas, 3 (13%) misclassified as ionian paterae, and 2 (8%) misclassified as lunar impact craters.

Principal component, or “eigenshape”, analysis was performed on the Z-R shape descriptors as shown in Fig. 11. Eigenshape analysis uses the normalized (\(\phi\)) Z-R shape function values. The first principal axis (PC-1), or first “eigenshape”, represents 79.9% of shape variance among the groups of craterforms and the second principal axis (PC-2) is the second dimension of greatest variation representing 10.5% of shape variance. This indicates ~90% of the total shape variance among the craters in this study can be represented in two dimensions. This demonstrates that principal component analysis is a valuable tool for visualizing major shape variation trends among craterform shapes.

Fig. 11 shows that shapes of ionian paterae vary the most among all of the groups and, predictably, lunar impact craters vary the least in shape. Similarly, large ellipses for ash-flow calderas and ionian paterae show these groups have the greatest diversity in shape. The 95% ellipses for basaltic shield calderas and martian calderas have similar placements in the plot, suggesting these craterform shapes vary similarly.

7. Discussion

This study shows that craterforms can be classified using rim outlines alone. Furthermore, the study has some implications for the formation processes of paterae on Io. The analyses demonstrate that the shapes of paterae on Io are most similar to terrestrial ash-flow calderas, the shapes of basaltic shield calderas are most similar to martian calderas, and the shapes of lunar impact craters differ significantly from all other compared groups.

Scoring metrics for the classifiers produced in this study are shown in Table 9 and classifier performance (accuracy) by craterform group is displayed in Table 10. The CVA-EFA10 classifier, produces the higher overall accuracy rate (90.4%) and provides a robust model able to differentiate craters from simple to moderately complex.

The EFA shape descriptors are most powerful in their ability to achieve high accuracy classification among craterforms of moderate complexity, especially ash-flow calderas and basaltic shield calderas (Tables 9–10). Classifiers trained using EFA shape descriptors, when
of volatiles in Io’s crust, that results in more irregular crater shapes like terrestrial ash-flow calderas, or the presence of volatiles in the crust enables failure of the crust in a piecemeal fashion, resulting in more irregular margins (Keszthelyi et al., 2004). Similarly, it is possible that erosion modifies the shapes of terrestrial ash-flow calderas over time, increasing shape variability. Therefore, the volcanic craters may form by different geologic processes, but then they are modified over time by processes that make their shapes similar.

Only minor differences in the shapes of basaltic shield calderas and martian calderas were found in this study. Given the similarities in magma compositions, lava flow morphologies and probable viscosities, and shield shapes, this is consistent, and supports other comparative studies of the two classes of features. Most martian calderas are widely considered to be analogous to calderas on terrestrial basaltic shield volcanoes, with the main difference being size; both result from basaltic magmas and are thought to share similarities resulting from collapse style and post-formation processes (Mougins-Mark and Robinson, 1992; Crumpler et al., 1996; Mouginis-Mark et al., 2007; Howard, 2010) as well as traditional morphometric similarities (Plescia, 2004) from rim height and crater width.

It is common for the bounding morphology of these craters to be influenced by volcanic nesting, or post-formation intra-caldera collapse events, and these events are also notable in ionian paterae. Qualitative similarities between basalt shield calderas and paterae on Io exist, such as both having steep walls and flat floors (Radebaugh et al., 2001) that can sustain lava lakes (Lopes et al., 2004) as indicated by dark floor materials inferred by Geissler et al. (1999) and others to be mafic silicate lavas (Keszthelyi et al., 2007). However, results from the statistical models in this study indicate that the shapes of paterae on Io in general lack similarity to those of the basaltic shield calderas or martian calderas examined here.

The shape differences between the two most dissimilar groups, ionian paterae and lunar impact craters, are likely due to differences in the nature and complexity of the geologic formation processes. Ionian patera formation probably requires several different stages of collapse through evacuation of a shallow magma chamber through successive volcanic eruptions paired with volatile sublimation and removal of wall support (Keszthelyi et al., 2004) and may evolve over millions of years. In contrast, lunar and other impact craters form over short time periods by simple (although violent) adjustment to the surface through the localized impact event.

Three ionian paterae: Atar, Bochica, and an unnamed ionian patera at 11.23°N, 84.74°W (shown in Fig. 12) are classified as lunar impact craters by all statistical models. These patera could be the result of entirely different processes (e.g., impact cratering) or a variation of the “normal” mechanism of patera formation. It is also possible that the near circular shapes are the result of image artifacts and that the low resolution of the images (~1 km/pixel at best) reduces the perceived landform to a shape with less irregularity than it has in reality. While models using a higher number of harmonics are the most successful, successful classification also depends on sufficient image resolution and outline points to represent crater boundaries. While images of most paterae measured had high resolutions, the circularity of the unnamed ionian patera at 11.23°N, 84.74°W could be the result of experimental error due to insufficient image resolution. As a general rule, digitization of crater rims in this study was completed with images of sufficient resolution to resolve the complete morphologic boundary and all features were normalized to a total of 99 points. Thus, the magnitude of the deviations from circular represent real variations in landform morphology and are not thought to be artifacts of measurement.

While the shapes of highly circular paterae are well-correlated with impact-produced craters, other mechanisms common in volcanic settings, such as pit crater formation, could explain these features. They are all found at high latitudes, where the crust has been interpreted to be thicker compared with the equatorial regions that undergo the greatest tidal heating (Radebaugh et al., 2001). Thus, crustal collapse...
may only occur as simpler, large-scale features, rather than piecemeal as may be more likely to occur in regions with thinner, weaker crust. The interpretation of the origin of these circular paterae requires a larger craterform dataset to support causal inference, or additional data, such as higher-resolution images or in-situ observations. This would greatly enhance the capabilities of shape-based classifiers to interpret underlying processes and variables. A dedicated mission to Io, such as the Io Volcano Observer (IVO; McEwen, 2019), that would collect topographic data and higher resolution images would further our understanding of how and why patera shapes are most similar with terrestrial ash-flow calderas and what implications this could have for existing models of terrestrial volcanism.

8. Conclusion

This study introduces a multidisciplinary approach to the quantification of planetary landforms that combines methods from systematic biology and geomorphology, two fields of the natural sciences that rely substantially on form to interpret processes of origin and evolution. Outline-based approaches to geomorphologic analysis are supported by the classifiers shape quantities using canonical variate analysis (CVA) of these shape quantities using canonical variate analysis (CVA) easily distinguishes lunar impact craters from the other craterforms examined. The shapes of ash-flow calderas on Earth and paterae on Io differ from all other groups in similar ways, and terrestrial basaltic shield calderas and martian calderas differ the least from each other of all crater types examined.

This study successfully demonstrates the application of outline-based shape analysis in the classification of planetary craterforms and establishes that statistical classification of geomorphological features using outline information can enhance our understanding of the relationship between geologic form and process. The Z-R model is robust in its ability to classify end-member differences in crater complexity while the EFA model is more robust in its ability to correctly classify end-member differences in crater complexity.

The interpretation of planetary form and process. The Z-R model is robust in its ability to classify end-member differences in crater complexity while the EFA model is more robust in its ability to correctly classify end-member differences in crater complexity.

This methodology will enhance our ability to understand how different planetary landforms and evolve over time and throughout the solar system.

Author Contributors

Thomas J. Slezak: Conceptualization, Methodology, Software, Data curation, Visualization, Formal analysis, Investigation, Writing – Original Draft. Jani Radebaugh: Writing – Reviewing and Editing, Methodology, Visualization, Formal analysis, Data Curation. Eric H Christiansen: Writing – Reviewing and Editing, Methodology, Visualization, Formal analysis, Data Curation. Mark C. Belk: Methodology, Visualization, Formal analysis.

Data availability

Datasets related to this article can be found at doi:10.17632/hcmrk2sf2xg.1, an open-source repository hosted at Mendeley Data (Slezak et al., 2020).

Declaration of Competing Interest

None.

References


Christensen, P., et al., 2009. JAMS–A Planetary GIS. AGU Fall Meeting Abstracts ads:2009AGUFM122A.06C.


