



Using Elliptical Fourier Descriptor Analysis (EFDA) to Quantify Titan Lake Morphology

Rajani D. Dhingra¹, Jason W. Barnes¹, Matthew M. Hedman¹, and Jani Radebaugh²
¹Dept. of Physics, University of Idaho, Moscow, ID, USA; rhapsodyraj@gmail.com, rdhingra@uidaho.edu
²Department of Geological Sciences, Brigham Young University, Provo, UT, USA

Received 2019 February 1; revised 2019 September 26; accepted 2019 September 28; published 2019 November 18

Abstract

We use the elliptical Fourier descriptor analysis (EFDA) to quantify the shapes of Titan's lakes to technically demonstrate the use of this methodology in planetary morphometry. We map the lakes on Titan's north pole and find that the equivalent radii of 224 lakes follow a relatively narrow log-normal distribution like Earth's thermokarst lakes and Io's volcanic paterae, indicating a limited number of formation processes. Then, we quantify the shapes of lakes using EFDA. The Fourier analysis decomposes the shape of a lake into multiple Fourier series, and the corresponding coefficients represent a fingerprint of the lake shape. After testing the methodology on synthetic lakes and two kinds of terrestrial lakes, we analyze 67 Titan lake shapes on the north pole of Titan. We find that the majority of shape variation in Titan's lakes is from circular to elliptical followed by lakes with significant asymmetries along their short axis and long axis. We also find that a few lakes on Titan like Myvatn, Xolotlan, Sotonera, Viedma, Muggel, and Neagh Lacus have very distinctive shapes. Letas Lacus is an extreme outlier among the shapes of Titan lakes with an intruding island. This demonstration shows the promise of the elliptical Fourier descriptor approach for testing hypotheses for Titan lake formation. Our statistical analysis divides the Titan north polar lakes into four clean shape-based groups hereby indicating possible four formation mechanisms or four stages of formation of Titan's lakes. Uneven subsequent modification of the lakes could be another reason for the differences, which might be a result of different ages.

Key words: Planetary polar regions

1. Introduction

Titan is the only moon in our solar system with a thick atmosphere (Sagan & Dermott 1982; Lunine et al. 1983; Squyres et al. 1984; Baines et al. 2005; Niemann et al. 2005; Griffith 2009) and stable bodies of surface liquids owing to the methane-based hydrological cycle (Elachi et al. 2005, 2006; Porco et al. 2005; Tomasko et al. 2005). Most of the liquids on Titan are located on the north pole in the form of three large seas (Kraken Mare, Ligeia Mare and Punga Mare) and circular or irregularly shaped lakes. While the presence of a hydrological cycle explains how the depressions are filled with liquid methane, the formation mechanisms of the depressions still remain a mystery.

Post-Cassini, we have gathered some understanding of the liquids on Titan (Burr et al. 2009; Wall et al. 2010; Hofgartner et al. 2014) via the lake and sea compositions (Brown et al. 2008; Le Gall et al. 2016), depths (Mastrogiuseppe et al. 2014), formation mechanism hypotheses (Stofan et al. 2007), sediments left behind after evaporation (MacKenzie et al. 2014; MacKenzie & Barnes 2016), and the geographic asymmetry in their distribution (Dhingra et al. 2018; Tokano 2019). Birch et al. (2017) carried out the most complete geomorphological mapping of Titan's poles and suggest that the present-day landscape may be an erosional remnant that is being lowered in elevation through time. Hayes et al. (2017) used Titan's most complete topographic map generated by Corlies et al. (2017) and conclude that, similar to Earth, Titan's largest seas and lakes have a common equipotential surface probably connected through an aquifer or subsurface hydrocarbon reservoirs.

The leading models that explain the formation of these depressions are impact cratering (Stofan et al. 2007), volcanism, karstic processes, and sublimation or dissolution of a volatile substrate (Hayes 2016; Birch et al. 2017). No single

lake formation mechanism can completely address the challenge of explaining the variedly differing lake shapes on Titan.

Lopes et al. (2007) attribute volcanic calderas and cryovolcanism as probable lake formation mechanisms due to the circular shape of several north polar lakes. However, many irregularly shaped lakes at Titan's north pole do not reflect the caldera hypothesis. While liquid-filled calderas are common on the Earth, no interior model for Titan has yet explained the latitudinal dependence and poleward location of caldera-producing volcanic processes. Impact craters similarly do not address the preferential density of impacts on the pole.

Many researchers have put forward karst dissolution as a possible explanation for the irregularly shaped depressions. On Earth, karstic lakes are formed due to the dissolution of carbonates by water. Growing evidences also suggest that an organic sedimentary bedrock on Earth, like fossiliferous limestone and coal, could also likely undergo karstic dissolution (Ford & Williams 2007). On Titan, water ice—the bedrock—is not readily soluble in liquid methane or ethane. In addition, if we consider that the limnological timescales on Titan are longer (Tokano 2009) than on Earth and the constant solid hydrocarbon particle rain from the atmosphere is the solute (Atreya et al. 2006), the topography and depth of these lakes still do not compare with that of terrestrial karst lakes that are shallower than the Titan lakes (Mastrogiuseppe et al. 2014). The sharp rims around some of the Titan lakes seem constructional in morphology and hence imply a constructional mechanism of lake formation.

Cornet et al. (2015) show that in polar regions, karst topography and resultant lakes could form due to dissolution processes and would be 30 times slower on Titan than on Earth due to the seasonality of precipitation (Turtle et al. 2011, 2009; Dhingra et al. 2019). They estimate the timescales of the

development of the lacustrine depressions as few as tens of millions of years at polar latitudes higher than 70° north and south while at the drier latitudes it could take hundreds of millions of years.

While karst lakes result from the dissolution of a soluble rock in a solvent, thermokarst lakes result from the thawing of permafrost regions in terrestrial Arctics (for example: Alaska and Canada). CH_4 and C_2H_6 permafrost are unlikely to occur on Titan's surface due to freezing point depression from N_2 dissolution (Hanley et al. 2017). Although, Furfaro et al. (2010) do claim that “goo-permafrost” and “goo-glaciers” could occur and could be composed of a thick layer (~ 100 m) of acetylene or ethylene. This goo-permafrost requires mean Titan-annual temperatures below the freezing point to extend a continuum between glacier and permafrost terrains, like on Earth, suggesting collapsed karst or thermokarst as potential lake formation mechanisms on Titan.

Pelletier (2005) proposes a new explanation for the orientation, shape, and speed of growth of terrestrial oriented thaw lakes. The lakes' unusual characteristics could result from seasonal slumping of the banks when the permafrost thaws abruptly. The lakes grow when rapid warming melts a lake's frozen bank, then the soggy soil loses its strength and slides into the water. Such lakes are found in the permafrost zone in Alaska, northern Canada, and northern Russia.

Sharp-edged depressions (SEDs)—lakes with raised rims of heights ~ 100 m—present the most baffling case in the already puzzling challenge of lake formation mechanisms. Hayes (2016) and Hayes et al. (2017) discuss the complexity of fitting any formation mechanism to the SEDs. The favored hypothesis for the formation of lakes to explain the SEDs is the sublimation or dissolution of a volatile substrate (karst-like mechanism on Earth); however, this does not necessarily explain the formation of rims. Rims argue back for an explosive cryovolcanic origin (Hayes 2016).

An underlying assumption that we make on Earth and extraterrestrial bodies when we observe features is that form follows process. Qualitative differences in form can be identified using observations. For example, simple craters on the Moon have bowl-shaped depressions and have diameters of less than ~ 15 km. Complex craters have more complicated forms, diameters larger than ~ 15 km, shallow, relatively flat floors, central uplifts, and slump blocks and terraces on the inner wall of the crater rim (Neukum et al. 2001). Relatively complex forms like the Ionian paterae or the Titanian lakes are difficult to qualitatively differentiate using mere visual information.

Quantifying shapes as multivariate descriptors to extract the information stored in the shape of a structure is commonly used in evolutionary biology and paleontology (Elewa 2004; Smith & Turner 2005). However, such methods of boundary morphometrics have yet to be applied to the analysis of forms in planetary studies. Using modern methods of multivariate statistics we can quantitatively classify geologic morphology and compare form among other variables. As the limnological studies of terrestrial lakes indicate, similar lake shapes might indicate similar formation mechanisms. In this work we use outline-based shape analysis or morphometry—the measurement of morphological characteristics of a geologic feature—to quantify the lake shapes and provide constraints on lake commonalities on Titan.

There have been a few previous studies quantifying the outline shapes and morphometry of Titanian lakes. Sharma & Byrne (2011) carried out an in-depth study comparing 114

terrestrial lakes with 190 Titan lakes using fractal analysis. None of the terrestrial lake formation mechanisms conformed to Titan lake outlines by fractal analysis, according to their study. However, collapsed karst or thermokarst lakes—lakes formed by subsumption of the frozen ice—were excluded from the Sharma & Byrne (2011) analysis. An important limitation of fractal analysis is the empirically determined fractal dimension that depends on the image resolution.

Another study carried out a shoreline analysis of Titan lakes and compared them with terrestrial lakes using ruler methodology. Their comparison of Titan lakes with Minnesotan, Siberian, and African Rift Valley lakes suggested that Titanian lakes are most similar to Siberian and Minnesotan lakes. The Siberian lake packets used for the analysis are formed by a thermokarst mechanism and the Minnesotan lakes by glacial recession. We found that the previous outline-based studies either excluded or partially included karst, thermokarst, or a combination of both as the possible processes in action on Titan for forming the lakes.

In this article, we aim to constrain the probable lake formation mechanisms on Titan using morphometrics. We explore two different types of morphometric analysis in this article, a more classical size and orientation analysis (Section 2) and a novel shape analysis using the elliptical Fourier descriptor analysis (EFDA) method (Section 3).

We then apply the EFDA method on lakes in Section 3.2. Sections 3.2.1 and 3.2.2 discuss the methodology on synthetic lakes and a suite of Earth lakes to validate our technique. We finally apply the EFDA method on Titan lakes in Section 4. Sections 4.1 and 4.1.1 discuss the observational data used for Titan lakes and the results obtained. Finally, we discuss our observations and results concluding the manuscript in Section 5.

2. Size Distribution and Orientation

Viktorov & Kapralova (2012) analyze the spatial regularity of the morphological structures associated with terrestrial thermokarst lakes. They find that on Earth, the diameters of thermokarst lakes follow a log-normal distribution, under the assumptions that the lakes form over small intervals of time and the spread in the sizes of lakes is small. We reviewed the literature for other planetary processes that follow log-normal distribution and find that the size distribution of Ionian paterae that have volcanic origin (Radebaugh 1999) also follow a log-normal distribution as well as solar system calderas (Radebaugh 1999). Intrigued by these distributions, the non-inclusion of thermokarst lakes in previous analyses and the abovementioned mathematical morphological studies carried out on terrestrial thermokarst lakes, we were motivated to determine the size distributions for Titan's lakes.

Oriented lakes are lakes that demonstrate a preferred long axis orientation. The orientation could result from substrate topography and structural controls like development along joints or folds.

Sections 2.1 and 2.2 discuss the observational data, methodology, and results for the analysis of the size distribution and orientation of the lakes.

2.1. Observational Data

We use RADAR (Elachi et al. 2005), the Visual and Infrared Mapping Spectrometer (VIMS; Brown et al. 2004), and the Imaging Science Subsystem (ISS; Porco et al. 2004) data derived from multiple Cassini flybys to generate a comprehensive north

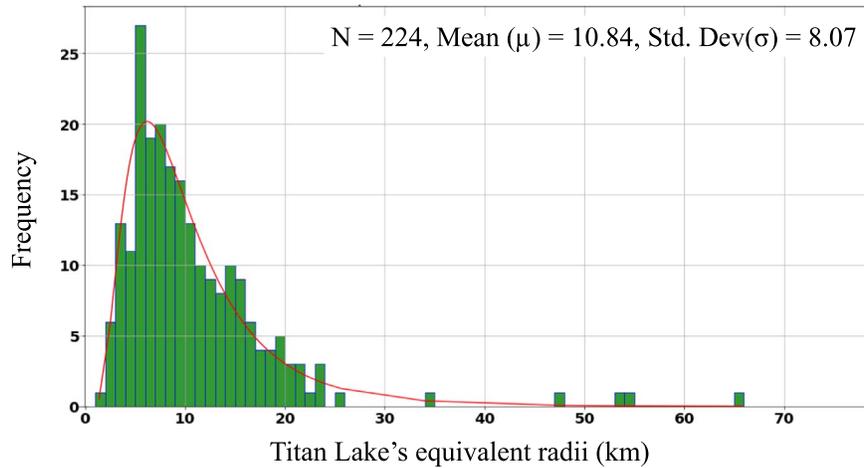


Figure 1. Size distribution of the equivalent radii ($\sqrt{\text{Area}/\pi}$) of the north polar Titan lakes. The size distribution (green) of the equivalent radii of lakes on the north pole of Titan follow a log-normal distribution (red). Terrestrially, the diameter of thermokarst lakes follow a log-normal distribution (Viktorov 1998; Viktorov & Kapralova 2012). The volcanic paterae on Io (Radebaugh 1999) also follow log-normal distribution.

polar map. We map the north polar lakes in the Geographic Information System (ArcGIS) using all the three data sets to cover the entire area. We use ArcGIS to retrieve morphometric information from the geographically projected digitized lake outlines of the 224 Titanian lakes that we analyze for size and orientation on the north pole. These are the lakes observed as filled lakes in the RADAR, ISS, and VIMS data sets.

2.2. Method and Results

We trace the lakes on the north pole of Titan using varied data sets (RADAR, VIMS, and ISS), allowing for an accurate calculation of lake areas and resulting effective radii ($\sqrt{\text{Area}/\pi}$). This measurement reveals the high degree of sinuosity and irregularity in Titan lakes' outlines. We then plot the equivalent radii of the lakes in a histogram, fit it with log normal distribution, and statistically test the fit.

As mentioned before, terrestrial thermokarst lake diameters and calderas follow a log-normal distribution. The log-normality motivated us to plot the equivalent radii of Titan's lakes in a histogram and fit it with a log-normal distribution ($N = 224$, mean = 10.84 km, standard deviation = 8.07 km) using the Anderson Darling test. Figure 1 shows the distribution of equivalent radii of 224 Titanian lakes mapped using the different data sets (ISS, VIMS, and RADAR). We do a statistical fit (red line in Figure 1) to this distribution and validate the log-normal fit of the distribution. The Anderson Darling test statistic for this fit is 0.34, which is less than the critical value (0.64) at a significance level of 10. Thus the distribution follows a log-normal distribution indicating a good fit to the distribution of Titan lake's equivalent radii.

Previously Hayes et al. (2008) divided Titan lakes into dark (filled), bright (empty), and granular (mid-way) lakes and also found that all three lake areas were distributed log-normally. To contrast the Titan lake distribution with terrestrial lakes, Cael & Seekell (2016) find that Earth's lake areas are power-law distributed—many small lakes, but few large lakes—for lakes $\geq 8.75 \text{ km}^2$.

Figure 2(A) shows the RADAR image of lakes on the north pole of Titan and adjacently (B) shows lake orientation in a polar histogram. We define the orientation angle as the angle between the line joining the antipodal points for the longest axis of the lake and the north pole. We use this morphometric measurement and

plot it in a rose diagram to verify if the north polar lakes demonstrate any preferred orientation. We can see that while the north polar lakes do not necessarily have a preferred orientation visually (in A) and in the plot (in B), there is a slight increase in the frequency of orientation in the northeast/southwest directions. We carry out the Rayleigh test for circular uniformity and find that the p -value is less than 0.01, thus we reject the null hypothesis that there is uniformity in lake orientation. We conclude that non-uniformity in the lake orientation does exist. The mountain ridge belts, the oriented valleys in Xanadu, and more crucially, at the north pole, and the really straight river valleys emptying into Ligeia Mare are a few of the pieces of evidence of the tectonic processes on Titan (Cook-Hallett et al. 2015). Our rose diagram and the slight increase in the frequency of orientation in the northeast/southwest directions are further evidence of probable tectonic processes at the north pole. The tectonic lake formation mechanisms that do have a preferred long axis orientation seem to be slightly plausible in this scenario.

3. Elliptical Fourier Descriptor Analysis for Outline-shape Quantification

Fourier-based approaches are powerful tools to extract the geometry information from the outline shapes. They work on the basis of Fourier series—decomposing a complex periodic function in terms of simple trigonometric functions like sine and cosine. The simple trigonometric functions have frequencies that are integer multiples or harmonics of one another. Lower order harmonics help explain the coarse features in the outline while higher order harmonics are required to explain the fine scale sinuosities in the outline. The lake outlines are periodic functions in the sense that if we start traversing a closed outline we will cross a reference point repeatedly periodically, thus making a closed outline a periodic function (Bonhomme et al. 2014). Hence we use a Fourier-series-based approach to quantify the geometry of the closed outlines of Titan lakes (Figure 3(A)).

The EFDA has many advantages over the other Fourier-based approaches and tends to be a reliable method to quantify shapes because of the following reasons.

1. Equally spaced points are not required.
2. Any shape can be accurately represented, even ones that fold back on themselves.

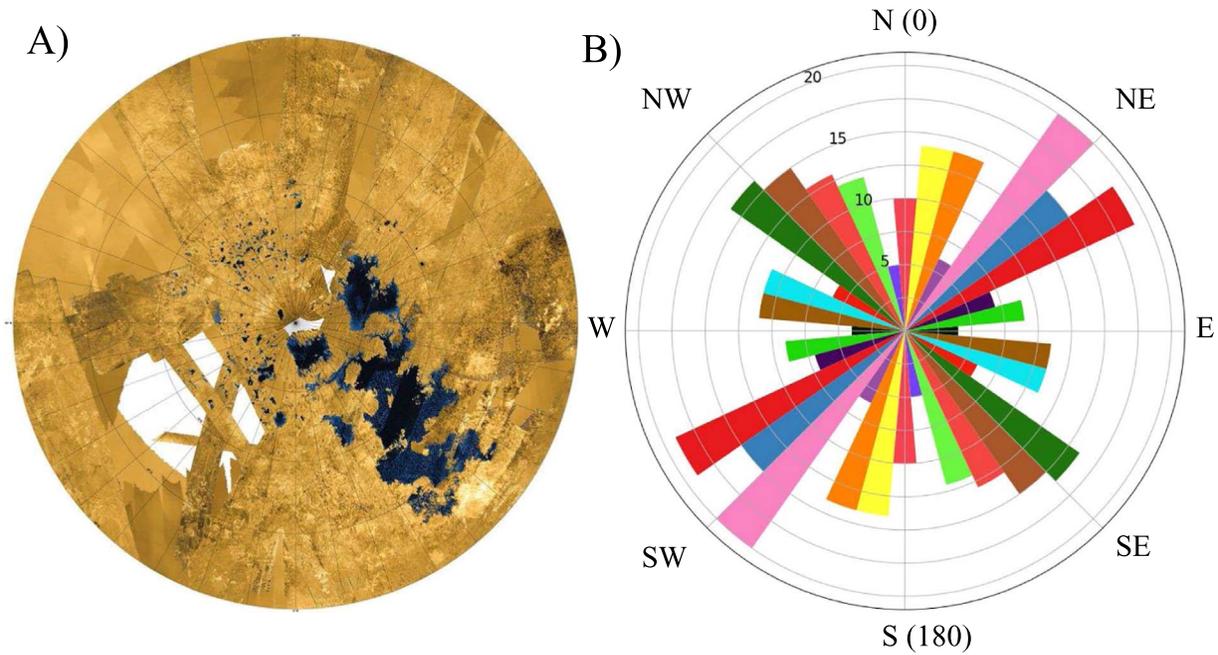


Figure 2. (A) RADAR (false colored) image of Titan’s north pole. (B) Rose diagram indicating the orientation of the Titan lakes (angle between the line joining the antipodal pair of a lake and the north pole). The orientation of lakes show a northeast–southwest or (45°W – 225°W) direction in B. The colors in this plot are just the bar colors for ease of viewing.

3. The outline can be quantified without having a need for a synchronous point (twig on a leaf or the thumb on a hand for comparing leaves’ or hands’ outlines) on every shape.
4. The Fourier coefficients can be made independent of outline position and normalized for size and rotation.

Section 3.1 describes in detail the methodology for EFDA and its advantages as a method to quantify and compare the outlines of lake shapes.

3.1. Method

EFDA (Giardina & Kuhl 1977; Kuhl & Giardina 1982) is a method that fits the x and y coordinates of an outline separately. Morphometric analyses need x , y coordinates sampled on each outline as an input. So our first step is to extract the x and y coordinates of the lake outlines. We use the tpsUtil (Rohlf 2015) software to convert our images (.jpgs/.pngs) to .tps files that can be handled by the tpsDig software for placing landmarks on the outline. In order to reduce inconsistencies we follow the same set of rules for digitization of every lake.

1. Last point overlaps the first point making every outline a closed loop.
2. All points are laid clockwise.
3. To decrease the angularity in digitization and obliterate the need for higher frequency (harmonics) to fit in the outline, we heavily oversample the edge position.
4. For Titan lakes: RADAR’s darkest pixels are only considered in the lake outline unless other data sets (ISS, VIMS) show differently.
5. For Titan lakes: if there is a granular shelf-like region in the RADAR data, it is not included in the lake outline.
6. For Titan lakes: lakes like Myvatn Lacus, Abaya Lacus, and Ranoch Lacus with prominent extensions are all considered one lake.

We then use the x and y coordinates generated by tpsDig into a R-based package called MOMOCS (modern morphometrics; Bonhomme et al. 2014). The coordinate information is used to regenerate the outline shape. This code determines a series of harmonic coefficients, A_n , B_n , C_n , and D_n , with which these positions can be written as: $x_n = \sum_{n=1}^N A_n \cos(nt) + B_n \sin(nt)$

$$y_n = \sum_{n=1}^N C_n \cos(nt) + D_n \sin(nt)$$

where, n = harmonic amplitude.

N = maximum no of harmonic amplitudes used in the construction.

t = evaluation angle (varies from 0 to 2π).

Higher order harmonics (larger n) better reproduce the fine details in the shape, but also have less power associated with them and very minutely improve the shape of a feature. We estimate the number of harmonics required to best fit the shape (after examining the spectrum of harmonic Fourier power). In this analysis we selected the number of harmonics to be used, so that their cumulative power gathers $\sim 95\%$ – 97% of the total cumulative power, thereby representing the shape of the lake by $\sim 95\%$ – 97% . Once the right number of harmonics is determined, we perform the elliptic Fourier analysis.

Figure 3 explains the big picture EFDA methodology in panel A as a general concept. The first harmonic is shown as a red ellipse while the tenth harmonic can be seen fitting the sinuosities of the leaf and hence its outline shape. In panel (B), we show the EFDA methodology working for one of the most peculiarly shaped lakes, Myvatn Lacus near Titan’s north pole. Myvatn Lacus can be coarsely regenerated by the ninth harmonic and is faithfully matched by the twentieth harmonic. We then extract the Fourier coefficient for each harmonic for every shape and normalize the coefficients for the lake shape’s size and rotation that affect the results.

While the above harmonics clearly describe the shape of any given lake, their values also depend on the overall size and

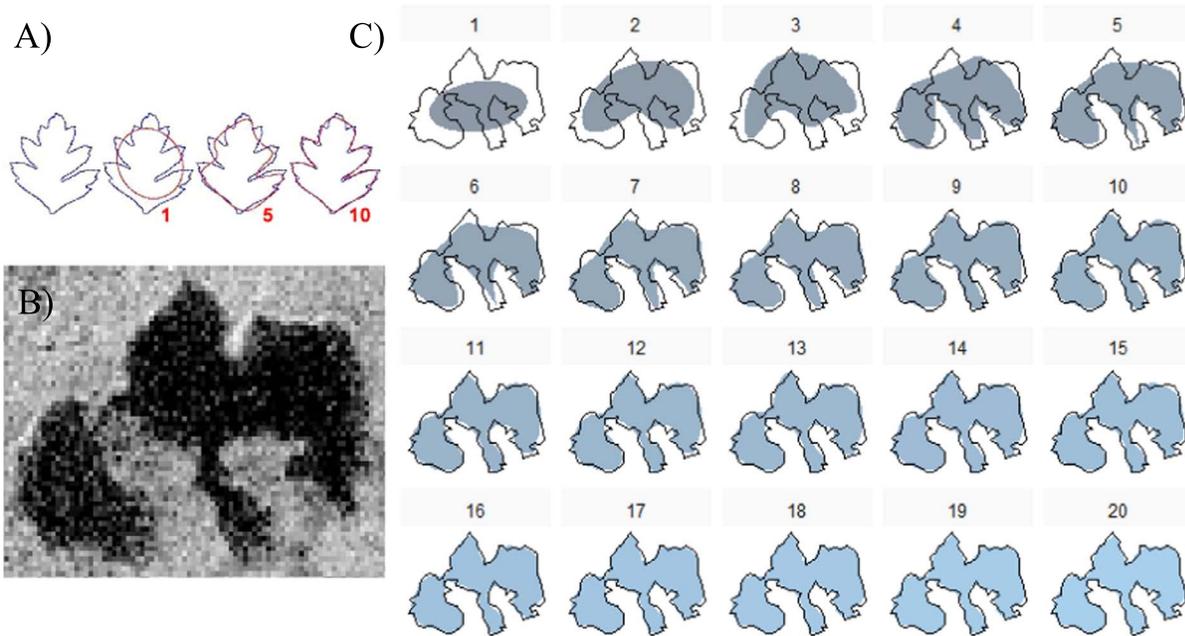


Figure 3. (A) (modified from Wikipedia) shows the methodology of elliptical Fourier analysis. Leaf outline is shown in blue. The red ellipse over the blue leaf outline is the first harmonic. The fifth and tenth harmonics regenerate the sinuosities of leaf’s edge and shape. (B) RADAR image of Myvatn Lacus (78°N, 135°W) (peculiarly shaped). Figure (C) shows that by twentieth harmonic the extremely complicated shape of even Myvatn Lacus can be explained quantitatively by the elliptical Fourier analysis. After examining the spectrum of harmonic Fourier power we select the number of harmonics to be used, so that their cumulative power gathers ~95%–97% of the total cumulative power, thereby representing the shape of the lakes by ~95%–97%.

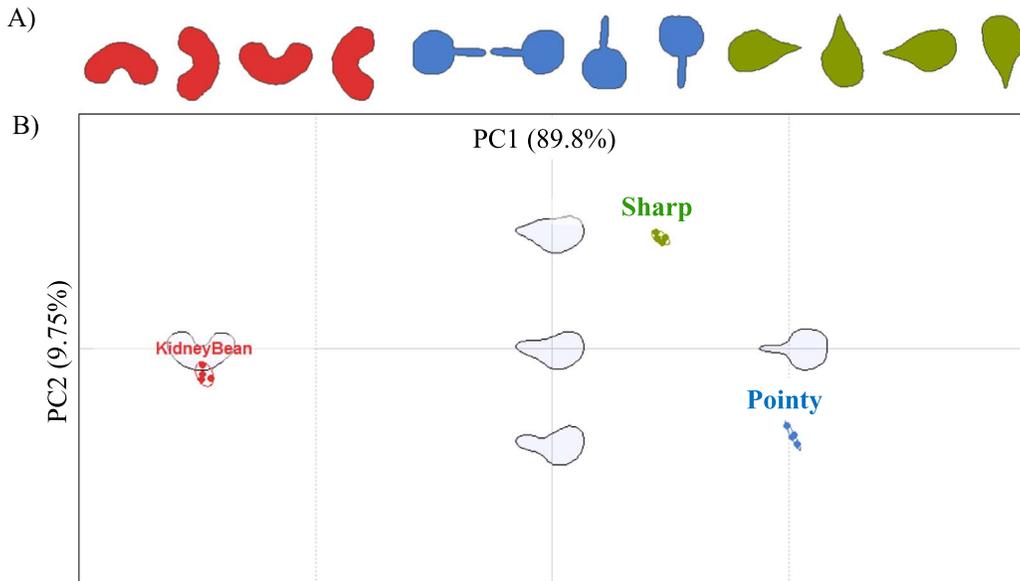


Figure 4. (A) Synthetic lake shapes we use for analysis. The kidney bean lake shape is in red, pointy lake shape is in blue, and sharp lake shape is in green. (B) Principal component plot with the geomorphological shapes. The shaded gray shapes in the morphological space are the shape variations along the principal component axes. PC1 explains 90% of the variation in shape while PC2 explains ~10% of the variation in shape. The remaining 0.455% of the variance is in the higher principal components. The outer four figures represent extrema of the principal components and are located at places where the normalized principal components are ± 1 . The red dots indicate the four kidney beans with different rotation angles (90°, 180°, 270°, 360°). The ellipse around the red points represents the 95% confidence interval. Similarly the green and blue dots indicate the pointy lakes and sharp lakes with the ellipses representing the 95% confidence interval.

orientation of the lake. Hence, in order to compare the shapes of different lakes, the harmonic coefficients need to be properly normalized to remove their dependence on these parameters. The standard procedure for normalizing the coefficients to remove dependence on size and orientation is given in Bonhomme et al. (2014) and uses the $n = 1$ harmonics. However, this procedure assumes there is a synchronous point for the profiles, which is not the case for Titan’s lakes.

When we apply the methodology to lakes, however, a rotation issue arises. With the standard methods the same lake shape can be reoriented two different ways that differ by 180° in rotation, which get misinterpreted as two very different shapes if the lake has a prominent feature on one side (see Figure 4(A))

To understand this issue of rotation we referred back to the original methodology manuscript (Kuhl & Giardina 1982).

The Fourier coefficients A_n , B_n , C_n , and D_n of the Fourier approximation to a closed contour are used as the classification of the contour. Since the Fourier coefficients vary according to the starting point of a trace of the contour and the spatial rotation, magnitude, and translation of the contour, self-consistent normalization procedures based only on the intrinsic shape properties of the contour are important. The rotating phasors provide normalization modes easily when the locus of the first harmonic phasor is elliptic, yielding two related classifications corresponding to the positions at either end of the major axis of the ellipse.

The contour classification associated with one semimajor axis obtained through starting point and spatial angle rotations of θ_1 and ψ_1 radians, respectively, (where $\theta_1 = 2\pi\lambda_1/T$ and λ_1 is the displacement of the starting point) is different from the contour classification of other semimajor axis obtained by a further rotation of both the starting point and spatial angles through π radians. For both the contour classifications, the odd harmonics remain the same for all n , but the even harmonics change sign.

We took this into account by first identifying the harmonic coefficient that had the maximum loading. We find that the C_2 coefficient has the maximum loading in most cases. To account for the rotation discrepancy, we multiply all the even harmonics of the outlines with negative C_2 coefficients by -1 .

After we make the changes in our algorithm to take care of the rotation we find that similar lake shapes/outlines irrespective of however they are rotated group together as shown in Section 3.2.

The Fourier coefficients can be used to carry out statistical multivariate analysis. In this article we do principal component analysis, hierarchical clustering analysis, and k -means clustering on the Fourier coefficients and these three methods are explained below.

Principal component analysis (PCA) is the simplest of the multivariate analyses. The observations are orthogonally transformed in such a way that the first principal component accounts for much of the variability in the data (Abdi & Williams 2010). Each succeeding component has the highest variance possible under the constraint that it is orthogonal to the preceding components. It can be used to collapse the data to fewer dimensions that explain the variance in the data.

We use hierarchical clustering as another statistical method to see what lake shapes cluster together, how they stand with respect to the PCA, and human visual semantics. Hierarchical clustering is a statistical methodology that groups similar data into clusters using hierarchy. The metric for deciding the similarity between two lake shapes in order to be clustered together or measure the dissimilarity between two clusters is usually the distance between vectors of Fourier coefficients (Rokach & Maimon 2005). The result is a group of clusters with objects in one cluster being similar to each other and each cluster being uniquely distinct.

The k -means clustering mechanism is another statistical methodology to group the data set into a user-defined number (k) of clusters. The data cluster into k numbers of clusters even if the k is not the right number of clusters for the data. In order to address this issue, we use the elbow method to determine the optimal number of clusters (MacQueen et al. 1967; Bradley & Fayyad 1998). The elbow method runs k -means clustering on the data set for a range of values of k (say, k from 1 to 10), and for each value of k calculates the sum of squared errors. If the

plot of the sum of squared errors for each value of k resembles an arm, then the elbow of the arm is the value of k that suits the data set the best. As we increase the number of k , the sum of squared errors approaches zero because each data point in the data set gets its own cluster and hence there is no error between the data point and center of cluster. We want a small value of k that still has a low sum of squared errors and the elbow approximately represents the point where increasing k has less returns (Ketchen & Shook 1996).

3.2. Testing EFDA on Lake Forms

Conventional morphometric approaches have been applied to symmetric lunar craterforms (Eppler et al. 1977) to study the circularity and axial ratio of craters. However, the use of morphometrics to asymmetric, complex, and irregular natural forms have been scarcely used in planetary research. We apply the Fourier-based approach to Titan's lakes to assess if distinctions in their morphology can be classified using shape information alone. An additional complexity in lake shape analysis is the absence of a synchronous point. However, the strength of the Fourier analysis lies in the methodology that it can be applied to practically any kind of outline that the myriad shapes of Titan lakes exemplify the best.

3.2.1. Synthetic Lakes

We run our analysis methodology on synthetic lakes. Figure 4 shows the three different shapes we use in our methodology testing (kidney bean (red), sharp (green), and pointy (blue)). We observe that in the morphological space in Figure 4, all similar lakes group together irrespective of the rotation. Similar synthetic lake outlines also clump together in three clean clusters (kidney bean, pointy, and sharp) in our hierarchical clustering and k -means cluster analysis. With that proof of technique and validation of our new normalization scheme we progress into performing the analysis on some real lakes.

3.2.2. Earth Lakes

Next, we carry out the analysis on terrestrial lakes. We pick two distinct types of terrestrial lakes to see if our methodology can separate their shapes and hence provide insight into their formation mechanism. We use two types of lakes whose origins as volcanic or tectonic are already known before our analysis. We select them specifically knowing their formation mechanism so that we could test the formation mechanism against their shapes. The volcanic lakes are fairly circular in shape while the tectonic lakes are fairly elongated in shape. Although volcanic lakes can be separated from tectonic lakes intuitively (visually), we intend to test how our analysis works on separating them. We use the lakes as tabulated in the book Gierlowski-Kordesch (2004). The EFDA methodology can quantitatively separate the two types of shapes indicating two plausibly distinct formation mechanisms.

Figure 5, panel (A) shows the terrestrial lakes on which we chose to run EFDA. Panel (B) shows their location in the principal component plot. We can clearly separate volcanic lakes from tectonic lakes in the morphological space. However, lakes like Lake Tahoe, which is tectonic in origin yet more circular than the other tectonic lakes can be seen in the circular space in the principal component plot.

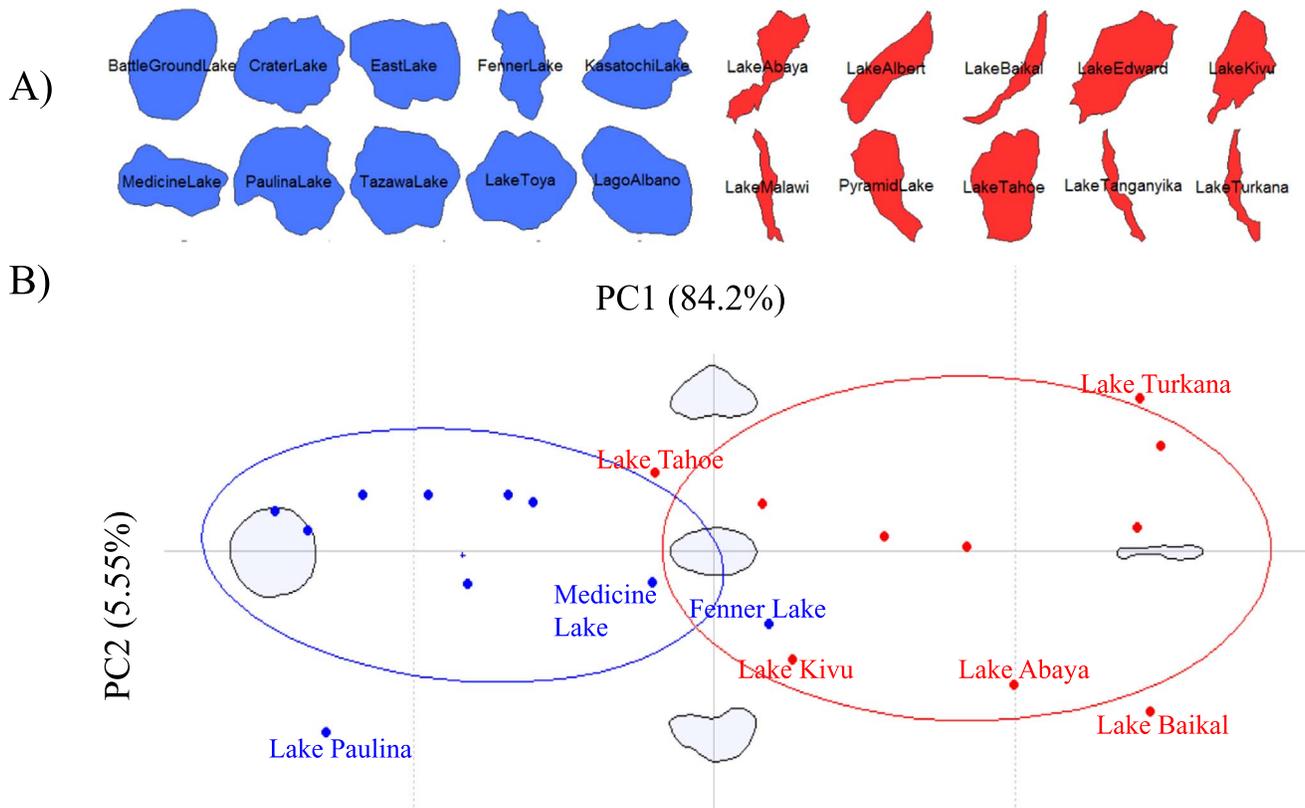


Figure 5. (A) Earth lakes we use for analysis. The lake shapes blue in color are classified as volcanic lakes. Similarly the lake shapes in red are classified as tectonic lakes. Since the shapes of these two kinds of lakes are discretely different we chose these two lake types to see if they separate in the principal component space. The shaded colors indicate the group that the lakes belong to according to our k -means cluster analysis. (B) Principal component plot with the end member shapes. PC1 explains 84.2% of the variation in shape while PC2 explains $\sim 5.55\%$ of the variation in shape. We can see that principal component 1 explains the variation of lake shapes from circular to elliptical (fat to fit). Principal component 2 explains the curvature in the shape. The red dots indicate the distribution of the tectonic lakes. As expected they clump more toward the skinny or elliptical side of shape variation in the principal component space. Similarly the blue dots represent the terrestrial volcanic lakes (more circular in shape). These cluster toward the circular end of the principal component plot. The ellipse shows the 95% confidence region for the mean of the lake shape parameters.

4. Application to Titan Lakes

We use the methodology described in Section 3.2.2 to quantify the outlines of the 67 IAU named north polar lakes of Titan. We exclude the larger bodies of fluids (Kraken, Ligeia, Punga, Jingpo Lacus) present on the north pole because the connections between them complicate the identification of discrete lakes. The top panel of Figure 6 shows the outlines of the 67 lakes from the north pole of Titan that we use for our analysis. We omitted lakes that have diameters less than 2 km.

4.1. Observational Data and Method

For the shape analysis we only consider the IAU named 67 lakes on the north pole of Titan. In order to determine the errors on the outline determinations, we randomly chose five lakes (Feia Lacus, Oneida Lacus, Rukwa Lacus, Sparrow Lacus, and Vanern Lacus) that we digitize before starting the digitization of other lakes. The outlines of these lake and their distributions in the morphometric space indicates the error of outlining the lakes. The error here refers to repeatability of the mapping of the lake boundary on different days by the same researcher. The lakes on which we carry out the shape analysis are shown in Figure 6. We digitize the 67 lakes for shape analysis in the software, tpsDig. The lake outlines with their x and y coordinates in the text files are then used in a R package called MOMOCS (modern morphometrics) for the EFDA as explained in the Section 3.1.

4.1.1. Results

Once we have the outlines of the lakes and the Fourier coefficients extracted, we statistically analyze the Fourier coefficients in order to decipher if there is any pattern in the lakes' Fourier coefficients. The left principal component plot in Figure 6 shows the variation of principal component 1 w.r.t 2 in morphological space. The first principal component varies from a \sim circular to elliptical shape and explains $\sim 30\%$ of the variation in the shapes. This indicates that the majority of lake shapes on Titan's north pole vary from circular to elliptical. The circular end member is best represented by Mweru Lacus and Quilotoa Lacus. Hlawga Lacus and Roca Lacus have the maximum ellipticity and represent the elliptical end member on the principal component 1 (PC1) axis.

The second principal component explains $\sim 15\%$ of the variation in the lake shapes. Positive principal component 2 (PC2) corresponds to lakes with strong asymmetries along their short axis, while those with strong negative values of PC2 have strong asymmetries along their long axis. While Muggel Lacus represents the asymmetric lake end member on negative PC2, Xolotlan Lacus represents the end member for the (asymmetric) lake with a curvature on the positive PC2.

The outliers—Myvatn Lacus, Rwegura Lacus, Sotonera Lacus, Muggel Lacus, and Feia Lacus—explain that those lake shapes are peculiarly odd and rare on Titan's surface. The principal component 3 (PC3) represents an increasing strong kidney-bean-like shape. Indeed, lakes with strong negative

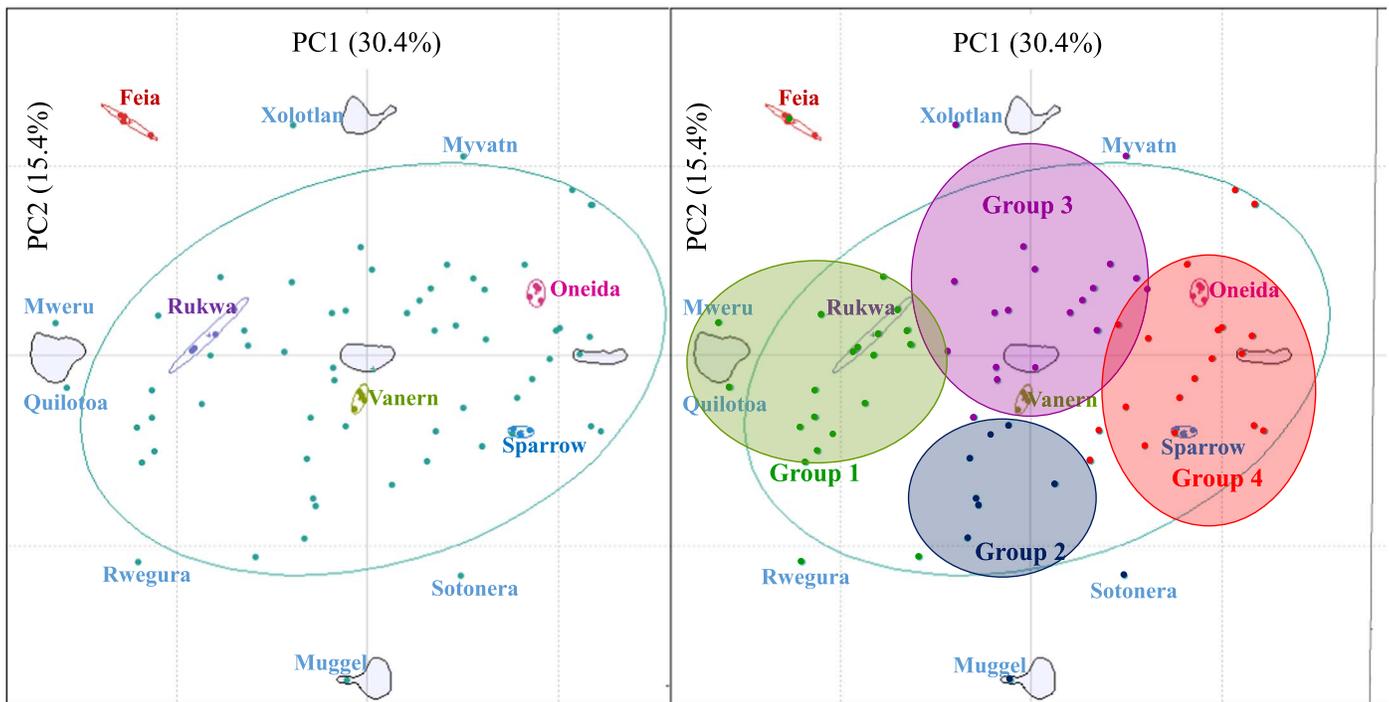
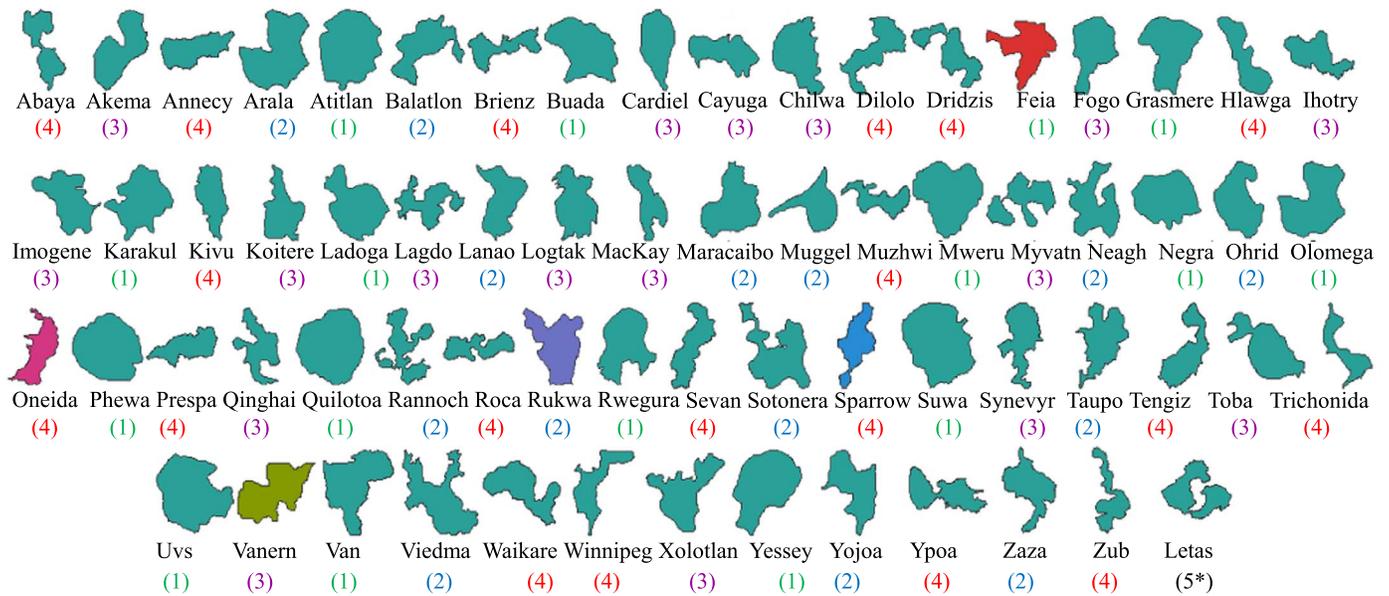


Figure 6. Top panel represents the lake shapes with their respective names on the north pole of Titan. The numbers in the brackets indicate the group (using our *k*-means cluster analysis) that the lake belongs to. Bottom left panel indicates the first two principal component’s plot for Titan lakes (from the top panel). We randomly chose five standard lakes (Feia Lacus, Oneida Lacus, Rukwa Lacus, Sparrow Lacus, and Vanern Lacus) that we digitized everyday before starting the digitization of other lakes. The cloud of spread of the standard lakes in the principal component space is the error in the digitization of other lakes. PC1 shows the shape variation from circular to elliptical and explains 30.4% of the variation in the shapes of Titan lakes considered in this study. PC2 shows the shape variation from pointy extensions to a curvature in the pointy extensions and explains 15.6% of the variation in the data. The ellipse shows the 95% confidence region for the mean of the Titan lake shape parameters. Bottom right panel shows the four groups of lakes that we derived from the *k*-means analysis on the principal component plot. As can be seen, several lake shapes, e.g., Muggel Lacus, Myvatn Lacus, Xolotlan Lacus, Mweru Lacus, Viedma Lacus, and Quilotoa, are outliers—different from 95% of the other lake shapes on Titan.

values of PC3 (Sotonera Lacus and Neagh Lacus) often have broad peninsulas jutting into them. We have not included Letas Lacus in this analysis. Letas Lacus has an atypical island that gives it an appearance of a two lobed lake, which is prominently different from all the other lakes on Titan’s north

pole. Since it is difficult to clump lakes into separate groups according to the PCA in the geomorphic space we use hierarchical and *k*-means cluster analyses (Hartigan & Wong 1979).

Figure 7 shows the results of the hierarchical clustering of lake shapes based on the euclidean distance between clusters.

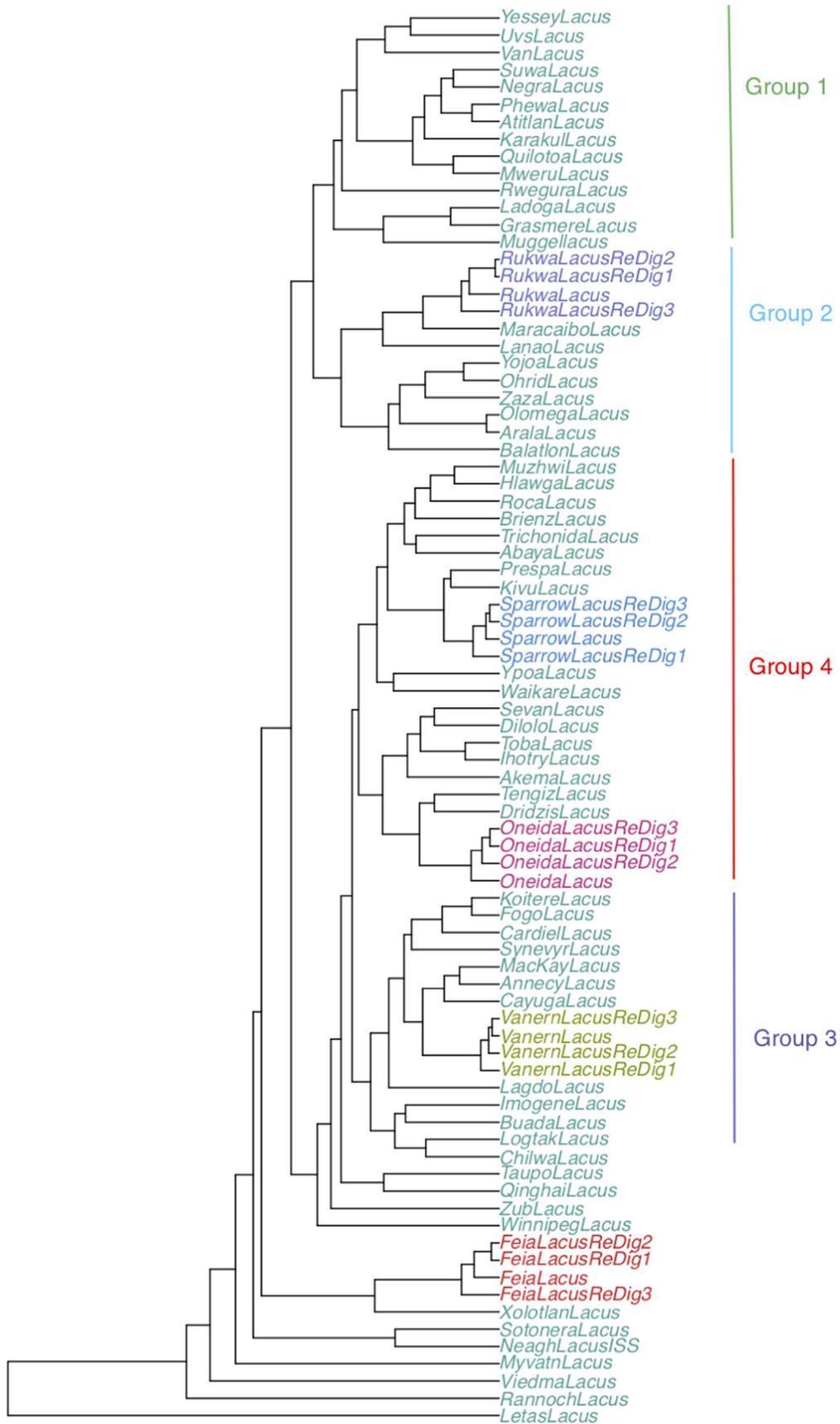


Figure 7. Phylogram for the hierarchical clustering of Titan north polar lakes. Letas Lacus exists by itself like in the PCA. The same lakes re-outlined to determine the error are all together in the same clusters. Other clusters of lakes when compared with their shapes can be seen clustering well in similar shapes. The hierarchical cluster helps us group similar shapes together thereby indicating similar formation mechanisms. The groups 1 to 4 marked in Figure 7 are not the exact same clusters as in Figure 8. They do overlap a lot, yet are not exactly same.

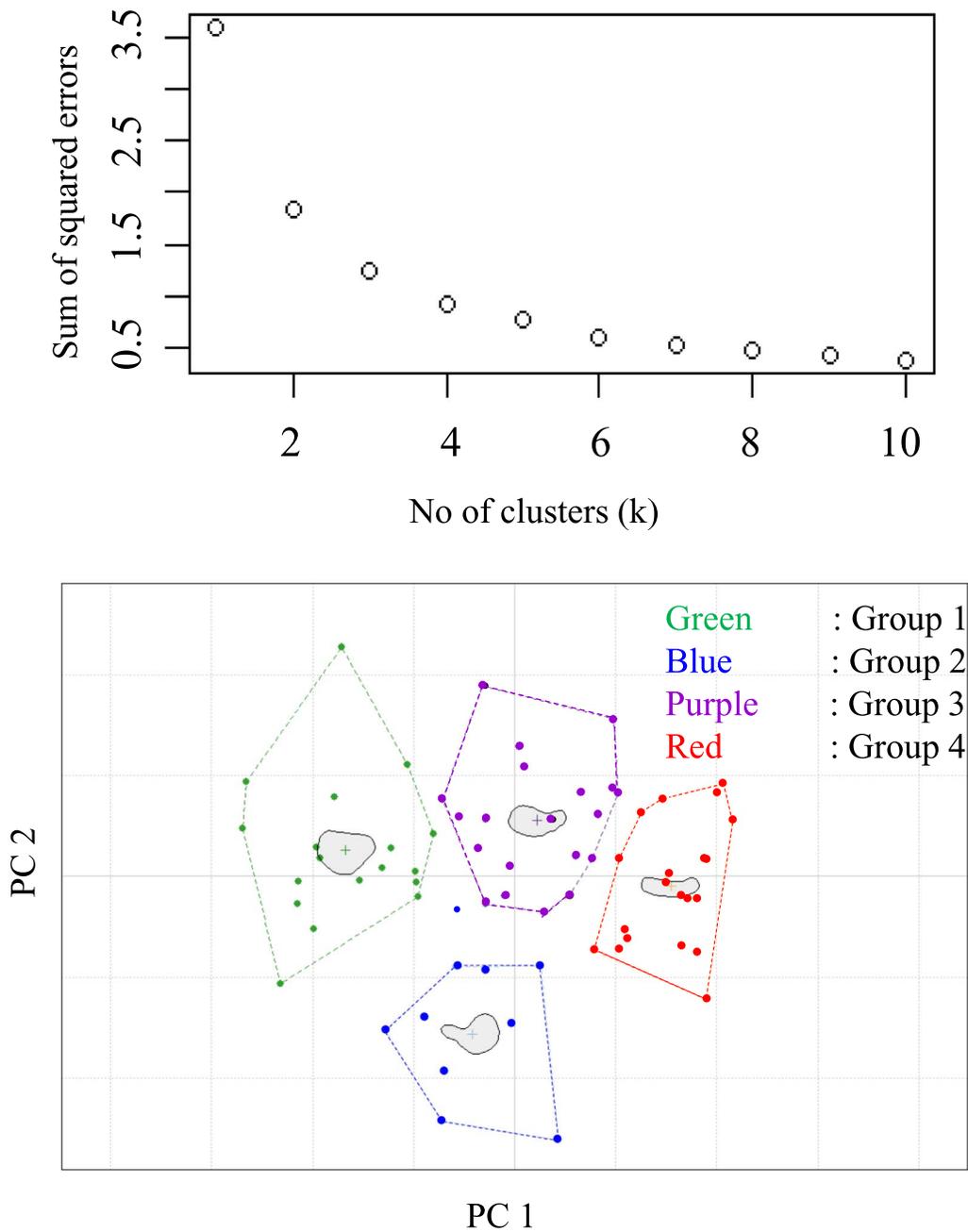


Figure 8. Upper panel shows the elbow method for the determination of number of clusters in the k -means clustering analysis. The elbow of the arm roughly indicates the right number of clusters to be used. Bottom panel shows the four clusters of lake shapes in the PCA space. The distribution of points in 6B and 8B do not match exactly because this clustering analysis excludes the multiple realizations of selected lakes which causes slight shifts in the principal components that change the exact positions of the points in the map, but do not affect the groupings significantly. Also note that in this plot, unlike in Figure 6(B) the illustrated shapes are the average lake outlines for the four clusters from our K -means analysis and are located in the middle of each group.

The clustering tree is generated by finding all possible pairwise distances for points belonging to two different clusters and then calculating the average, also called average linkage clustering. We can see that the repeated measurements of our random lakes that we use to determine the approximate errors of digitization are all clustered together, lending confidence in the repeatability of our measurements. The clustering tree indicates four major clusters indicating four probable formation or evolution mechanisms.

Treading up the cluster tree, Rannoch, Viedma, and Myvatn Lacus all have high sinuosities in their shorelines and are stand-alone cases as clusters of their own. Neagh Lacus and Sotonera

Lacus in fact look like twin lakes, both having distinctive peninsulas. Overall big picture lake clusters are as expressed by the PCA, circular-ish (group 1)—the upper limb of the cluster tree (including Yessey, Uvs, Van, Suwa, Negra, Phewa, Atitlan, Karakul, Quilotoa, and Mweru) to elliptical (group 4)—the middle limb of the cluster tree (including Sparrow, Ypoa, Waikare, Sevan, and Dilolo).

In our k -means analysis, the optimum number of clusters were determined by the elbow method to be four (Figure 8, top panel). Similar to our earlier results Letas Lacus indicates its own cluster. Removing Letas Lacus from the analysis results in four clean clusters as shown in Figure 7. We also show in

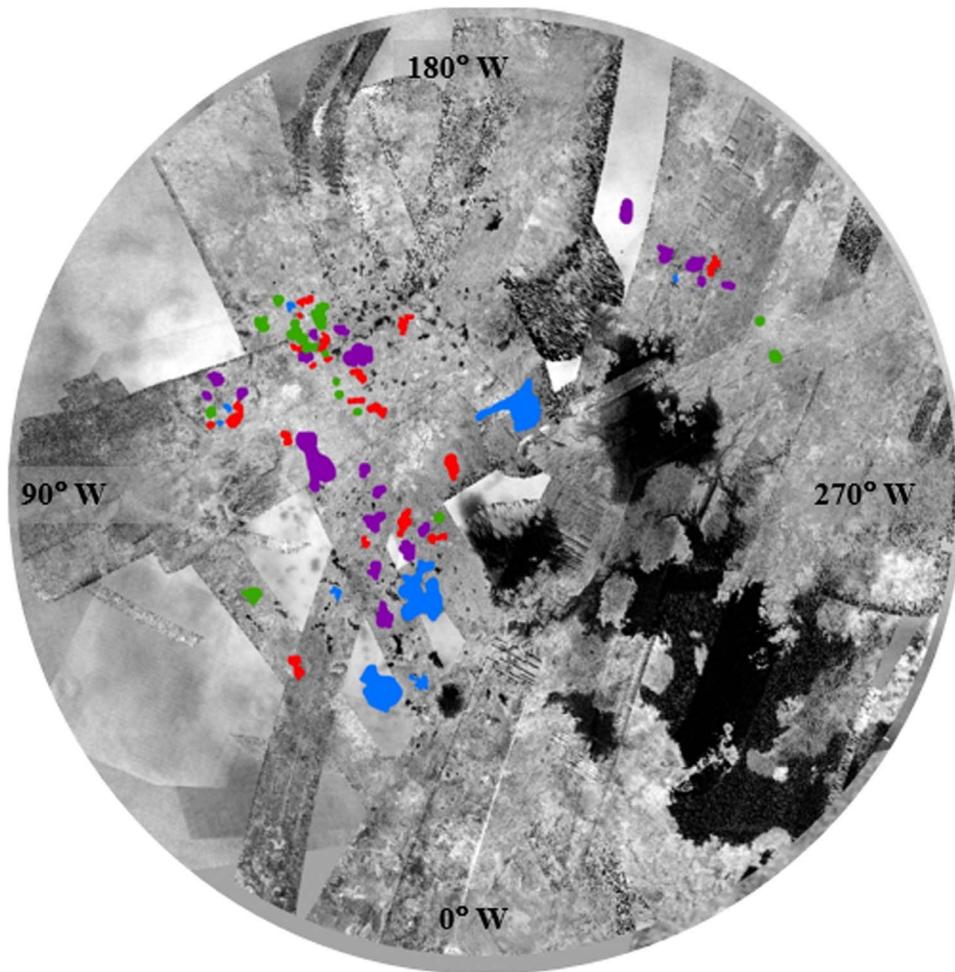


Figure 9. Geographic location of the four clusters of lake shapes (derived from k -means clustering) on the north pole of Titan. The base images are ISS and RADAR north polar images of Titan. The colors are group 1: round, e.g., Yessey Lacus; group 2: long with asymmetries along their long axis, e.g., Rukwa Lacus; group 3: long with asymmetries along their short axis, e.g., Vanem Lacus, and group 4: long lakes, e.g., Oneida Lacus.

brackets in Figure 6 (top panel) the group (using our k -means cluster analysis) that the lake belongs to. In the bottom panel of Figure 6 we show in the principal component plot the groups to which the lakes belong in shaded ellipses. The distribution of points in 6B and 8B do not match exactly because this clustering analysis excludes the multiple realizations of selected lakes, which causes slight shifts in the principal components that change the exact positions of the points in the map, but do not affect the groupings significantly. Also note that the illustrated shapes in these two figures show different things. In Figure 6(B) the shape outlines are end members located at the end of the principal component axis. By contrast, in Figure 8(B) the shapes show the average lake outlines for the four clusters derived from our K -means analysis and are located within each group, rather than as end members on the principal component axis in contrast to Figure 6(B).

Finally, in Figure 9 we plot the geographic location of the four groups of lakes (in their respective colors) on the north polar map of Titan to understand if there is any geographic similarity in groups' locations. Our visual examinations suggest that most of the round lakes in group 1 seem to be in the upper left quadrant (180°W to 90°W). The long (group 4) lakes might have a preferred orientation (more seem to be up-down (line joining 0°W to 180°W) than left-right (line joining 90°W to

270°W)). These are preliminary observations that can be verified by further analysis of more lakes in a future work.

The k -means and hierarchical clustering analysis give roughly four groups:

- group 1: round lakes;
- group 2: moderately long lakes with a strong asymmetry along their long axis;
- group 3: moderately long lakes with a strong asymmetry across their short axis; and
- group 4: long lakes.

5. Discussion and Conclusion

We show that the EFDA methodology for outline shape analysis is a robust way to decipher differences and similarities in lake outlines. We validate through our synthetic lake example that the size and rotation-normalized algorithm helps in clustering similarly shaped outlines together. The Earth lake examples of clustering of Type I and Type II lakes bolster the algorithm's strength.

Since the algorithm can be used on any shape without a synchronous point, it is very useful for shapes like lakes or other planetary geomorphic landforms which do not necessarily have a synchronous point.

Our statistical analyses of the Fourier coefficients demarcate the Titan lake shapes into four groups. Our methodology determines that the maximum variation of shape on Titan lakes are from circular to elliptical. The second component shows that lake shapes vary from moderately long lakes with a strong asymmetry along their long axis to moderately long lakes with a strong asymmetry across their short axis. We also find that Letas Lacus is a total outlier among the (presently analyzed) Titanian lakes. This brings into focus the lake shapes that are dissimilar from the other lakes on Titan directing us to another challenge of how these outliers were formed or evolved into present-day lakes.

Terrestrial lakes show similar formation mechanisms based on their geographic location. The Siberian lake packets are formed by thermokarst processes and lakes in Minnesota are formed due to glacial recession. Similarly the lakes in the East African Rift are mostly tectonic (Morley et al. 1990). In our analysis we see that the round and elliptical lakes cluster in the upper left quadrant (180°W to 90°W) of the north pole. We also see that the group 2 and 3 lakes cluster along a line near the borders of the bigger seas/lakes.

According to our size analysis Titan lakes follow log-normality like the terrestrial thermokarst lakes. However, temperature differences to sustain thermokarst mechanisms on Titan are improbable according to our current understanding. Thermokarst can be intriguing as a putative lake formation hypothesis but posits the weak thermal insolation on Titan's north pole as an issue. Kapralova (2007) in their model assume that the terrestrial thermokarst lakes form over a small interval of time. We cannot assume this for Titan's lakes owing to the 30 times slower dissolution rates on Titan (Cornet et al. 2015). However, a combination of karst and thermokarst processes cannot be ruled out wherein the dissolution of a permafrost equivalent with the hydrocarbon combinations raining from the atmosphere is highly probable.

Seasonal thawing of permafrost with the fluctuation of temperatures leading to the formation of lakes in a few freeze and thaw cycles could be a possibility. Yet Cassini surface temperatures from 2004–2014 decrease by only 2000 for southern hemisphere and increase by 1000 in northern hemisphere (Jennings et al. 2016) making an active permafrost thawing an unlikely mechanism. Titan lakes also show a slight preferred orientation for a northeast–southwest or 45°W – 225°W direction, indicating possible tectonic lake formation mechanisms.

We observe a size relationship between the groups in Figure 9. Group 2 is associated with big lakes as opposed to group 1 which is more roundish lakes. A common feature of Titan's lakes is that the smaller lakes are mostly circular with simple outlines while the larger ones are of myriad shapes with sinuous outlines. This could be telling us that the initial lake formation mechanism could be something simple but as soon as the lake starts to grow there must be different processes developing to cause the sinuosities and growth.

A permafrost thawing lake formation mechanism may only work on Titan if the temperature of hydrocarbons that make up the permafrost is below the mean surface temperature under the consideration that the freezing point is depressed by dissolved nitrogen. Also, the seasonal temperature variation range must be sufficient for the thawing of the permafrost.

An alternative theory of surface-ice cracking where smaller evaporation puddles formed from seasonal precipitation likely leave a bottom ice layer that sublimates after the liquid puddle evaporates (evidenced by the equatorial brightening events; Barnes et al. 2013) is plausible. Such a process of seasonal yet cyclic evaporation–sublimation over millions of years will likely have an erosional impact on the immediate water-ice bedrock that could carve out and deepen a surface depression.

Our present study considers the IAU named lakes on Titan. We wish to expand the lake population to consider all the lakes present on Titan's north pole and the south pole. We intend to carry out the same analysis over the other larger seas to understand how or where they lie in the formation mechanism groups. We also intend to use the equatorial Hotei and Tui Regio apart from the south polar paleo basins to understand how the filled lakes in the north pole contrast the paleo-lakes and basins. Apart from this, we would like to test the tectonic lake formation hypothesis by comparing our orientations with the straight rivers in the region. A statistical size correlation with the geographic location is under consideration once we have the analysis for all the lakes on the north pole.

R.D. acknowledges Dr. Sarah Jacobs for sharing her experience in using geomorphology for biology and helping R.D. get started with the EFDA. R.D. is grateful to the anonymous reviewer whose comments markedly helped the manuscript improve.

ORCID iDs

Rajani D. Dhingra  <https://orcid.org/0000-0002-3520-7381>
 Jason W. Barnes  <https://orcid.org/0000-0002-7755-3530>
 Matthew M. Hedman  <https://orcid.org/0000-0002-8592-0812>

References

- Abdi, H., & Williams, L. J. 2010, *WIREs Comp. Stat.*, 2, 433
 Atreya, S. K., Adams, E. Y., Niemann, H. B., et al. 2006, *P&SS*, 54, 1177
 Baines, K. H., Drossart, P., Momary, T. W., et al. 2005, *EM&P*, 96, 119
 Barnes, J. W., Buratti, B. J., Turtle, E. P., et al. 2013, *PLSci*, 2, 1
 Birch, S. P. D., Hayes, A. G., Dietrich, W. E., et al. 2017, *Icar*, 282, 214
 Bonhomme, V., Picq, S., Gaucherel, C., et al. 2014, *J. of Stat. Softw.*, 56, 1
 Bradley, P. S., & Fayyad, U. M. 1998, in Proc. Fifteenth Int. Conf. on Machine Learning, ICML 98, ed. J. W. Shavlik (San Francisco, CA: Morgan Kaufmann), 91
 Brown, R. H., Baines, K. H., Bellucci, G., et al. 2004, *SSRv*, 115, 111
 Brown, R. H., Soderblom, L. A., Soderblom, J. M., et al. 2008, *Natur*, 454, 607
 Burr, D. M., Jacobsen, R. E., Roth, D. L., et al. 2009, *GeoRL*, 36, L22203
 Cael, B. B., & Seekell, D. A. 2016, *NatSR*, 6, 29633
 Cook-Hallett, C., Barnes, J. W., Kattenhorn, S. A., et al. 2015, *JGRE*, 120, 1220
 Corlies, P., Hayes, A., Birch, S., et al. 2017, *GeoRL*, 44, 11754
 Cornet, T., Cordier, D., Bahers, T. L., et al. 2015, *JGRE*, 120, 1044
 Dhingra, R. D., Barnes, J. W., Brown, R. H., et al. 2019, *GeoRL*, 46, 1205
 Dhingra, R. D., Barnes, J. W., Yanites, B. J., & Kirk, R. L. 2018, *Icar*, 299, 331
 Elachi, C., Wall, S., Allison, M., et al. 2005, *Sci*, 308, 970
 Elachi, C., Wall, S., Janssen, M., et al. 2006, *Natur*, 441, 709
 Elewa, A. M. 2004, *Morphometrics: Applications in Biology and Paleontology*, Vol. 14 (Berlin: Springer)
 Eppler, D. T., Nummedal, D., & Ehrlich, R. 1977, Impact and explosion cratering: Planetary and Terrestrial Implications, ed. D. J. Roddy, R. O. Pepin, & R. B. Merrill, (New York: Pergamon Pres), 511
 Ford, D. C., & Williams, P. W. 2007, *Karst Hydrogeology and Geomorphology*, Vol. 576 (New York: Wiley)
 Furfaro, R., Kargel, J. S., Lunine, J. I., Fink, W., & Bishop, M. P. 2010, *P&SS*, 58, 761
 Giardino, C. R., & Kuhl, F. P. 1977, *Comput. Graph. Image Process.*, 6, 277

- Gierlowski-Kordesch, E. 2004, *PALAIOS*, 19, 184
- Griffith, C. A. 2009, *RSPTA*, 367, 713
- Hanley, J., Grundy, W. M., Thompson, G., et al. 2017, AGU Fall Meeting, 2017, P13D-2590
- Hartigan, J. A., & Wong, M. A. 1979, *J. Royal Stat. Soc. Ser. C (Appl. Stat.)*, 28, 100
- Hayes, A., Aharonson, O., Callahan, P., et al. 2008, *GeoRL*, 35, L9204
- Hayes, A. G. 2016, *AREPS*, 44, 57
- Hayes, A. G., Birch, S. P. D., Dietrich, W. E., et al. 2017, *GeoRL*, 44, 11745
- Hofgartner, J. D., Hayes, A. G., Lunine, J. I., et al. 2014, *NatGe*, 7, 493
- Jennings, D., Cottini, V., Nixon, C., et al. 2016, *ApJL*, 816, L17
- Kapralova, V. 2007, *Landform Analysis*, 5, 35
- Ketchen, D. J., & Shook, C. L. 1996, *Strateg. Manage. J.*, 17, 441
- Kuhl, F. P., & Giardina, C. R. 1982, *Comput. Graph. Image Process.*, 18, 236
- Le Gall, A., Malaska, M. J., Lorenz, R. D., et al. 2016, *JGRE*, 121, 233
- Lopes, R. M. C., Mitchell, K. L., Stofan, E. R., et al. 2007, *Icar*, 186, 395
- Lunine, J. I., Stevenson, D. J., & Yung, Y. L. 1983, *Sci*, 222, 1229
- MacKenzie, S. M., & Barnes, J. W. 2016, *ApJ*, 821, 17
- MacKenzie, S. M., Barnes, J. W., Sotin, C., et al. 2014, *Icar*, 243, 191
- MacQueen, J., et al. 1967, in *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob.*, Vol. 1, ed. L. M. Le Cam & J. Neyman (Berkeley, CA: Univ. California Press), 281
- Mastrogiuseppe, M., Poggiali, V., Hayes, A., et al. 2014, *GeoRL*, 41, 1432
- Morley, C., Nelson, R., Patton, T., & Munn, S. 1990, *AAPG Bulletin*, 74, 1234
- Neukum, G., Ivanov, B. A., & Hartmann, W. K. 2001, in *Chronology and Evolution of Mars*, Vol. 12, ed. R. Kallenbach, J. Geiss, & W. K. Hartmann (Dordrecht: Springer), 55
- Niemann, H. B., Atreya, S. K., Bauer, S. J., et al. 2005, *Natur*, 438, 779
- Pelletier, J. D. 2005, *JGRF*, 110, F02018
- Porco, C. C., Baker, E., Barbara, J., et al. 2005, *Natur*, 434, 159
- Porco, C. C., West, R. A., Squyres, S., et al. 2004, *SSRv*, 115, 363
- Radebaugh, J. 1999, PhD thesis, Brigham Young Univ. Department of Geology
- Rohlf, F. J. 2015, *Hystrix It. J. Mamm.*, 26, 9
- Rokach, L., & Maimon, O. 2005, in *Data Mining and Knowledge Discovery Handbook*, ed. O. Maimon & L. Rokach (Boston, MA: Springer), 321
- Sagan, C., & Dermott, S. F. 1982, *Natur*, 300, 731
- Sharma, P., & Byrne, S. 2011, *GeoRL*, 38, L24203
- Smith, N. D., & Turner, A. H. 2005, *Systematic Biology*, 54, 166
- Squyres, S. W., Thompson, W. R., & Sagan, C. 1984, *BAAS*, 16, 664
- Stofan, E. R., Elachi, C., Lunine, J. I., et al. 2007, *Natur*, 445, 61
- Tokano, T. 2009, *AsBio*, 9, 147
- Tokano, T. 2019, *Icar*, 317, 337
- Tomasko, M. G., Archinal, B., Becker, T., et al. 2005, *Natur*, 438, 765
- Turtle, E. P., Perry, J. E., Hayes, A. G., et al. 2011, *Sci*, 331, 1414
- Turtle, E. P., Perry, J. E., McEwen, A. S., et al. 2009, *GeoRL*, 36, L2204
- Viktorov, A. 1998, *Mathematical Morphology of Landscape* (Moscow: Tratek)
- Viktorov, A. S., & Kapralova, V. N. 2012, *Water Resources*, 39, 790
- Wall, S., Hayes, A., Bristow, C., et al. 2010, *GeoRL*, 37, L5202